IOT, Data and Healthcare: How do we get it right?

Wendy J. Nilsen, PhD
Program Director, Smart and Connected Health,
Directorate for Computer & Information Systems,
National Science Foundation
Health how we think of it.
U.S. Health in International Perspective
Shorter Lives, Poorer Health

Mortality from Non-Communicable Disease

Mortality from Communicable Disease
U.S. Health in International Perspective

Shorter Lives, Poorer Health

FIGURE: Causes of Death for U.S. Men Before Age 50, Compared with Average of Peer Countries, 2006-2008

NOTE: CVD is cardiovascular disease

U.S. Health in International Perspective
Shorter Lives, Poorer Health


Maternal conditions
Communicable and nutritional conditions
Intentional injuries
Drug-related causes
Perinatal conditions
Unintentional injuries
Cardiovascular disease (CVD)
Noncommunicable diseases, excluding CVD
All causes

Years of Life Lost Before Age 50, 2006-2008

Average of Peer Countries
United States

NOTE: CVD is cardiovascular disease

Children in the US have the **highest probability of dying before age 5** of any of the peer countries.

- In 2004, 11% of US deaths before age 5 were from injuries.
- In 2006, the US had the highest rate of child deaths due to negligence, maltreatment, or physical assault.
- The violent death rate among US boys aged 1-4 has exceeded the OECD average since the late 1960s.
- The US is ranked 24th of 30 (OECD) and 21 of 21 (UNICEF) on selected measures of children’s well-being.
Half of Premature Deaths are Preventable

Poor Diet, Lack of Exercise Impede Progress on Reducing Early Deaths.

PERCENT OF EARLY DEATHS (BEFORE AGE 80) BY CAUSE, 1990 AND 2010

NOTE: Deaths due to medical errors not calculated in 1990.

Aging Demographics
Percentage over age 60

Shifting the Health Development Curve to Shift the Cost Curve

Neil Halfon, 2012 Transforming Child Health
Evolving Healthcare System

Health 1.0
- Germ Theory
- Acute Illness & Infectious Disease
- Reduce Deaths

Health 2.0
- Multiple Risk Factors
- Chronic Disease
- Reduce Disability

Health 3.0
- Complex systems
- Prevention of Disease
- Optimal Health

Adapted from Neil Halfon 2012
How do we get to Health 3.0 and beyond?
Siloed Sciences

- Computing
- Psychology
- Neurosciences
- Medicine
- Engineering
- Mathematics
Reasons to Collaborate

- “Wicked problems” - can’t be solved by a single discipline
- Access to expertise or particular skills
- Access to equipment, resources, or funding
- Enhancing trainee education
- Impact
Levels of Team Science

- **Unidisciplinary**
  - Researchers from a single discipline work together to address a shared research problem

- **Multidisciplinary/Interdisciplinary**
  - Researchers from different disciplines work independently, sequentially from own disciplinary perspective to address a shared research problem

- **Transdisciplinary**
  - Researchers from different disciplines work together to integrate theory, methods or concepts that extend discipline-specific language and models
Competence & Synergy

Hierarchy of Competence

Unconscious Competence

Conscious Competence

Conscious Incompetence

Unconscious Incompetence

Right Intuition

Right Analysis

Wrong Analysis

Wrong Intuition

1974 Gordon Training Intl
What did you say?

Language of Interdisciplinary Teams
Research increasingly done by investigator teams

“Sometimes I think the collaborative process would work better without you.”
Growth of Funding for Interdisciplinary Science

**NSF Total Funding** ($6.9 bill 2010)

<table>
<thead>
<tr>
<th></th>
<th>1982</th>
<th>2010</th>
</tr>
</thead>
<tbody>
<tr>
<td>Single investigators</td>
<td>88%</td>
<td>38%</td>
</tr>
<tr>
<td>Multiple Investigators</td>
<td>12%</td>
<td>62%</td>
</tr>
</tbody>
</table>

**NIH Multi-PI proposals – Increasingly annually**

- 2010 - 10.33%
- 2013 - 15.33%
- 2016 – 20.1%
Next Steps
What Parts of IoT make sense in health?
Questions

- Is this the right problem?
- Will the potential solution be better than what we have?
- Is there additional information that would make this more effective?
- How often does it need to be done?
- Who will be your user?
Example: Six Minute Walk

- The six-minute walk test measures the distance an individual is able to walk over a total of six minutes on a hard, flat surface. The goal is for the individual to walk as far as possible in six minutes. The individual is allowed to self-pace and rest as needed as they traverse back and forth along a marked walkway.

Additional Information - available and useful?

- What other data are available? Can the Internet of Things be useful?
- Not all data has to come from medical devices.
- What about patient-reported data on outcomes, symptoms or context?
- Where does the data go and how does it get there?
- How does the data appear to the user?
How often to measure?

- Example: A common dilemma in clinical medicine is whether to treat asymptomatic patients who present with bacteria in their urine. Because of increasing antimicrobial resistance, it is important not to treat patients with asymptomatic bacteriuria unless there is evidence of potential benefit.
Who will be your user?

- Now and in 2020....
Technology and Health

<table>
<thead>
<tr>
<th>Table. The Digitally Connecting World 2010-2020</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>World population, billion</td>
</tr>
<tr>
<td>No. connected</td>
</tr>
<tr>
<td>Devices, billion</td>
</tr>
<tr>
<td>Devices per person</td>
</tr>
<tr>
<td>No. of smartphone subscriptions, billion</td>
</tr>
<tr>
<td>No. of wireless hotspots, million</td>
</tr>
<tr>
<td>No. of transistors, million/chip, nm</td>
</tr>
<tr>
<td>No. of sensors</td>
</tr>
<tr>
<td>No. of individuals sequenced</td>
</tr>
</tbody>
</table>

Digital Medical Tools and Sensors: Topol, Steinhubl and Torkamani, JAMA, 2015
The Pathway to Patient Data Ownership and Better Health

Digital health data are rapidly expanding to include patient-reported outcomes, patient-generated health data, and social determinants of health. Measurements collected in clinical settings are being supplemented by data collected in daily life, such as data derived from wearable sensors and smartphone apps, and access to other data, such as genomic data, is rapidly increasing. One projection suggests that a billion individuals will have their whole genome sequenced in the next several years. These additional sources of data, whether patient-generated, genomic, or other, are critical for a comprehensive picture of an individual’s health.

Enabling access to personal health data, clinical or patient-generated, may benefit patients and health care professionals. Research is beginning to show that providing patients with their complete health data may help improve their health. For example, timely access to laboratory results can increase patient engagement. Access to physician notes after appointments appears to encourage individuals to improve their health and participate in decision-making, with electronically engaged patients demonstrating more successful medication adherence, quality outcomes, and symptom management. Economic benefits may include the avoidance of duplicative imaging or laboratory tests. Clinicians may also benefit from more informed patients. For example, they may score higher in quality performance programs because patients who are more informed may better adhere to treatment plans and hence may improve clinician health data. For this to proceed, control of health data must be transferred to the patient or the patient’s authorized representative.

More specifically, to obtain active patient engagement and health system improvement, three components are necessary: (1) common data elements that enable the sharing and merging of health data from multiple sources; (2) a patient encounter data receipt, comprised of relevant health data from each health care encounter, automatically pushed to the patient’s complete digital health record; and (3) a contract between patients and third-party health data managers (eg, health care organizations and commercial entities) that enables individuals to control their longitudinal digital health record. Most of these components already exist in some form, requiring only minor adjustments to effect health system transformation.

Clinicians, patients, and health care systems need a way to efficiently receive, integrate, understand, compute, and use digital health data from other practitioners and health encounter locations. This requires the merging of what is often disparate data from multiple sources, and the most effective way to do this is to establish common data elements agnostic of any particular vendor’s electronic health record (EHR) system.

With widespread implementation of common data elements and value sets, semantic and clinical interoperability can be achieved, and health information can be merged, while maintaining data integrity. New initiatives, such as the Standard Health Record, that focus on standardizing data within health records instead of solely on exchange standards enable the development of one complete, digital health record per patient containing health data merged from all of a patient’s clinicians and related health data sources.

Patients need and deserve the opportunity to control their health data. Despite growing evidence of such benefits, all
IoT and Data Challenges and Opportunities

New methods should be:

- Efficient
- Standardized
- Personalizable
- Scalable
- Domain-Sensitive
- Explanatory
- Able to deal with Uncertainty
- Visually Representable
Challenges and Opportunities

Data Quality

- Methods to identify batch effects, “missingness” and non-representativeness
- Simulate missing data based on redundancies in health data sets
- Methods that would allow real-time assessment and information on problematic data
- Dealing with biases

Data Validity

- Does the data represent what we think it does?
- Is the variance important – signal or noise?
Challenges and Opportunities

Data Acquisition

• Unobtrusive and minimally invasive
• Capitalize on the Internet of Things
• Capture new modalities – such as dynamic imaging

Accessibility

• Databases allow efficient data access and selection, providing clarity and transparency of the relational aspects of the data
• Interventions may require nearly real-time data availability and new approaches that integrate database technology with embedded algorithms
Challenges and Opportunities

Performance Metrics

- Performance needs or metrics in practice and in research, across the spectrum of health from basic to translational to clinical to population realms
- How do we assess when a metric is “good enough”?
- Quantified with benchmark datasets?

Scaling Inferential Methods

- Dealing with heterogeneity at scale
- Thresholds when existing methods no longer meets the needs (not new for new sake)
- Can one partition problems and have algorithms work on sub-problems and bring them together?
- Can parallelism work in health?
Challenges and Opportunities

Reasoning with Heterogeneous Data

• Acceptable error rates
• Reasoning across scales and time frames

Modeling Complex Temporal Data

• What are systemic methods to combine (“join”) such data to find relationships amongst them, or access this data?
• How do we discover discontinuities in such streams, especially collection protocol changes?
Challenges and Opportunities

Privacy and Security

- Access control and authentication
- Confidentiality and anonymity
- Trustworthy control
- Accountability
- Medical device security
Challenges and Opportunities

Interoperability

- Technical needs for interoperability - platform?
- Standards needed for harmonizing data and making it interoperate smoothly

Data Infrastructure

- Development of data structures and database schemas for multiscale representation.
- Graphical databases across domains and specialties that allow analysis of relationships and discovery of new relationships.
Challenges and Opportunities

Collaboration & Training

• What levels of training or kinds of tools will bridge the communication gap between domain experts and data science?

• Coordination requirements tailored to different types of users as well as data summarizations for clinical decision making.
• Health, medical and rehabilitation research can be found in many areas in NSF and within the mission of several cross-directorate initiatives
• It is a case of use-inspired basic research. The scientific advances in basic science can be in computing, information science, engineering or social or behavioral science. The benefit to health research is important, but second to the advances in basic science.
• Major homes for this research:
  • Smart and Connected Health
  • Cyber-physical Systems
  • National Robotics Initiative 2.0
  • Smart and Autonomous Systems
  • Core Programs
Pasteur’s Quadrant

Donald E. Stokes, Pasteur’s Quadrant – Basic Science and Technological Innovation, Brookings Institution Press, 1997
Goal: Seek improvements in safe, effective, efficient, equitable, & patient-centered health through innovations in fundamental computer & information sciences, engineering & social, behavioral & economic sciences

- Funded work must include & address:
  - A key health problem
  - Fill in research gaps that exist in science & technology in support of health & wellness
  - Include a research team with appropriate expertise in the major areas involved in the work

- Activities should complement rather than duplicate core programs of NSF & NIH as well as those of other agencies
To transform health:

- from individual data to connected people and systems
- from experienced-based to data- and evidence-driven
- from health care to **health** that extends to the **home**, workplace and community
- Move focus from devices to connections between data, devices, systems and people
Smart Health Research Areas

**Health Information Infrastructure**

*Infrastructure to enable connections*

- Integration of EHR, contextual, clinical and patient data
- Access to information, data harmonization
- Tools to enhance smart health research

**Connected Data**

*Reasoning with heterogeneous data under uncertainty*

- Heterogeneous and messy data
- Datamining, machine learning, deep learning
- Inference, visualization, decision support system

**Connected Systems**

*Multifunctional devices connected to systems*

- Systems for empowering patients
- Models of readiness to change
- State assessment from images, video

**Connected People**

*Effective, multidirectional flows of information and support*

- Enhancing communication between providers, patients, and caregivers
- Assistive technologies embodying computational intelligence
Cyber-physical systems (CPS) are engineered systems that are built from, and depend upon, the seamless integration of computational algorithms and physical components. Advances in CPS will enable capability, adaptability, scalability, resiliency, safety, security, and usability that will far exceed the simple embedded systems of today.

- Closed Loop
- Human in the Loop
The NRI-2.0 program focuses on issues of scalability: how teams of multiple robots and multiple humans can interact and collaborate effectively; how robots can be designed to facilitate achievement of a variety of tasks in a variety of environments, with minimal modification to the hardware and software; how robots can learn to perform more effectively and efficiently, using large pools of information from the cloud, other robots, and other people; and how the design of the robots' hardware and software can facilitate large-scale, reliable operations.
Large-Scale Probabilistic Phenotyping Applied to Patient Record Summarization

PI: Noemie Elhadad, SCH

Technical Approach:
• Uphenome: unsupervised probabilistic graphical model, where phenotypes (latent variables) are learned as a mixture of observed patient observations (free-text notes, meds, labs, diagnosis codes...)
• Gphenome: Uphenome + grounded phenotypes through informative priors, as determined through existing knowledge bases
• Temporal phenotyping: joint survival analysis that models multiple patients and multiple diseases; scalable variational inference algorithm applied to 13,000 patients and 8,800 conditions
• Interactive visualization of phenotypes in patient timeline designed and validated through user-centered design

Motivation:
• The electronic health record (EHR) contains a wealth of observations. There is a critical need for high-throughput phenotyping algorithms to help (1) study diseases, their progressions, and their interactions through time; (2) visualize and make sense of patient’s longitudinal records

Transformative:
• First attempt to learn jointly common phenotypes across large patient population (1,000 phenotypes; 350,000 patient records)
• Phenotypes are validated and used in a clinically meaningful task – patient record summarization and visualization, which mitigates information overload at the point of patient care

Broader Impacts:
• The HARVEST patient record summarizer deployed and used daily by 1,000 clinicians monthly at NewYork-Presbyterian Hospital (NYPH)
• Beyond point of care use scenario, users include quality abstraction nurses; in a study focused on their workflow, use of HARVEST yielded 20-minute average time savings on patient review (2-hour process otherwise); 90% of quality abstractors in study adopted HARVEST as part of their workflow

Contacts:
• Noémie Elhadad, Columbia University

Example of learned phenotype for SLE

lupus ana sle complement rheum anti mg ab rash absent esr ulcers igg
rheum anti dna alerosis mt antibody urina systemic disaes neg rheumatolegy c4 complement
random urine protein random antibodies urine protein random urine creatinine
710.0 systemic lupus erythematosus
Untangling Complex Diseases into Subtypes

PI: Suchi Saria, SCH

Motivation:
- Scleroderma, Lupus are examples of phenotypically heterogeneous diseases, and there are few proven solutions that predict an individual’s course.
- Data accumulated from routine clinical visits may help to better understand disease subtypes.
- Mature machine learning and statistical inference tools for leveraging electronic health data do not yet exist.

Transformative:
- Subtypes can help to improve delivery of care to individuals.
- Differences between subtypes can motivate new directions for basic medical research.

Broader Impacts:
- Developing computational tools for improving the effectiveness and efficiency of care in complex diseases through the use of electronic health data.
- Training PhD students and postdoctoral fellows in the application of statistical machine learning to medicine and health.
- Exposing undergraduates to high-impact, cross-disciplinary research.

Contacts:
- Suchi Saria
- Laura Hummers, Fredrick Wigley, Antony Rosen
- www.cs.jhu.edu/~ssaria/autoimmunedisease.html

Technical Approach:
- Probabilistic graphical models encode domain knowledge, and model uncertainty and dependencies between heterogeneous data.
- Inference algorithms allow reasoning about unobserved phenomena conditioned on the available data.

Progress:
- Developed the Probabilistic Subtyping Model (PSM) for discovering subgroups of individuals with similar disease activity trajectories using longitudinal clinical markers.
- Results from analyses using PSM have motivated new clinical investigations into differences in antibody expression across discovered subgroups.
From Critique to Collaboration: Rethinking Computerized Clinical Alerts

PI: Davide Bolchini, SCH

**Motivation:**
- Drug safety alerts are critical for patient safety but largely ignored by doctors during medication prescribing.
- Despite efforts to improve design and reduce alert fatigue, physicians continue to distrust computerized recommendations.

**Transformative Outlook:**
- To improve alerts, we must first look at how to improve the trust between physician and computerized advice.
- We explore the foundational principles of what physicians consider important when taking advice from peers.
- We use this knowledge to create novel designs for drug safety guidance that elicit physician trust and a sense of collaboration.

**Broader Impacts:**
- Potentially reduce the over 2M adverse drug events per year by improving the safety of drug prescribing.
- Translating findings to inform EMR systems design through the Regenstrief Institute.
- Spread adoption to industry with EHR vendors and NIST.

**Technical Approach:**
- **Formative field studies** in clinical settings to unearth key factors in sharing trusted advice among doctors when prescribing drugs.
- **Invent, design and deploy** novel drug safety alert interfaces to convey drug safety information to providers in a more trusted manner.
- **Evaluative sessions** in the lab and with physicians in central Indiana hospitals to assess the physician’s user experience (UX) and effect of the proposed designs on alert compliance.

**Contacts:**
- Davide Bolchini (PI), Indiana University
- Jon Duke (Co-PI), Regenstrief Institute
- Eskenazi Health
- bit.do/trusted-alerts

**Progress:**
- Completed field studies with physician hospital teams.
- Developed and evaluated 10 types of novel trust-based alerts.
- Discovered novel, more holistic metrics for alert effectiveness beyond binary compliance, aimed at supporting physician’s reflection and consideration of alternative courses of actions.
Replicating Clinic Physical Therapy at Home: Touch, Depth, and Epidermal Electronics in an Interactive Avatar System

PI: Pamela Cosman, SCH

Motivation:
• Physical therapy is crucially important for stroke patients; often slowed by poor adherence to home therapy regimens
• **Goal:** develop a home-based system that integrates unobtrusive wireless sensors with avatar rendering and machine learning
• Provide effective guidance for patients and caregivers

Transformative:
• Epidermal electronics like a temporary tattoo
• Hand and body pose estimation and tracking algorithms
• Cloud-based machine learning and avatar rendering algorithms

Broader Impacts:
• Assistive system useful for many healthcare training applications
• Finger tracking useful for sign language recognition
• Flexible electronics valuable for many types of physiologic monitoring

Contacts:
• Pamela Cosman, UC San Diego
• Todd Coleman, Sujit Dey, Sri Kurniawan, Truong Nguyen, Carter McElroy
• http://esdat.ucsd.edu/Mobile_Health_v2.html

Technical Approach:
• Creating a cloud-based system for motion monitoring and a user interface for guidance
• Comparison of spatiotemporal trajectory of a given motion with that of a reference motion
• Development of flexible, adhesive-integrated antennas and sensors

Progress:
• Hand articulation tracking using an adaptive hand model
• Demonstration of adhesive-integrated flexible antenna on skin transmitting Bluetooth over 150 ft
• Dynamic alignment of motion sequences to evaluate motion correctness
• Study of finger force steadiness and repeatability

Patient exercises at home with cloud-based guidance
Crafting a Human-Centric Environment to Support Human Health Needs
PI: Diane Cook, SCH

Technical Approach:
• We perform real-time activity recognition of smart home sensor data “out of the box”
• Machine learning techniques map activity parameters to assessment values
• Activity forecasting drives activity prompting intervention

Motivation:
• Design smart environment technologies to perform automated health assessment and intervention

Transformative:
• Our team combines expertise from machine learning, pervasive computing, and clinical neuropsychology
• We are designing and clinically validating methods to perform automated functional assessment and intervention

Broader Impacts:
• Data collected in the smart homes is cleaned, anonymized, visualized, and disseminated
• Research was integrated into a multi-disciplinary Gerontechnology class

Contacts:
• Diane J. Cook  Washington State Univ, cook@eecs.wsu.edu
• Sajal K. Das Missouri S&T, sdas@mit.edu
A Formalism for Customizing and Training Intelligent Assistive Devices

PI: Brenna Argall, SCH

Technical Approach:
- Techniques to customize the arbitration functions that govern control sharing.
- Techniques to customize the confidence measures that govern the prediction of the user's goal.
- Subject studies to evaluate these techniques, and their generalization across users and tasks.

Motivation:
- Challenge: Control of high-dimensional assistive machines (e.g. robotic arm) inaccessible to those with severe motor impairments (e.g. using a sip-and-puff).
- Solution: Introduce robotics autonomy.
- Question: How to share control between the robotics autonomy and human?

Transformative:
- Addresses user-driven customization and training of assistive devices.
- Introduces a formalism for how to customize control sharing to individual users or tasks, and over time.

Broader Impacts:
- Public health: Increase the independence of people with severe motor impairments.
- Industry: Collaboration with Kinova Robotics.
- Public and K-12 outreach, in local schools and museums.

Contacts:
- Brenna Argall, Northwestern University
- Siddhartha Srinivasa, University of Washington

Progress:
- Control sharing customization study with Spinal Cord Injured (SCI) and uninjured participants.
- Time-optimal mode switching study with uninjured participants.
- Continued hardware and software development.

An SCI participant using customized control sharing.
Triple Aim

Institute for Health Improvement, 2007
Effective Research is a Relay between basic and applied science
Join our Listserv

SMARTHEALTH_COMMUNITY

Join the electronic mailing list (LISTSERV) for forthcoming announcements by — Sending an e-mail message to LISTSERV@LISTSERV.NSF.GOV from the mailing address at which you want to receive announcements.

The body of the message should read Subscribe SMARTHEALTH_COMMUNITY [your full name].

The message is case sensitive; so capitalize as indicated!

- Don't include the brackets.
- The Subject line should be blank
- For example, for Robin Smith to subscribe, the message would read Subscribe SMARTHEALTH_COMMUNITY Robin Smith.

You will receive a confirmation of your subscription along with instructions on using the listserv.
Useful Website: www.nsf.gov
Questions or Comments?

Wendy Nilsen
Program Director, Smart and Connected Health
Directorate for Computer & Information Science & Engineering
National Science Foundation
Tel: 703-292-2568
Email: wnilsen@nsf.gov