

Rapid Diagnosis of Biological Warfare Agent Exposure

Olivia Peters

703-983-6797 • opeters@mitre.org

MITRE Sponsored Research



MITRE
Technology
Program

Problem

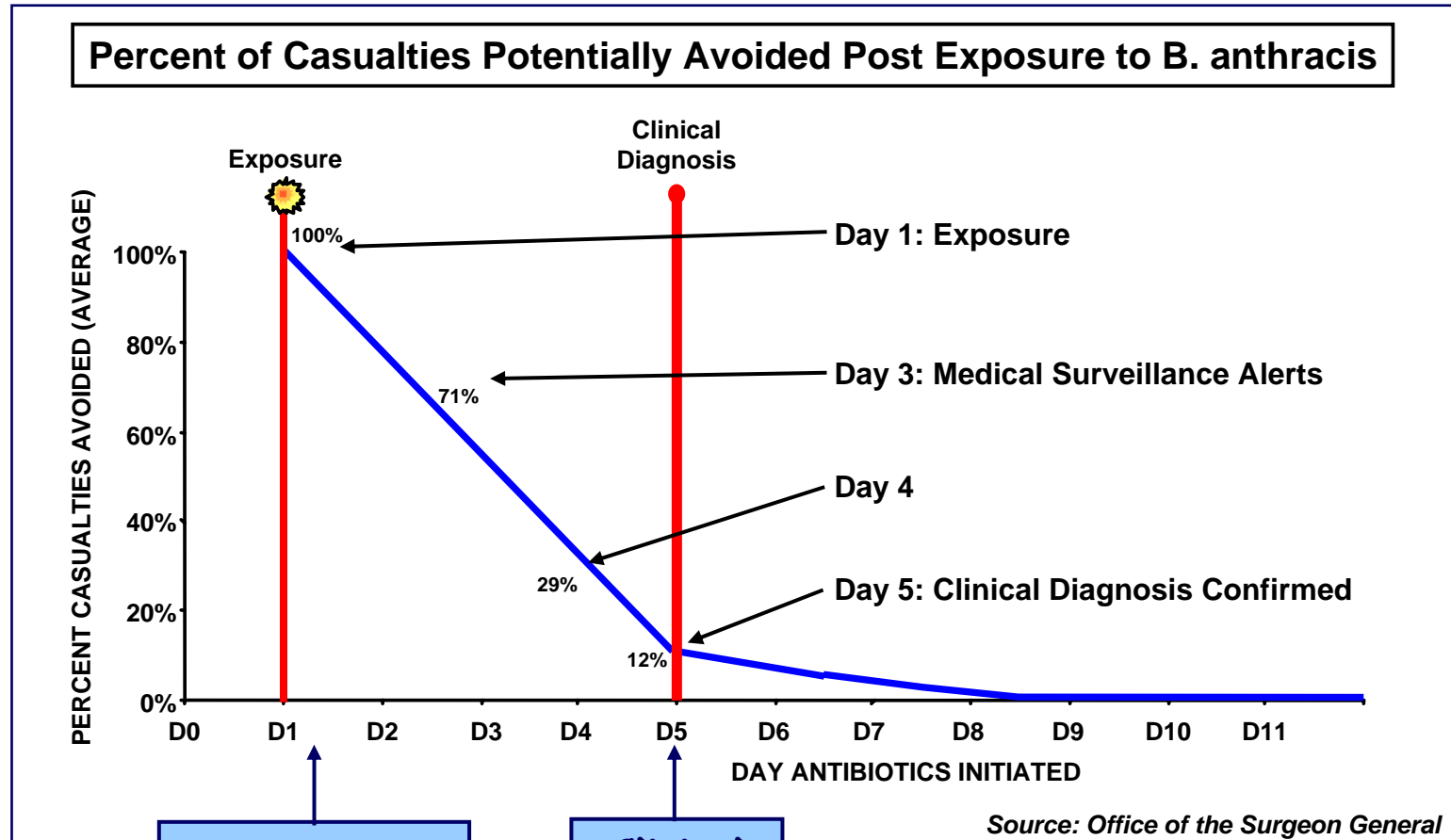
- Difficult to determine if a biological attack has taken place

<u>Disease</u>	<u>Agent</u>	<u>Symptoms</u>
Aflatoxin	<i>Aspergillus flavus</i>	Nausea, vomiting, then acute liver failure or cancer
Anthrax	<i>Bacillus anthracis</i>	High fever, labored breathing, rapid heartbeat
Botulism	<i>Clostridium botulinum</i>	Nausea, fatigue, cramps, headache, respiratory paralysis
Plague	<i>Yersinia pestis</i>	Lung infection, pneumonia, hemorrhage
Ricin	<i>Ricinus communis</i>	Convulsions, stupor, vomiting, bloody diarrhea

- Need to develop a way to test:
 - has the warfighter been exposed and to which agent?
 - what is the estimated time since exposure?
- Technology does not currently exist for rapid and early diagnosis of BW attack
 - No information infrastructure for efficient data storage
 - No identification of genes correlated to attack
 - No algorithms for analysis and classification

Background

Survival rates post exposure to *B. anthracis*



**Detection
by host gene
profiling**

**Clinical
diagnosis**

Objective

- **Develop a classifier to quickly identify if a warfighter has been exposed to a BW agent, which agent, and a timeframe for this exposure**
 - **Coordinate and manage a large amount of microarray data**
 - **Perform feature extraction and dimensionality reduction to find specific genes and/or combinations of genes indicative of exposure to a specific pathogen**
 - **Design a classifier that can determine if a cell has been exposed to a known or even unknown pathogen**

Activities

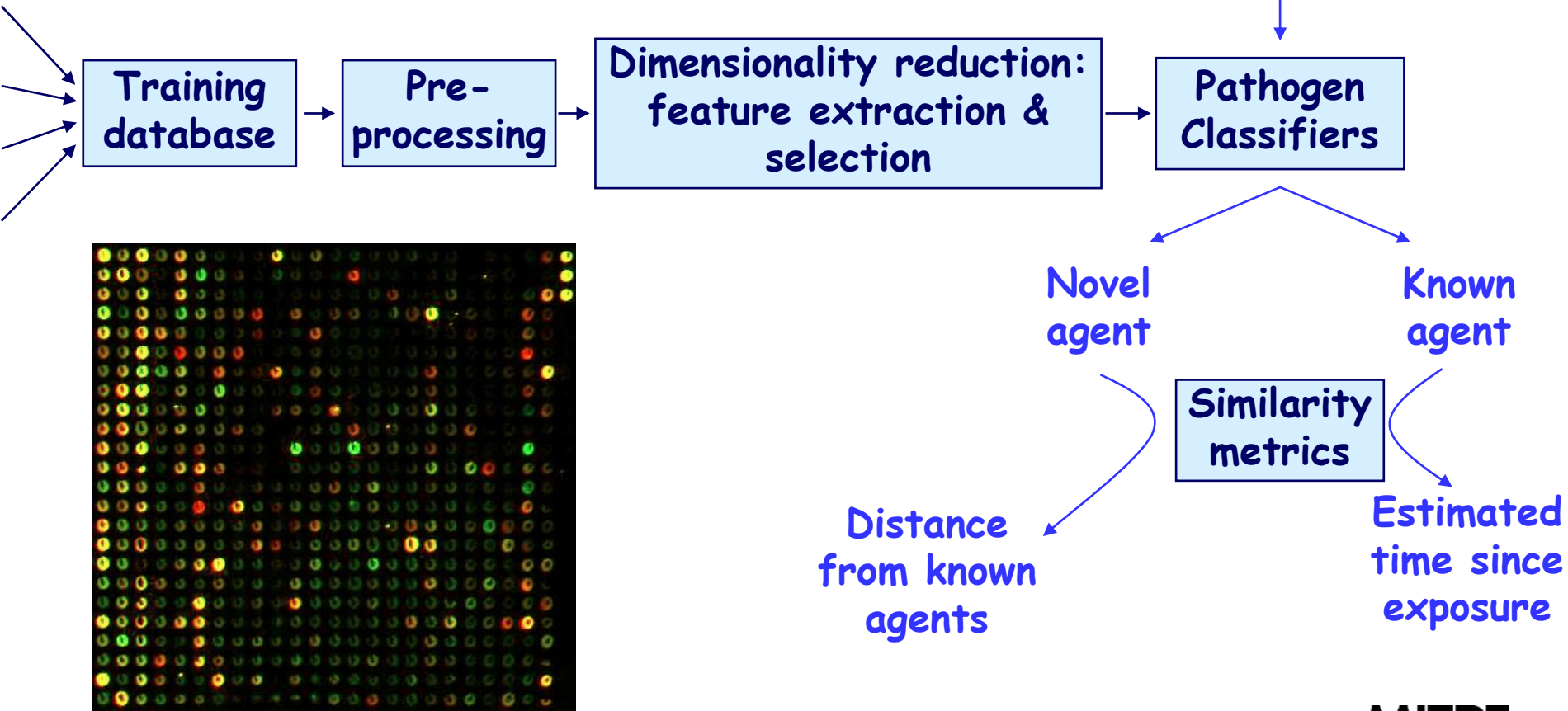
- Golub data set - differentiate between two cancers:
 - acute myeloid leukemia (AML)
 - acute lymphoblastic leukemia (ALL)

- Cyclosarin (CW agent) data set – differentiate between control and 3 levels of exposure:
 - 0.004 mg/m³
 - 0.0134 mg/m³
 - 0.0251 mg/m³

<u>Method</u>	<u>Detection rate (%)</u>
MTT/CSMFC	97.1
MTT/PCAPM	91.2
MTT/ICAPM	76.5
ML/NBC	88.2
GEC/ANN	95.6

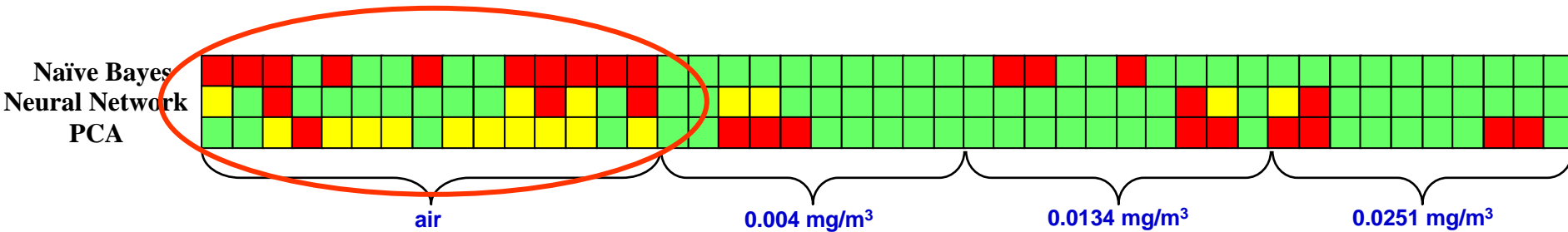
<u>Method</u>	<u>Detection rate (%)</u>		
	<u>0.004 mg/m³</u>	<u>0.0134 mg/m³</u>	<u>0.0251 mg/m³</u>
MTT/CSMFC	92.0	84.0	84.0
MTT/PCAPM	64.0	60.0	84.0
MTT/ICAPM	72.0	80.0	88.0
ML/NBC	83.3	81.6	83.4
GEC/ANN	97.7	95.3	97.2

Highlight



Highlight

- More errors are occurring in the control (air) group than anywhere else
- Air samples account for approximately 60% of misclassifications, but only 30% of all the data
- Almost without exception the air samples that are misclassified classify within group, i.e., an air sample that is from group 8 will classify as 0.004



	air : 0.004mg/m ³	air : 0.0134mg/m ³	air : 0.0251mg/m ³	0.004mg/m ³ : 0.0134mg/m ³	0.004mg/m ³ : 0.0251mg/m ³	0.0134mg/m ³ : 0.0251mg/m ³
Naïve Bayes	83.3%	81.6%	83.4%	99.2%	94.4%	99.0%
Neural Network	97.7%	95.3%	97.2%	99.9%	99.5%	100%

Impacts

- **Developing collaborative relationships within the biotechnology community**
 - ECBC, WRAIR, USAMRIID
 - CDC
 - NMRC
- **Software implementation of a classifier**
 - differentiate between cells exposed to different pathogens
 - include a suite of similarity metrics
- **Database of host cell responses when exposed to different biowarfare agents**

Future Plans

- **Study genes selected and their involvement in host response to exposure**
- **Explore similarity metrics**
- **Develop more meaningful classifier output than just winner**
- **Apply methods to additional chemical and biological warfare agents**
- **Extend application to various data sources**
 - **CDC chronic fatigue syndrome**
 - **HHS environmental and health hazard studies**