Smart Health – CSE 40816

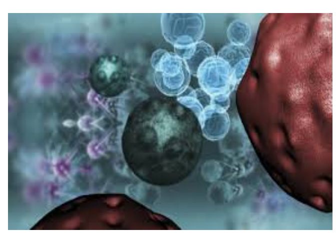
University of Notre Dame

Spring 2020

UNIVERSITY OF NOTRE DAME

What is a Biomarker?

- A **biological marker**, better known as a **"biomarker"**, is a characteristic that is objectively measured and evaluated as an indicator of normal biological processes, pathogenic processes, or pharmacological responses to a therapeutic intervention.



BIOMARKERS DEFINITIONS WORKING GROUP: BIOMARKERS AND SURROGATE ENDPOINTS: PREFERRED DEFINITIONS AND CONCEPTUAL FRAMEWORK. CLIN PHARMACOL THER 2001;69:89-95.

Examples of Biomarkers

- Glucose levels that are used as a biomarker in managing diabetes.
- Cholesterol levels that are used to detect coronary artery disease.
- Brain images such as MRI (Magnetic Resonance Imaging) that can provide information about the progression of Multiple Sclerosis.
- Biological substances such as enzymes (which may be found in the blood or tissue samples) are used in cancer research.
- Genetic (DNA) changes.
- Medical images and X-rays.

What is a Biomarker?

- Term biomarker first coined in the 1980s.
- Biomarkers were developed as a response to understand the relationship between environmental factors and disease.
- Biomarkers include tools and technologies that can aid in understanding the
  - Prediction
  - Cause
  - Diagnosis
  - Progression
  - Regression
  - Outcome

of various diseases.

What is a Biomarker?

- Biomarker variety
  - Each body system has specific biomarkers (e.g., cardiovascular, respiratory, neurological, psychological, …).
  - Each biomarker is relatively easy to measure.
  - Each biomarker forms a piece of routine medical examinations (e.g., weight and BMI measurements to predict obesity).

What is a Biomarker?

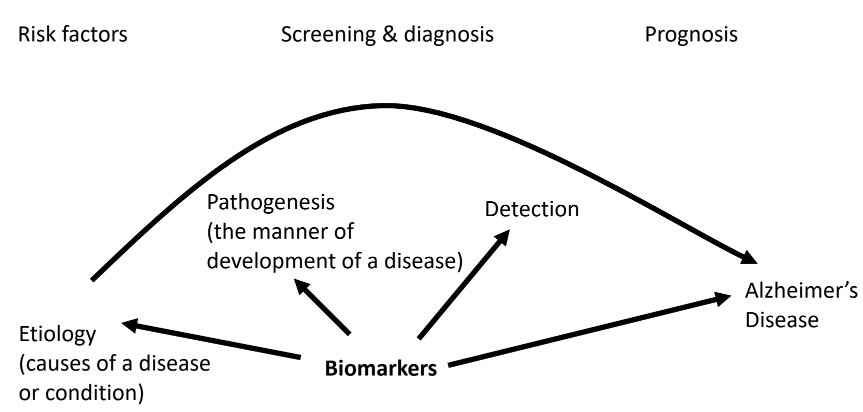
- Ideal biomarker characteristics:
  - Safe and easy to measure.
    - Should create as little discomfort for patient as possible (e.g., blood sample).
  - Cost-effective.
  - Rapid return of results for early initiation of treatment and monitoring effectiveness.
  - Modifiable with treatment.
  - Highly reproducible.
  - Consistent across gender and ethnic groups.

Computer Science and Engineering - University of Notre Dame

Biomarker Types

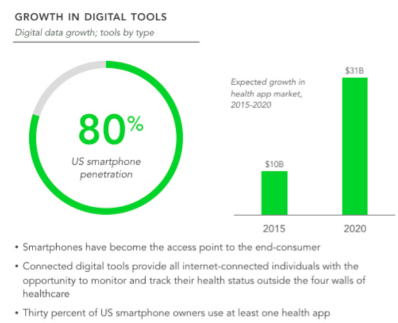
- **Biomarkers of exposure:**
  - Reconstruct and predict past exposure to risk factors.
    - Carbon monoxide in breath, blood or metals in blood, bones, hair, etc.
- **Biomarkers of risk or susceptibility:**
  - Identify individuals at increased risk for development of a disease.
    - Elevated low-density lipoprotein (LDL) cholesterol levels (coronary artery disease).
- **Biomarkers of disease:**
  - Used for screening or diagnosis, progression, or regression assessment, etc.
    - Blood sugar or hemoglobin A1c (HbA1c) may be used to identify patients with Type 2 diabetes mellitus (DM).

Disease Pathway

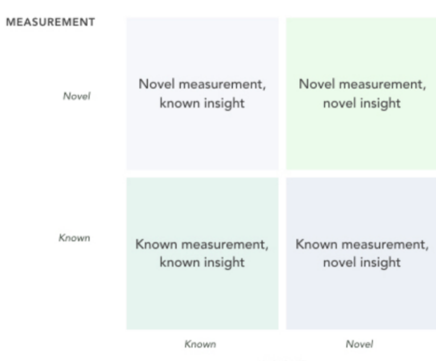


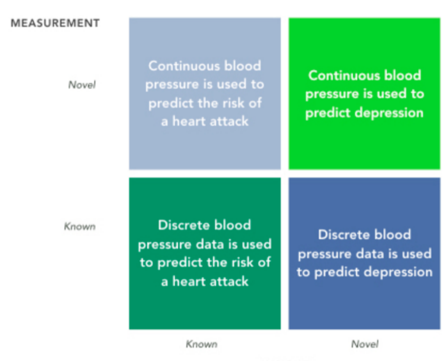
Digital Biomarker

- "Digital biomarkers are consumer-generated physiological and behavioral measures collected through connected digital tools that can be used to explain, influence and/or predict health-related outcomes." [Wang et al. 2019]

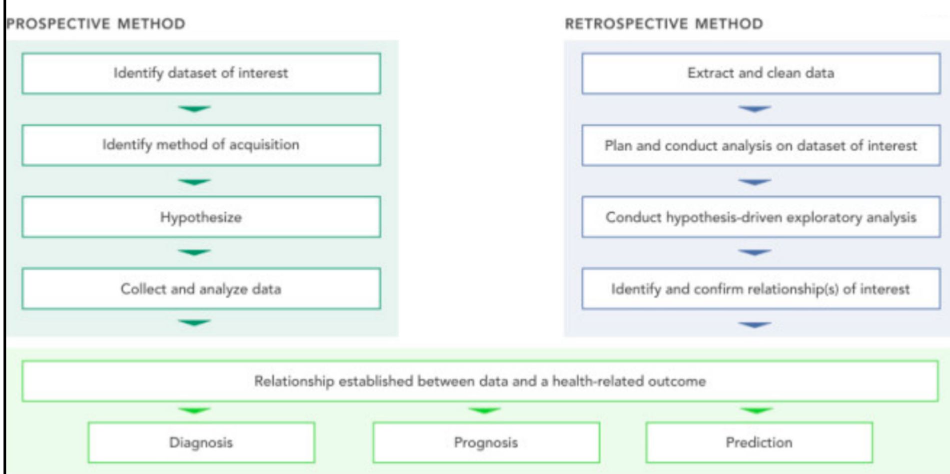


Digital Biomarkers



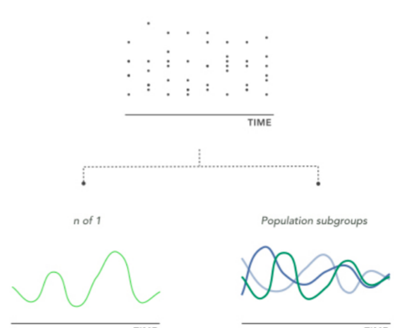


Digital Biomarkers

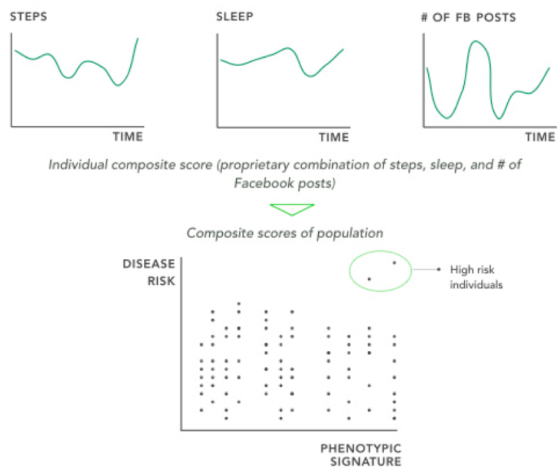


Digital Biomarkers

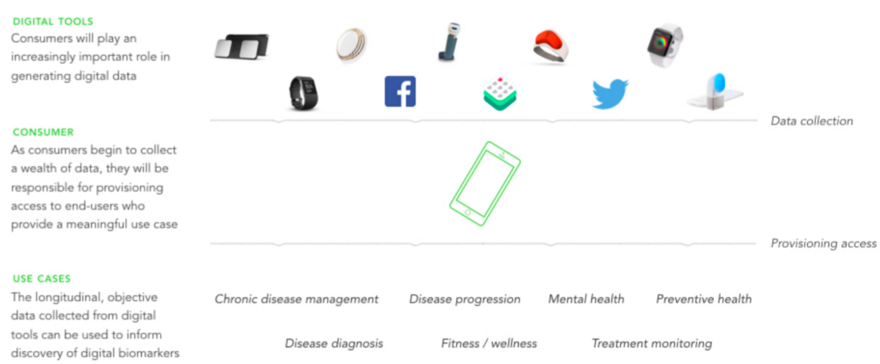
- n of 1
  - Longitudinal individual-level data
  - Personalized baselines
- Population subgroups
  - Longitudinal population-level data
  - Control for previous/existing disease states
  - More likely to find evidence of causality



Digital Biomarkers



Digital Biomarkers



Types of Research

- **Descriptive**: describe a group of individuals on a set of variables or characteristics (understanding and classification).
  - Case study
  - Cross-sectional study
  - Qualitative study
- **Exploratory**: examine a phenomenon of interest and explore its dimensions, including how it relates to other factors (relationships can lead to predictive models).
  - Cohort study
  - Case control study
- **Experimental**: basis for comparing two or more conditions; controls or accounts for the effects of extraneous factors; draw meaningful conclusions about observed differences.
  - True experimental designs
  - Quasi-experimental designs

Case Study Design

- Often a description of a individual case's condition or response to an intervention
  - Can focus on a group, institution, school, community, family, etc.
  - Data may be qualitative, quantitative, or both.
  - **Case series**: observations of several similar cases are reported.

Case Study

Example

- In 1848, young railroad worker, Phineas Gage, was forcing gun powder into a rock with a long iron rod when the gun powder exploded. The iron rod shot through his cheek and out the top of his head, resulting in substantial damage to the frontal lobe of his brain. Incredibly, he did not appear to be seriously injured. His memory and mental abilities were intact, and he could speak and work. However, his personality was markedly changed. Before the accident, he had been a kind and friendly person, but afterward he became ill-tempered and dishonest.

- Phineas Gage's injury served as a case study for the effects of frontal lobe damage. He did not lose a specific mental ability, such as the ability to speak or follow directions. However, his personality and moral sense were altered. It is now known that parts of the cortex (called the association areas) are involved in general mental processes, and damage to those areas can greatly change a person's personality.

Cross-Sectional Study

- Researcher studies a stratified group of subjects at one point in time.
- Draws conclusions by comparing the characteristics of the stratified groups.
- Well-suited to describing variables and their distribution patterns.

Cross-Sectional Study

- Example:

  Let's say we want to investigate the relationship between daily walking and cholesterol levels in the body. We recruit walkers and non-walkers at the same time and compare cholesterol levels among these different populations.

Qualitative Study

- Seeks to describe how individuals perceive their own experiences within a social context.
- Emphasizes in-depth, nuanced understanding of human experience and interactions.
- Methods include in-depth interviews, direct observations, examining documents, focus groups.
- Data are often participants' own words and narrative summaries of observed behavior.

Qualitative Study

- Example
  - A researcher wants to understand how provision of healthcare to undocumented persons affects the people and institutions involved.
  - In multiple communities, information is gathered from undocumented patients, primary care clinicians, specialists, and hospital administrators.
  - Methods: in-depth interviews, key informant interviews, participant observations, case studies, focus groups.

Cohort Study

- A group of individuals who do not yet have the outcome of interest are followed together over time to see who develops the condition.
- Participants are interviewed or observed to determine the presence or absence of certain exposures, risks, or characteristics.
- May identify risk by comparing the incidence of specific outcomes in exposed and not exposed participants.

Cohort Study

- Example
  - To determine whether exercise protects against coronary heart disease (CHD).
    - Assemble the cohort: 16,936 Harvard alumni were enrolled.
    - Measure predictor variables: administer a questionnaire about activity and other potential risk factors, collected data from college records.
    - 10 years later, sent a follow-up questionnaire about CHD and collected data about CHD from death certificates.

Cohort Study

- Strengths
  - Powerful strategy for defining incidence and investigating potential causes of an outcome before it occurs.
  - Time sequence strengthens inference that the factor may cause the outcome.
- Weaknesses
  - Expensive – many subjects must be studied to observe outcome of interest.
  - Potential **confounders**: e.g., cigarette smoking might confound the association between exercise and CHD.

Case-Control Study

- Generally retrospective.
- Identify groups with or without the condition.
- Look backward in time to find differences in predictor variables that may explain why the cases got the condition and the controls did not.
- Assumption is that differences in exposure histories should explain why the cases have the condition.
- Data collection via direct interview, mailed questionnaire, chart review.

Case-Control Study

- Strengths
  - Useful for studying rare conditions.
  - Short duration & relatively inexpensive.
  - High yield of information from relatively few participants.
  - Useful for generating hypotheses.
- Weaknesses
  - Increased susceptibility to **bias**:
    - Separate sampling of cases and controls.
    - Retrospective measurement of predictor variables.
  - Only one outcome can be studied.

Case-Control Study

- Example
  - Purpose: To determine whether there is an association between the use of aspirin and the development of Reye's syndrome in children.
    - Draw the sample of cases – 30 patients who have had Reye's syndrome.
    - Draw the sample of controls – 60 patients from the much larger population who have had minor viral illnesses without Reye's syndrome.
    - Measure the predictor variable: ask patients in both groups about their use of aspirin.

Longitudinal Studies

- A **longitudinal** study is a **research** design that involves repeated observations of the same variables over longer periods of time (i.e., uses **longitudinal** data).
- Example:
  - How do friendships change from freshman to senior year on a college campus?
  - How do friendships impact wellness and health?

Statistical Power Analysis

- Prior to conducting a study, it is advisable to conduct a statistical power analysis.
- Power is the probability that a statistical test will detect a significant effect that exists.
  - Null hypothesis: hypothesis that there is no significant difference between specified populations, any observed difference being due to sampling or experimental error.
  - Power is the probability that you will reject the null hypothesis when you should.
- The power analysis will suggest an adequate sample size for the study.

Statistical Power Analysis

- Four parameters:
  - Significance level (α)
    - Difference (p-value) between two groups or more based on some variable
  - Sample size (n)
    - Number of participants in study
  - Effect size (ES)
    - Magnitude of the difference between populations
  - Power (1 - β)

$$Power \propto \frac{Sample\ size\ (n)}{Effect\ size\ (\Delta), Alpha(\alpha)}$$

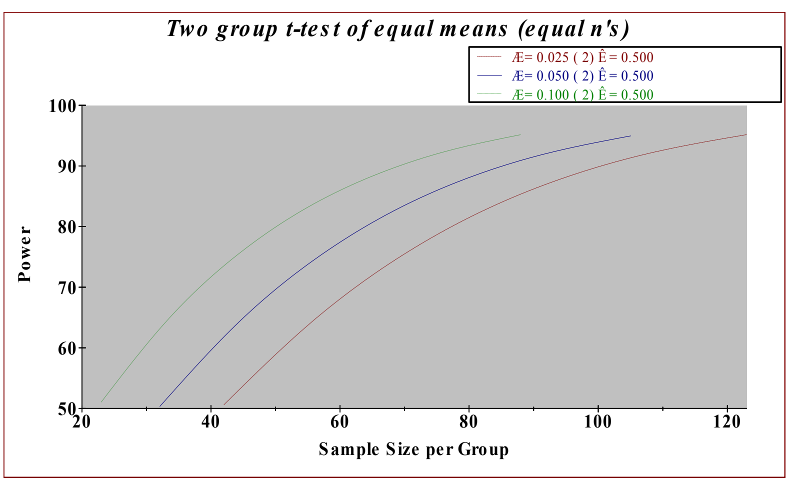
Computer Science and Engineering - University of Notre Dame

Statistical Power Analysis

▪ Significance level (α):	.05	*
	.01	
	.001	
▪ Effect size:	"small"	
	"medium"	*
	"large"	
▪ Power:	.80	*
	.90	

* Typical values for social/behavioral/health sciences

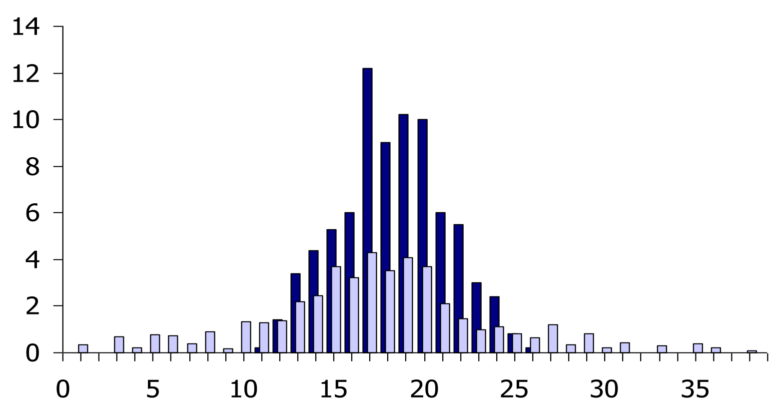
Relationship Between Alpha(α), Sample Size (n), and Power (1-β)



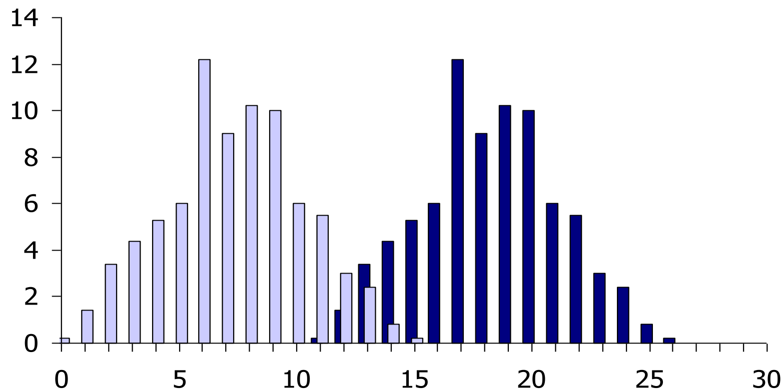
Bias

- **Bias**: Deviation of results or inference from truth, or processes leading to such deviations. Any trend in the collection, analysis, interpretation, publication, or review of data that can lead to conclusions that are systematically different from the truth.
- Bias is an **error**.
- Two types of errors:
  - **Random**: use of invalid outcome measure that equally misclassifies cases and controls.
  - **Systematic**: use of invalid measures that misclassify cases in one direction and controls in another.

Random Error (Chance)



Systematic Error (Bias)



Chance vs. Bias

- Chance is caused by random error.
- Bias is caused by systematic error.

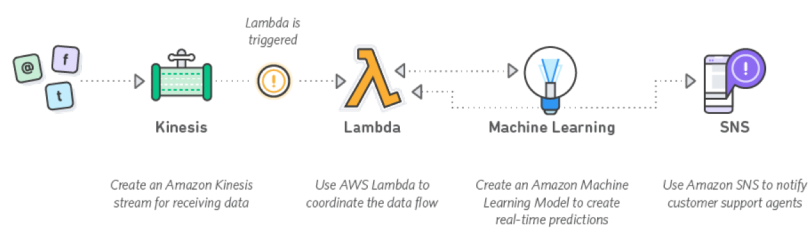
- Errors from chance will cancel each other out in the long run (large sample size).
- Errors from bias will not cancel each other out whatever the sample size.

- Chance leads to **imprecise** results.
- Bias leads to **inaccurate** results.

Examples

- Selection Bias: errors in the process of identifying the study population.
- Recall Bias: differences in the accuracy or completeness of the recollections retrieved by study participants.
- Confirmation Bias: often unconscious act of referencing only those perspectives that fuel our pre-existing views.

Example of Sensing/Processing Pipeline



- Kinesis: collect real-time, streaming data.
- Lambda: event-driven server-less computing platform.
- Machine Learning: ML models and predictions.
- Amazon SNS: Simple Notification Service.

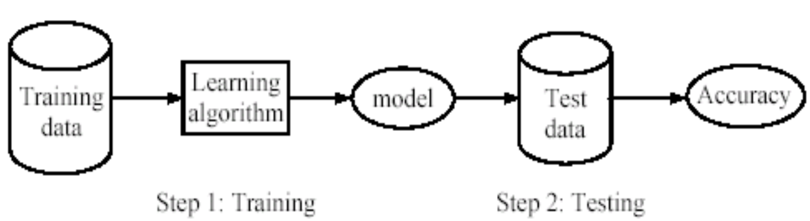
Machine Learning Options

- **Supervised learning**: classification is seen as supervised learning from examples.
  - Supervision: The data (observations, measurements, etc.) are **labeled** with pre-defined classes. It is like that a "teacher" gives the classes (supervision).
  - Test data are classified into these classes too.
- **Unsupervised learning** (clustering)
  - Class labels of the data are unknown.
  - Given a set of data, the task is to establish the existence of classes or clusters in the data.

Supervised Learning

- Learning (training): Learn a model using the **training data**.
- Testing: Test the model using unseen test data to assess the model accuracy

$$Accuracy = \frac{\text{Number of correct classifications}}{\text{Total number of test cases}},$$



Supervised Learning

- Given
  - a data set *D*,
  - a task *T*, and
  - a performance measure *M*,

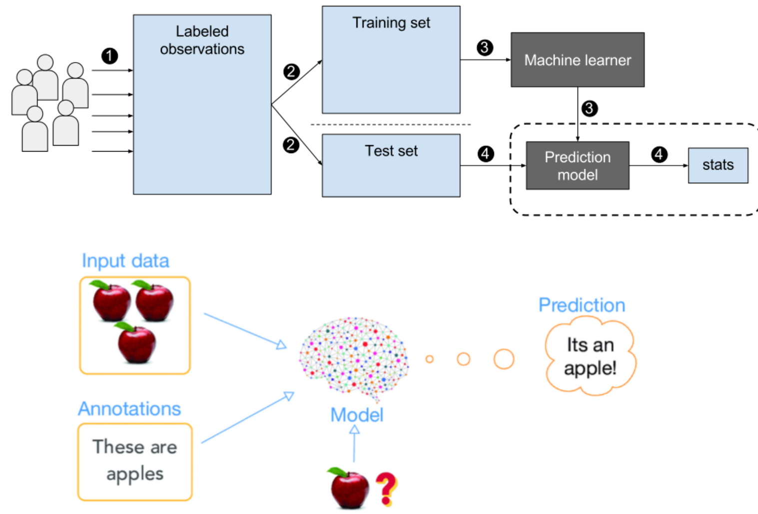
  a computer system is said to **learn** from *D* to perform the task *T* if after learning the system's performance on *T* improves as measured by *M*.
- In other words, the learned model helps the system to perform *T* better as compared to no learning.

Supervised Learning

Assumption: The distribution of training examples is identical to the distribution of test examples (including future unseen examples).

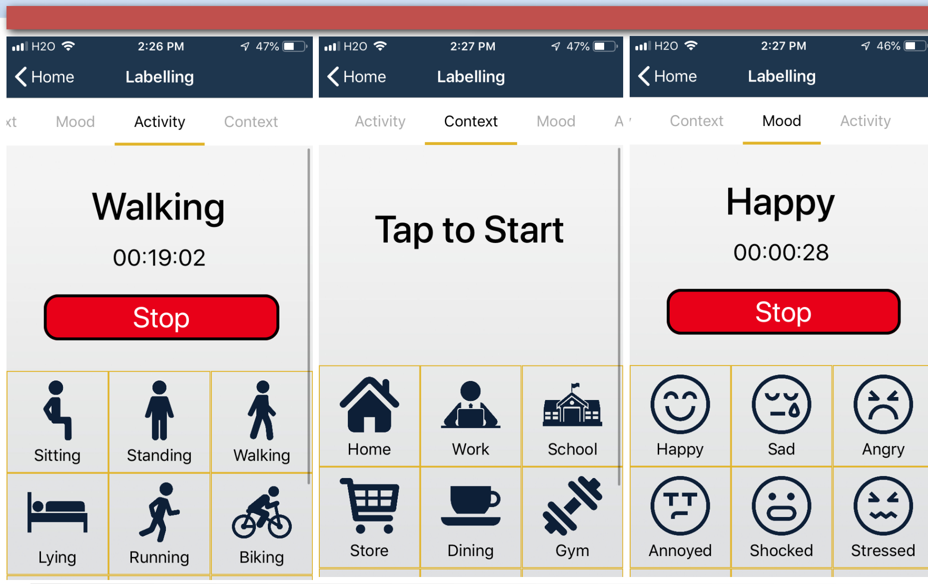
- In practice, this assumption is often violated to a certain degree.
- Strong violations will clearly result in poor classification accuracy.
- To achieve good accuracy on the test data, training examples must be sufficiently representative of the test data.

Supervised Learning



Computer Science and Engineering - University of Notre Dame

Labels (Annotations)



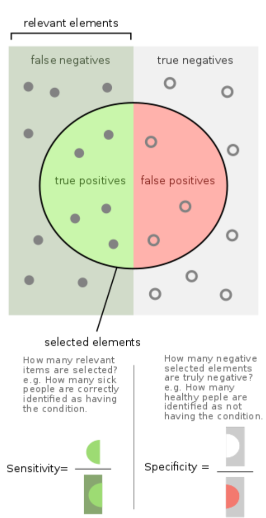
Computer Science and Engineering - University of Notre Dame

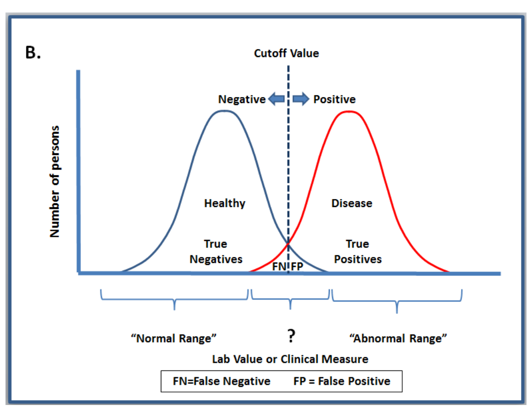
Validating Biomarkers

- Select candidates relevant to disease pathway.
- Identify and quantitate the association between the marker and the disease.

	Diseased (TD)	Healthy (TH)	
Test Positive	True positive (T_P)	False Positive (F_P)	PPV $= \frac{(TP)}{(TP)+(FP)}$
Test Negative	False Negative (F_N)	True negative (T_N)	NPV $= \frac{(TN)}{(FN)+(TN)}$
	Sensitivity $= \frac{(TP)}{(TP)+(FN)}$	Specificity $= \frac{(TN)}{(FP)+(TN)}$	

TP, FP, TN, FN



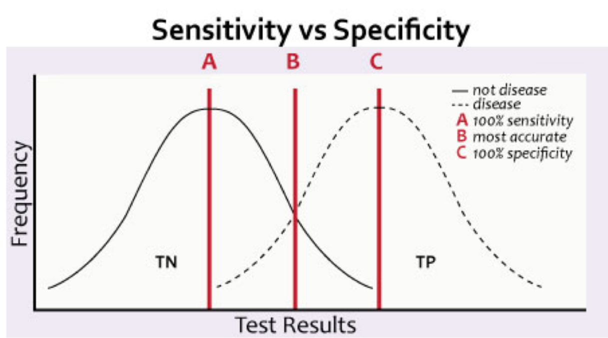


PPV, NPV

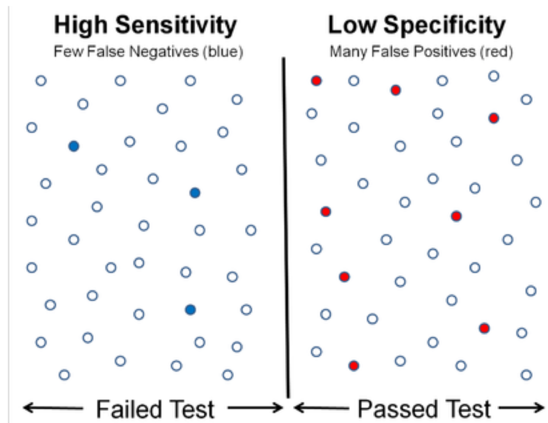
	PPV	NPV
Definition	% that a person with positive test is actually diseased.	% change that a person with negative test is actually disease free.
Use	Proceed with a patient with positive test	Proceed with a patient with negative test
Relation to prevalence	Low prevalence low PPV High prevalence high PPV	High prevalence low NPV Low prevalence High NPV

Sensitivity & Specificity

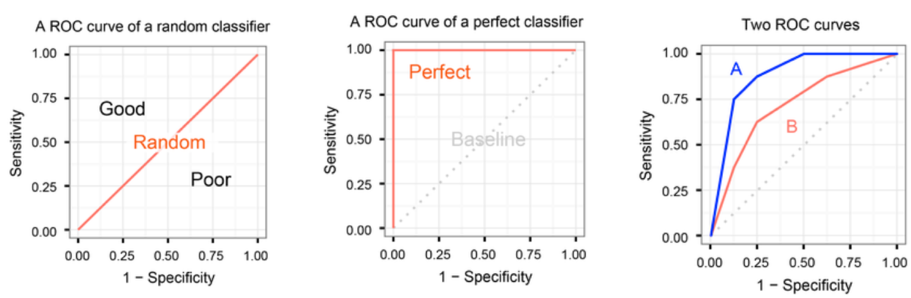
- Sensitivity: true positive rate; tests pick all diseased plus some without, i.e., they won't miss the disease.
- Specificity: true negative rate; tests pick only the diseased ones, but may miss some.



Sensitivity & Specificity

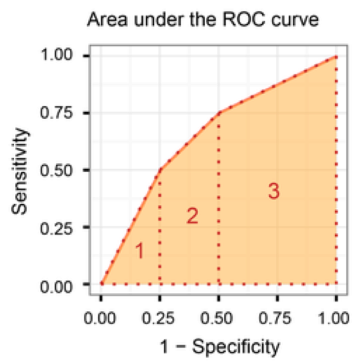


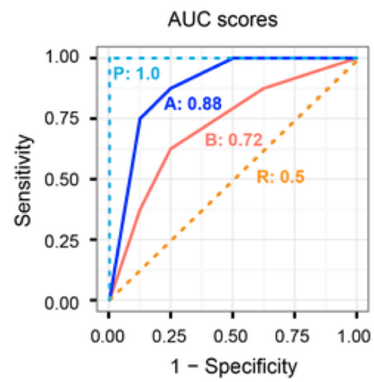
Receiver Operating Characteristics



Receiver Operating Characteristics

- Area under the ROC curve: AUC score





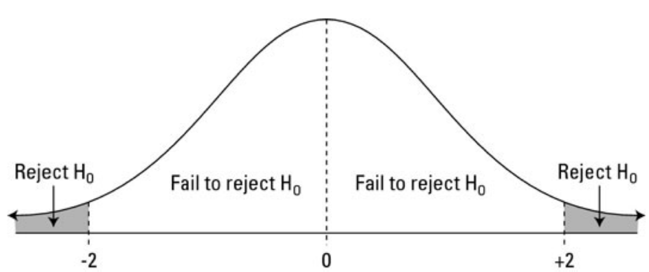
P-values

- Convenient & popular summaries of experimental results.
- P-value measures a sample's compatibility with a hypothesis.
- Example:
  - Does a disease affect a biomarker?
  - Take mean biomarker levels in healthy versus diseased samples and compute p value.
  - Indicates the probability that a difference in means at least as large as the one observed can be generated from random samples if the disease does not affect the mean biomarker level.

P-values

- Null hypothesis: “no association between a biomarker and a disease”.
  - A small *p*-value (typically ≤ 0.05) indicates strong evidence against the null hypothesis, so you reject the null hypothesis.
  - A large *p*-value (> 0.05) indicates weak evidence against the null hypothesis, so you fail to reject the null hypothesis.
  - *p*-values very close to the cutoff (0.05) are considered to be marginal (could go either way). Always report the *p*-value so your readers can draw their own conclusions.

P-values



Statistical Tests

Number of groups	Level of Measurement		
	Nominal	Ordinal	Interval/Ratio
1 group	χ^2 test	Kolmogorov-Smirnoff 1 sample test	t-test of sample mean vs. known population value
2 independent groups	χ^2 test	Mann-Whitney U test	Independent samples t-test
2 dependent groups	McNemar test	Wilcoxon test	Paired t-test
>2 independent groups	χ^2 test	Kruskal-Wallis ANOVA	ANOVA
>2 dependent groups	Cochran Q test	Friedman ANOVA by ranks	Repeated measures ANOVA

Independent Samples t-Test

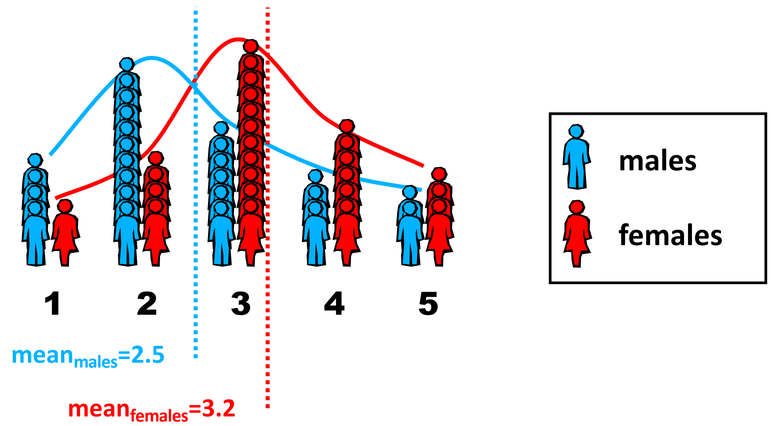
Males and females are asked a question that is measured on a five-point Likert scale:

To what extent do you feel that regular exercise contributes to your overall health?

1 **Strongly agree**
2 **Agree**
3 **Neither agree nor disagree**
4 **Disagree**
5 **Strongly disagree**

Do males and females differ in their response to this question?

Independent Samples t-Test



Independent Samples t-Test

- Use tools like SPSS, R, SAS, Excel, Matlab, Minitab, …

Group Statistics

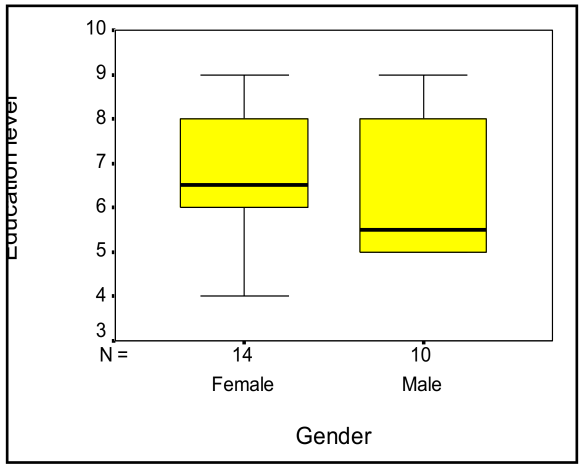
	GENDER	N	Mean	Std. Deviation	Std. Error Mean
EXERCISE	1 male	25	2.56	1.158	.232
	2 female	25	3.24	1.012	.202

Independent Samples Test

	t-test for Equality of Means			
	t	df	Sig. (2-tailed)	Mean Difference
EXERCISE	-2.212	48	.032	-.68

Mann-Whitney U Test

9	A graduate degree
8	Some graduate work
7	Completed college
6	Some college
5	Completed high school
4	Some high school
3	Completed grade school
2	Some grade school
1	No formal education



Mann-Whitney U Test

For each group, the Sum and mean of ranks Is computed.

Ranks

	GENDER	N	Mean Rank	Sum of Ranks
EDUC Education level	1 Female	14	13.46	188.50
	2 Male	10	11.15	111.50
	Total	24		

The test statistics suggest that males' and females' education levels do not differ in this population.

Test Statistics[b]

	EDUC Education level
Mann-Whitney U	56.500
Wilcoxon W	111.500
Z	-.807
Asymp. Sig. (2-tailed)	.420
Exact Sig. [2*(1-tailed Sig.)]	.437[a]

a. Not corrected for ties.

b. Grouping Variable: GENDER

Computer Science and Engineering - University of Notre Dame

Seminar Topic Selection

- To be done individually
- Identify a topic of interest, e.g.,:
  - Identify a technology and explore its medical use
  - Identify a medical challenge and explore how technology is used to address it
- Proposal:
  - Title of proposed topic
  - Your name
  - One paragraph (less than ½ page) describing focus of your chosen topic

- Find 3-5 relevant publications for your topic
- Prepare oral report in class, about 5-7 minutes
- Prepare written report (up to 3 pages)