Small, n=me, data
from mobile health to immersive recommendations

Deborah Estrin
Cornell Tech: Professor of Computer Science, Founder of Jacobs Institute Health Tech Hub
Weill Cornell Medical College: Professor of Healthcare Policy and Research
Open mHealth: Co-Founder
destrin@cs.cornell.edu, @deborahestrin, http://destrin.smalldata.io
harness previously-unmeasured function and behavior
to fuel personalized and evidence-producing care and behavior
The Reality

Chronic diseases -- 7 of 10 deaths and 86% healthcare costs (US,CDC)
Precision medicine, advanced genomics alone will not address the most preventable mortalities whose primary determinants are behavioral.

Research objective
Develop small-data driven techniques for user modeling, phenotyping, preference profiling, social-support, and behavioral incentives to help us feel and be ‘better’

To fill the gap between what we are born with... and what we die suffer from...
Participant self-care
*How is this new medication working for me?*

Clinical care
*How is the patient responding to new care plan?*

Research evidence
*What works best in different contexts?*
small data: diverse data sources for the individual, n=1

Passively-recorded activity, location

“Real” Sensors, wearables

Novel, Visual, self-report

Digital traces: purchases, media
Smart(er) self report:
Photographic Affect Meter (PAM), Your Activities of Daily Living (YADL)

Pollak, et al

small data lab
Beyond mobile and beyond health

- **Purchases**
  - What I bought
  - When I bought it
  - Where I bought it

- **Mobility**
  - How much I moved
  - Where I was
  - Where I got to

- **Finances**
  - How much I spend
  - How much I save

- **Location**
  - Where I am
  - Where I was
  - When I was where

- **Email**
  - Who I write
  - Who writes me
  - How I write
  - When I write

- **Calendars**
  - When I'm busy
  - What I'm busy doing
Challenge: moving up the information food chain

- patient function (behavioral biomarkers)
- summarization, fusion
- raw measurements
Industry leaders: Ginger.io Check Engine Light

Continuous & Passive
Ginger.io fills in data gaps that are often missed when using self-report measures. This reduces patient burden and results in improved data quality.

Protecting Privacy
Ginger.io is HIPPA compliant, and collects statistics, not specific contacts, locations, or content.

Check Engine Light
Identify at-risk patients based on objective patient data, and alert a provider or caregiver.

Assess intervention
Track outcomes more explicitly, and determine efficacy of outreach approaches

Data Collection

Behavior Patterns

Health Status

Patient Smartphone App

Continuous & Passive
Ginger.io fills in data gaps that are often missed when using self-report measures. This reduces patient burden and results in improved data quality.

Protecting Privacy
Ginger.io is HIPPA compliant, and collects statistics, not specific contacts, locations, or content.

Check Engine Light
Identify at-risk patients based on objective patient data, and alert a provider or caregiver.

Assess intervention
Track outcomes more explicitly, and determine efficacy of outreach approaches

http://ginger.io
Actionable by whom?

Solving patient’s problems using data-driven social facilitation
MoodRhythm(TM)
Bipolar patients to measure and manage their own daily routines

Tanzeem Choudhury, HealthRhythms
Why isn't there more evidence???

It takes time to generate evidence; but not this much time!!

have the marketplace and current academic processes underperformed?

Lets learn from all (other) things digital

    rapid iteration (continuous improvement using analytics)
    modularity (generalizable learning)

There is hope

    Sage Bionetworks Bridge and Synapse projects: collaborative research studies and data commons
    Open mHealth: standardized, clinical representation and open source components
Emerging “evidence engines”: Mobile Research Study Frameworks

ResearchKit™

DIGITAL HEALTH

Bringing iPhone-style Medical Research to the Android World
By STEVE LOHR  NOVEMBER 12, 2015 11:00 AM  6 Comments

ResearchStack (April 2016)
What is ResearchKit™?

- iOS framework for creating mobile research study apps
- Manages the security, privacy, and informed consent considerations typical in IRB-backed research
- Opinionated UI makes building out components easy, keeping a consistent experience across ResearchKit apps
- Consent and survey components can be abstracted into JSON, making maintenance of these pieces easy for non-programmers
- Open source and backend-agnostic
Research Kit example: Parkinson mPower Study App

App launched on 3/9 with Apple Research Kit

Consented Participants (as of 4/24)
- 8485 to share with all qualified researchers
- 2360 to share with study team only

Scientific Advisors
- Karl Kieburtz, U Rochester
- Ray Dorsey, U Rochester
- Caroline Tanner, UCSF
- Bas Bloem, Radboud U
- Max Little, Aston U

Foundation Partners
- Michael J Fox Foundation
- ParkinsonNet
ResearchStack: Goals

- **Make it as little work as possible for programmers**
- Drag and drop RK assets directly into an RS app
- Provide an opinionated UI/style guidelines for those without much knowledge of Android design conventions
- Make it secure and private by default
- Provide a community for creating plugins and extensions, like custom tasks

*Initial development generously funded by a grant from the Robert Wood Johnson Foundation.*
Open source framework that helps developers integrate health app & device data using a common language

@openmhealth  http://openmhealth.org
CrowdSignals.io

http://crowdsignals.io

HOW IT WORKS

Crowdfunding → Recruiting + Onboarding → Data Collection + Management → Post-Processing + Sharing

Companies and researchers sponsor CrowdSignals.io

Admins recruit and onboard data collection subjects

Subjects run our app and are paid in proportion to the data they upload

Collected data is watermarked and shared with sponsors
measure, motivate, modify, maintain aspirational health behaviors

small-data fueled personal assistance/ts

nutrition
(e.g., household shopping)

PUSHCART
May 17, 2014
Whole Foods Market

Your Goal
I want to eat more fruits and vegetables (5-7 servings per day)

Your Cart
30 items $147.70

Food Group Breakdown (% of Cart)

pain management
(e.g., lower back exercises)

student health
(e.g., time management)
Food Purchase recommendation: Pushcart

1. Member Household Signs Up
2. Member Sets A Goal
3. Online Grocery Purchases Auto-Collected Via Email
4. Groceries Auto-Analyzed
5. Member Receives Swap Suggestions From Coach
6. Member Receives Optimized Shopping Lists

http://gopushcart.com/
Limbr: an intervention for chronic pain

- FT Exercise
- Daily walks
- Meds/Coping

YADL
MEDL
Message
Mobility
FT Apps

Limbr

report
analytics
variation

(w/Vijay Vad, Force Therapeutics, UHG)
Fueling personalized recommendation

model preferences w/ location, language, activity, photos; selective sharing

Small Data Streams Analysis App

Text and visual digital traces

Locations Landmarks Mobility Patterns

Collaborative Filtering

Topics and Interests Inference

Item Rankings by preference and context

the right suggestions for me
User-centric Preference-learning using small data

- Online Posts
- Private Communication
- Shared Images
- Personal Image Gallery

Preference Profile

- News
- Search Engine
- Dietary
- Entertainment

Yang, Hsieh, et al
PlateClick: Bootstrapping your food preferences

Visual pair-wise comparisons
image and metadata analysis, online learning

- Low cognitive load, personalized and legible.
- Preference Elicitation, no history required, no ratings
- Completed within a minute.

Personal diet-preference profile engine for food recommendation and health applications

bit.ly/plateclick

Longqi Yang, et al
News and Meetup Recommendation

Cold start problem

Immersive rec.

[Diagram showing a user profile connected to the internet, with arrows pointing to different recommendations like Gmail, Twitter, YouTube, and Slack, leading to another user profile connected to the internet.]
News recommendation using Medium, Twitter

Publicly available version running to support better user studies in future

Modular—available to try out improved algorithms, new application contexts

Newsfie
newsfie.org
An honest guide to the San Francisco startup life
Aug 21
Stationed on the West Coast of the United States, flanked by the blue hues of the Pacific Ocean, it's a city that's home to some of the...

The Rise of the Punk Rock B-boy
Aug 24
When hardcore and hip-hop collided in 1980s New York City, the cultural crossover was uncomfortable, dangerous and exhilarating.

Why Twitter’s Dying (And What You Can Learn From It)
Oct 15
It wasn’t Professor Plum in the Library with the Candlestick. So what killed Twitter?

Smooth Opinion
Sep 9
A long remembrance of a short-lived band.
Two phase algorithm to match individual to target content

- **Phase 1 User Profiling**
  Create user profile from users’ digital traces

- **Phase 2 Recommendation**
  Hybrid collaborative filtering algorithm combines user/item profiles and ratings
Future directions:
modeling user-state based on personal language patterns: sentiment, cognition
Future directions:
Creating selective-sharing analytics environment

Small Data Streams Analysis App + Text and visual digital traces
Selective Sharing + Analytics Environment
Location Landmarks Mobility Patterns
Collaborative Filtering
Topics and Interests Inference
Item Rankings by preference and context

the right suggestions for me
Precision Resource management

Future directions:
Creating selective-sharing analytics environment
Aspiration: create an ecosystem around small data

1. modular APIs shims for accessing data from popular apps and services
2. simple storage patterns for ephemeral raw data processing
3. reusable data processing and analytics modules on raw data primitives
4. plug-and-play visualizations and UIs on different outputs of linked small data analysis
5. specific applications tailored to key use cases and patient care models
6. consumer, developer and research communities built around small data building blocks and applications
2 Year MS in Health Tech @ Jacobs Institute

- Focus on tech innovation for individual and population health
- Technologies and applications for health status monitoring and prediction, chronic disease management, consumer health behavior
- Commercial engagement in course projects and student internships
- Developed in collaboration with leaders from relevant healthcare sectors: Business, Clinical, Research, Entrepreneurial, and with faculty.
Creating tech that solves “people’s/patients’ problems”

- **Internet of Things** that sources signals from heart beats to traffic to serve the Internet of People
- **Personalization** that promotes healthy eating and other preventive and protective activities
- **Social media engagement** and exposure used to understand and improve patient state of mind
- **High Touch care giving** made smart and scalable by high tech that informs human caregivers
- **Security** that balances individual privacy and sharing for science
For more information

http://smalldata.tech.cornell.edu
http://destrin.smalldata.io
http://tech.cornell.edu/programs/masters-programs/ms-in-is-health-tech

@deborahestrin, @openmhealth

Faculty: Deborah Estrin, JP Pollak,
Clinical Advisors in Residence: Daniel Stein, Thomas Tsang
Technologists-in-Residence: Aaron Baum, Michael Carroll, Diana Freed
PhD Students: Faisal Alquadoomi, Andy Hsieh, Fabian Okeke, Longqi Yang

Funding: NSF, NIH, AOL, Google, UHG, Pfizer, RWJF, Intel, Cornell Tech
Co-authors and Collaborators at small data lab, Cornell Tech

PhD students

Researchers and Technologists in Residence

Faculty Collaborators: Yoav Artzi, Serge Belongie, Tanzeem Choudhury, Ari Juels, Mor Naaman, Arnaud Sahuguet