Bio-Medical Computing
(6.872/HST.950)

Peter Szolovits, PhD (psz@mit.edu)
Gil Alterovitz, PhD (gil@mit.edu)
+ guest lecturers
Course Secretary: Fern DeOliveira (fernd@mit.edu)

http://stellar.mit.edu/S/course/6/fa13/6.872J/

Tue, Thu 9:30-11:00am, 34-304

Who Am I?

• Background is in Artificial Intelligence (AI)
• Got interested in medicine as an application of AI, in 1974
• Knowledge-based systems for diagnosis, therapy selection & management
• Early emphases
  • Reasoning based on causal, pathophysiological models
  • time, space and probability
• 1990’s revolution brought tons of data: e.g., MIMIC contains terabytes of ICU data
  • Shift from knowledge-based to empirical models
  • Data mining and machine learning: logistic regression, naive Bayes, Bayes networks, neural nets, support vector machines, conditional random fields, etc.
  • Challenge: combining general knowledge with experience
• Natural Language Processing to enable use of clinical data locked up in narratives
• Shift from institutional to personal systems: Guardian Angel (http://ga.org)
• NRC studies on patient privacy, health care challenges to IT, CMS IT revitalization

Why do People Care?

• “The Learning Health Care System” — Institute of Medicine
  • Growing availability of clinical data on past patient care
    • E.g., Partners has a Research Patient Data Registry (RPDR) containing records of ~4M patients
    • E.g., Kaiser Permanente has ~32 peta bytes of data (much of it imaging)
    • 2009 Stimulus Bill subsidizes hospitals, clinics, doctors’ offices to install electronic health records to meet “meaningful use” criteria
    • Improving machine learning methods to exploit such data
• “The $1000 Genome” — spread of genomic medicine
  • Phenotype = f(Genotype, Environment); what is f?
    • Theories of genetic causation: (a) Mendelian, (b) Common Variants, (c) Rare Variants, (d) neo-Mendelian, (e) ???
  • Cost
    • E.g., a Chevy contains more in employee health care costs than steel!

Inevitable Growth of Healthcare Spending

![Graph showing healthcare spending as a percentage of GDP from 1982 to 2009 for different countries: USA, France, Japan, Taiwan.](image)
The Challenge of Productivity

- Productivity = goods & services produced / resources consumed
- Staying in place is not good enough

What is the Productivity of Health Care?

Hard to know what outputs to measure
- Life expectancy
- Childhood mortality
- Quality of life

Di Backward to estimate resources consumed

UK Healthcare productivity falls 3.3 per cent

Comparison of 1901 and 2003 survival curves for US white women. From http://www.state.gov/g/oes/rls/or/81537.htm

Less Efficient Sectors Come to Dominate

Hypothetical sector growth over 30 years, assuming constant demand for each sector. Productivity rates from OECD 2009.

Standing Still is Not Good Enough

Chris Dede, Harvard Ed

http://theincidentaleconomist.com/wordpress/the-health-care-productivity-problem/
Vioxx and Heart Attacks

"Trends in inpatient stay due to MI were tightly coupled to the rise and fall of prescriptions of COX-2 inhibitors, with an 18.5% increase in inpatient stays for MI when both rofecoxib and celecoxib were on the market (P<0.001). For every million prescriptions of rofecoxib and celecoxib, there was a 0.5% increase in MI (95%CI 0.1 to 0.9) explaining 50.3% of the deviance in yearly variation of MI-related hospitalizations."

What is Evidence-Based Medicine?

- Randomized Controlled Clinical Trials
  - E.g., is drug A more effective than drug B for condition X?
  - Narrow selection of patient cases and controls
  - Careful collection of systematically organized data
  - Statistical analysis of outcomes
  - \( \Rightarrow \) Statistically significant conclusions

- But:
  - **Heterogeneity:** Most cases to which RCT results are applied do not fit trial criteria
  - **Short Follow-Up:** Trials run for limited times, but use is longer
  - **Small Samples:** Some effects are rare but devastating

- Instead: consider every patient’s experience as a source of knowledge by which to improve health care

Universal Electronic Health Records

- Rapid adoption in hospitals, likely driven by 2009 "Stimulus Bill"
  - "Basic EHR systems include ... patient demographics, patient problem lists, patient medication histories, clinical notes, electronic orders for prescriptions, laboratory results viewing, and imaging results viewing." — ONC

  (Data from http://dashboard.healthit.gov/HITAdoption?view=0)

- Positive factors for adoption: large size, urban location and HMO penetration

EHR Adoption in Doctors’ Practices

![Bar Chart showing adoption rates for Overall, Primary Care, Small Practices, and Rural Practices from 2008 to 2011.](http://dashboard.healthit.gov/HITAdoption/?view=0)

E-Prescribing

Number of Users of Surescripts Network

![Line Graph showing E-prescribing growth from 2008-12 to 2013-2.](http://dashboard.healthit.gov/HITAdoption/?view=0)

EHR Challenges

- Interoperability
  - Standards vs. evolution
    - Core data to take care of a patient: Demographics, insurance, diagnoses & problem lists, medications, lab values, allergies, recent encounters, images, care plan
    - E.g., HL7-CCD or CCR: text + optional codes (SNOMED, LOINC, …)
    - Comprehensive data: HL7-RIM
    - Instance of Berners-Lee’s Semantic Web issues
- Impediments to health information exchange
- Workflow changes
  - Errors
    - E.g., mortality 2.8% → 6.6% after CPOE system @ Pitt Children’s
  - Waste of effort
    - “Cut and Paste” records
  - Proprietary lock-in

Han et al. Unexpected increased mortality after implementation of a commercially sold computerized physician order entry system. Pediatrics (2005) vol. 116 (6) pp. 1506-12
3. Meaningful Use for Quality Improvement

- 2009 Recovery Act defines “Meaningful Use”:
  - Use EHR in a meaningful manner (e.g., e-prescribing)
  - Electronic exchange of health information to improve quality
  - Submit clinical quality measures

<table>
<thead>
<tr>
<th>Stage 1</th>
<th>Stage 2</th>
<th>Stage 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Electronically capturing health information in a standardized format</td>
<td>More rigorous health information exchange (HIE)</td>
<td>Improving quality, safety, and efficiency, leading to improved health outcomes</td>
</tr>
<tr>
<td>Using that information to track key clinical conditions</td>
<td>Increased requirements for e-prescribing and incorporating lab results</td>
<td>Decision support for national high-priority conditions</td>
</tr>
<tr>
<td>Communicating that information for care coordination processes</td>
<td>Electronic transmission of patient care summaries across multiple settings</td>
<td>Patient access to self-management tools</td>
</tr>
<tr>
<td>Initiating the reporting of clinical quality measures and public health information</td>
<td>More patient-controlled data</td>
<td>Access to comprehensive patient data through patient-centered HIE</td>
</tr>
<tr>
<td>Using information to engage patients and their families in their care</td>
<td></td>
<td>Improving population health</td>
</tr>
</tbody>
</table>

Reducing Readmissions

- Starting Oct 1, 2012, hospitals are penalized by Medicare if a patient is re-admitted for the same condition within 30 days
  - Initial focus conditions: Acute Myocardial Infarction, Heart Failure, Pneumonia
  - Risk adjustments for demographics, co-morbidities, patient frailty
  - Predictive models are worth $$$
    - E.g., Potentially avoidable 30-day re-admissions
      - Retrospective study of 10,731 discharges from Partners HealthCare
      - 2,398 (22.3%) readmitted within 30 days, 879 (8.5%) judged avoidable
      - Eliminated planned re-admissions and those for other conditions
      - Logistic regression model predicts avoidable re-admission with AUC=0.71, good calibration
      - Important variables: hemoglobin at discharge, oncology service, Na+ level at discharge, procedures performed, non-elective admission, number of admissions in past 12 months, length of stay

Medical Informatics

- Intersection of medicine and computing
- Plus theory and experience specific to this combination
  - =Medical Computing, ~Health Informatics

- Science
- Applied science
- Engineering
Types of Bio-Medical Informatics

- Cellular level: *Bioinformatics, Systems Biology*
- Patient level: *Clinical Informatics, Health I., Medical I., …*
- Population level: *Public Health I.*
- Imaging Informatics

Bio-Medical Informatics

- Phenotype = Genotype + Environment
- In humans, we rely on “natural experiments”
- Measurements
  - Genotype: sequencing, gene chips, proteomics, etc.
  - Environment: longitudinal surveys, etc.
  - Phenotype: clinical records, assembled to longitudinal data

Outline (today)

- What is biomedical informatics?
- BMI is defined by goals and methods of health care
- The genomic revolution
- The science of health care
  - Genotype, phenotype, environment
  - From associations to mechanisms
- What is health?
- Practice of health care
- Challenges

Outline (semester)

- Clinical and Genomic Data
- Methods of modeling
- Combining clinical and genetic data
- “Translational medicine”
- Engineering the health care system
- Decision support to improve health care
- Personalized medicine
- Public health
- The developing world
- *Your projects*
The Medical Cycle

- Patient
  - Initial presentation
  - Therapy
  - Data
    - Interpret
  - Information
    - Formulate
  - Plan
    - Diagnosis

Care Processes

- Data: instrumentation, monitoring, telemetry
- Information: interpretation, filtering, sampling, smoothing, clustering
- Diagnosis: inference, model-based reasoning, classification
- Prognosis: prediction, natural course, experience
- Therapy: planning, predicting effects, anticipating

Meta-level processes

- Acquisition and application of knowledge
- Education
- Quality control and process improvement
- Cost containment
- Reference (library)

Enterprise-level Clinical Process Automation...
The “Learning Health Care System”

The “Learning Health Care System” involves the processes of Model, Plan, Intervene, Analyze, Observe/Measure, and making decisions based on these processes.

The diagram depicts various components such as Patient, Community Care, Self Care, Evaluate, Assess, Plan, Act, Schedule, Refer, Visit, Activation, Mgt., Team, Health Mgt., Health Record, Measure, Observe/Measure, Process, Medical content, and Dogma.

**Dogma**

- **Phenotype** = **Genotype** + Environment
  - Traits: Gene sequence
  - Diseases: SNP’s
  - Behaviors: Expression data
  - Don’t be exposed to Diet, smoking, drugs, ...
  - Insults and injuries: Exposures
  - What is the functional form?
  - How do we investigate these relationships?
  - Can we take advantage of the exponential growth of genomic data?

**Growth in Gene Expression Omnibus Measurements**

- Fall 2004: ~30,000 samples, ~5B measurements
- 9/2010: ~472,929 samples
- 9/2013: ~987,662 samples

Where are the Phenotype and Environment-related Data?

- Perform Controlled Experiments?
  - Unethical using human subjects!!!
  - OK on rats.

Figure 4: GEO submission statistics. Cumulative individual sample measurements submitted to GEO are shown. Data are presented by quarter since operational began on July 25, 2000.

http://ncbi.nlm.nih.gov/geo/
Experimental Subjects

High-throughput phenotyping at Medical College of Wisconsin

Where are the Phenotype and Environment-related Data?

- Environment
  - (Hardest to get)
  - Questionnaires,
    - e.g., Nurses’ Health Study, Framingham Heart Study
  - Monitoring
    - e.g., LDS hospital infectious disease monitors
  - Metagenomics
    - DNA of the stuff living in your guts

- Phenotype
  - “Natural Experiments”
  - Clinical Data

The fantasy: Informatics for Integrating Biology & Bedside

I2b2: I. Kohane, et al.
Plausibility
Butte & Kohane, *Nature Biotech* 2006

- Phenome-Genome Network
  - Gene Expression Omnibus
    - expression data
    - annotations: tissue, disease, exp. conditions, …
  - Interpret annotations to UMLS
  - Differential expression vs. condition
- Interesting relations:
  - 11 genes & aging
  - DDX24 and leukemia
  - 2 genes & injury

Clinico-Genomic Research

- Identify a highly specific clinical population, and controls
- Gene-wide association studies (GWAS)
- Hope that notable differences appear between those with/those without disease
- Disease models:
  - Mendelian
  - Single-nucleotide polymorphisms
  - Private variation
  - ?

Time scale in medicine

- Cure—usually acute illness
- Manage—long-term, chronic illness
- Prevent
- Predict (especially based on genetics)

WHO Constitution defines “health”

“a state of complete physical, mental and social well-being and not merely the absence of disease or infirmity”

- Physical
- Mental
- Social
  —very hard to measure
Distribution of Ages

- Life table deaths by year (Japan, 1989)

Life table death rates by age

Measures of Health

- Longevity at birth (CIA World Fact Book, 2001)

<table>
<thead>
<tr>
<th>Country</th>
<th>Male</th>
<th>Female</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rwanda</td>
<td>38.35</td>
<td>39.65</td>
</tr>
<tr>
<td>Kenya</td>
<td>46.57</td>
<td>48.44</td>
</tr>
<tr>
<td>South</td>
<td>47.64</td>
<td>48.56</td>
</tr>
<tr>
<td>Cambodia</td>
<td>54.62</td>
<td>59.12</td>
</tr>
<tr>
<td>Brazil</td>
<td>58.96</td>
<td>67.73</td>
</tr>
<tr>
<td>Russia</td>
<td>62.12</td>
<td>72.83</td>
</tr>
<tr>
<td>Albania</td>
<td>69.01</td>
<td>74.87</td>
</tr>
<tr>
<td>USA</td>
<td>74.37</td>
<td>80.05</td>
</tr>
<tr>
<td>Japan</td>
<td>77.62</td>
<td>84.15</td>
</tr>
</tbody>
</table>
Causes of death
(industrialized countries, 1989)

<table>
<thead>
<tr>
<th>Cause</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Circulatory system</td>
<td>48%</td>
</tr>
<tr>
<td>Malignant neoplasms</td>
<td>19%</td>
</tr>
<tr>
<td>Accidents</td>
<td>7%</td>
</tr>
<tr>
<td>Others</td>
<td>26%</td>
</tr>
</tbody>
</table>

Quality of life

- Value of a total life depends on
  - Length (assume now is \( N \))
  - Quality \( q(t) \) over time
  - Discounts \( g(t) \) for future or past
    (depends very much on what the value is to be used for)

\[
V_N = \int_{t=0}^{T} q(t)g(t-N)dt
\]

Modeling life quality

Mortality, Disability, Morbidity

![Graphs showing mortality, disability, and morbidity over age](image-url)
Top 10 Chronic Conditions
Persons aged ≥ 65

<table>
<thead>
<tr>
<th>Condition</th>
<th>Both</th>
<th>Male</th>
<th>Female</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arthritis</td>
<td>49.6</td>
<td>40.7</td>
<td>55.7</td>
</tr>
<tr>
<td>Hypertension</td>
<td>39.0</td>
<td>33.0</td>
<td>43.2</td>
</tr>
<tr>
<td>Hearing impairment</td>
<td>30.0</td>
<td>35.2</td>
<td>26.3</td>
</tr>
<tr>
<td>Heart disease</td>
<td>25.7</td>
<td>26.9</td>
<td>24.9</td>
</tr>
<tr>
<td>Orthostatic impairment</td>
<td>16.8</td>
<td>15.7</td>
<td>17.8</td>
</tr>
<tr>
<td>Cataracts</td>
<td>15.5</td>
<td>11.3</td>
<td>18.4</td>
</tr>
<tr>
<td>Chronic sinusitis</td>
<td>15.2</td>
<td>13.7</td>
<td>16.2</td>
</tr>
<tr>
<td>Visual impairment</td>
<td>10.1</td>
<td>12.0</td>
<td>8.8</td>
</tr>
<tr>
<td>Genitourinary</td>
<td>9.9</td>
<td>11.3</td>
<td>8.9</td>
</tr>
<tr>
<td>Diabetes</td>
<td>8.9</td>
<td>7.8</td>
<td>9.7</td>
</tr>
</tbody>
</table>


Societal quality of life

- Aggregation of individual qualities
- Equity (distributions)
- Is more better? (Population control.)
- Is less better?
- How much to spend?

Who makes decisions?

“In those days there was no bureaucratic regimentation, there were few forms to fill out, malpractice premiums were affordable, and the overhead costs of running a practice were reasonable. Our bills were simple, spelled out so anybody could understand them without the use of codes. Patients usually paid their own bills, promptly too, for which an ordinary receipt was given. Hospital charges were set by the day, not by the aspirin. Medical care was affordable to the average person with rates set by the laws of the marketplace, and care was made available to all who requested it regardless of ability to pay. Doctors were well respected; rarely were we denigrated by a hostile press for political reasons. Yes, in the days before government intervention into the practice of medicine, doctor's fees were low, but the rewards were rich; those were truly the ‘golden years’ for medicine.”

Edward Annis, past President of AMA
*Code Blue*, 1993

Aggregation

- Trend: social aggregation leads to decisions at a larger scale
  - Multi-specialty providers
  - Government guarantees and mandates
  - Risk sharing
  - Oregon-wide spending “optimization”;
  - British NHS
Changing Context of Health Care

- Fee-for-service
- CMS (Center for Medicare and Medicaid Services pays for Medicare, ~40% of US $
- Capitation
  - HMO’s (Health Maintenance Organizations) take overall responsibility to care for patient for fixed fee
  - Pushing risk down to the physician or group

Determining Factors:

$£€¥R$

Exponentially growing expense of health care

- More healthcare than steel in GM cars
- Increased demand
  - Much more possible
  - Better tests, therapies
  - High human motivation
- No pushback
- Waste
  - Unnecessary procedures
    - ½ of health expenses in last year of life
  - Marginally useful procedures
    - Defensive medicine
  - Bad Medicine
    - IOM: 48-98K “unnecessary” deaths/year

Managed Care

“Decisions that were once the exclusive province of the doctor and patient now may be examined in advance by an external reviewer—someone accountable to an employer, insurer, health maintenance organization (HMO), or other entity responsible for all or most of the cost of care. Depending upon the circumstances, this outside party may be involved in discussions about where care will occur, how treatment will be provided, and even whether some treatments are appropriate at all.”

Controlling Costs and Changing Patient Care IOM, 1989
**How is care managed?**

- Active case management:
  - Preadmission review
  - Continued-stay review
  - Second surgical opinion
- Selective case management—high-cost cohorts
- Institutional
  - Capitation
  - Institutional arrangements (referrals, hospitals, pharmacies, …)
  - Control “leakage”

**Managed Care Scorecard**

- “U.M. has helped to reduce inpatient hospital use and to limit inpatient costs…”
- “The impact of U.M. on net benefit costs is less clear. Savings on inpatient care have been partially offset by increased spending for outpatient care and program administration.”
- “U.M. … does not appear to have altered the long-term rate of increase in health care costs.”
  
  IOM, 1989

**ObamaCare**

- Universal coverage: everyone must get insurance
  - Employer
  - Insurance collaborative
  - Government (?) — rejected
- Insurance companies cannot deny insurance, cancel coverage, impose reimbursement limits based on illness, past or present
- Government assistance to poor people, small companies
- Health Information Technology (HIT) to smooth info flow
- Cost savings from avoiding billing disputes, ceasing to reimburse only procedures, evidence-based medicine.

**Quality Improvement**

- IOM Study: 96,000 US deaths/year from medical error (perhaps half preventable?)
- Information intervention at the point of decision making can improve decisions
- DPOE: Direct Physician Order Entry allows such intervention
- Leapfrog Group: Large employers ($$$) require DPOE from providers
- Patient Involvement: Indivo Health, Google Health, Microsoft Healthvault
### Implications of Health Care Organization for Informatics

- Money determines much
  - Medicine spends 1-2% on IT, vs. 6-7% for business overall, vs. 10-12% for banking
  - “Bottom line” rules, therefore emphasis on
    - Billing
    - Cost control
    - Quality control, especially if demonstrable cost savings
    - Retention and satisfaction (maybe)
  - Management by accountants

### Challenges

- Computerized Medical Records (EMR/CPR/…)
- Usability of systems in the workflow of health care
- Large-scale “Engineering Systems” problem
- Genomic Medicine