Computational Sciences at the FDA

Stephen J. Merrill
Department MSCS and
Visiting Professor at
FDA/CDRH/OSEL/DB

Organization of The FDA

- Divided into "Centers"
  - Center for Biologic Evaluation and Research (CBER)
  - Center for Devices and Radiological Health (CDRH) *
  - Center for Drug Evaluation and Research (CDER) *
  - Center for Food Safety and Applied Nutrition (CFSAN) food and cosmetics
  - Center for Tobacco Products (CTP) – started in 2009
  - Center for Veterinary Medicine (CVM)
  - National Center for Toxicology Research (NCTR) – in Arkansas (think Bill Clinton as President)
- And enforcement "Offices"
  - Office of Regulatory Affairs (ORA)
  - Office of Criminal Investigations (OCI)

Centers are divided into “Offices” (think Colleges) and Offices into Divisions (think Departments).

Since April of 2012, and continuing through August, 2014, I have been serving as a visiting professor in the Center for Devices and Radiological Health (CDRH) in the Office of Scientific and Engineering Laboratories (OSEL) in the Division of Biology (DB).

Federal Food, Drug, and Cosmetic Act (FD&C) -- 1938

- Provided the authority for the FDA (which was founded in 1906 to insure public hygiene) to oversee the safety of food, drugs, and cosmetics
- Amended in 1968 to include Electronic Product Radiation Control
- In 1968, the FDA started the Drug Safety Implementation to incorporate recommendations from the National Academy of Sciences who had investigated the effectiveness of drugs on the market at that time.
- In 1976, FD&C Act was amended to include regulation for medical devices. In 1982, CDRH was founded to regulate devices ("from Band-Aids to MRI machines").


A STRATEGIC PLAN
AUGUST 2011
www.fda.gov/regulatoryscience
Advancing Regulatory Science at FDA

FDA Science Priority Areas

- 1. Modernize Toxicology to Enhance Product Safety *
- 2. Stimulate Innovation in Clinical Evaluations and Personalized Medicine to improve Product Development and Patient Outcomes
- 3. Support New Approaches to Improve Product Manufacturing and Quality *
- 4. Ensure FDA Readiness to Evaluate Innovative Emerging Technologies
- 5. Harness Diverse Data through Information Sciences to Improve Health Outcomes
- 6. Implement a New Prevention-Focused Food Safety System to Protect Public Health
- 7. Facilitate Development of Medical Countermeasures to Protect Against Threats to U.S. and Global Health and Security *
- 8. Strengthen Social and Behavioral Science to Help Consumers and Professionals Make Informed Decisions about Regulated Product

More FDA Details

- Headquarters of the FDA are in White Oak, MD (which is also the home of CDER, CDRH, and soon CBER) between Baltimore and Washington, DC. The campus was originally the site of the Naval Ordnance Laboratory and is now named the Federal Research Center at White Oak.
- About 7000 will work in the complex by this Fall.

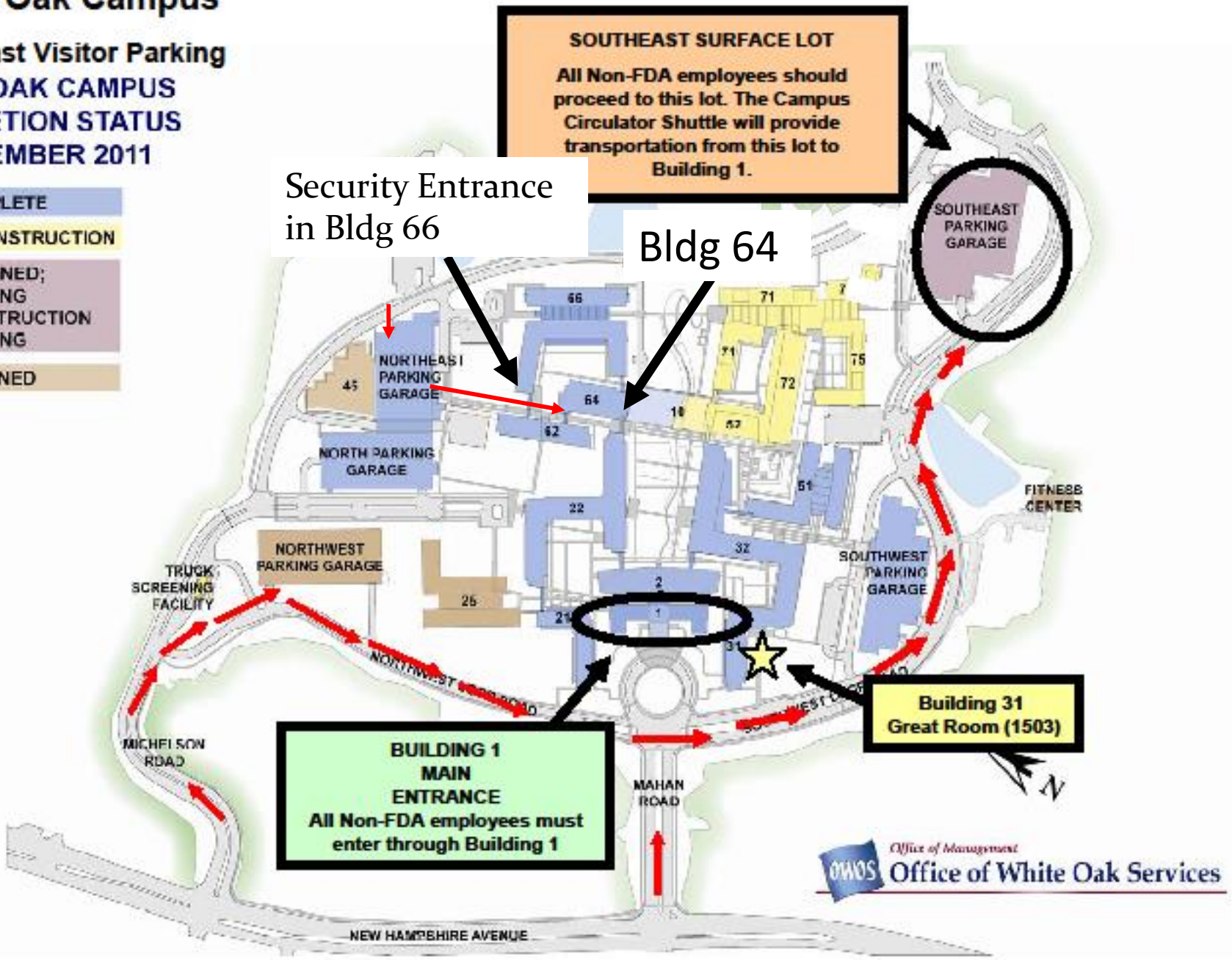

FDA
U.S. Department of Health and Human Services
Food and Drug Administration

White Oak Campus

Southeast Visitor Parking

WHITE OAK CAMPUS
COMPLETION STATUS
SEPTEMBER 2011

- COMPLETE
- IN CONSTRUCTION
- DESIGNED; PENDING CONSTRUCTION FUNDING
- PLANNED



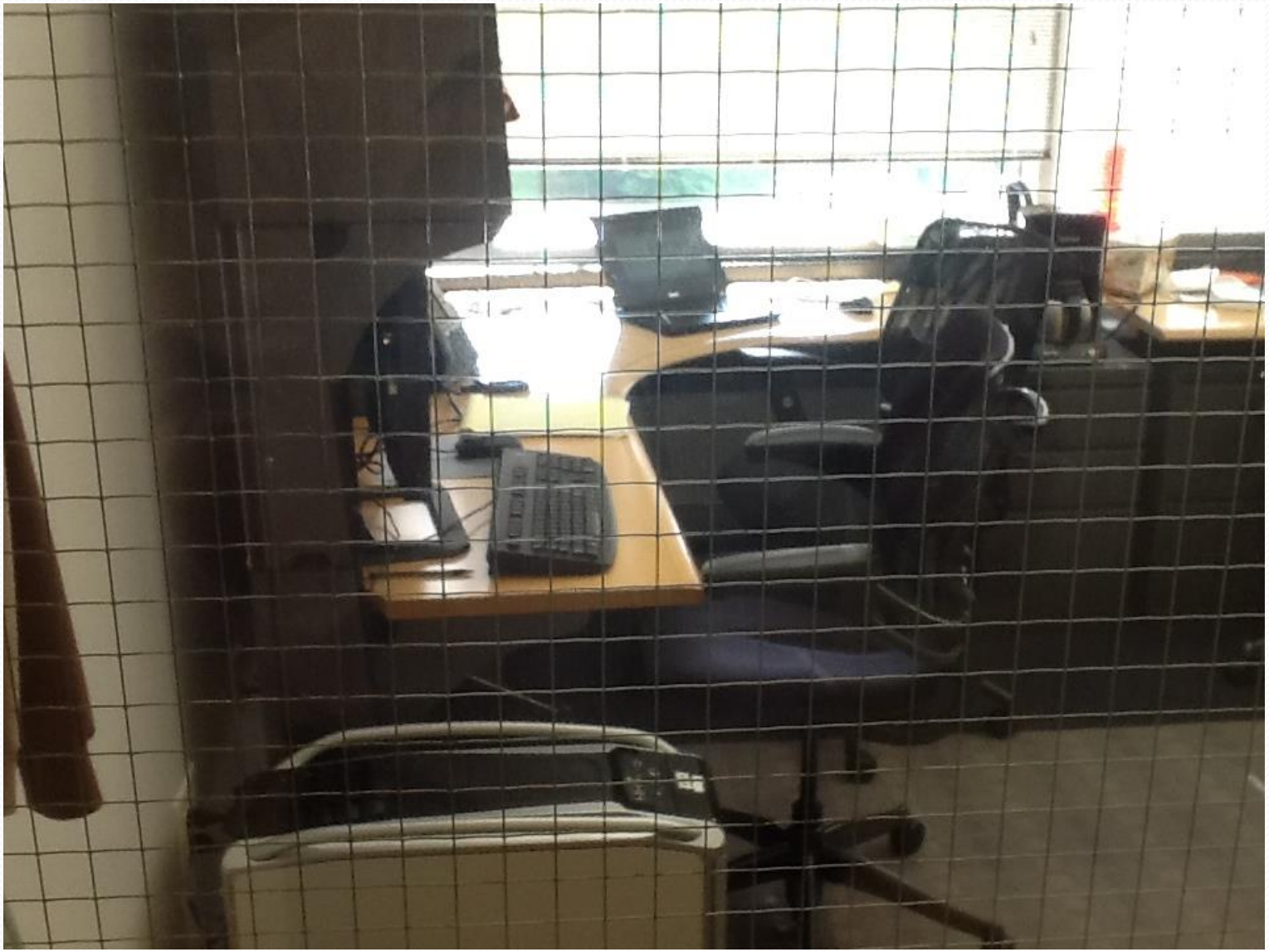
Office of Management
Office of White Oak Services

Building 64 (Life Sciences Lab)

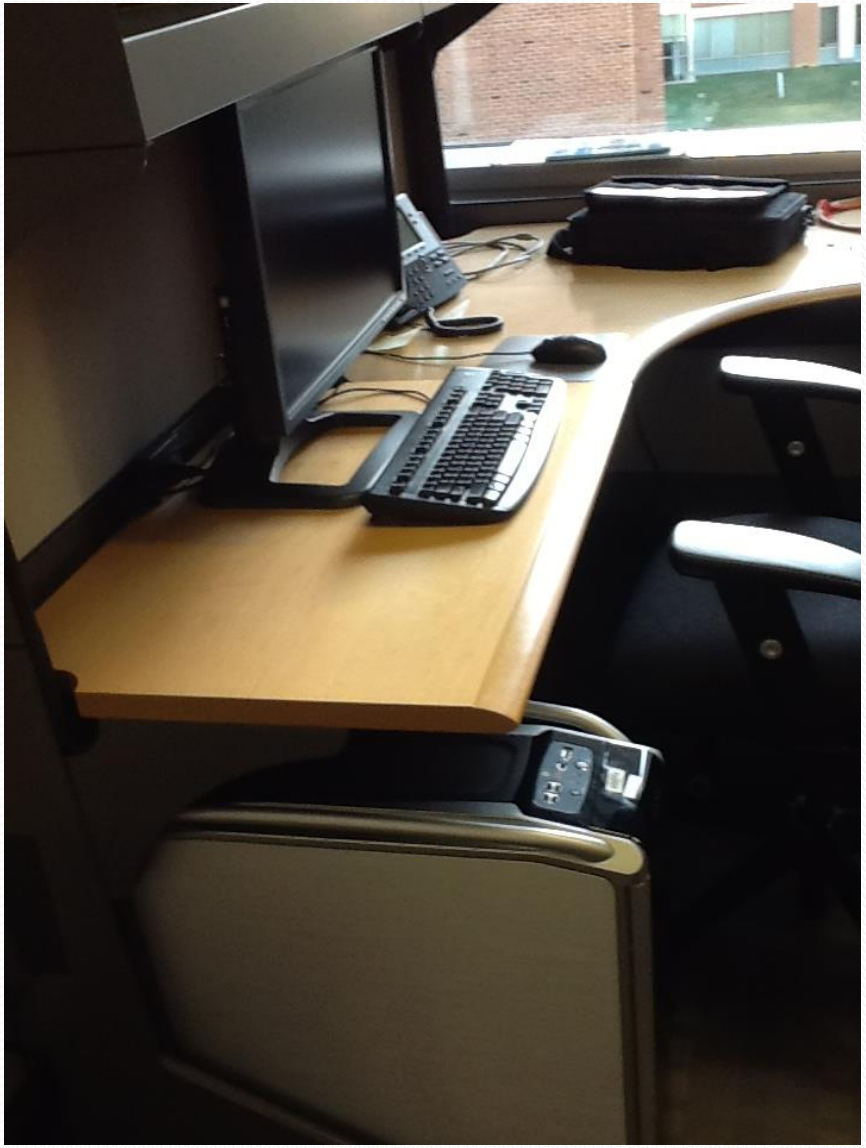
Looking back towards the parking ramp

3rd Floor of 64 – Office on left

Office WO 64-3032



Adib's Computer (8 processor Xeon)



Signing on to this computer puts you on the HPC network and provides direct access to the two research clusters -- Blue Meadows and Betsy

Third Floor Bridge to Building 62 (Engineering and Physics Laboratory)

Scientific Computing Lab in 62

What has been involved

- 29 trips to the Labs, about 90 work days on site in the past 1.5 years.
- 11 of the 29 trips were with with doctoral students:

  Mohammad Adibuzzaman

  Prachi Pradeep

  Casey O'Brien
- Summer support for Mr. Adibuzzaman and Ms. Pradeep through the Oak Ridge Institute for Science and Education (ORISE)
- Anticipated support for Mr. O'Brien in the Spring and/or Summer of 2014 through ORISE.

Assignment

- Identify and participate in research in the Division of Biology and the Division of Physics that could benefit from sophisticated computational models and approaches.
- Participate in Office-wide efforts in using models in a regulatory setting. This has two main aspects:
  - Building of a digital library of images and models used in device evaluation and research.
  - VVUQ –verification, validation, and uncertainty quantification. This has involved representing the FDA at 3 meetings at the National Academy of Sciences in DC.
- Evaluate and make recommendations regarding needed infrastructure for computational needs and data management requirements.

Some of the recent conferences

- Verification, Validation, and Uncertainty Quantification in Regulation, National Academy of Sciences, April 23, 2013.
- FDA/NIH/NSF Workshop on Computer Models and Validation for Medical Devices, June 11-12, 2013
- ASME 2013 Frontiers in Medical Devices: Applications of Computer Modeling and Simulation, September 11-13, 2013
- FDA/Critical Path Institute/ISOP Modeling & Simulation for Medical Products Workshop, September 26, 2013

Some Specific Projects

- "Improving toxicology prediction of QSAR tools"

Prachi Pradeep (MU), Ron Brown (DB), Peter Goering (DB), Shannon White (DB)

- "Methods for early detection of hypotensive events"

Md. Adibuzzaman (MU), Loriano Galeotti (DP), George Kramer (U Texas Medical Branch), David Strauss (DP)

- "Cleanability of medical devices"

Casey O'Brien (MU), Vicki Hitchins (DB), Anne Lucas (DP)

- "Role of UV in HPV-associated cancer"

Steve Merrill (MU), Dianne Godar (DB)

- There are four more

Improving toxicology predictions

- Primary interest are compounds released from medical devices: dyes, coatings, results of the body interacting with the material.
- Many of the compounds studied are not found in common toxicology or carcinogenicity databases (most come from pesticides, drugs, and compounds used in manufacturing).
- Basic idea is QSAR – quantitative structure activity relationship. “Similar” compounds should have similar properties. One issue is with the descriptors used to describe the compound. Finding the best descriptors for a given endpoint (like carcinogenicity or liver toxicity) is trial and error.
- The second issue is that each available QSAR package uses different methodology and is created with different training datasets.

An Ensemble Model of *In Silico* Tools to Improve Toxicity Prediction

Prachi Pradeep[1,2], Richard J. Povinelli[1], Ronald Brown[2], Stephen J. Merrill[1,2], Shannon White[2], Artem Korolev[2]
[1]Marquette University, Milwaukee, WI
[2]US FDA Center for Devices and Radiological Health, Silver Spring, MD

INTRODUCTION

Quantitative structure activity relationship (QSAR) models relate a quantitative measure of chemical structure (e.g. a physicochemical property) to a physical property or to a biological effect (e.g. a toxicological endpoint). Such models are of particular interest in regulatory settings for toxicity profiling of drugs or compounds that may be released from medical devices.

In silico methods of toxicity prediction provide a faster alternative to otherwise time-consuming laboratory and clinical testing methods. Several commercial and freely available *in silico* quantitative structure activity relationship (QSAR) models are available that can make predictions on various toxicological endpoints for new chemical compounds.

REGULATORY SIGNIFICANCE

Studies show that these tools vary in the accuracy of predictions across different datasets. Handling multiple predictions and determining the best prediction are difficult, especially when the predictions differ. Current protocols are restricted to using a majority voting scheme in making these consensus decisions.

The ensemble learning approach in machine learning paradigm can be used for combining these expert systems based on the understanding that consensus predictions from a set of models, is better than an individual prediction. Limited research in this direction urges an inspection of techniques to combine multiple predictions and ways to improve overall predictive ability.

We present the application of an ensemble learning technique to combine predictions from multiple *in silico* tools for improving the toxicity prediction accuracy with an impact in regulatory settings.

OBJECTIVE

The objective of this study is to investigate the Bayes combiner approach for combining predictions from multiple *in silico* tools to improve toxicity prediction with an impact in regulatory settings.

METHODS

***In Silico* Tools**
- Toxtree
- Vega
- Danish QSAR
- OECD Toolbox

Toxicity Endpoint Predicted – Carcinogenicity
- Ability to cause cancer
- Can be genotoxic or non-genotoxic

Datasets
- Dataset 1 (Inhalation compounds, 332 compounds)
  Actives: Inactives = 114:218
- Dataset 2 (Subset of Carcinogenic Potency Database, 199 compounds)
  Actives: Inactives = 105:94

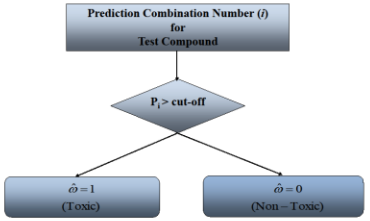
ALGORITHM

Bayes Combiner Model
The prior probabilities (P_i) for combinations (s) of predictions (0: non-toxic, 1: toxic) are calculated and the final prediction ($\hat{\omega}$) is estimated based on the value of P_i

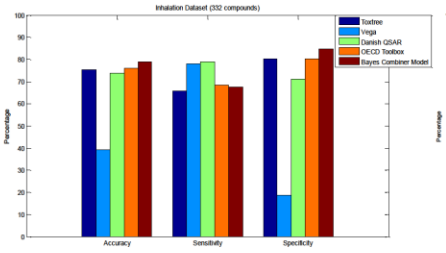
Leave One Out Cross Validation
N experiments performed with (N-1) training chemicals and 1 test chemical.

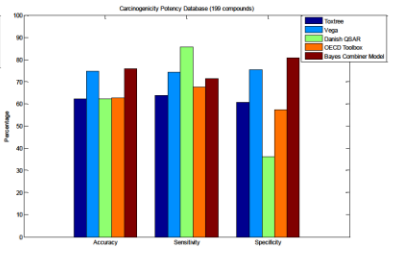
Combination Number	Tool 1	Tool 2	Tool 3	Tool 4	Prior Probability
1	0	0	0	0	P_1
2	0	0	0	1	P_2
3	0	0	1	0	P_3
4	0	1	0	0	P_4
5	1	0	0	0	P_5
6	0	0	1	1	P_6
7	0	1	0	1	P_7
8	0	1	1	1	P_8
9	1	0	0	0	P_9
10	1	0	0	1	P_{10}
11	1	0	1	1	P_{11}
12	0	1	1	1	P_{12}
13	1	0	0	0	P_{13}
14	1	0	0	1	P_{14}
15	1	0	1	1	P_{15}
16	1	1	1	1	P_{16}

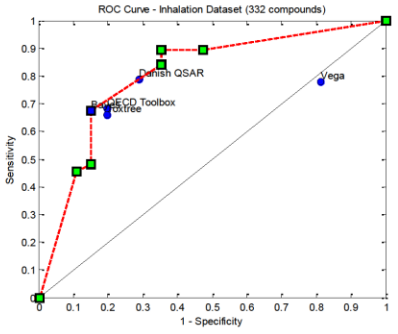
$$P_i = \frac{\text{Number of Toxic Compounds}}{\text{Total Number of Compounds with Prediction Combination } i}$$

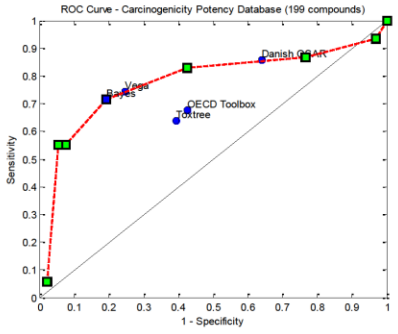


RESULTS









DISCUSSION

- It is observed that *in silico* predictions differ based on the model and dataset used.
- Predictions from the ensemble approach using Bayes Combiner Model are consistent across both datasets.
- The ROC curve can be used to select a trade-off between sensitivity and specificity of the Bayes Combiner Model.

Validation of Models

Model / Dataset	Toxtree	Vega	Danish QSAR	OECD Toolbox	Bayes Combiner Model
1	0.4568	-0.0237	0.4627	0.4795	0.5285
2	0.2443	0.4971	0.2245	0.2515	0.5193

Table 1: Kappa (K, proportion of specific agreement)

- Bayes Combiner Model improves the accuracy and sensitivity of predictions across both datasets.
- Bayes Combiner Model improves the Kappa coefficient across both datasets.

CONCLUSION

- *In silico* models show promise as tools for faster design and regulation of new medical products.
- Most training datasets are composed of pharmaceutical or environmental chemicals. The chemical space may not adequately represent a new test chemical.
- The accuracy of predictions depends on the domain of applicability of a given model.
- Using an ensemble of *in silico* models and arriving at a consensus prediction can improve the quality of predictions and help in developing a regulatory strategy for early product development.

FUTURE WORK

Improving the ensemble algorithm
- Defining prior probabilities in terms of *similarity* of the test chemical to the training dataset.
- Incorporation of in-vitro data as descriptors to the model.

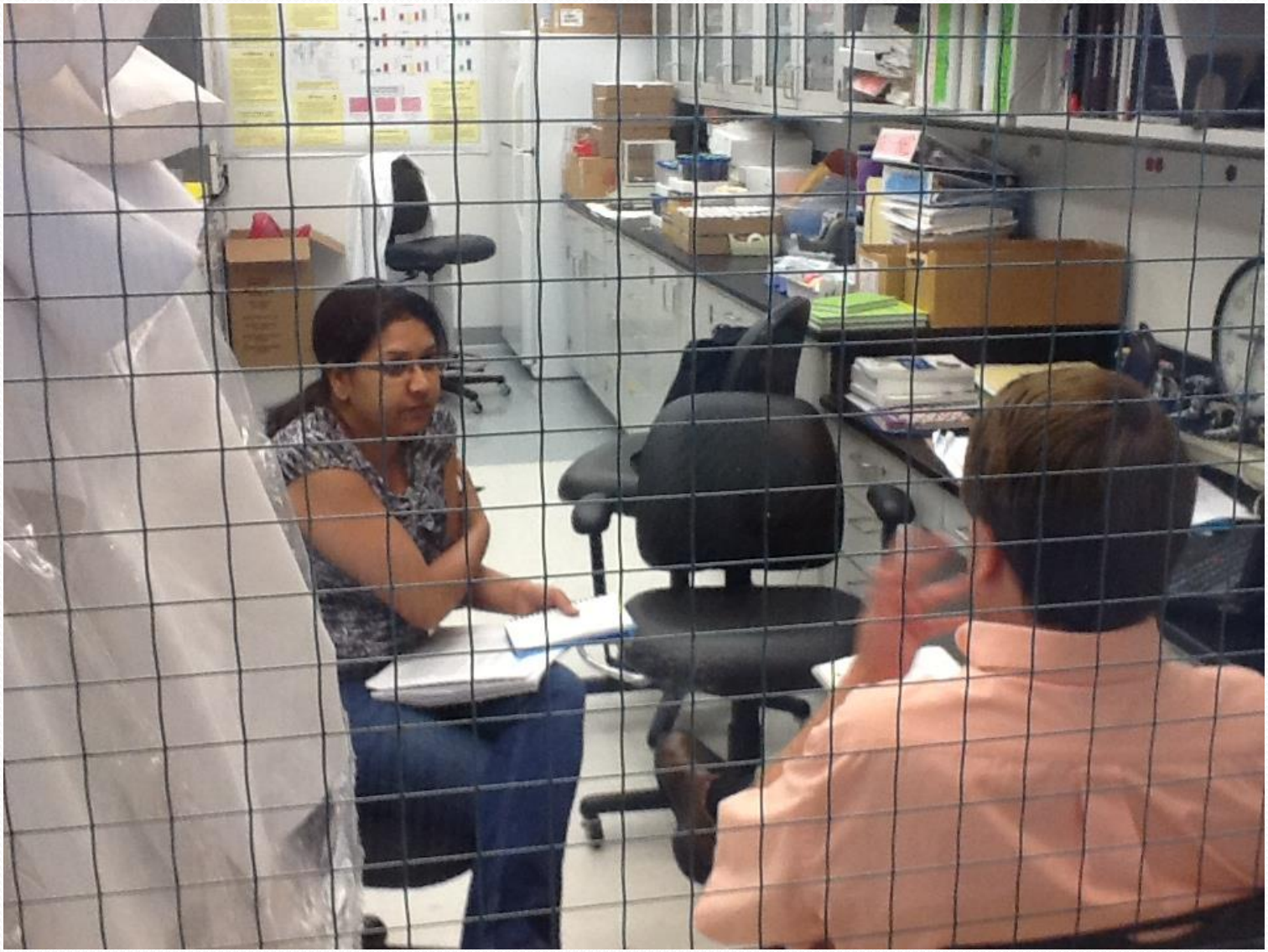
Validation on other/larger datasets
- *e.g.* Dental leachables

Validation on other toxicity endpoints
- *e.g.* Skin sensitivity

ACKNOWLEDGEMENT

- This project was supported in part by an appointment to the Research Participation Program at the Center for Devices and Radiological Health administered by the Oak Ridge Institute for Science and Education through an interagency agreement between the U.S. Department of Energy and the U.S. Food and Drug Administration.
- Ed Gordon for his assistance with poster production.

Prachi with Ron Brown in his Lab



Early Predictors of Hemorrhage

- Motivation is in a mass casualty setting (earthquake, bombing, building collapse).
- As there are many more patients then medical professionals, how can sensors placed on the patients, be used effectively to give an early signal of shock (often as a result of internal hemorrhage).

Data from U. Texas Medical Branch

- A number of pigs are slowly bled and vital signs are monitored – some by noninvasive methods.
- Size, rate of bleed, number of bleeds varies, as does the interventions during the process (IV's primarily)
- Because of the slow nature of the loss of blood, early markers for failure of the natural compensation systems should be able to be identified (if you knew what to look for)
- Method: create Markov chain empirical models by sampling from the high-frequency data.
- Optimize aspects of the sampling rate and the number of states through HPC clusters.

Novel physiological monitoring strategies for early detection of hemorrhage

Mohammad Adibuzzaman[1,2], Loriano Galeotti[2], Vicente Jose[2], Richard A. Gray[2], George Kramer[3], David G. Strauss[2], Stephen Merrill[1]

1.Department of Mathematics, Statistics, and Computer Science, Marquette University, Milwaukee, WI

2.Food and Drug Administration, CDRH/OSEL/ Division of Physics

3.The University of Texas Medical Branch, Department of Anesthesiology, Galveston, TX

Background

- In a mass casualty situation, one key problem is identifying patients that need immediate care. In a combat situation this is much more important since this risks the life of the paramedics.
- We seek to **identify early markers of hemorrhage** for triage.
- We propose a smart monitoring algorithm to identify patients that need immediate care which would work as a decision support system.
- Existing algorithms that use mean arterial pressure or heart rate variability has **limitations** such as:
  - Mean arterial pressure does not change due to the compensating mechanism until at a later stage.
  - Heart rate variability changes may depend on individual responses.

Our Approach

- Markov chain modeling for **detecting early change in arterial blood pressure**, that indicates change in system dynamics, rather than mean arterial pressure.
- System dynamics deals with internal feedback loops and time delays that affect the system.
- Use of **heart rate variability as a second index** to increase sensitivity and specificity of the decision support system.

Methods

Eigenvalue of Markov chain

- We used empirical Markov Chain modeling of arterial blood pressure
- Markov assumptions include the next step depends only on the current state.
- The states are defined as ranges in the time series data.

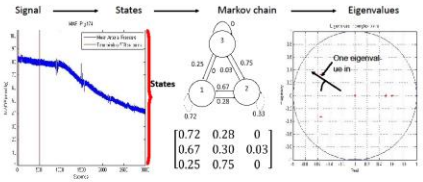


Figure 1: (a) Mean Arterial Pressure of a pig. (b) A sample Markov Chain Model with the transition probabilities and the corresponding matrix. (c) Eigenvalues of a transition probability matrix in complex plane with angle and absolute value.

- Eigenvector of a matrix is a vector that would be scaled if multiplied by the Matrix. The scaling factor is called the eigenvalue corresponding to the eigenvector. If λ is an eigenvalue, *v* is the eigenvector and A is the matrix, then

$$Av = \lambda v$$

- Eigenvalues of a matrix describe are defined as the roots of the characteristic polynomial,

$$\det(A - \lambda I) = 0$$

Detection of changes in Eigenvalue

Mean and standard deviation (SD) are computed using a 50-data points moving window (40 data points overlap). Windows with SD > 0.003 are probably noisy and discarded. The threshold was selected as the one corresponding to maximum value of sensitivity.

Heart rate variability analysis

Heart rate variability is measured as the standard deviation of 100 samples of RR intervals, calculated from electrocardiogram.

Results

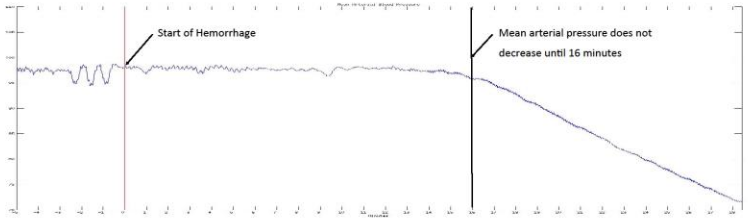


Figure 2: Mean arterial pressure of a pig subject to slow hemorrhage. The red vertical line indicates start of hemorrhage (time zero).

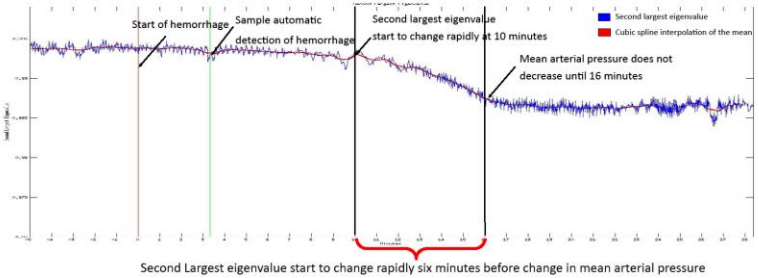


Figure 3: The second largest eigenvalue or the mixing time of the Markov chain tends to decrease after hemorrhage. The green vertical line represents a sample detection of hemorrhage using change in mean and standard deviation.

Second Largest eigenvalue start to change rapidly six minutes before change in mean arterial pressure

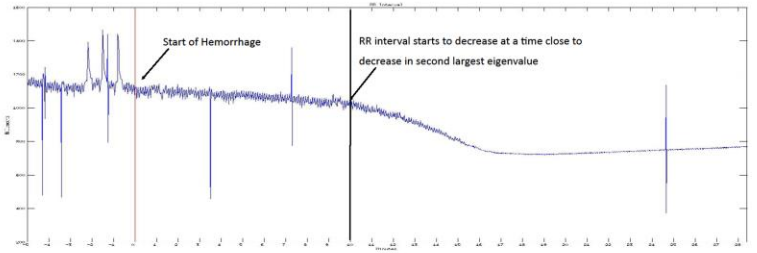


Figure 4: RR interval tends to decrease as heart rate increases at similar time the second largest eigenvalue starts to decrease.

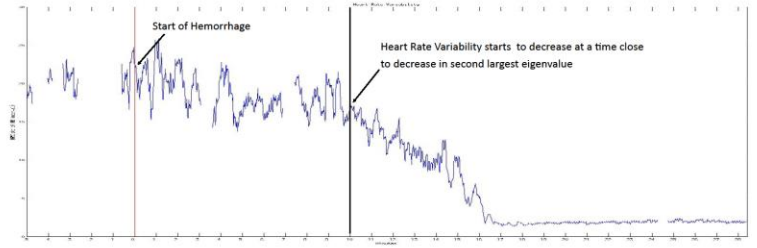


Figure 5: Heart rate variability, measured as the standard deviation of every 100 samples of RR intervals, also starts to decrease at similar time second largest eigenvalue from mean arterial pressure starts to decrease.

Results

- The approach is tested on the hemorrhage data of three pigs with different hemorrhage rates.
- The hemorrhages are identified as 8 consecutive, non noisy windows with sum of the differences less than –0.0001.
- The means for which the standard deviation is greater than 0.003 are probably noisy and are discarded.

Pig		Sensitivity	Specificity
174	ABP	0.72	0.76
	Non invasive ABP	0.71	0.89
515	ABP	0.70	0.80
	Non invasive ABP	Not Available	Not Available
517	ABP	0.83	0.76
	Non invasive ABP	0.66	0.84

Table 1: Sensitivity and specificity.

Summary

The second largest eigenvalue of the Markov chain, also known as the Mixing rate of Markov chain, constructed from Arterial Blood Pressure, changes during hemorrhage indicating faster convergence rate to the limit distribution.

Developed an algorithm to automatically detect the changes in the second largest eigenvalues to detect hemorrhage.

We have also found similar changes in the second largest eigenvalue for non invasive blood pressure during hemorrhage.

Future Work

- We propose to use heart rate variability as a second index to confirm hemorrhage.
- In the literature, there is a significant work to correlate heart rate variability with hemorrhage using low frequency and high frequency components of the power spectra. We are working on using these methods to detect hemorrhage from heart rate variability using changes in the high frequency component of the RR interval. This may give us a better accuracy for detecting hemorrhage from multiple markers.

References

[1]. S. J. Merrill, "Markov chains for identifying nonlinear dynamics," *Nonlinear dynamical systems analysis for the behavioral sciences using real data*, pp. 401-423, 2011.

[2]. S. J. Merrill, "Markov chain methods in the analysis of heart rate variability," *Fields institute communications*, vol. 11, 1997.

[3]. W. Cooke, C Rickards, K. Ryan, V Convertino," Autonomic compensation to stimulated hemorrhage with heart period variability," *Critical care medicine*, vol. 36, 2008.

[4]. L. Mayaud, P. Lai, G. D. Clifford, L. Tarassenko, L. Celi, D. Annane, "Dynamic data during hypotensive episode improves mortality predictions among patients with sepsis and hypotension," *Critical care medicine*, vol. 41, 2013.

[5]. J. Thayer, S. Yamamoto, J. Brosschot, "The relationship of autonomic imbalance, heart rate variability and cardiovascular disease risk factors," *International journal of cardiology*, vol. 141, 2010.

Acknowledgement

This project is supported by the Medical Countermeasures Initiative (MCMi) and by an appointment to the Research Participation Program at the Center for Devices and Radiological Health administered by the Oak Ridge Institute for Science and Education through and interagency agreement between the U.S. Department of Energy and the U.S. Food and Drug administration.

Cleanability of Reusable medical devices

- Recognition of the tendency to have devices that are cleaned and reused.
- Difficulty in validating that the device is indeed cleaned (usually by a third party reprocessor).
- From the regulatory point of view, difficulty in regulating the cleanability of the device.
- Next slides are from Edward Gordon (FDA/CDRH/OSEL/DB)

What is "clean"?

- **"Yesterday"** – clean was visibly clean
- **"Today"** – we can't see if inside the device is "visibly clean" due to:
  - complex devices with long and/or narrow opaque lumens
  - lumens with acute angles
  - junctions between insulating sheaths
  - stopcocks that can't be disassembled, etc.

Design Features Offering Difficulty in Cleaning

AAMI TIR30:2011

- **Lumens**, especially lumens of flexible design, multiple internal lumens, lumens that arenot freely accessible, bifurcated lumens, and elevator-wire lumens inside- viewing duodenoscopes and some endoscopic ultrasound endoscopes
- **Valves**
- **Crevices**
- **Fittings** with very close tolerances
- **Clamps** that cannot be fully opened for cleaning (e.g., pylorus clamps)
- **Small internal parts** (e.g., springs, magnets)
- **Forceps**, especially their articulations and grooves and especially forceps that cannot be readily disassembled (e.g., arthroscopy forceps)

Design Features Offering Difficulty in Cleaning

AAMI TIR30:2011

- **Rough, irregular, discontinuous surfaces** that can entrap or retain bioburden and debris
- **Hinges, depressions, joints with gaps**, overlapping or butted joints that result in acute angles, or ribbed or otherwise "roughened" surfaces (e.g., jaws)
- **Capillary gaps**
- **Luer locks**
- **Porous materials** (smooth surfaces are desirable, where possible)
- **Junctions between insulating sheaths and activating mechanisms** (as in certain laparoscopic instruments)

Casey's problem

- Based on work done shapes motivated by those found in arthroscopic shaver handles.

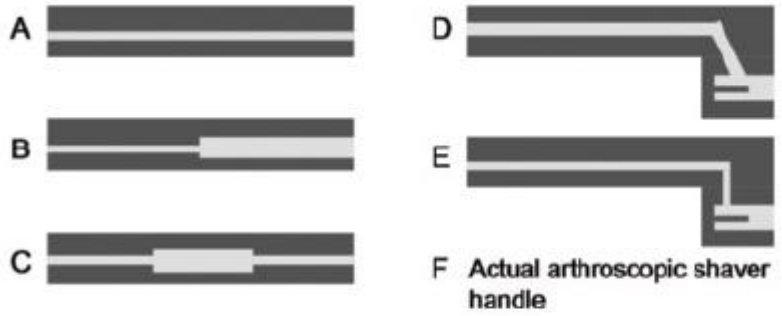


Figure 3A. Mass of debris extracted from device and device complexity, with illustrations (not to scale) of cross-sections of model devices. Dark gray indicates metal, light gray indicates the fluid path.

Victoria M. Hitchins, PhD, is a research microbiologist with the Office of Science and Engineering Laboratories, CDRH/FDA. E-mail: :toria.hitchins@fda.hhs.gov

- Haugen, Duraiswamy, and Hitchins, Quantification by mass of residual debris in reusable medical devices, AAMI *Horizons* Spring 2012.
- These were made in aluminum – new ones are in plastic in which amount and location of material (artificial bone) after cleaning can be quantified (through dyes) and location determined through micro CT images.

What this experience tells us of the Computational Sciences program

- Our Computational Sciences students are prepared to work in a large variety of real-world scientific problems.
- These problems are of central import to governmental agencies and companies.
- Their work is seen to be of very good quality – both the University and the Program are showcased.