UNIVERSITY OF CALIFORNIA,
IRVINE

Increasing Adoption of Deep Learning Models in Medicine and Circadian Omic Analyses through Interpretability and Data Availability

DISSERTATION

submitted in partial satisfaction of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

in Computer Science

by

Muntaha Samad

Dissertation Committee:
Professor Pierre Baldi, Chair
Professor Marco Levorato
Doctor Maxime Cannesson

2023

DEDICATION

To my wonderful parents for their unconditional love and support.
To my older siblings for setting a perfect example for me to follow.
To my fiance for adding joy and laughter into every day.

TABLE OF CONTENTS

LIST OF FIGURES

Page

LIST OF TABLES

ACKNOWLEDGMENTS

I would like to thank my advisor Pierre Baldi for his guidance and support.

I would like to thank my labmates for their support and friendship.

I wish to acknowledge publishers of my previous publications for permission to incorporate the material for the sole purpose of writing this thesis. Portions of Chapter 2 were previously published as "Development and validation of an interpretable neural network for prediction of postoperative in-hospital mortality, January 2021. Christine Lee, Muntaha Samad, Ira Hofer, Maxime Cannesson, Pierre Baldi". Permissions to reuse the text were granted by NPJ digital medicine. Portions of Chapter 4 were previously published as "CircadiOmics: circadian omic web portal, Nucleic Acids Research, June 2022. Muntaha Samad, Forest Agostinelli, Tomoki Sato, Kohei Shimaji, Pierre Baldi". Permissions to reuse the text were granted by Nucleic Acids Research.

I wish to thank all co-authors of those publications in enabling the research which forms the basis for this dissertation.

I wish to acknowledge financial support from grant NIH GM123558, NSF NRT 1633631, and NIH R01-EB029751 issued to Pierre Baldi.

VITA

Muntaha Samad

EDUCATION

Doctor of Philosophy in Computer Science **2023**
University of California, Irvine *Irvine, CA*

Master of Science in Computer Science **2019**
University of California, Irvine *Irvine, CA*

Bachelor of Science in Computer Science and Engineering **2017**
University of California, Davis *Davis, CA*

RESEARCH EXPERIENCE

Graduate Research Assistant **2017–2022**
University of California, Irvine *Irvine, California*

TEACHING EXPERIENCE

Teaching Assistant **2017–2018**
University of California, Irvine *Irvine, California*

REFEREED JOURNAL PUBLICATIONS

The hidden link between circadian entropy and mental health disorders. **2022**
Translational psychiatry

Brain histone beta-hydroxy-butyrylation couples metabolism with gene expression. **2022**
bioRxiv

Tryptophan metabolism is a physiological integrator regulating circadian rhythms. **2022**
Molecular metabolism

Bioinformatics and systems biology of circadian rhythms: BIO_CYCLE and Circadiomics. **2022**
Circadian Regulation

Circadiomics: circadian omic web portal. **2022**
Nucleic Acids Research

Atlas of exercise metabolism reveals time-dependent signatures of metabolic homeostasis. **2022**
Cell Metabolism

Interrogating metabolic interactions between skeletal muscle and liver circadian clocks in vivo. **2022**
bioRxiv

Integration of feeding behavior by the liver circadian clock reveals network dependency of metabolic rhythms. **2021**
Science Advances

Exercise opens a 'molecular memory window' to facilitate changes in gene expression, synaptic plasticity and memory. **2021**
Neuropsychopharmacology

Development and validation of an interpretable neural network for prediction of postoperative in-hospital mortality. **2021**
NPJ digital medicine

S-adenosyl-l-homocysteine hydrolase links methionine metabolism to the circadian clock and chromatin remodeling. **2020**
Science Advances

Mir-29 coordinates age-dependent plasticity brakes in the adult visual cortex. **2020**
EMBO report

Fkbp10 regulates protein translation to sustain lung cancer growth. **2020**
Cell report

Deep learning analysis of vibrational spectra of bacterial lysate for rapid antimicrobial susceptibility testing. **2020**
ACS nano

Reshaping circadian metabolism in the suprachiasmatic nucleus and prefrontal cortex by nutritional challenge. **2020**
Proceedings of the National Academy of Sciences

Defining the independence of the liver circadian clock. **2019**
Cell

Time of exercise specifies the impact on muscle metabolic pathways and systemic energy homeostasis. **2019**
Cell metabolism

ABSTRACT OF THE DISSERTATION

Increasing Adoption of Deep Learning Models in Medicine and Circadian Omic Analyses through Interpretability and Data Availability

By

Muntaha Samad

Doctor of Philosophy in Computer Science

University of California, Irvine, 2023

Professor Pierre Baldi, Chair

There are numerous applications for deep learning in a healthcare setting including: providing more accurate diagnoses, recommending treatment plans, predicting patient outcomes, tracking patient engagement and adherence, and reducing the burden of administrative tasks. This plethora of applications has resulted in the widespread publication of deep learning algorithms applied to healthcare data. Despite numerous publications showing deep learning to be very successful in retrospective healthcare studies, very few of these algorithms are then actually incorporated into clinical practice. While there are many factors influencing the lack of algorithm deployment, one of the major reasons is a lack of trust in deep learning. This lack of trust stems in part from a lack of model interpretability and an inability to independently verify published results due to a lack of data availability. In this work, we explore generalized additive models with neural networks (GAM-NNs) as a method of improving model interpretability and we propose MOVER: Medical Informatics Operating Room Vitals and Events Repository a publically available repository of medical data designed to improve visibility into deep learning algorithms in healthcare.

Similarly, deep learning can be used to analyze circadian omic (e.g. transcriptomic, metabolomic, proteomic) time series data. Several studies have shown that a disruption to circadian

rhythms have been linked to health problems such as cancer, diabetes, obesity, and premature aging. In order to gain clinician trust in the conclusions drawn from circadian omic analyses we propose CircadiOmics: the largest annotated repository of circadian omic time series data analyzed using deep learning. Clinicians and researchers can use CircadiOmics to not only validate the findings of their circadian omic experiments, but also to analyze multiple circadian omic experiments in aggregate.

Chapter 1

Introduction

Deep learning models have tremendous potential to greatly improve patient care and outcomes and improve our understanding of circadian rhythms. Despite numerous publications showing deep learning to be very successful in retrospective healthcare studies, very few of these algorithms are then actually incorporated into clinical practice. Similarly, clinical adoption of the findings from circadian omic experiments analyzed using BIO-CYCLE can be an arduous process due to a lack of clinician trust in findings. Two factors contributing to a lack of trust in the output of deep learning models are: a lack of interpretability and a lack of publicly available data.

1.1 Increasing Adoption of Deep Learning Models in Medicine

In healthcare, deep neural networks (DNNs) and other machine learning models often have higher accuracy than simpler models like logistic regression (LR) however, they are often considered to be "black box" models and this lack of interpretability and transparency is

considered a challenge for clinical adoption. Additionally, the lack of trust in deep learning models stems in part from an inability to independently verify published results. Although the publication of AI algorithms in the healthcare space has become pervasive, the data used to develop these algorithms is often not published alongside the algorithm. This lack of visibility into the analyzed dataset prevents independent researchers not only from reproducing published results but also from improving upon existing algorithms and developing novel algorithms. In this work, we will demonstrate that our proposed generalized additive neural network (GAM-NN) architecture is able to (1) leverage a neural network's ability to learn nonlinear patterns in the data, which is more clinically intuitive, (2) be interpreted easily, making it more clinically useful, and (3) maintain model performance as compared to previously published DNNs. We also demonstrate the necessity for publicly available medical data using MOVER: Medical Informatics Operating Room Vitals and Events Repository.

1.2 Increasing Adoption of Deep Learning Models in Circadian Omic Analyses

The advance of modern high-throughput technologies has made it possible to investigate circadian rhythms on the molecular level. This increase in available circadian omic (e.g. transcriptomic, metabolomic, proteomic) time series data has made it possible to analyze circadian rhythms using deep learning. BIO-CYCLE is a deep-learning-based program that can analyze omic time series and statistically assess their periodic nature and, when periodic, accurately infer the corresponding period, amplitude, and phase. To increase the clinical adoption of the findings from circadian omic experiments analyzed using BIO-CYCLE, we propose CircadiOmics: the largest annotated repository of circadian omic time series processed using BIO-CYCLE. CircadiOmics contains over 290 experiments and 100 million individual measurements, across 20 unique tissues, and across 11 different species. Via

CircadiOmics, users can process their circadian omic data using BIO-CYCLE, verify their experimental results by comparing their results to other datasets found in the repository, and analyze all of the datasets in aggregate. CircadiOmics makes the BIO-CYCLE output easily interpretable and enables powerful bioinformatics and systems biology analyses.

Chapter 2

Development and Validation of an Interpretable Neural Network for Prediction of Postoperative in-hospital Mortality

2.1 Introduction

We and others have recently shown that deep neural networks (DNNs) and random forest algorithms, using only readily available information extracted from the electronic health record before or at the end of surgery, can successfully predict postoperative in-hospital mortality with area under the curve (AUC) ranging from 0.87 to 0.93 [1, 2, 3]. While DNNs and other machine learning models often have higher accuracy than simpler models like logistic regression (LR), they are often considered to be "black box" models and this lack of interpretability and transparency is considered a challenge for clinical adoption [4]. In

healthcare, intelligible models not only help clinicians to understand the problem and create more targeted action plans, but also help to gain the clinicians' trust. Thus, LR models remain popular in the healthcare space, as they are easily interpretable, robust, easy to implement, and usually have good performance, as previously observed in our work comparing DNNs to LR3 . However, LR can be limited by the fact that it is a shallow model with no ability to create new feature representations, such as with DNNs. An LR model can only combine the input features linearly before passing that combination through a logistic function, and this linear combination of features may not reflect clinical intuition. For example, both hypervolemia and hypovolemia have been shown to increase the risk of postoperative complications, reflecting a nonlinear relationship between a patient's volume status and the risk for complications [5]. Nonlinear relationships can be captured by LR, but only through extra featurization and analyses, which may result in an infinite number of possible relationships and combinations of features. While DNNs are capable of learning nonlinear relationships between features on their own, they lack the interpretability of LR. One method of overcoming the limited interpretability of more complex models is to use Generalized Additive Models (GAMs). Standard GAMs simply model the target response as a sum of univariate models. Caruana et al. demonstrated that GAMs which also included pairwise interactions of features could be applied to real healthcare problems such as pneumonia risk with interpretability and high accuracy [6]. Through a graphical representation of each model feature's learned contribution to the predicted risk, the interpretable GAMs help to visualize learned patterns and identify new patterns in the data or confirm what clinicians already know. Inspired by GAMs, the same idea can be applied to neural networks through an architecture referred to as Generalized Additive Models with Neural Networks (GAM-NNs) [7]. In GAMNNs, a network is built on top of each input feature (or each group of input features) and the output of these networks are linearly combined to produce the final regression or classification output. To incorporate a modest number of pairwise interactions, additional networks processing the corresponding pairs can also be

included. Pairing of features was not assessed in this study to avoid cluttering the final interpretation. Bras-Geraldes et al. showed GAM-NNs could be used to predict mortality in the ICU with an AUC of 0.83, using 19 features from vital signs, lab values, demographics, admission information, and comorbidities [8]. In short, models like DNNs allow for learning the more complex relationship between the input and class label. However, they are not as easily interpretable as LR. In this manuscript, we present the development and validation of a model applying the concept of GAM-NNs to allow for interpretability by visualizing the learned feature patterns related to risk of in-hospital mortality for patients undergoing surgery under general anesthesia.

2.2 Results

2.2.1 Patient Characteristics

The data consisted of 59,985 surgical records, and the percent of occurrence of in-hospital mortality was 0.81% (n = 389) in the training set and 0.72% (n = 87) in the testing set. Patient demographics and characteristics of the training and testing datasets are summarized in Table 2.1.

2.2.2 Development of the Model

The final hyperparameters for the GAM-NN model with Healthcare Cost and Utilization Project (HCUP) features consist of one hidden layer with 50 neurons hyperbolic tangent (tanh) activations (Table 2.2). The model was trained with dropout probability of 0.5 and L2 weight decay of 0.0001. The final hyperparameters for the GAMNN model without HCUP features were the same except for an L2 weight decay of 0.001.

	Train	Test
No. of patients	47,988	11,997
No. of patients with in-hospital mortality (%)	389 (0.81%)	87 (0.73%)
Age (years)	56 ± 17	56 ± 18
Estimated blood loss (cc)	95 ± 540	94 ± 410
Presence of arterial line (%)	8585 (17.9%)	2135 (18.0%)
Presence of pulmonary artery line (%)	1641 (3.4%)	430 (3.6%)
Presence of central line (%)	2444 (5.1%)	635 (5.3%)
ASA score (%)		
1	3023 (6.3%)	762 (6.4%)
2	17930 (37.4%)	4477 (37.3%)
3	23960 (49.9%)	5986 (49.9%)
4	2911 (6.1%)	735 (6.1%)
5	144 (0.3%)	30 (0.3%)
6	4 (0.01%)	0 (0%)
HCUP code description (%)		
UPPER GASTROINTESTINAL ENDOSCOPY, BIOPSY	3864 (8.05%)	965 (8%)
COLONOSCOPY AND BIOPSY	1693 (3.53%)	388 (3.2%)
LAMINECTOMY, EXCISION INTERVERTEBRAL DISC	1029 (2.14%)	287 (2.4%)
OTHER THERAPEUTIC PROCEDURES, HEMIC AND LYMPHATIC SYSTEM	1013 (2.11%)	247 (2.1%)
OTHER OR THERAPEUTIC PROCEDURES ON RESPIRATORY SYSTEM	985 (2.05%)	254 (2.1%)
INCISION AND EXCISION OF CNS	942 (1.96%)	255 (2.1%)
OTHER OR PROCEDURES ON VESSELS OTHER THAN HEAD AND NECK	932 (1.94%)	207 (1.7%)
OTHER THERAPEUTIC ENDOCRINE PROCEDURES	904 (1.88%)	258 (2.2%)
HIP REPLACEMENT, TOTAL AND PARTIAL	792 (1.65%)	186 (1.6%)
ARTHROPLASTY KNEE	768 (1.6%)	193 (1.6%)
OTHER OR THERAPEUTIC NERVOUS SYSTEM PROCEDURES	750 (1.56%)	181 (1.5%)
THYROIDECTOMY, PARTIAL OR COMPLETE	737 (1.54%)	172 (1.4%)
SPINAL FUSION	735 (1.53%)	150 (1.3%)
OTHER OR THERAPEUTIC PROCEDURES ON BONE	722 (1.5%)	195 (1.6%)
CONVERSION OF CARDIAC RHYTHM	720 (1.5%)	184 (1.5%)
HEART VALVE PROCEDURES	715 (1.49%)	186 (1.6%)
CHOLECYSTECTOMY AND COMMON DUCT EXPLORATION	700 (1.46%)	216 (1.8%)
ENDOSCOPIC RETROGRADE CANNULATION OF PANCREAS (ERCP)	663 (1.38%)	155 (1.3%)
KIDNEY TRANSPLANT	659 (1.37%)	194 (1.6%)
OTHER OR THERAPEUTIC PROCEDURES ON NOSE, MOUTH AND PHARYNX	653 (1.36%)	173 (1.4%)
OTHER HERNIA REPAIR	652 (1.36%)	178 (1.5%)
HYSTERECTOMY, ABDOMINAL AND VAGINAL	641 (1.34%)	155 (1.3%)
APPENDECTOMY	634 (1.32%)	147 (1.2%)
OTHER THERAPEUTIC PROCEDURES ON MUSCLES AND TENDONS	629 (1.31%)	154 (1.3%)
COLORECTAL RESECTION	609 (1.27%)	127 (1.1%)
INSERTION, REVISION, REPLACEMENT, REMOVAL OF CARDIAC PACEMAKER OR CARDIOVERTER/DEFIB RILLATOR	601 (1.25%)	128 (1.1%)
ABORTION (TERMINATION OF PREGNANCY)	587 (1.22%)	162 (1.4%)
TREATMENT, FRACTURE OR DISLOCATION OF HIP AND FEMUR	570 (1.19%)	155 (1.3%)
OTHER OR GASTROINTESTINAL THERAPEUTIC PROCEDURES	569 (1.19%)	124 (1%)
OPEN PROSTATECTOMY	554 (1.15%)	140 (1.2%)
DIAGNOSTIC BRONCHOSCOPY AND BIOPSY OF BRONCHUS	550 (1.15%)	131 (1.1%)
NEPHRECTOMY, PARTIAL OR COMPLETE	526 (1.1%)	124 (1%)

Table 2.1: Training and testing dataset patient characteristics reported as number of patients (%) or mean ± standard deviation. HCUP code description and distribution is shown for those representing >1% of the training dataset

	No. of features	No. of hidden layers	No. of neurons	Hidden layer activation	Dropout probability	L2 lambda
With HCUP features	88	1	50	tanh	0.5	0.0001
Without HCUP features	55	1	50	tanh	0.5	0.001

Table 2.2: Final model parameters for each Generalized Additive Models with Neural Networks (GAM-NNs) model with and without HCUP category description features

Feature set	Model	AUC	AP
With HCUP features	GANN	0.921 (0.895–0.95)	0.176 (0.109–0.26)
	LR	0.912 (0.879–0.94)	0.207 (0.127–0.3)
Without HCUP features	GANN	0.912 (0.883–0.94)	0.197 (0.124–0.29)
	LR	0.906 (0.873–0.94)	0.217 (0.136–0.31)

Table 2.3: Area under the receiver operating characteristic curve (AUC ROC) and average precision (AP) with 95% CIs for the Generalized Additive Models with Neural Networks (GAM-NNs) and logistic regression (LR) models, with and without HCUP category description features.

Patient Example 1 (top 10 contributions)			Patient Example 2 (top 10 contributions)		
Feature	Value	Contribution	Feature	Value	Contribution
ART_LINE_YN	1	0.993	HCUP_cat_1_YN (Incision and excision of CNS)	1	1.080
ASA_SCORE	4	0.939	ART_LINE_YN	1	0.993
MIN_DBP	22	0.269	MIN_DBP	19	0.271
AGE	81	0.259	MIN_HB	7.6	0.184
AVG_DBP	68	0.234	PHENYLEPHRINE_CURRENT_RATE_MCG_MIN	43	0.177
PHENYLEPHRINE_CURRENT_RATE_MCG_MIN	45	0.191	PHENYLEPHRINE_MAX_RATE_MCG_MIN	43	0.174
PHENYLEPHRINE_MAX_RATE_MCG_MIN	45	0.176	MIN_MAP	17	0.132
MIN_MAP	30	0.122	AGE	69	0.094
AVG_HR	95	0.104	AVG_DBP_10_min	72	0.043
AVG_DBP_10_min	74	0.060	ANES_CASE_HOURS	3.9	0.001

Table 2.4: Top 10 neural network contributions learned from the best-performing Generalized Additive Models with Neural Networks (GAM-NNs) model with HCUP features, for two in-hospital mortality patient examples from the test set

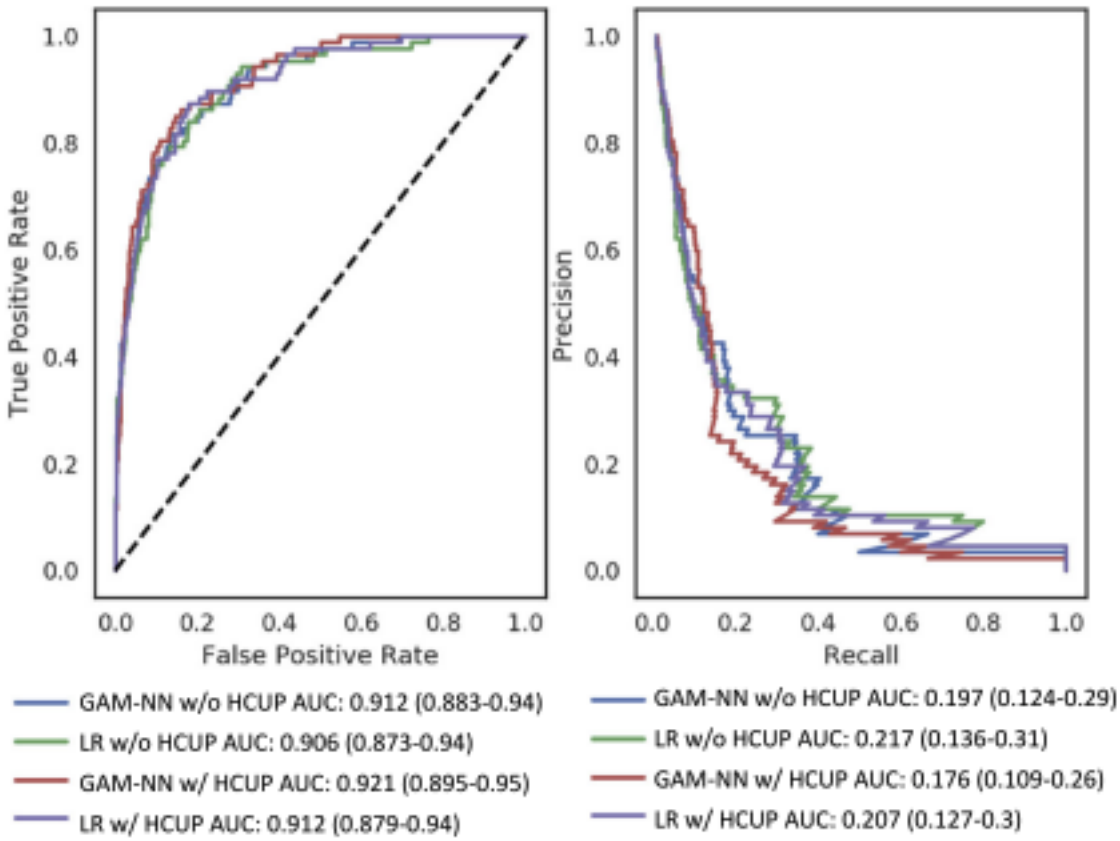

Figure 2.1: Receiver Operator characteristic curves and precision-recall curves for LR models and GAM-NN for prediction of mortality with and without HCUP features. GAM-NN: Generalized Additive Models with Neural Networks; HCUP: Healthcare cost and Utilization Project; LR: Logistic Regression

2.2.3 Performance Metrics

All performance metrics reported below refer to the testing set (n = 11,997). Area under the receiver operating characteristic curve (AUC ROC) and average precision (AP) are summarized in Table 2.3. The GAM-NN model with HCUP features had the highest AUC 0.921 (0.895–0.95). Overall, both GAM-NN models had higher AUCs than LR models, however had lower APs. The LR model without HCUP features had the highest AP 0.217 (0.136–0.31).

2.2.4 Interpretability: Visualizing Feature Contributions.

To assess the interpretability of the GAM-NNs, we visualized the learned contributions of the GAM-NNs and compared against the learned contributions of the LRs for the models with HCUP features. In Fig. 2.1, we visualize these contributions for a select sample of the top nine contributing features in the GAM-NN model. The top nine were chosen by selecting the features with the highest mean GAM-NN contribution. We did not include any binary features in this example, such as presence of arterial line, as their visualization would not be as interesting, since there would only be two values to plot. We see that, overall, the direction of the learned contributions from both the GAM-NN and LR models were similar, i.e., as MAX_DES increases, the contributions for both models decreased. However, while the LR model will always have a linear relationship, the GAM-NN learned nonlinear relationships that were unique to each feature. For example, for the feature AVG_MAP_10_MIN we see a nonlinear function where GAM-NN contributions increase for mean arterial pressure (MAP) $<$ 60 mmHg and MAP $>$ 60 mmHg. One odd relationship is the one observed between ANES_CASE_HOURS and mortality risk, where, with less hours spent under anesthesia there was more contribution to mortality risk. This could be a reflection of the infrequency of extremely high anesthesia case hours (>10 h), and that in-hospital mortality patients may

not spend significantly longer amounts of time under anesthesia compared to non-mortality patients. In addition, while risk contribution increased with lower MIN_DBP, there was the opposite relationship for AVG_DBP_10_MIN and AVG_DBP, which could indicate that not all summary measures of vital signs are the same, and that these should be taken into consideration when selecting features. Both of these examples demonstrate that the effect of a particular feature may not always represent an underlying physiological phenomena, and that modification of a particular feature for a particular patient may not necessarily produce a reduction in risk. For an up-close comparison of interpretability at the patient specific level, we can also look at the top GAM-NN contributors to the risk of mortality (Table 2.4). If we look at the top 10 GAM-NN contributions from the best-performing GAM-NN with HCUP features for two unique in-hospital mortality patients from the testing set, we can see that the features that contributed most were different. ASA was a top contributor for Patient Example 1 but not for Patient Example 2. Surgery-related features like presence of HCUP category 1 (HCUP_cat_1_YN) (Incision and excision of CNS), minimum case hemoglobin (MIN_HB), and time of anesthesia (ANES_CASE_HOURS) were top contributors for Patient Example 2, not found in Patient Example 1. While five of the shared top contributing features between Patient Examples 1 and 2 were blood pressure and phenylephrine-related features, Patient Example 1's top contributing features also included an additional blood pressure and heart rate related feature. These differences could indicate that while vital signs were top contributors for both patients, the surgery type contributed more to risk for Patient Example 2 than for 1.

2.3 Discussion

Despite their popularity and success in many applications such as speech recognition and computer vision, DNNs still face challenges to being fully accepted in the healthcare data

space. There has been growing interest and success in the application of DNNs for healthcare tasks due to the availability of large and complex electronic biomedical data, such as genomic data, biomedical images, and electronic medical records (EMRs) [1, 2, 3, 9, 10, 11]. In addition, in many cases, DNNs have shown better predictive performance than traditional models such as LR, however, a significant perceived problem with DNNs has been their "black box" reputation [4]. Clinicians are interested in not only the probability of an adverse event, such as in-hospital mortality, occurring but also need to understand what variables contributed to the increased risk so that they can change and target their therapies to potentially avoid an adverse event altogether. The inability of a model to allow for this level of transparency and interpretability is a potential barrier to positive clinical perception and can decrease trust and subsequently usability [12, 13, 14]. A small survey of ICU and ED clinicians found that clinicians viewed interpretability of a machine learning model as justification for clinical decision making following a model's prediction, and so models should be built with enough transparency around the clinical features driving the model's decision that clinicians could validate model outputs with their clinical knowledge and judgment [13]. Ginestra et al. found that when evaluating the real-time hospital implementation of their ML-based sepsis prediction alert, only 16% of providers found the alert helpful 6 h after an initial alert and only 9% reported that the alert changed management [14]. In addition, the most frequent suggestion by clinicians was transparency regarding factors leading to a sepsis alert. LR is often preferred in the medical field due to its easy implementation and interpretability. The learned coefficients can easily be extracted and interpreted as relative significance, and odds ratios calculated from those coefficients are routinely used in the medical research community to interpret a feature's contribution to increased odds of an adverse event. However, LR is a shallow model with no ability to create new feature representations and can only combine the features linearly before passing through a logistic function to represent the probability of response labels, such as in-hospital mortality. Neural networks have the ability to self-learn new and significant linear and nonlinear features that

are combinations of the original input features. However, these features can be thought of as "hidden" in the network layers. In this study, we were able to demonstrate that our proposed generalized additive neural network (GAM-NN) architecture is able to (1) leverage a neural network's ability to learn nonlinear patterns in the data, which is more clinically intuitive, (2) be interpreted easily, making it more clinically useful, and (3) maintain model performance as compared to our first study [3]. It should be noted that LR models can still incorporate nonlinear feature representations, but this requires extra featurization. For example, hypotension and hypertension are both of concern during surgery. If we use the average MAP as a feature, an LR model would only learn a coefficient that indicates either risk increases with increased MAP or risk decreases with increased MAP, as we see in Fig. 2.2. To incorporate the domain knowledge that risk should increase with both high MAP and with low MAP, the MAP feature would have to be transformed into new features, i.e., binning the MAP values and creating multiple binary features. However, neural networks minimize the need for this type of tedious feature engineering and preprocessing, and they can effectively learn this clinically intuitive relationship without the domain knowledge or extra featurization. Two limitations to our current study are that we were only able to develop and validate our model on (1) a single institution and (2) from the years 2013 to 2016, potentially limiting the generalizability of our results. Clinical practice not only varies from institution to institution, but also can change year to year with the emergence of new clinical evidence. While the difficulty in having large enough retrospective medical datasets to effectively train very complex models such as DNNs is no longer a limitation, developing the infrastructure to be able to not only gain access to the data, but to also obtain the data and process it for research use is a tremendous task. Obtaining past and new data from the same institution can itself be limiting, and the ability to access and integrate other institutions' data for validation can also be difficult. One benefit of this model is that the features needed from other institutions to validate our model are not only common across all institutions, but are also commonly used. For example, MAP is a commonly observed vital sign, however, features

like central venous pressure (CVP) and pulmonary arterial pressure (PAP) require invasive catheters and are only standard of care in more critical patients. Features like the bispectral index (BIS) do not require invasive catheters, however, it is not standard of care practice to monitor it. Thus, we expect the features in our model to be applicable to all patients across all institutions. However, it should also be noted that standard-of-care practice also varies from institution to institution, and so patterns discovered in this single-institution dataset may not be generalizable to other institutions and may require re-training of the model to individual institutions or more variety of institutions. As mentioned before, developing the infrastructure for such data extraction is a difficult process. The Perioperative Data Warehouse at UCLA used in the data extraction for this study exemplifies how this can be done successfully, however, replicating the process at another institution with a different electronic health record system and standardizing the disparate medical dataset to be able to merge it with our current one is a well-known issue in the medical data community [15]. Despite the difficulties mentioned above in obtaining new data for validation, we are currently working to address the limitations of our current validation results by collaborating with other institutions to replicate the data extraction used in this study as well as working within our own institution to access more recent data to validate the generalizability of our model. In addition, while the models in this study were made to be interpretable, it should be emphasized that the interpretation is not necessarily causation, and the modification of a highly contributing specific feature would not necessarily decrease the patient's risk of mortality. For example, in Table 2.4, both patient examples have high contribution related to arterial line placement, but deciding to not place an arterial line would not necessarily result in avoiding mortality. This is also true of other models such as LR. Although our model is transparent and the extraction of feature contributions described here explains how the model made the predictions, the relationship between the features and the risk of in-hospital mortality should still be thought of as correlation. These relationships would likely change with the removal of various features or addition of new features. However, the relationships

learned in this model appear to be clinically intuitive and they are still important in that they provide new or confirm known insight that is not usually available with DNNs. While we are no longer limited to using more traditional methods such as LR due to availability of data when developing more complex models, we should consider the needs for clinical adoption and impact. DNNs, such as ours, can be automated and incorporated with real-time EMR data. For example, with our model, all the model input features described can easily be automatically extracted or calculated at the end of surgery and our model would then be used to provide a probability for in-hospital mortality. If the probability is high, a summary of which features contributed the most to an increased risk of mortality (Table 2.4) and where the patient lies relative to other patients (Fig. 2.2) can also be displayed for the clinician. Thus, our model can serve as clinical decision support tool helping to identify patients in need of more postoperative resources and potentially informing therapeutic actions. For example, if a patient's minimum DBP being very low contributed the most to that patient's high risk of in-hospital mortality, the clinician may consider hypotension and associated risks such as acute kidney injury. A very different application of our model would be to re-train and apply it to a single institution to understand what areas of a patient's care during surgery clinicians could be paying more attention to moving forward, if they are not already. For example, in Fig. 2.2, low average MAP below 50 and high average MAP above 80 are both associated with increased risk of in-hospital mortality. Clinicians at this institution could then target therapies during surgery to never leave that range of MAP. However, at a different institution, the learned relationship could be different and the targeted range of MAP may change based on the current practices of that institution. In either application, our model could be used to quickly assess a large amount of data and provide actionable insight, a task that may otherwise be time-consuming for clinicians. In summary, this study shows that DNNs can be made to be not only accurate, but also interpretable. Any complex predictive model needs both to build enough trust that a clinician can interpret and act on a model's decision over or complementary to their own clinical intuition. Future work includes

not only validating the performance and generalizability of this model on other hospitals' datasets, but also assessing how clinicians interact with the interpretability of the model.

2.4 Methods

This manuscript follows the "Guidelines for Developing and Reporting Machine Learning Predictive Models in Biomedical Research: A Multidisciplinary View" [16].

2.4.1 Electronic Medical Record Data Extraction and Description.

All data used in this study came from the UCLA Medical Center's Perioperative Data Warehouse, a custom data warehouse built on top of the EMR (EPIC Systems, USA) and has been described in a previous paper [15]. All data used for this study were obtained from this data warehouse and IRB approval (UCLA-A IRB#15-000518) has been obtained for this retrospective review. Patients' written approval was waived because of the retrospective nature of this study. Data included all surgical procedures performed between March 1, 2013 and July 16, 2016, and excluded cases not performed under general anesthesia, ambulatory cases, and patients older than 89 or less than 18 years of age.

2.4.2 Model Endpoint Definition

The definition for in-hospital mortality was defined in the same way as described in our previous work [3]. The occurrence of an in-hospital mortality was extracted as a binary event [0, 1] based upon either the presence of a "mortality date" in the EMR between surgery time and discharge, or a discharge disposition of expired combined with a note associated with

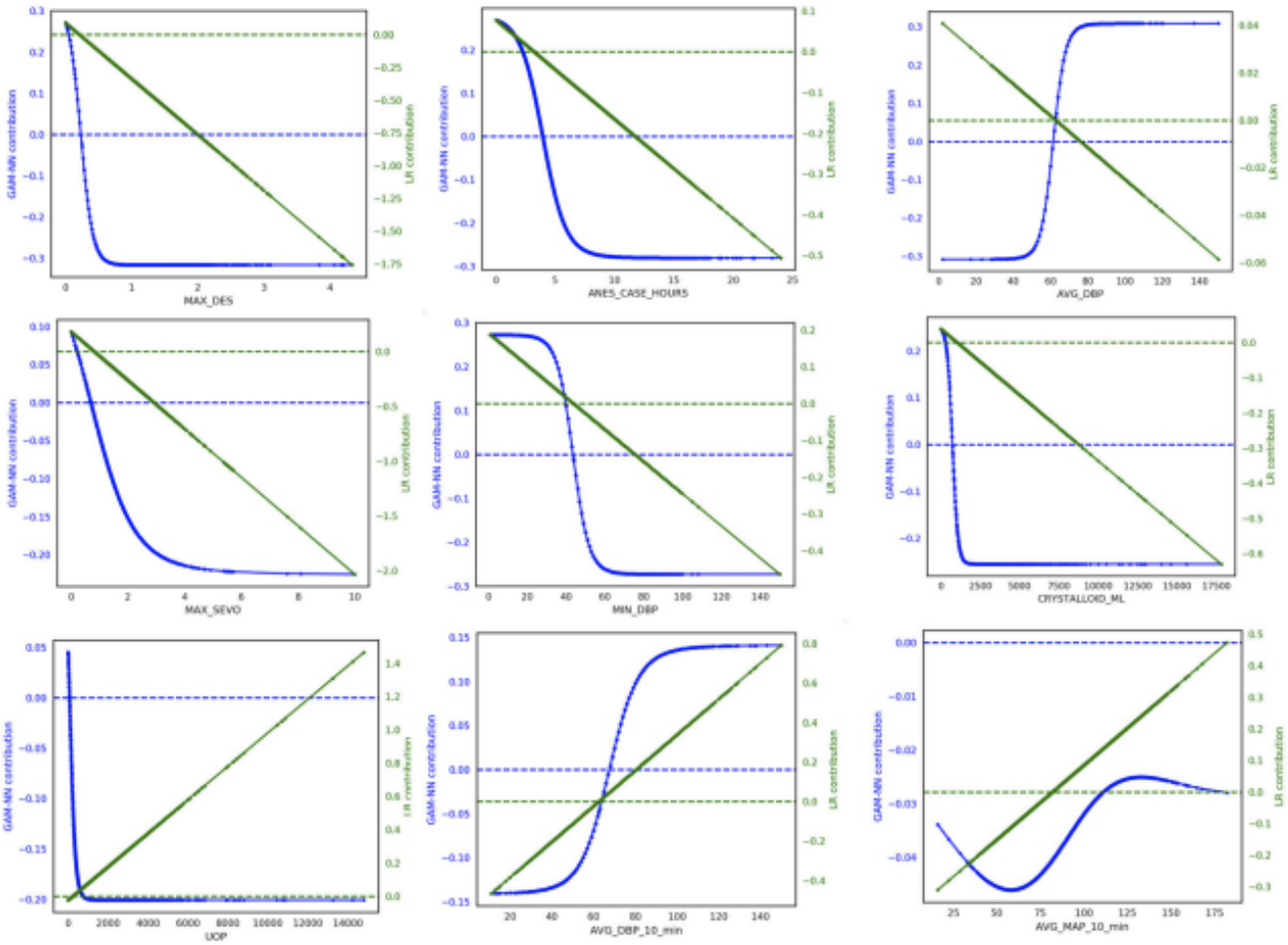

Figure 2.2: Sample of nine continuous features that had the highest mean mortality risk GAM-NN contributions across all patients, in order of highest to lowest. These features in order are maximum desflurane (MAX_DES), total anesthesia case hours (ANES_CASE_MINUTES), average diastolic blood pressure (AVG_DBP), maximum sevoflurane (MAX_SEVO), minimum diastolic blood pressure (MIN_DBP), total crystalloid administered (CRYSTALLOID_ML), urine output (UOP), average diastolic blood pressure of the last 10 min of the case (AVG_DBP_10_min), and average mean blood pressure of the last 10 min of the case (AVG_MAP_10_min). The feature's values for all patients are plotted on the x-axis and the respective GAM-NN contribution (blue) on the primary y-axis and LR contribution (green) on the secondary y-axis. The more negative the risk contribution, the less contribution the respective value has to the risk of mortality

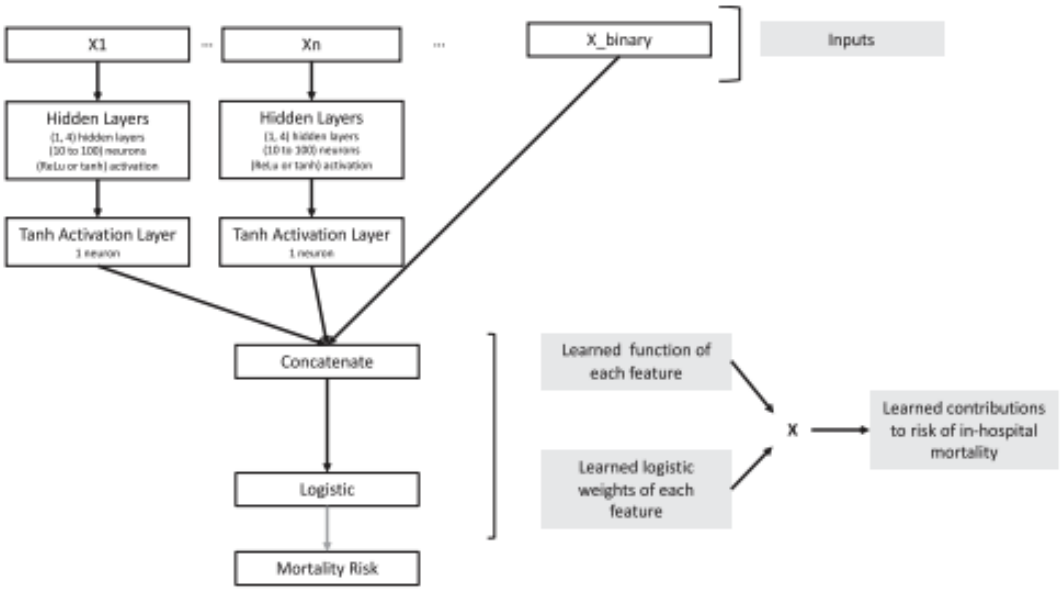

Figure 2.3: Proposed generalized additive models with neural networks architecture. This figure describes feature contributions calculation, for n individual continuous features (X1,. . . ,Xn) vs binary features (X_binary).

the death (i.e., death summary, death note). The definition of in-hospital mortality was independent of length of stay in the hospital.

2.4.3 Model Input Features

The data and features used in this study are from our previous work modeling in-hospital mortality [3]. The data consists of 59,985 patients with an original feature set of 87 features extracted at the end of surgery. These features included demographics, labs, ASA score, intraoperative vital signs, total case time, medication administration, and anesthesia events. These original 87 features were reduced to 45 features in our previous work, and ASA was added as a feature in the final model (46 features) that improved model performance [3]. In this study, we used the same 46 features, and also added previously not included features: total anesthesia case time (1 feature); the time in minutes spent with MAP below 40, 45, 50, 55, 60, and 65 mmHg (6 features); and HCUP Code Descriptions of the Primary current

procedure terminology (CPT) codes (33 features) (Table 2.5). There were 183 unique HCUP Code Descriptions in our dataset, and we selected 33 HCUP Code Descriptions that were present in at least 1% of the total data. These HCUP Code Descriptions were then encoded as 33 binary features.

2.4.4 Data Processing

Before model development, missing values for ASA scores were filled with the most common value (ASA 3); missing values for medications administration features indicated that no medication was actually administered and so were filled with 0; and all other missing values were filled with the means for that feature. Values that were greater than a clinically normal maximum were set to a maximum possible, as described in previous work [3]. Finally, all training data were rescaled to have mean 0 and standard deviation 1 per feature. Testing data were rescaled with the training data mean and standard deviation.

2.4.5 Development of Model and Feature Contribution Extraction

In this work, we were interested in classifying patients at risk of in-hospital mortality utilizing a proposed generalized additive neural network (GAMNN) architecture (Fig. 2.3). All data were randomly split into 80% for training (n = 47,988) and 20% for testing (n = 11,997) prior to model development. The loss function used in training was cross-entropy and to deal with the highly unbalanced classes, we applied class weights to the loss function by assigning the positive class 100x weight compared to the negative class to reflect the <1% occurrence of in-hospital mortality in our dataset. To optimize hyperparameters, a grid search across varying hyperparameter combinations was performed, where each model was trained on 80% of the data with 5-fold cross validation. The model with the highest mean 5-fold validation AUC was chosen as the one with the best hyperparameter combination, and

Feature	Feature description
AGE	Age of the patient in years (note we exclude ages <18 and >89)
ANES_CASE_HOURS	Case time under anesthesia in hours
ART_LINE_YN	Presence of arterial line
ASA_SCORE	ASA score
AVG_SBP, AVG_DBP, AVG_MAP, AVG_HR, AVG_PULSE_OX	Average systolic BP, diastolic BP, mean BP, heart rate, and pulse oximetry for the case
AVG_SBP_10_MIN, AVG_DBP_10_MIN, AVG_MAP_10_MIN, AVG_HR_10_MIN, AVG_PULSE_OX_10_MIN	Average systolic BP, diastolic BP, mean BP, heart rate, and pulse oximetry for the last 10 min of the case
BASELINE_GFR	Most recent GFR prior to surgery (only within 365 days)
COLLOID_ML	Total colloid administered
CRYSTALLOID_ML	Total crystalloid administered
CURRENT_HB, STARTING_HB	Most recent hemoglobin prior to surgery, Starting hemoglobin
CVC_ANES_YN	Presence of a central venous line
EBL	Total estimated blood loss
EPINEPHRINE_CURRENT_RATE_MCG_KG_MIN, EPINEPHRINE_MAX_RATE_MCG_KG_MIN	End of case rate of epinephrine, Highest infusion rate of epinephrine during the case
ESMOLOL_CURRENT_RATE_MCG_KG_MIN, ESMOLOL_MAX_RATE_MCG_KG_MIN	End of case rate of esmolol, Highest infusion rate of esmolol during the case
HCUP_CAT_x_YN	33 binary features for HCUP Category Descriptions IDs: [1. 3. 9. 10. 12. 33. 37. 42. 43. 48. 61. 67. 70. 76. 78. 80. 82. 84. 86. 99. 104. 105. 114. 124. 126. 146. 152. 153. 158. 160. 161. 172. 225]
MAX_DBP, MAX_DES, MAX_GLUCOSE, MAX_HR, MAX_ISO, MAX_MAP, MAX_PULSE_OX, MAX_SBP, MAX_SEVO	Maximum diastolic BP for the case, Maximum MAC of desflurane during the case (note this is not age adjusted), Maximum glucose for the case, Maximum heart rate for the case, Maximum MAC of isoflurane during the case (note this is not age adjusted), Maximum mean BP for the case, Maximum pulse oximetry for the case, Maximum systolic BP for the case, Maximum MAC of sevoflurane during the case (note this is not age adjusted)
MILRINONE_CURRENT_RATE_MCG_KG_MIN, MILRINONE_MAX_RATE_MCG_KG_MIN	End of case infusion rate of milrinone, Highest infusion rate of milrone during the case
MIN_DBP, MIN_GLUCOSE, MIN_HB, MIN_HR, MIN_MAP, MIN_PULSE_OX, MIN_SBP	Minimum diastolic BP for the case, Minimum glucose for the case, Minimum hemoglobin during the case, Minimum heart rate for the case, Minimum mean BP for the case, Minimum pulse oximetry for the case, Minimum systolic BP for the case
MIN_MAP_LT_40,MIN_MAP_LT_45,MIN_MAP_LT_50,MIN_MAP_LT_55, MIN_MAP_LT_60 min_MAP_LT_65	Minutes MAP < 40 mmHg, Minutes MAP < 45 mmHg, Minutes MAP < 50 mmHg, Minutes MAP < 55 mmHg, Minutes MAP < 60 mmHg, Minutes MAP < 65 mmHg
NICARDIPINE_CURRENT_RATE_MG_HR, NICARDIPINE_MAX_RATE_MG_HR	End of case rate of nicardipine, Highest infusion rate of nicardipine during the case
NITRIC_OXIDE_YN	Nitric oxide used for the case
NITROGLYCERIN_CURRENT_RATE_MCG_MIN, NITROGLYCERIN_MAX_RATE_MCG_MIN	End of case rate of nitroglycerin, Highest infusion rate of nitroglycerin during the case
NITROPRUSSIDE_CURRENT_RATE_MCG_KG_MIN, NITROPRUSSIDE_MAX_RATE_MCG_KG_MIN	End of case rate of nitroprusside, Highest infusion rate of nitroprusside during the case
PA_LINE_YN	Presence of pulmonary artery catheter
PHENYLEPHRINE_CURRENT_RATE_MCG_MIN, PHENYLEPHRINE_MAX_RATE_MCG_MIN	End of case rate of phenylephrine, Highest infusion rate of phenylephrine during the case
UOP	Total urine output
VASO_CURRENT_RATE_UNITS_HR, VASOPRESSIN_MAX_RATE_UNITS_HR	End of case rate of vasopressin, Highest infusion rate of vasopressin during the case
XFUSION_RBC_ML	Total red blood cells transfused

Table 2.5: Description of model input features

retrained on all of the training data prior to being tested. All hyperparameter values that were assessed are shown in parentheses. All models were trained with a batch size of 256 and Adam optimization with default parameters and reduced the learning rate by a factor of 10 when the validation loss stopped improving for five consecutive epochs, a batch size of 256 and a maximum of 100 epochs [17]. Dropout (0.25, 0.5, 0.9) and L2 regularization (0.001, 0.0001) were also used to prevent overfitting [18,19]. In our GAM-NN architecture, each feature had its set of hidden layers (1–4) with layer sizes of 10, 40–50, 90, 100 neurons with all activations being either rectified linear unit (ReLu) or hyperbolic tangent (tanh) (Fig. 2.3). These hidden layers are followed by a last layer with just one neuron with a tanh activation. This last tanh layer transforms the previous layer's output into one value and forces the feature's neural network final output to be between 1 and 1. The outputs of all the features' tanh layers are then concatenated prior to being input into the final logistic layer (Fig. 2.3). The feature contributions are calculated as their tanh layer outputs multiplied by their respective logistic weights. Binary features only had a direct connection from the input layer to the final logistic layer, and so their feature contributions are calculated as the input value multiplied by their respective logistic weights.

2.4.6 HCUP Feature Experiment

HCUP codes provide informative groupings in regard to a patient's surgery and are also uniformly coded, making them easy to use as model inputs. However, they are not immediately available at the end of surgery, and so their inclusion could limit our model's practical use. Thus, we also assessed developing a model without HCUP features to assess the impact on performance.

2.5 Model Performance

All model performances were assessed on the 20% of the data held out from training as a testing set. The same training and testing sets were used in this work as our previous work on in-hospital mortality for comparison [3]. Model performance was compared to a standard LR model using the same features as the GAM-NN.

Performance Metrics

Model performance was assessed using area under the receiver operating characteristic curve (AUC ROC) and average precision (AP), and 95% confidence intervals were calculated using bootstrapping with 1000 samples.

Interpretability: Visualizing Feature Contributions

As previously described, the learned contribution of the GAM-NNs for each feature is its last tanh layer's output multiplied by its respective weight from the logistic layer. Since the binary features have a direct connection from input to the logistic layer, the binary features' learned contributions would be their input values multiplied by their respective weight from the logistic layer. For every data sample, each individual feature's value was plotted on the x-axis vs its respective contribution on the y-axis. Individual feature contributions in the LR model were calculated as the individual feature's value multiplied by its learned coefficient. For both models, the more negative the risk contribution, the less contribution the respective value has to the risk of mortality. All neural network models were developed using Keras [20]. LR models and performance metrics were calculated with scikit-learn [21].

2.6 Data Availability

The datasets generated during and/or analyzed during the current study are not publicly available due to institutional restrictions on data sharing and privacy concerns. However, the data are available from the corresponding author on reasonable request.

2.7 Code Availability

Code is available from the corresponding author on reasonable request. All neural network models were developed using Keras. LR models and performance metrics were calculated with scikit-learn.

Chapter 3

MOVER: Medical Informatics Operating Room Vitals and Events Repository

3.1 Introduction

In 2009, the Health Information Technology for Economic and Clinical Health (HITECH) Act was enacted to promote the adoption of healthcare information technology in hospitals. This act includes incentives for using electronic health record (EHR) systems [22]. The passage of the HITECH Act has resulted in widespread hospital EHR adoption with 80.5% of hospitals in the US using an EHR system, as of 2015 [23]. The increased adoption of EHR systems and subsequent rise in digitally available healthcare data has resulted in a newfound ability to perform predictive modeling on healthcare data using artificial intelligence (AI). Applications of AI in a healthcare setting include: providing more accurate diagnoses, recommending treatment plans, predicting patient outcomes, tracking patient engagement

and adherence, and reducing the burden of administrative tasks [24]. The increase in data availability and numerous applications for AI have resulted in the widespread publication of AI algorithms applied to healthcare data [25, 26, 27, 28, 29, 30, 31, 32]. Despite numerous publications showing AI algorithms to be very successful in retrospective healthcare studies, very few of these algorithms are then actually incorporated into clinical practice [33]. While there are many factors influencing the lack of algorithm deployment, one of the major reasons is a lack of trust in AI models. This lack of trust stems in part from an inability to independently verify published results. Although the publication of AI algorithms in the healthcare space has become pervasive, the data used to develop these algorithms is often not published alongside the algorithm. This lack of visibility into the analyzed dataset prevents independent researchers not only from reproducing published results but also from improving upon existing algorithms and developing novel algorithms [34, 35]. To address this issue, several repositories hosting medical data have been created including: MIMIC-IV, the UCI Machine Learning Repository, and n2c2: National NLP Clinical Challenges [36, 37, 38]. While the publication of these repositories has helped address some of these problems, the repositories themselves are extremely specialized and therefore only improve visibility for very specific AI algorithms.

To address the overarching issue of lack of visibility and the issue of excessively specific data, we propose MOVER: Medical Informatics Operating Room Vitals and Events Repository. This repository contains real hospital visits for patients undergoing surgery at a prestigious medical center in Orange County, California. The data included in MOVER was collected over seven years and contains comprehensive EHR records and waveforms for patients who underwent surgery. These records include general information about each patient and their medical history, and specific information regarding the surgical procedure being performed including: medicines used, lines or drains used, and post-operative complications. The repository includes 58,799 unique patients with data from 83,468 surgeries. MOVER is freely available for download for all researchers who sign a data usage agreement and is

intended to advance a wide variety of healthcare research.

3.2 Materials and Methods

3.2.1 Patient Population

This first release of MOVER database includes all adult patients who underwent surgery at the University of California, Irvine Medical Center from 2015–2022. The University of California, Irvine Medical Center is a level I trauma center, a burn treatment center, and a National Cancer Institute-designated comprehensive cancer center. In addition, the UCI Douglas Hospital has some of the most technologically advanced surgical suites including state-of-the-art endovascular hybrid suites and intraoperative computed tomography (CT) and magnetic resonance imaging (MRI) suites.

3.2.2 Database Development

The data acquisition process did not interfere with the clinical care of patients or methods of monitoring. Data for patients who underwent surgery were captured from two different sources. First, waveforms (EKG, pulse oximetry, and arterial line if present) from all of the operating rooms were captured in real-time using Bernoulli Health's hardware software platform. All of the waveforms were saved to a server on the medical center's network organized by source location (operating room) and datetime. Subsequently, the medical center's IT team delivered a data extract from the hospital EHR system from 2015-2022. For the years 2015-2017 the EHR system used was Surgical Information Systems (SIS) and from 2017-2022 the EPIC EHR system was used. For this reason MOVER is separated into two datasets: the first contains two years of data from the SIS EHR system (SIS dataset)

and the second contains five years of data from the EPIC EHR system (EPIC dataset). Although the data captured by these EHR systems is similar, there are enough significant differences that keeping the data separate was necessary in order to preserve as much data as possible.

3.2.3 Clinical and Physiological Data

The SIS dataset includes 19,114 patients and is separated into 9 tables: patient information, patient I/O, patient vitals, patient observations, patient medications, patient labs, patient procedure events, patient ventilator, and patient a-line. These tables contain patient demographics, information regarding the surgical procedure and anesthesia, laboratory data, and administered medications. This dataset is unique because in addition to waveforms, it contains high temporal resolution vital signs including: cardiac output, blood pressure, and stroke volume variation.

The EPIC dataset is the larger of the two datasets, containing 39,685 patients, and is separated into 10 tables: patient information, patient history, patient visit, patient medications, patient LDA (lines, drains, and airway devices), patient labs, patient measurements, patient postoperative complications, patient procedure events, and patient coding. Similar to the SIS dataset, the EPIC dataset includes patient demographics and specific information regarding the surgical procedure being performed including: medicines used, surgical events, and laboratory data. Although similar, a major difference between these two datasets is that the EPIC dataset contains outcome information in the form of postoperative complications, mortality, and if the patient was admitted to the ICU. Additionally the EPIC dataset includes information about a patient's medical history prior to surgery and their ASA status, while the SIS dataset does not. The final major difference is that the EPIC dataset includes billing codes. Table 3.1 outlines the contents of each of the EPIC dataset tables.

EPIC Dataset Table	Description
Patient Information	patient demographic information including: age, sex, height, and weight; information regarding the surgery being performed including: the type of surgery, the start and end time of both the surgery and anesthesia, ASA status, and discharge disposition.
Patient History	patient's health history including all patient diagnoses available in the EHR
Patient Visit	diagnosis and diagnosis code for a particular visit
Patient Medications	medications which were prescribed to patients including: the dose of the medication, the time the medication was administered or prescribed, and the medication route
Patient LDA	description of lines, drains, and airway devices used on the patient, the time of placement, the time of removal, and location of placement
Patient Labs	labs ordered for a patient, the corresponding observed measurements, and the reference measurements for each lab.
Patient Measurements	all preoperative and postoperative patient measurements including: vitals, intake and output, pain levels, complications, and disposition
Patient Postoperative Complications	the type of complication, when the complication was observed, and a free text note field outlining more specific details
Patient Procedure Events	preoperative, perioperative, and postoperative procedural events and the corresponding time of the event.
Patient Coding	patient billing codes

Table 3.1: Description of the 10 tables included in the EPIC dataset

3.2.4 Database Merger and Postprocessing

The second stage in developing MOVER involved significant data post processing and database organization. Following delivery of the data extract, the start and end time of each case was used to extract the appropriate waveform data for that case (based on location and date/time) and link it to the case. The EHR data for both datasets was then restructured and organized into logical tables for simplicity and to help facilitate data analysis. The raw EHR data for the EPIC dataset contained several redundant identifiers for patients and patient visits that differed between tables. To simplify this, the number of patient/event identifiers in the data were reduced to just two: the patient identifier and the patient visit identifier. This identification system allows patients to be tracked over time if they have multiple surgeries. The SIS dataset only has a single identifier representing each surgery, it is not possible to track patients temporally using this dataset. Due to the countless number of ways this data can be analyzed, we chose not to clean it to ensure that all elements of the clinical dataset were preserved. We do however provide explanations for the meaning of each column included in the data as well as the column's unit of measurement to assist researchers in properly cleaning that data for themselves.

3.2.5 Data Deidentification and HIPAA Compliance

Under HIPAA Privacy Rule, all patient identifiers were removed or deidentified. For deidentification, all patient identifiers and patient visit identifiers were encoded via one-way hash functions. Additionally, PHI was removed from free-text using regular expressions and manually reviewed to ensure that all PHI was removed. Patient ages were capped at 90, so any patient with a recorded age of more than 90 years old was set to 90 years old. Ages were capped to protect patient anonymity because extreme ages are considered identifying. Finally, dates were shifted by a random number of days. The number of days by which to

shift the data is linked to each patient to ensure that the data for a single patient is internally consistent. For example, if a patient had two surgeries two months apart in the raw data then the deidentified data would also reflect the surgeries as being two months apart. It is important to note, that this temporal consistency can only be observed for a single patient and not across distinct patients. For example in the deidentified data two patients who are listed as having surgery on the same day in reality did not necessarily have surgery on the same day or even in the same year.

3.2.6 Database Distribution and Documentation

MOVER is available for download at https://mover.ics.edu and can be downloaded by anyone who signs a data usage agreement, to restrict traffic to legitimate researchers. The website outlines the content of each downloadable table including: the meaning of each column, an explanation of the possible values of each column (where applicable), and the unit of measurement for each column (where applicable).

3.2.7 Database Characterization

We outline the characteristics of version 1.0 of MOVER so that investigators can understand possible applications of the data and see if the data fits their needs. Summarized data includes: demographics, ASA status, and most frequently performed procedures. Additionally, to illustrate some of the analyses possible with the database we summarize outcomes including: post-operative complications, admission to ICU, and mortality. In addition to using clinical data, investigators can utilize the time-series physiological data to develop real-time predictors to assist anesthesiologists. For example, using MOVER investigators could attempt to predict post-induction hypotension (mean arterial pressure $<$ 60mmHG) which is a well-known risk factor for adverse postoperative outcomes [39].

3.3 Results and Discussion

MOVER includes 58,799 unique patients with data from 83,468 surgeries. The EPIC dataset makes up the majority of the repository, with 39,685 patients and 64,354 surgeries. Table 3.2 shows summary statistics and patient demographics for all surgeries in the EPIC dataset. Of the 64,354 surgeries we can see that the average age of patients was 55 and that the average length of stay was 7 days. Additionally, looking at the ASA scores we can see that the mode is an ASA score of 3 but the dataset has a diverse distribution of scores. The ASA score is a system used to represent a patient's pre-anesthesia medical comorbidities, with a higher score representing a patient in worse health. Having a diverse distribution of ASA scores in this dataset shows that patients undergoing surgery are in varying degrees of health, which makes this dataset more generalizable than datasets exclusively containing patients in critical condition. Table 3.3. characterizes the outcomes available in the EPIC dataset in MOVER. This characterization is useful to investigators to get an idea of what predictions they can make using MOVER. In the EPIC dataset, 45.3% of patients are transferred to the ICU after surgery and there is a 1.6% mortality rate. Table 3.3 also shows the percentages of the 11 classes of post-operative complications. Each postoperative complication is assigned to a class and more specific details surrounding the complication can be found in the associated free-text. Investigators would be able to use these outcomes individually for specific outcome prediction, or use them in combination to understand what factors contribute to a bad outcome of any kind. While the SIS dataset does not contain outcome information, it does include high temporal resolution vital signs which would be invaluable for making real-time predictions to assist anesthesiologists.

Characteristic	EPIC Dataset
Gender	
Female	30,139 (46.8%)
Male	34,214 (53.2%)
Age	55 ± 17
ASA Rating	
1	2,960 (4.6%)
2	18,068 (28.1%)
3	29,449 (45.8%)
4	6,370 (9.9%)
5	657 (1.0%)
6	41 (0.06%)
Length of Stay (in days)	7 ± 14
10 Most Performed Procedures	
CATHETERIZATION, HEART, LEFT, WITH INTERVENTION IF INDICATED	1,521 (2.3%)
CHOLECYSTECTOMY, LAPAROSCOPIC	1,189 (1.8%)
LAPAROSCOPY, DIAGNOSTIC	1,040 (1.6%)
LAPAROTOMY, EXPLORATORY	985 (1.5%)
DILATION AND EVACUATION, UTERUS	887 (1.4%)
DEBRIDEMENT, WITH SPLIT-THICKNESS SKIN GRAFT APPLICATION	741 (1.2%)
ARTHROPLASTY, KNEE	680 (1.1%)
IRRIGATION AND DEBRIDEMENT, LOWER EXTREMITY	623 (1.0%)
ORIF, FRACTURE, FEMUR	608 (0.9%)
AV FISTULOGRAM, WITH ANGIOPLASTY IF INDICATED	573 (0.9%)

Table 3.2: Characterization of the EPIC Dataset reported as number of records (%) or Mean ± SD

Outcome	EPIC Dataset
Transfer to the ICU	29,131 (45.3%)
Death	1,023 (1.6%)
Postoperative Complications	
Other	1,093 (1.7%)
Cardiovascular	861 (1.3%)
Respiratory	735 (1.1%)
Airway	373 (0.6%)
Metabolic	154 (0.2%)
Neurological	147 (0.2%)
Administrative	118 (0.2%)
Injury/Infection	117 (0.2%)
Medication	94 (0.1%)
Regional	60 (0.1%)
Chronic Pain	32 (0.05%)

Table 3.3: Characterization of the EPIC Dataset Outcomes reported as number of records (%)

3.3.1 Comparison With Other Databases

To the best of our knowledge MOVER is the only freely available public database that contains EHR and waveforms data for patients undergoing surgery. The only other similar database we could find is The Veterans Affairs Surgical Quality Improvement Program (VASQIP) database but this database is not publicly available. There are several other kinds of medical datasets that have been made publically available including: MIMIC-IV, the UCI Machine Learning Repository, and n2c2: National NLP Clinical Challenges [36, 37, 38]. While the publication of these repositories has helped address some of the problems surrounding a lack of trust in AI models, the repositories themselves are extremely specialized and therefore only improve visibility for very specific AI algorithms. The patient population included in MIMIC-IV are patients who were admitted to the ICU or emergency department, limiting algorithms using this data to focus on patients in critical condition. Similarly, the UCI Machine Learning Repository has datasets focusing on very specific health issues, such as diabetes, and does not include complete EMR data. Finally, n2c2 exclusively contains unstructured text and is used for natural language processing applications.

3.4 Conclusion

AI has innumerable possible applications in the healthcare field and has tremendous potential to greatly improve patient care and outcomes. Our hope in publishing this data is to forward the development of novel AI algorithms, allow for the improvement of previously published AI algorithms, and to help validate AI algorithms in the healthcare space. This improved visibility should in turn help reduce the gap between research and deployment of AI algorithms in a clinical setting.

Chapter 4

CircadiOmics: Circadian Omic Web Portal

4.1 Introduction

Circadian rhythms are found in plants, animals, fungi, and cyanobacteria and are fundamental to biology [40, 41, 42, 43]. They date back to the first cyanobacteria and the origin of life on earth and, since then, through approximately 2 trillion revolutions of the earth on its axis, they have been deeply etched in the molecular machinery of all cells. Disruptions of circadian rhythms have been linked to health problems such as cancer, diabetes, obesity, and premature aging [41, 44, 45, 46, 47, 48, 49, 50, 51]. The advance of modern high-throughput technologies has made it possible to investigate circadian rhythms at the molecular level. Measuring the concentrations of molecular species across time has shown that circadian oscillations are pervasive in all living cells [43, 52, 53]. Circadian oscillations are generated by feedback loops which are regulated in part by the 'core clock' [54]. The core clock is a classical inhibitory transcription-translation feedback loop which is highly

conserved from animals to plants. While the core clock comprises a dozen genes including CLOCK, BMAL1, PER1, PER2, CRY1, and CRY2, in any given experiment 10% of all transcripts and metabolites display circadian oscillations [55, 56, 57, 58, 59, 60, 61]. However the complement of molecular species exhibiting circadian oscillations greatly varies with genetic, epigenetic, tissue/organ, health, age, and environmental conditions. This circadian reprogramming is a major target of active investigations aimed at understanding how environmental conditions, such as drug treatments or diets, affect circadian oscillations, and how oscillations in different cells/tissues are coordinated and interact with each other [58, 62, 63, 64, 65, 66, 67, 68, 69]. The large repository of omic data available via CircadiOmics, serves as an invaluable resource to analyze the complexity of circadian mechanisms and their downstream implications. The CircadiOmics interface is especially advantageous and unique because it allows users to easily perform comparative analyses and aggregated inferences about circadian rhythms at the molecular level across species, tissues/organs, and genetic, epigenetic, and environmental conditions.

4.2 Materials and Methods

4.2.1 Datasets

CircadiOmics currently contains over 290 omic datasets with over 100 million individual measurements, across more than 20 unique tissues/organs, and 11 different species. For simplicity we group the unique tissue/organ types into 13 categories: liver, brain, digestive, skin, serum, muscle, adipose, glands, cells, kidney, heart, eye, and other. Figure 4.1b shows a breakdown of the number of datasets contained within the repository for each tissue/organ category. The species currently represented in CircadiOmics include: Aedes aegypti, Anopheles gambiae, Arabidopsis thaliana, Danio rerio, Drosophila melanogaster, Homo sapiens, Mus musculus,

Source	Datasets	Tissues	Species
CircadiOmics	299	25	11
CircaDB	43	15	2
BIOCLOCK	2	2	2
CirGRDB	99	<20	2

Table 4.1: Comparison of CircadiOmics with other circadian web servers

Neurospora crassa, Papio anubis, Rattus norvegicus, and Rhesus macaques. Figure 4.1a shows a breakdown of the number of datasets per species. As such, CircadiOmics is the most extensive, comprehensive, and current repository for circadian data. For comparison purposes, Table 4.1 shows a breakdown of the number and types of datasets currently available in the most prominent circadian data repositories [70, 71, 72, 73]. The majority of datasets in CircadiOmics are collected from the species Mus musculus (mouse) and Papio anubis (baboon) and from liver and brain tissues. In addition to a wide variety of species and tissues, CircadiOmics also has a diverse set of experimental conditions represented. Some of the experimental conditions represented include: knock-downs, knock-outs, diet changes, exercise, and drug treatments. In addition, CircadiOmics uniquely contains data from different omic experiments, including transcriptome, metabolome, proteome, and acetylome experiments. Figure 4.1 summarizes the number of available datasets by detailed categories. The full table summarizing all of the datasets is available on the CircadiOmics web portal with a short explanation of the dataset, a brief description of the experimental protocol, the citation, the GEO accession number, and other summary information.

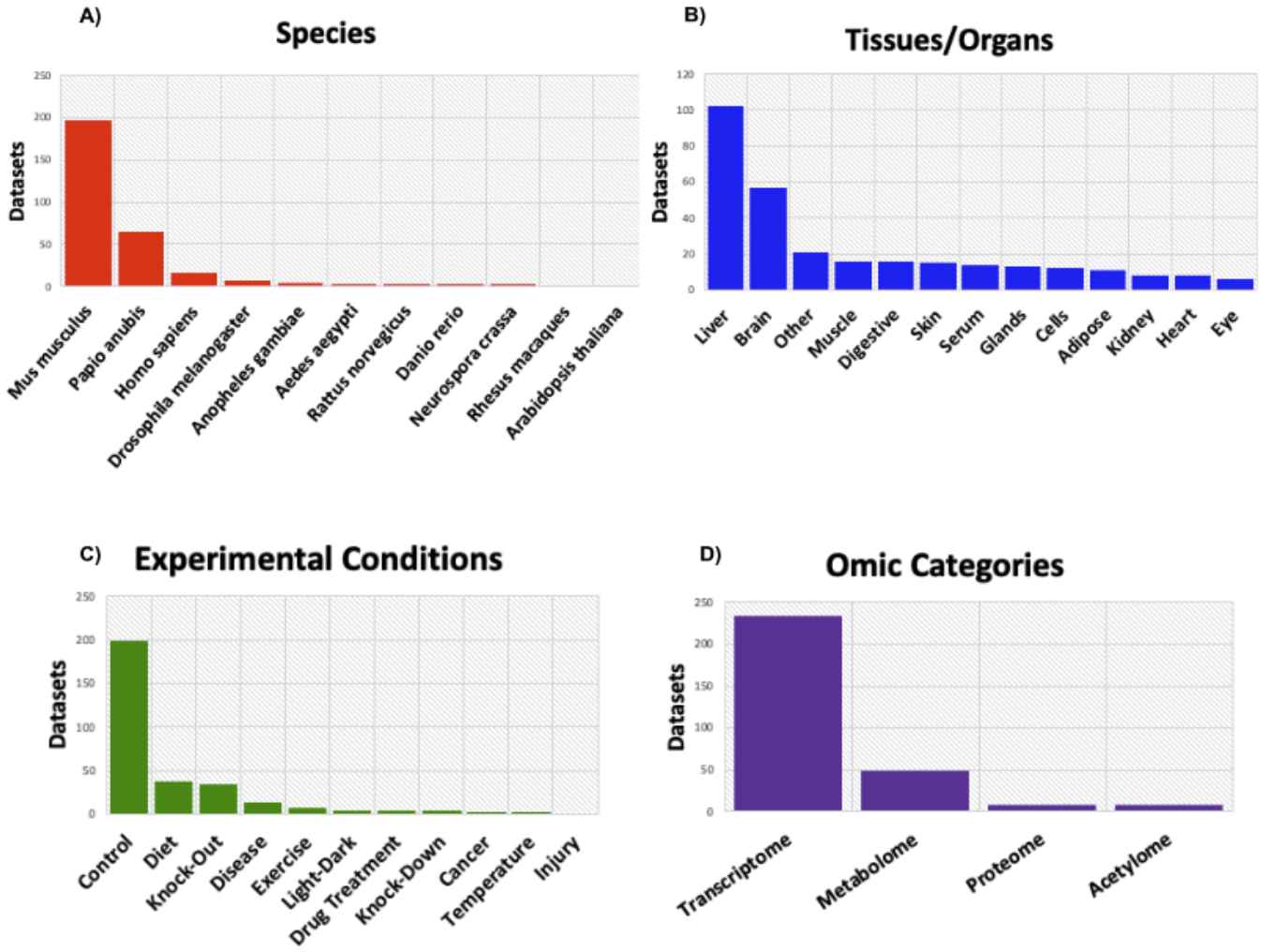

Figure 4.1: Breakdown of datasets by species, tissue, experimental conditions, and omic categories

4.2.2 Dataset Collection

The datasets in CircadiOmics are collected from research collaborations, automated discovery, and manual discovery. The two main automated approaches used to identify newly available circadian dataset are a web crawler developed in-house and the publicly available web service PubCrawler [74]. The web crawler developed in-house uses the Python packages scholarly and geotools to search the literature to discover new circadian omic studies and their affiliated datasets. To find new datasets, the crawler performs keyword searches on published abstracts, extracts various features from the published articles, and then uses logistic regression on the extracted features to classify whether or not a dataset is a good candidate for inclusion in CircadiOmics. The datasets discovered by the crawler are then manually vetted and processed to be included into CircadiOmics. Using this crawler in tandem with PubCrawler, which sends a daily email containing a list of possible publications of interest, we are able to keep CircadiOmics current by continuously adding the latest cutting-edge research in circadian rhythms to the repository. Additionally, the CircadiOmics team and collaborating biologists include datasets from collaborative research projects and perform periodic manual searches on recent publications to further complement the data obtained through the automated discovery tools.

4.3 Results

4.3.1 Features

The main focus of CircadiOmics is the search function, which allows users to compare and visualize the oscillation trends of molecular species. The user can select a single dataset, or multiple datasets, from within the repository and search for any molecular species. Circa-

diOmics allows for the overlay of multiple searches together to enable comparative studies and normalizes the output for easy visual comparison. For each query, a table of periodicity statistics including: period, amplitude, phase, p-value, and q-value is displayed. These statistics are calculated using BIO_CYCLE and JTK_CYCLE [75, 76]. Molecular species are determined to have circadian oscillations by using p-values and accompanying q-values at a user selected threshold. In addition to selecting whether to view statistics from BIO_CYCLE or JTK_CYCLE, users can also filter datasets based on species, tissue/organ, and experimental conditions. In addition to the search functionality, to assist with the analysis of circadian experiments we have created the BIO_CYCLE web server: http://circadiomics.igb.uci.edu/biocycle. The web server runs the latest version of the BIO_CYCLE software on user-uploaded omic time-series datasets and provides the user with easy-to-use analysis tools which include: histograms of periods, phases, amplitudes, and offsets, querying of molecular species based on a user-selected p-value or q-value cutoff, visualization of molecular concentrations across time, and analysis at 24, 12, and 8 hour periods. To use the web server, users must upload a file containing the measurements related to the concentrations of molecular species across time points (e.g. transcript levels measured every 4 hours). Each row must contain the ID of the molecular species, followed by the concentration measurement at each time point. Each column corresponds to a different time point or replicate. After the file is uploaded, the server will run BIO_CYCLE on the uploaded file for three separate period ranges: 20 through 28 hours, 10 through 14 hours, and 7 through 9 hours. A separate deep neural network (DNN) is trained for each set of timepoints and for each period range. If the DNNs are already trained, then the results should be ready within about 1 minute. If the DNNs are not trained, BIO_CYCLE will automatically train them and the results will be ready within about 2 minutes. The user can then visualize the results using various drop-down menus to select the period of interest and p-value and q-value cutoffs. As shown in Figure 4.2, given the selected period to investigate, and a p-value or q-value threshold, the web server will produce histograms of periods, lags, amplitudes,

and offsets. Another feature provided by CircadiOmics is the The Metabolic Atlas web portal. The Metabolic Atlas web portal (http://circadiomics.ics.uci.edu/metabolicatlas), allows researchers to generate and visualize interactive metabolic networks. These networks are derived from the KEGG database and can be filtered using BIO_CYCLE statistics [77]. To create a metabolic network, users start by selecting a dataset and a particular metabolite. From there, the user can select options to create a network. For example, one option is to display a network of all metabolites that are oscillating in-phase with the selected metabolite. Another possible option is to display a network of all metabolites that are involved in the same pathways as the selected metabolite. There are six possible options for the user to select from for the network creation. Once the network is displayed, the user can choose to filter out edges based on BIO_CYCLE statistics.

4.3.2 Improvements

Since its last publication, CircadiOmics has undergone substantial improvements including a significant increase in the number of datasets and the diversity of datasets available to its users. The number of available queryable datasets has increased from 227 to over 290, and the number of species and experimental conditions included in the web server has also increased. In addition to the significant increase in available data, the latest version of CircadiOmics and the corresponding automated data discovery pipeline have received several improvements. To optimize the automated dataset discovery process, we have started utilizing open source web-crawlers in addition to our web crawler developed in-house to make sure we are capturing as many relevant datasets as possible, and we have improved our in-house web crawler by broadening the keyword searches performed on published abstracts to allow us to discover more species and tissue types that had not previously been represented in CircadiOmics. This improvement to the web crawler is what allowed us to find circadian experiments performed on Drosophila melanogaster, Danio rerio, and Neurospora crassa, species not

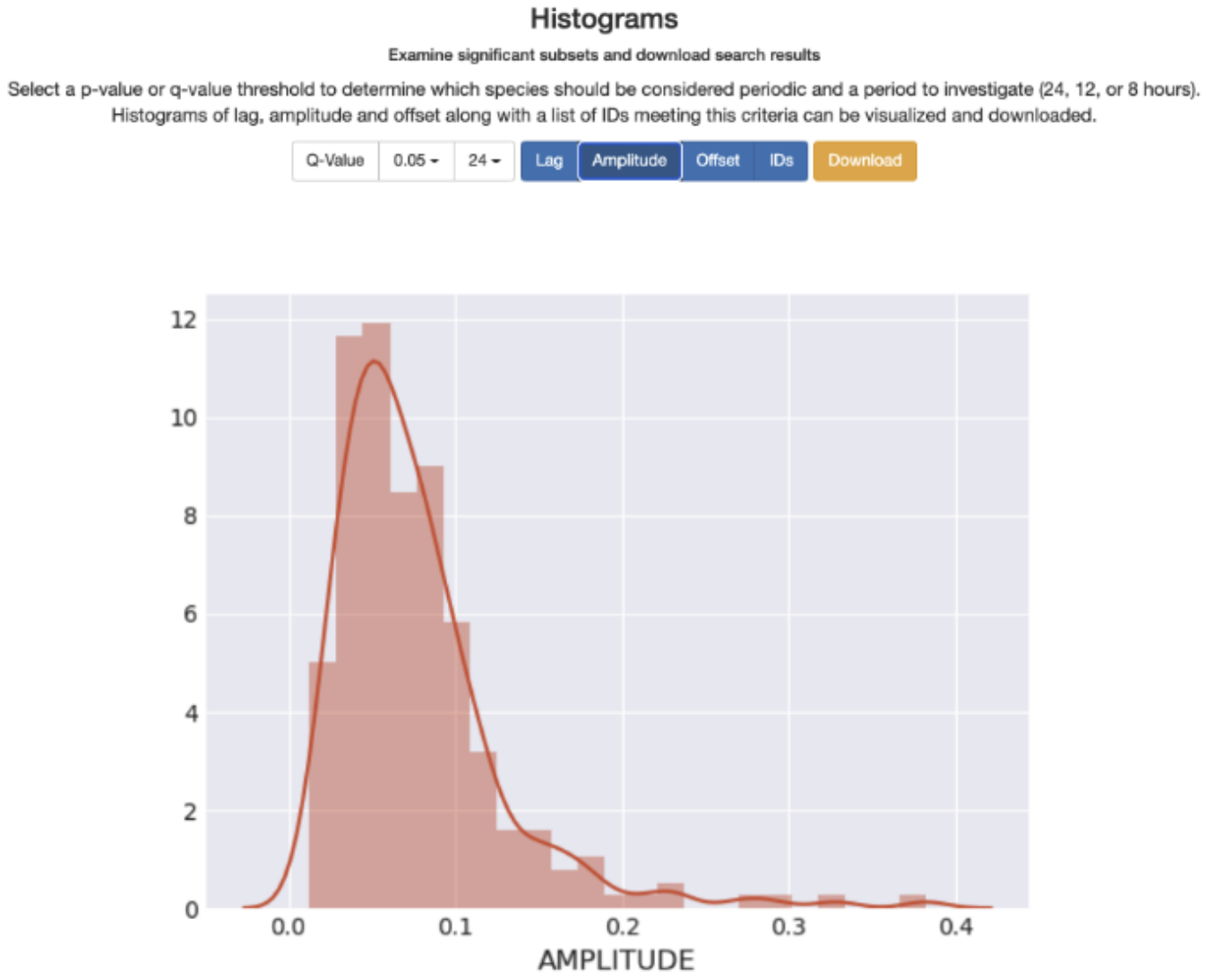

Figure 4.2: The BIO_CYCLE web server interface

previously included in CircadiOmics. In addition, the latest version of BIO_CYCLE, which has undergone significant improvement, has been made available via the CircadiOmics web portal. The improvements to BIO_CYCLE include: implementation in Python to take advantage of state-of-the-art deep learning software, ability to handle missing timepoints, improved amplitude estimation, the addition of offset estimation, and modeling real-world replicated experimental data to produce more realistic p-values. The previous version of BIO_CYCLE was implemented in R, which does not have convenient access to deep learning libraries that allow users to utilize Graphics Processing Units (GPUs) to increase the speed of training and testing DNNs. In the previous version, we were restricted to training on slower Central Processing Units (CPUs). As a result, we only trained a small 3 layer network with 100 hidden units per layer. Since the new version of BIO_CYCLE is implemented in Python, we take advantage of the PyTorch deep learning library to train significantly larger DNNs on GPUs [78]. The increased size of the DNNs substantially helps in handling missing data (e.g. missing replicates). The latest BIO_CYCLE also utilizes real-world experimental data available via CircadiOmics not only to evaluate performance and fairly compare the new BIO_CYCLE algorithm to other available algorithms, but also to better fine tune the algorithm and make p-value estimations more accurate. In combination, these new features allow researchers to perform end-to-end circadian analyses of their data and to compare and combine their data with other available datasets.

4.3.3 Applications

CircadiOmics has numerous and diverse applications. To name just a few: users can analyze a single dataset, analyze multiple datasets of the same omic type across different tissues or species,and analyze relationships between datasets of different omic types. This flexibility to perform comparative analyses has proven to be highly effective for biological discovery and hypothesis generation and as such has contributed to numerous studies that have been

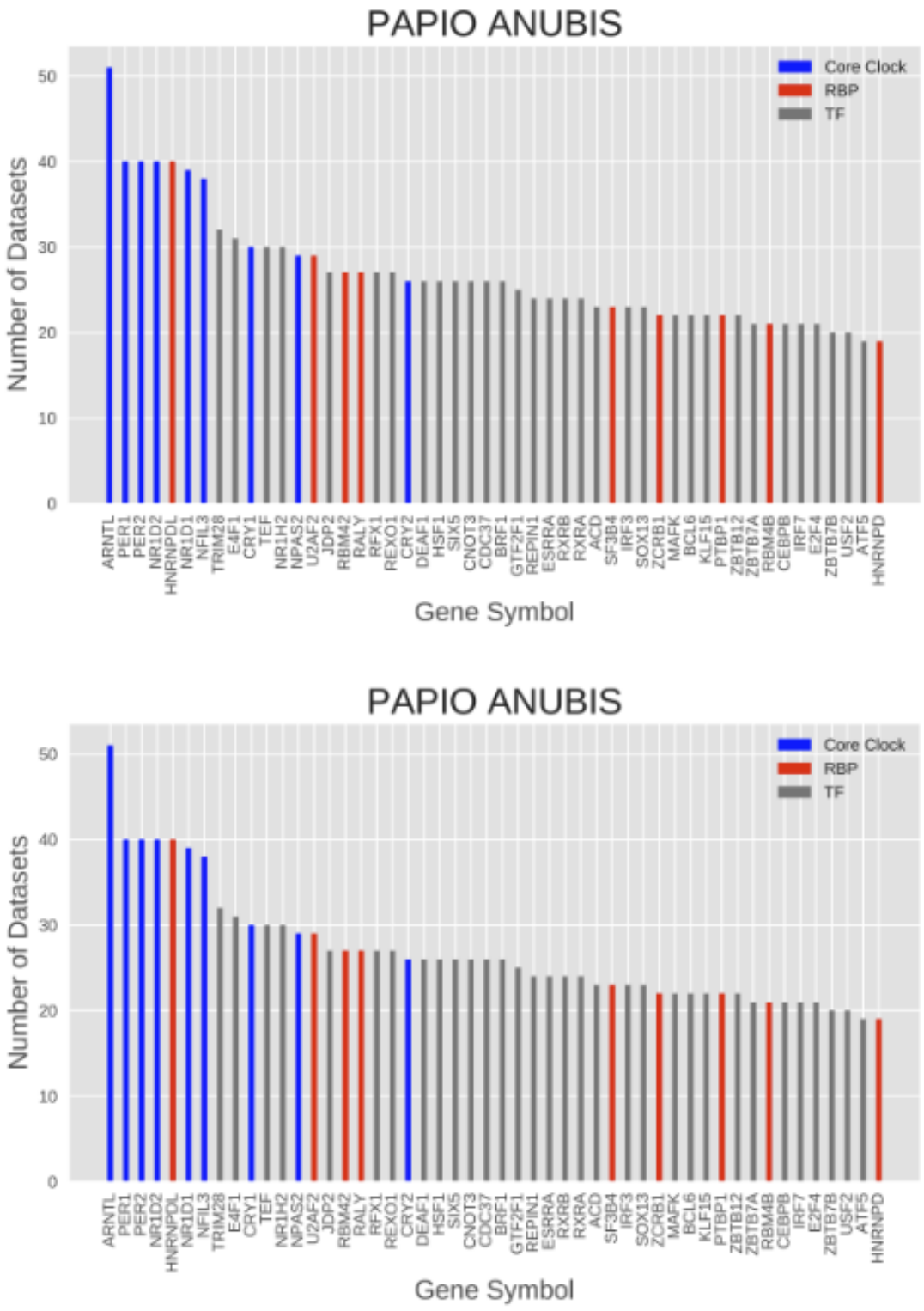

Figure 4.3: Frequency analysis rediscovers core clock as well as a few novel circadian regulatory TFs and RBPs. Highlighted genes are those validated in the in vivo experiments found in Figure 4.4.

published in high impact journals [79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94]. For example, in Koronowski et al. CircadiOmics was used to better understand the independence of the liver circadian clock. Using high-throughput transcriptomic and metabolomic data, they showed that the liver has independent circadian functions specific for metabolic processes, however full circadian function in the liver depends on signals from other clocks [95]. In Tognini et al. CircadiOmics was used to analyze metabolomic data in the suprachiasmatic nucleus (SCN) under various experimental conditions to discover a sensitivity of brain clocks to nutrition [96]. Finally, in Masri et al. CircadiOmics contributed to showing that lung cancer has no effect on the core clock but rather specifically reprograms hepatic metabolism, proving that a pathological condition in a given tissue can influence the circadian homeostasis in other tissues [97].

Additionally, we performed our own analysis using the data available via CircadiOmics in aggregate to better understand the overall hierarchical architecture of transcriptomic circadian regulation. To perform this analysis we looked at the frequency at which important regulators, such as transcription factors (TFs) or RNA-binding proteins (RBPs) are found to oscillate across all mouse and baboon datasets to quantify their importance in circadian regulation. The top oscillating TFs and RBPs in mice and baboons can be seen in Figure 4.3. We found that the circadian core clock appears with the highest frequency, and is closely followed by TFs and RBPs with known interactions to the core clock. Aside from the core clock, this analysis identified multiple TFs and RBPs important to circadian regulation, some of which are corroborated by evidence in the literature and others which are novel. We were able to validate some of the novel findings with animal experiments. For example, the TFs FUS and EIF4B were identified in our analyses as having the potential for being strong circadian regulators. Consistent with this result, Figure 4.4 shows that the reverse transcription-quantitative PCR (RT-qPCR) and western blot analyses showed that mRNA and protein levels of both of these genes have diurnal rhythm in the liver in certain experiments and FUS was also shown to have a diurnal rhythmicity in the SCN. In addi-

tion, the TF MXI1 was identified in our analysis as a novel circadian factor and RT-qPCR analyses were able to detect rhythmic gene expressions of MXI1 in mice livers (Figure 4.4). This analysis was repeated on a few other TFs to validate the computational experimental findings. In short, in vivo experiments confirmed the circadian expression of important genes predicted in computo by analyzing the data in CircadiOmics. Together, these findings show that CircadiOmics provides a strong foundation for understanding the organization of the circadian transcriptome on a large scale.

4.4 Conclusion

CircadiOmics allows users to seamlessly compare and analyze multiple omic time-series data sets simultaneously. For example, a user can compare transcripts across species or tissues, or map out relationships between metabolites, proteins, and transcripts to identify underlying oscillatory trends. CircadiOmics has proven to be highly effective for performing end-to-end circadian analyses from hypothesis generation to publication-ready figures creation. This web server has contributed to numerous studies that have been published in high impact journals and in aggregate has been cited in over 190 publications. The server receives approximately 1,000 queries per week from around the world and to the best of our knowledge is the largest single repository of circadian omic data available. With the quantity and breadth of its growing, high-quality, circadian omic data, CircadiOmics continues to be an invaluable resource for understanding the fundamental landscape of circadian rhythms and how these rhythms are programmed, and can be re-programmed, in cells, tissues, organs, and organisms with significant implications for medicine and therapeutic interventions.

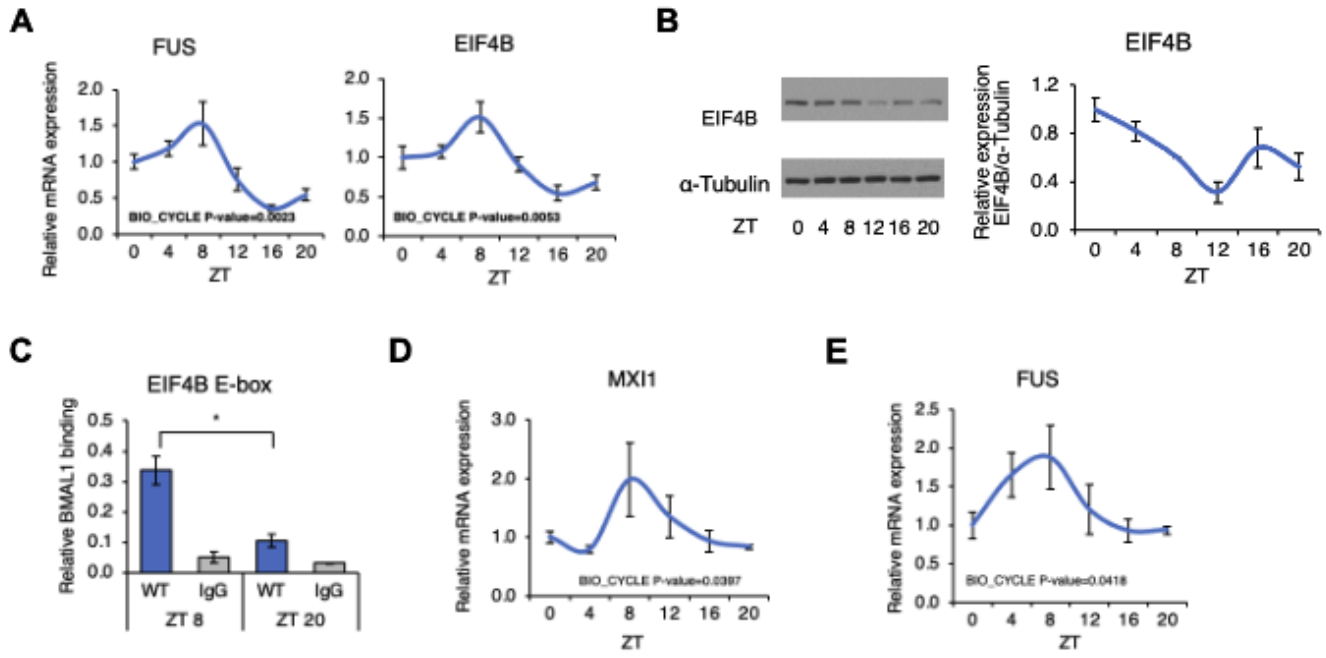

Figure 4.4: Validation of computational analysis results by in vivo experiments. Wild type (WT) mice samples were obtained under ad lib conditions. (A) RT-qPCR were used to determine expression of novel circadian factors detected by computational analysis in the mouse liver. The results are displayed as percent increase/decrease, from the level of mRNA expressed in the mice at ZT 0. (B) Daily rhythms in protein expression of EIF4B in the whole cell lysate from the liver (n=2). Representative image of immunoblot analysis of EIF4B are shown. Line graph shows quantification from EIF4B normalized to α-tubulin. Values are expressed as a percentage of the value for ZT 0. (C) Chromatin recruitment of BMAL1 at the E-box motif contained in the EIF4B promoter. ChIP-qPCR assays were done utilizing dual cross-linked livers at ZT 8 and 20 with antibodies against BMAL1 (n=3 at ZT 8, n=2 at ZT 20). $*p < 0.05$ in Student's t test. (D) RT-qPCR was used to determine mRNA expression of the novel circadian factors detected by computational analysis in the liver (n=5). The results are displayed as percent increase/decrease, from the level of mRNA expressed in the mice at ZT 0. (E) RT-qPCR was used to determine mRNA expression of novel circadian factors detected by computational analysis in the SCN (n=2 at ZT 0, n=3 at ZT 4, 8, 12, 16, 20). The results are displayed as percent increase/decrease, from the level of mRNA expressed in the mice at ZT 0.

Chapter 5

Conclusion

Deep learning has tremendous potential to greatly improve patient care and outcomes and improve our understanding of circadian rhythms. Despite numerous publications showing deep learning to be very successful in retrospective healthcare studies and molecular studies, there is a lack of clinical adoption due in part to a lack of trust in deep learning models. Our hope in developing an interpretable deep learning model, publishing MOVER, and publishing CircadiOmics is to help close the gap between development and deployment of deep learning models in healthcare and circadian omic analyses.

Bibliography

[1] Hill, B. L. et al. An automated machine learning-based model predicts postoperative mortality using readily-extractable preoperative electronic health record data. Br. J. Anaesth. 123, 877–886 (2019).

[2] Fritz, B. A. et al. Deep-learning model for predicting 30-day postoperative mortality. Br. J. Anaesth. 123, 688–695 (2019).

[3] Lee, C. K., Hofer, I., Gabel, E., Baldi, P. & Cannesson, M. Development and validation of a deep neural network model for prediction of postoperative inhospital mortality. Anesthesiology 129, 649–662 (2018).

[4] Xiao, C., Choi, E. & Sun, J. Opportunities and challenges in developing deep learning models using electronic health records data: a systematic review. J. Am. Med. Inform. Assoc. 25, 1419–1428 (2018).

[5] Vincent, J. L. et al. Perioperative cardiovascular monitoring of high-risk patients: a consensus of 12. Crit. Care 19, 224 (2015).

[6] Caruana, R. et al. Intelligible models for HealthCare. In Proc. 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining - KDD '15 1721–1730 (ACM Press, 2015), https://doi.org/10.1145/2783258.2788613.

[7] Potts, W. J. E. Generalized additive neural networks. In Proc. Fifth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining 194–200 (Association for Computing Machinery (ACM), NY, 1999), https://doi.org/10.1145/ 312129.312228.

[8] Brás-Geraldes, C., Papoila, A. & Xufre, P. Generalized additive neural network with flexible parametric link function: model estimation using simulated and real clinical data. Neural Comput. Appl. 31, 719–736 (2019).

[9] Baldi, P. Deep learning in biomedical data science. Annu. Rev. Biomed. Data Sci. 1, 181–205 (2018).

[10] Urban, G. et al. Deep learning localizes and identifies polyps in real time with 96% accuracy in screening colonoscopy. Gastroenterology 155, 1069–1078.e8 (2018).

[11] Chang, P. et al. Deep-learning convolutional neural networks accurately classify genetic mutations in gliomas. Am. J. Neuroradiol. 39, 1201–1207 (2018).

[12] Asan, O., Bayrak, A.E. & Choudhury, A. Artificial intelligence and human trust in healthcare: focus on clinicians. J. Med. Internet Res. 22, e15154 (2020).

[13] Tonekaboni, S., Joshi, S., McCradden, M.D. & Goldenberg, A. What clinicians want: contextualizing explainable machine learning for clinical end use. Proc. Mach. Learn. Res. 106, 359–380 (2019)

[14] Ginestra, J. C. et al. Clinician perception of a machine learning-based early warning system designed to predict severe sepsis and septic shock. Crit. Care Med. 47, 1477–1484 (2019).

[15] Hofer, I. S., Gabel, E., Pfeffer, M., Mahbouba, M. & Mahajan, A. A systematic approach to creation of a perioperative data warehouse. Anesth. Analg. 122, 1880–1884 (2016).

[16] Luo, W. et al. Guidelines for developing and reporting machine learning predictive models in biomedical research: a multidisciplinary view. J. Med. Internet Res. 18, e323 (2016).

[17] Kingma, D. P. & Lei Ba, J. Adam: a method for stochastic optimization. Preprint at https://arxiv.org/abs/1412.6980.

[18] Srivastava, N., Hinton, G., Krizhevsky, A. & Salakhutdinov, R. Dropout: a simple way to prevent neural networks from overfitting. J. Mach. Learn. Res. 15, 1929–1958 (2014).

[19] Baldi, P. & Sadowski, P. The dropout learning algorithm. Artif. Intell. 210, 78–122 (2014).

[20] Cholett, F. Keras. https://keras.io (2015). Accessed 19 Dec 2020.

[21] Pedregosa, F. et al. Scikit-learn: machine learning in Python. J. Mach. Learn. Res. 12, 2825–2830 (2011).

[22] Everson, Jordan, et al. “Reconsidering Hospital Ehr Adoption at the Dawn of Hitech: Implications of the Reported 9% Adoption of a ‘Basic’ EHR.” *Journal of the American Medical Informatics Association*, vol. 27, no. 8, 2020, pp. 1198–1205., https://doi.org/10.1093/jamia/ocaa090.

[23] Adler-Milstein, Julia, et al. “Electronic Health Record Adoption in US Hospitals: The Emergence of a Digital ‘Advanced Use’ Divide.” *Journal of the American Medical Informatics Association*, vol. 24, no. 6, 2017, pp. 1142–1148., https://doi.org/10.1093/jamia/ocx080.

[24] Davenport, Thomas, and Ravi Kalakota. “The Potential for Artificial Intelligence in Healthcare.” *Future Healthcare Journal*, vol. 6, no. 2, 2019, pp. 94–98., https://doi.org/10.7861/futurehosp.6-2-94.

[25] Lee, Christine K., et al. “Development and Validation of a Deep Neural Network Model for Prediction of Postoperative in-Hospital Mortality.” *Anesthesiology*, vol. 129, no. 4, 2018, pp. 649–662., https://doi.org/10.1097/aln.0000000000002186.

[26] Lilot, M., et al. “Variability in Practice and Factors Predictive of Total Crystalloid Administration during Abdominal Surgery: Retrospective Two-Centre Analysis.” *British Journal of Anaesthesia*, vol. 114, no. 5, 2015, pp. 767–776., https://doi.org/10.1093/bja/aeu452.

[27] Hill, Brian L., et al. “An Automated Machine Learning-Based Model Predicts Postoperative Mortality Using Readily-Extractable Preoperative Electronic Health Record Data.” *British*

Journal of Anaesthesia, vol. 123, no. 6, 2019, pp. 877–886., https://doi.org/10.1016/j.bja.2019.07.030.

[28] Lee, Christine K., et al. "Development and Validation of an Interpretable Neural Network for Prediction of Postoperative in-Hospital Mortality." *Npj Digital Medicine*, vol. 4, no. 1, 2021, https://doi.org/10.1038/s41746-020-00377-1.

[29] Cannesson, Maxime, et al. "Machine Learning of Physiological Waveforms and Electronic Health Record Data to Predict, Diagnose and Treat Haemodynamic Instability in Surgical Patients: Protocol for a Retrospective Study." *BMJ Open*, vol. 9, no. 12, 2019, https://doi.org/10.1136/bmjopen-2019-031988.

[30] Kopitar, Leon, et al. "Early Detection of Type 2 Diabetes Mellitus Using Machine Learning-Based Prediction Models." *Scientific Reports*, vol. 10, no. 1, 2020, https://doi.org/10.1038/s41598-020-68771-z.

[31] Datta, Arghya, et al. "Machine Learning Liver-Injuring Drug Interactions with Non-Steroidal Anti-Inflammatory Drugs (NSAIDs) from a Retrospective Electronic Health Record (EHR) Cohort." *PLOS Computational Biology*, vol. 17, no. 7, 2021, https://doi.org/10.1371/journal.pcbi.1009053.

[32] Lauritsen, Simon Meyer, et al. "Explainable Artificial Intelligence Model to Predict Acute Critical Illness from Electronic Health Records." *Nature Communications*, vol. 11, no. 1, 2020, https://doi.org/10.1038/s41467-020-17431-x.

[33] Wiens, J. et al. Do no harm: a roadmap for responsible machine learning for health care. *Nat. Med.* 25, 1337–1340 (2019).

[34] Beam, A. L., Manrai, A. K. & Ghassemi, M. Challenges to the reproducibility of machine learning models in health care. *J. Am. Med. Assoc.* 323, 305–306 (2020).

[35] Gerke, S., Babic, B., Evgeniou, T. & Cohen, I. G. The need for a system view to regulate artificial intelligence/machine learning-based software as medical device. *NPJ Digit. Med.* 3, 53 (2020).

[36] Johnson, A., Bulgarelli, L., Pollard, T., Horng, S., Celi, L. A., & Mark, R. (2020). MIMIC-IV (version 0.4). *PhysioNet*. https://doi.org/10.13026/a3wn-hq05.

[37] Dua, D. and Graff, C. (2019). UCI Machine Learning Repository [http://archive.ics.uci.edu/ml]. Irvine, CA: University of California, School of Information and Computer Science.

[38] "DBMI Data Portal." *DBMI Portal*, https://portal.dbmi.hms.harvard.edu/.

[39] Bijker, Jilles B., et al. “Incidence of Intraoperative Hypotension as a Function of the Chosen Definition.” *Anesthesiology*, vol. 107, no. 2, 2007, pp. 213–220.

[40] Eckel-Mahan, K., & Sassone-Corsi, P. (2009). Metabolism control by the circadian clock and vice versa. *Nature Structural & Molecular Biology*, *16*(5), 462–467. https://doi.org/10.1038/nsmb.1595

[41] Froy, O. (2011). Circadian rhythms, aging, and life span in mammals. *Physiology*, *26*(4), 225–235. https://doi.org/10.1152/physiol.00012.2011

[42] Panda, S., Hogenesch, J. B., & Kay, S. A. (2002). Circadian rhythms from flies to human. *Nature*, *417*(6886), 329–335. https://doi.org/10.1038/417329a

[43] Patel, V. R., Ceglia, N., Zeller, M., Eckel-Mahan, K., Sassone-Corsi, P., & Baldi, P. (2015). The pervasiveness and plasticity of circadian oscillations: The coupled circadian-oscillators framework. *Bioinformatics*, *31*(19), 3181–3188. https://doi.org/10.1093/bioinformatics/btv353

[44] Karlsson, B. (2001). Is there an association between shift work and having a metabolic syndrome? results from a population based study of 27 485 people. *Occupational and Environmental Medicine*, *58*(11), 747–752. https://doi.org/10.1136/oem.58.11.747

[45] Antunes, L. C., Levandovski, R., Dantas, G., Caumo, W., & Hidalgo, M. P. (2010). Obesity and shift work: Chronobiological Aspects. *Nutrition Research Reviews*, *23*(1), 155–168. https://doi.org/10.1017/s0954422410000016

[46] Froy, O. (2009). Metabolism and circadian rhythms—implications for obesity. *Endocrine Reviews*, *31*(1), 1–24. https://doi.org/10.1210/er.2009-0014

[47] Kohsaka, A., Laposky, A. D., Ramsey, K. M., Estrada, C., Joshu, C., Kobayashi, Y., Turek, F. W., & Bass, J. (2007). High-fat diet disrupts behavioral and molecular circadian rhythms in mice. *Cell Metabolism*, *6*(5), 414–421. https://doi.org/10.1016/j.cmet.2007.09.006

[48] Sharifian, A., Farahani, S., Pasalar, P., Gharavi, M., & Aminian, O. (2005). Shift work as an oxidative stressor. *Journal of Circadian Rhythms*, *3*, 15. https://doi.org/10.1186/1740-3391-3-15

[49] Turek, F. W., Joshu, C., Kohsaka, A., Lin, E., Ivanova, G., McDearmon, E., Laposky, A., Losee-Olson, S., Easton, A., Jensen, D. R., Eckel, R. H., Takahashi, J. S., & Bass, J. (2005). Obesity and metabolic syndrome in circadian clock mutant mice. *Science*, *308*(5724), 1043–1045. https://doi.org/10.1126/science.1108750

[50] Shi, S.-qun, Ansari, T. S., McGuinness, O. P., Wasserman, D. H., & Johnson, C. H. (2013). Circadian disruption leads to insulin resistance and obesity. *Current Biology*, *23*(5), 372–381. https://doi.org/10.1016/j.cub.2013.01.048

[51] Kondratov, R. V., Kondratova, A. A., Gorbacheva, V. Y., Vykhovanets, O. V., & Antoch, M. P. (2006). Early aging and age-related pathologies in mice deficient in BMAL1, the core Componentof the circadian clock. *Genes & Development*, *20*(14), 1868–1873. https://doi.org/10.1101/gad.1432206

[52] Dibner, C., Schibler, U., & Albrecht, U. (2010). The mammalian circadian timing system: Organization and coordination of Central and peripheral clocks. *Annual Review of Physiology*, *72*(1), 517–549. https://doi.org/10.1146/annurev-physiol-021909-135821

[53] Koike, N., Yoo, S.-H., Huang, H.-C., Kumar, V., Lee, C., Kim, T.-K., & Takahashi, J. S. (2012). Transcriptional architecture and chromatin landscape of the core circadian clock in mammals. *Science*, *338*(6105), 349–354. https://doi.org/10.1126/science.1226339

[54] Ko, C. H., & Takahashi, J. S. (2006). Molecular components of the mammalian circadian clock. *Human Molecular Genetics*, *15*(suppl_2). https://doi.org/10.1093/hmg/ddl207

[55] Hughes, M. E., DiTacchio, L., Hayes, K. R., Vollmers, C., Pulivarthy, S., Baggs, J. E., Panda, S., & Hogenesch, J. B. (2009). Harmonics of circadian gene transcription in mammals. *PLoS Genetics*, *5*(4). https://doi.org/10.1371/journal.pgen.1000442

[56] Eckel-Mahan, K. L., Patel, V. R., de Mateo, S., Orozco-Solis, R., Ceglia, N. J., Sahar, S., Dilag-Penilla, S. A., Dyar, K. A., Baldi, P., & Sassone-Corsi, P. (2013). Reprogramming of the circadian clock by Nutritional Challenge. *Cell*, *155*(7), 1464–1478. https://doi.org/10.1016/j.cell.2013.11.034

[57] Eckel-Mahan, K. L., Patel, V. R., Mohney, R. P., Vignola, K. S., Baldi, P., & Sassone-Corsi, P. (2012). Coordination of the transcriptome and metabolome by the circadian clock. *Proceedings of the National Academy of Sciences*, *109*(14), 5541–5546. https://doi.org/10.1073/pnas.1118726109

[58] Panda, S., Antoch, M. P., Miller, B. H., Su, A. I., Schook, A. B., Straume, M., Schultz, P. G., Kay, S. A., Takahashi, J. S., & Hogenesch, J. B. (2002). Coordinated transcription of key pathways in the mouse by the circadian clock. *Cell*, *109*(3), 307–320. https://doi.org/10.1016/s0092-8674(02)00722-5

[59] Andrews, J. L., Zhang, X., McCarthy, J. J., McDearmon, E. L., Hornberger, T. A., Russell, B., Campbell, K. S., Arbogast, S., Reid, M. B., Walker, J. R., Hogenesch, J. B., Takahashi, J. S., & Esser, K. A. (2010). Clock and BMAL1 regulate MyoD and are necessary for maintenance of skeletal muscle phenotype and function. *Proceedings of the National Academy of Sciences*, *107*(44), 19090–19095. https://doi.org/10.1073/pnas.1014523107

[60] Masri, S., Rigor, P., Cervantes, M., Ceglia, N., Sebastian, C., Xiao, C., Roqueta-Rivera, M., Deng, C., Osborne, T. F., Mostoslavsky, R., Baldi, P., & Sassone-Corsi, P. (2014). Partitioning circadian transcription by SIRT6 leads to segregated control of cellular metabolism. *Cell*, *158*(3), 659–672. https://doi.org/10.1016/j.cell.2014.06.050

[61] Miller, B. H., McDearmon, E. L., Panda, S., Hayes, K. R., Zhang, J., Andrews, J. L., Antoch, M. P., Walker, J. R., Esser, K. A., Hogenesch, J. B., & Takahashi, J. S. (2007). Circadian and clock-controlled regulation of the mouse transcriptome and cell proliferation. *Proceedings of the National Academy of Sciences*, *104*(9), 3342–3347. https://doi.org/10.1073/pnas.0611724104

[62] Storch, K.-F., Lipan, O., Leykin, I., Viswanathan, N., Davis, F. C., Wong, W. H., & Weitz, C. J. (2002). Extensive and divergent circadian gene expression in liver and heart. *Nature*, *417*(6884), 78–83. https://doi.org/10.1038/nature744

[63] Yan, J., Wang, H., Liu, Y., & Shao, C. (2008). Analysis of gene regulatory networks in the mammalian Circadian Rhythm. *PLoS Computational Biology*, *4*(10). https://doi.org/10.1371/journal.pcbi.1000193

[64] Azzi, A., Dallmann, R., Casserly, A., Rehrauer, H., Patrignani, A., Maier, B., Kramer, A., & Brown, S. A. (2014). Circadian behavior is light-reprogrammed by plastic DNA methylation. *Nature Neuroscience*, *17*(3), 377–382. https://doi.org/10.1038/nn.3651

[65] Haspel, J. A., Chettimada, S., Shaik, R. S., Chu, J.-H., Raby, B. A., Cernadas, M., Carey, V., Process, V., Hunninghake, G. M., Ifedigbo, E., Lederer, J. A., Englert, J., Pelton, A., Coronata, A., Fredenburgh, L. E., & Choi, A. M. K. (2014). Circadian rhythm reprogramming during lung inflammation. *Nature Communications*, *5*(1). https://doi.org/10.1038/ncomms5753

[66] Li, X.-M., Delaunay, F., Dulong, S., Claustrat, B., Zampera, S., Fujii, Y., Teboul, M., Beau, J., & Lévi, F. (2010). Cancer inhibition through circadian reprogramming of tumor transcriptome with meal timing. *Cancer Research*, *70*(8), 3351–3360. https://doi.org/10.1158/0008-5472.can-09-4235

[67] Murakami, M., Tognini, P., Liu, Y., Eckel‐Mahan, K. L., Baldi, P., & Sassone‐Corsi, P. (2016). Gut microbiota directs PPAR γ‐driven reprogramming of the liver circadian clock by nutritional challenge. *EMBO Reports*, *17*(9), 1292–1303. https://doi.org/10.15252/embr.201642463

[68] Bellet, M. M., Deriu, E., Liu, J. Z., Grimaldi, B., Blaschitz, C., Zeller, M., Edwards, R. A., Sahar, S., Dandekar, S., Baldi, P., George, M. D., Raffatellu, M., & Sassone-Corsi, P. (2013). Circadian clock regulates the host response to salmonella. *Proceedings of the National Academy of Sciences*, *110*(24), 9897–9902. https://doi.org/10.1073/pnas.1120636110

[69] Gaucher, J., Kinouchi, K., Ceglia, N., Montellier, E., Peleg, S., Greco, C. M., Schmidt, A., Forne, I., Masri, S., Baldi, P., Imhof, A., & Sassone-Corsi, P. (2019). Distinct metabolic adaptation of liver circadian pathways to acute and chronic patterns of alcohol intake. *Proceedings of the National Academy of Sciences*, *116*(50), 25250–25259. https://doi.org/10.1073/pnas.1911189116

[70] Ceglia, N., Liu, Y., Chen, S., Agostinelli, F., Eckel-Mahan, K., Sassone-Corsi, P., & Baldi, P. (2018). Circadiomics: Circadian Omic Web Portal. *Nucleic Acids Research*, *46*(W1). https://doi.org/10.1093/nar/gky441

[71] Pizarro, A., Hayer, K., Lahens, N. F., & Hogenesch, J. B. (2012). CircaDB: A database of mammalian circadian gene expression profiles. *Nucleic Acids Research*, *41*(D1). https://doi.org/10.1093/nar/gks1161

[72] Li, D., Yang, R., Miao, Z., and Tao, W. (2010) BioClock: a web server and database aimed for interpreting circadian rhythm. Intelligent Systems for Molecular Biology 2010 meeting

[73] Li, X., Shi, L., Zhang, K., Wei, W., Liu, Q., Mao, F., Li, J., Cai, W., Chen, H., Teng, H., Li, J., & Sun, Z. (2017). CirGRDB: A database for the genome-wide deciphering circadian genes and regulators. *Nucleic Acids Research*, *46*(D1). https://doi.org/10.1093/nar/gkx944

[74] Hokamp, K., & Wolfe, K. H. (2004). PubCrawler: Keeping up comfortably with pubmed and GenBank. *Nucleic Acids Research*, *32*(Web Server). https://doi.org/10.1093/nar/gkh453

[75] Agostinelli, F., Ceglia, N., Shahbaba, B., Sassone-Corsi, P., & Baldi, P. (2016). What time is it? deep learning approaches for circadian rhythms. *Bioinformatics*, *32*(19), 3051–3051. https://doi.org/10.1093/bioinformatics/btw504

[76] Hughes, M. E., Hogenesch, J. B., & Kornacker, K. (2010). JTK_CYCLE: An efficient nonparametric algorithm for detecting rhythmic components in genome-scale data sets. *Journal of Biological Rhythms*, *25*(5), 372–380. https://doi.org/10.1177/0748730410379711

[77] Kanehisa, M. (2000). Kegg: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Research*, *28*(1), 27–30. https://doi.org/10.1093/nar/28.1.27

[78] Paszke A, Gross S, Massa F, Lerer A, Bradbury J, Chanan G, Killeen T, Lin Z, Gimelshein N, Antiga L et al. (2019) Pytorch: An imperative style, high-performance deep learning library. In: *Advances in neural information processing systems*. pp. 8026–8037.

[79] Alhassen, S., Chen, S., Alhassen, L., Phan, A., Khoudari, M., De Silva, A., Barhoosh, H., Wang, Z., Parrocha, C., Shapiro, E., Henrich, C., Wang, Z., Mutesa, L., Baldi, P., Abbott, G. W., & Alachkar, A. (2021). Intergenerational trauma transmission is associated with brain metabotranscriptome remodeling and mitochondrial dysfunction. *Communications Biology*, *4*(1). https://doi.org/10.1038/s42003-021-02255-2

[80] Baldi, P., Alhassen, W., Chen, S., Nguyen, H., Khoudari, M., & Alachkar, A. (2021). Large scale analysis reveals spatiotemporal circadian patterns of cilia transcriptomes in the primate brain. *Journal of Neuroscience Research*, *99*(10), 2610–2624. https://doi.org/10.1002/jnr.24919

[81] Greco, C. M., Koronowski, K. B., Smith, J. G., Shi, J., Kunderfranco, P., Carriero, R., Chen, S., Samad, M., Welz, P.-S., Zinna, V. M., Mortimer, T., Chun, S. K., Shimaji, K., Sato, T., Petrus, P., Kumar, A., Vaca-Dempere, M., Deryagin, O., Van, C., … Sassone-Corsi, P. (2021). Integration of feeding behavior by the liver circadian clock reveals network dependency of metabolic rhythms. *Science Advances*, *7*(39). https://doi.org/10.1126/sciadv.abi7828

[82] Greco, C. M., Cervantes, M., Fustin, J.-M., Ito, K., Ceglia, N., Samad, M., Shi, J., Koronowski, K. B., Forne, I., Ranjit, S., Gaucher, J., Kinouchi, K., Kojima, R., Gratton, E., Li, W., Baldi, P., Imhof, A., Okamura, H., & Sassone-Corsi, P. (2020). S-adenosyl- L -homocysteine hydrolase links methionine metabolism to the circadian clock and chromatin remodeling. *Science Advances*, *6*(51). https://doi.org/10.1126/sciadv.abc5629

[83] Napoli, D., Lupori, L., Mazziotti, R., Sagona, G., Bagnoli, S., Samad, M., Sacramento, E. K., Kirkpartick, J., Putignano, E., Chen, S., Terzibasi Tozzini, E., Tognini, P., Baldi, P., Kwok, J. C.

F., Cellerino, A., & Pizzorusso, T. (2020). Mir-29 coordinates age-dependent plasticity brakes in the adult visual cortex. *EMBO Reports*, *21*(11). https://doi.org/10.15252/embr.202050431

[84] Sato, S., Dyar, K., Treebak, J. T., Linde Basse, A., Schönke, M., Chen, S., Samad, M., Baldi, P., Lutter, D., Zierath, J., & Sassone-Corsi, P. (2020). Atlas of Exercise Metabolism reveals time-dependent signatures of metabolic homeostasis. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3707270

[85] Sato, S., Basse, A. L., Schönke, M., Chen, S., Samad, M., Altıntaş, A., Laker, R. C., Dalbram, E., Barrès, R., Baldi, P., Treebak, J. T., Zierath, J. R., & Sassone-Corsi, P. (2019). Time of exercise specifies the impact on muscle metabolic pathways and systemic energy homeostasis. *Cell Metabolism*, *30*(1). https://doi.org/10.1016/j.cmet.2019.03.013

[86] Brami-Cherrier, K., Lewis, R. G., Cervantes, M., Liu, Y., Tognini, P., Baldi, P., Sassone-Corsi, P., & Borrelli, E. (2020). Cocaine-mediated circadian reprogramming in the striatum through dopamine D2R and PPARΓ activation. *Nature Communications*, *11*(1). https://doi.org/10.1038/s41467-020-18200-6

[87] Dyar, K. A., Lutter, D., Artati, A., Ceglia, N. J., Liu, Y., Armenta, D., Jastroch, M., Schneider, S., de Mateo, S., Cervantes, M., Abbondante, S., Tognini, P., Orozco-Solis, R., Kinouchi, K., Wang, C., Swerdloff, R., Nadeef, S., Masri, S., Magistretti, P., … Sassone-Corsi, P. (2018). Atlas of circadian metabolism reveals system-wide coordination and communication between clocks. *Cell*, *174*(6). https://doi.org/10.1016/j.cell.2018.08.042

[88] Kwapis, J. L., Alaghband, Y., Kramár, E. A., López, A. J., Ciernia, A. V., White, A. O., Shu, G., Rhee, D., Michael, C. M., Montellier, E., Liu, Y., Magnan, C. N., Sassone-Corsi, P., Baldi, P., Matheos, D. P., & Wood, M. A. (2018). Epigenetic regulation of the circadian gene per1in the hippocampus mediates age-related changes in memory and synaptic plasticity. https://doi.org/10.1101/301135

[89] Masri, S., Patel, V. R., Eckel-Mahan, K. L., Peleg, S., Forne, I., Ladurner, A. G., Baldi, P., Imhof, A., & Sassone-Corsi, P. (2013). Circadian acetylome reveals regulation of mitochondrial metabolic pathways. *Proceedings of the National Academy of Sciences*, *110*(9), 3339–3344. https://doi.org/10.1073/pnas.1217632110

[90] Sato, S., Solanas, G., Peixoto, F. O., Bee, L., Symeonidi, A., Schmidt, M. S., Brenner, C., Masri, S., Benitah, S. A., & Sassone-Corsi, P. (2017). Circadian reprogramming in the liver identifies metabolic pathways of aging. *Cell*, *170*(4). https://doi.org/10.1016/j.cell.2017.07.042

[91] Asher, G., & Sassone-Corsi, P. (2015). Time for Food: The intimate interplay between nutrition, metabolism, and the circadian clock. *Cell*, *161*(1), 84–92. https://doi.org/10.1016/j.cell.2015.03.015

[92] Costello, Z., & Martin, H. G. (2018). A machine learning approach to predict metabolic pathway dynamics from time-series Multiomics Data. *Npj Systems Biology and Applications*, *4*(1). https://doi.org/10.1038/s41540-018-0054-3

[93] Braun, R., Kath, W. L., Iwanaszko, M., Kula-Eversole, E., Abbott, S. M., Reid, K. J., Zee, P. C., & Allada, R. (2018). Universal method for robust detection of circadian state from Gene Expression. *Proceedings of the National Academy of Sciences*, *115*(39). https://doi.org/10.1073/pnas.1800314115

[94] Hughey, J. J., Hastie, T., & Butte, A. J. (2016). Zeitzeiger: Supervised learning for high-dimensional data from an oscillatory system. *Nucleic Acids Research*, *44*(8). https://doi.org/10.1093/nar/gkw030

[95] Koronowski, K. B., Kinouchi, K., Welz, P.-S., Smith, J. G., Zinna, V. M., Shi, J., Samad, M., Chen, S., Magnan, C. N., Kinchen, J. M., Li, W., Baldi, P., Benitah, S. A., & Sassone-Corsi, P. (2019). Defining the independence of the liver circadian clock. *Cell*, *177*(6). https://doi.org/10.1016/j.cell.2019.04.025

[96] Tognini, P., Samad, M., Kinouchi, K., Liu, Y., Helbling, J.-C., Moisan, M.-P., Eckel-Mahan, K. L., Baldi, P., & Sassone-Corsi, P. (2020). Reshaping circadian metabolism in the suprachiasmatic nucleus and prefrontal cortex by nutritional challenge. *Proceedings of the National Academy of Sciences*, *117*(47), 29904–29913. https://doi.org/10.1073/pnas.2016589117

[97] Masri, S., Papagiannakopoulos, T., Kinouchi, K., Liu, Y., Cervantes, M., Baldi, P., Jacks, T., & Sassone-Corsi, P. (2016). Lung adenocarcinoma distally rewires hepatic circadian homeostasis. *Cell*, *165*(4), 896–909. https://doi.org/10.1016/j.cell.2016.04.039