HEALTHY CHICAGO

CHICAGO DEPARTMENT OF PUBLIC HEALTH
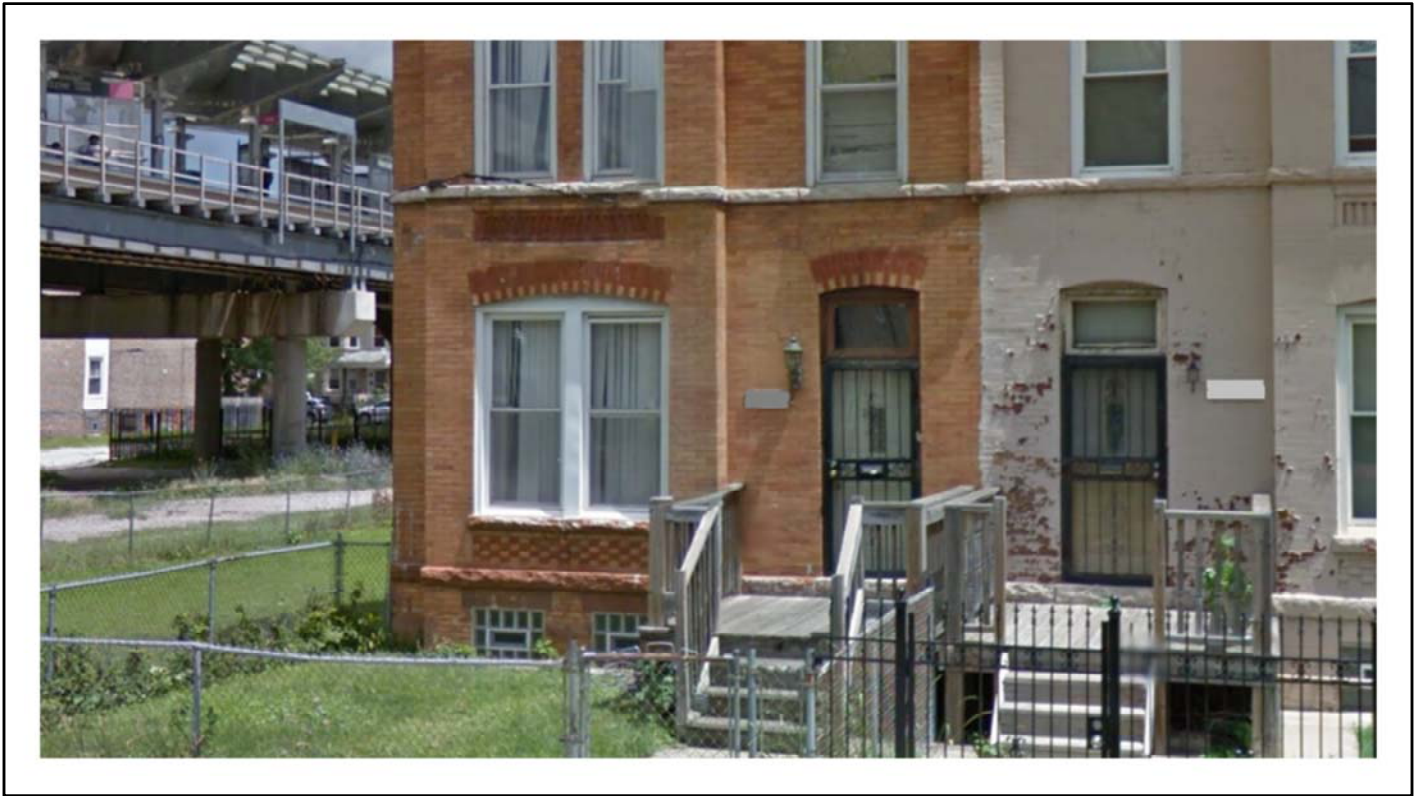
# Getting Ahead of Lead: Can Predictive Modeling Help Prevent Childhood Lead Exposure?

**(Janna)** In Chicago, almost 90% of the housing stock was built before the lead paint ban in 1978.

(Janna) Most of the risk for kids begins when they are able to crawl, walk, and become curious about their surroundings after 12 months old (touching, hands and toys in their mouth).

Children 0 to 6 years with BLL ≥10 mcg/dL in Chicago 1997 - 2015
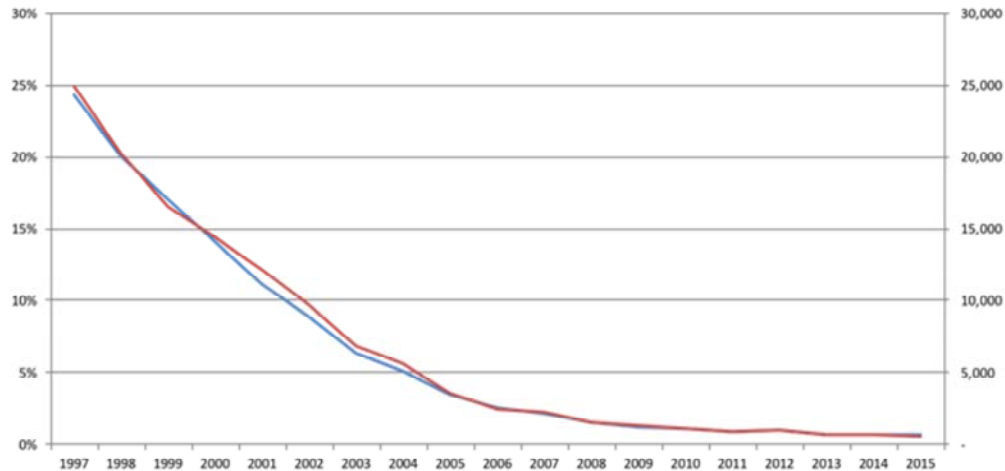
(Janna) When we were considering a lead predictive model back in 2014, the State was still at >=10mcg/dL. As is well known, while there are no safe levels of lead, we now use >=5mcg/dL. Over the years, Chicago's blood lead levels have markedly decreased from what used to be. In the 1990s, 1 in 4 Chicago children tested had a BLL of at least 10 μg/dL, that number is now less than 1 in 100 children.

Percent of Children under 3 with BLL>6 by Community Area in 2014
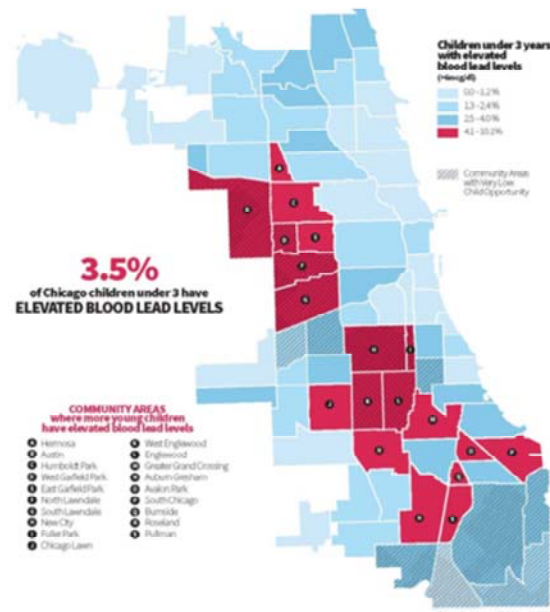
(Janna) But lead persists in the environment, especially in communities with older housing stock. And as with so many health problems, there are great disparities in Chicago. As you can see on this map, the neighborhoods in red have 3.5% of children living there with BLL above 6.

Couple this with very low Child Opportunity Index, we looked at how we can move to primary prevention of lead poisoning.

It's 3.5% in certain neighborhoods and is significantly better than the 25% we faced just 20 years ago citywide, but it is still not good enough.

And it is no surprise to those of us in public health that the neighborhoods most affected are those same neighborhoods that are in low economic empowerment areas, have low childhood opportunity indices, and whose residents face a disproportionate number of health issues.

So our plan is to focus our limited resources in these neighborhoods and move that number down even further.

Changing Our Model

Reactive — CHILD POISONED

Proactive — IDENTIFY HIGH RISK HOMES

LEAD MITIGATION

(Janna) Right now, a child is poisoned and we must inspect that home and mitigate/abate/lead safe, but can we predict where kids live that are at risk to be exposed to lead paint and mitigate the lead before they are poisoned and provide financial assistance to remove paint hazards and ensure children receive case management services to minimize impact of EBLLs on neurocognitive development?

DATA	YEARS	RECORDS	VARIABLES	OWNER
Blood Lead Level	1995 – 2017	2,700,000	First name, last name, date of birth, address, blood lead level, sample type, sample date	CDPH Lead Program
Home Inspection Records Summary	1989 -2017	66,000	Date of initial inspection, lead based paint hazard (yes/no), location of lead-based paint hazards (interior/exterior/both/), date complied, address	CDPH Lead Program
Building Permits	2006 - 2017	400,000	Address, issue date, permit type	Chicago Department of Buildings (Chicago Open Data Portal)
Building Violations	2006 – 2017	1,500,000	Address, violation Date, violation description, violation ordinance, inspection status	Chicago Department of Buildings (Chicago Open Data Portal)
Building Footprints	2015	800,000	Year of building construction, physical condition, number of units, stories (floors), vacancy status	Chicago Department of Buildings (Chicago Open Data Portal)
Cook County Assessor	2013	800,000	Address, assessed property values, building classifications, building characteristics, year built	2014 Cook County Assessor
Chicago Census Boundaries	2010	800	Shape File	Chicago Open Data Portal
Chicago Ward Boundaries	2015	50	Shape File	Chicago Open Data Portal
American Community Survey	2005 - 2014	800	Census tract variables including socio-demographics, education, health insurance, home ownership.	US Census Bureau
Frequently Occurring Surnames	2000	150,000	Census surname ethnicity	US Census Bureau

(Raed) In 2014, I met Rayid Ghani, and we pitched some ideas around whether building a lead predictive model can work. As you can see, we collect a lot of data, across various sources within antiquated and siloed IT infrastructures, with the governance to match. The proof of concept was built outside of the city, but we still had a lot to do before we can place it into practice. The implementation was one part, but we still needed to validate the model, move it into into the city, align it with the program, and all the while following a myriad of local, state, and federal regulations.

The Childhood Lead Paint Hazard Data Sharing Across Sectors of Health Project
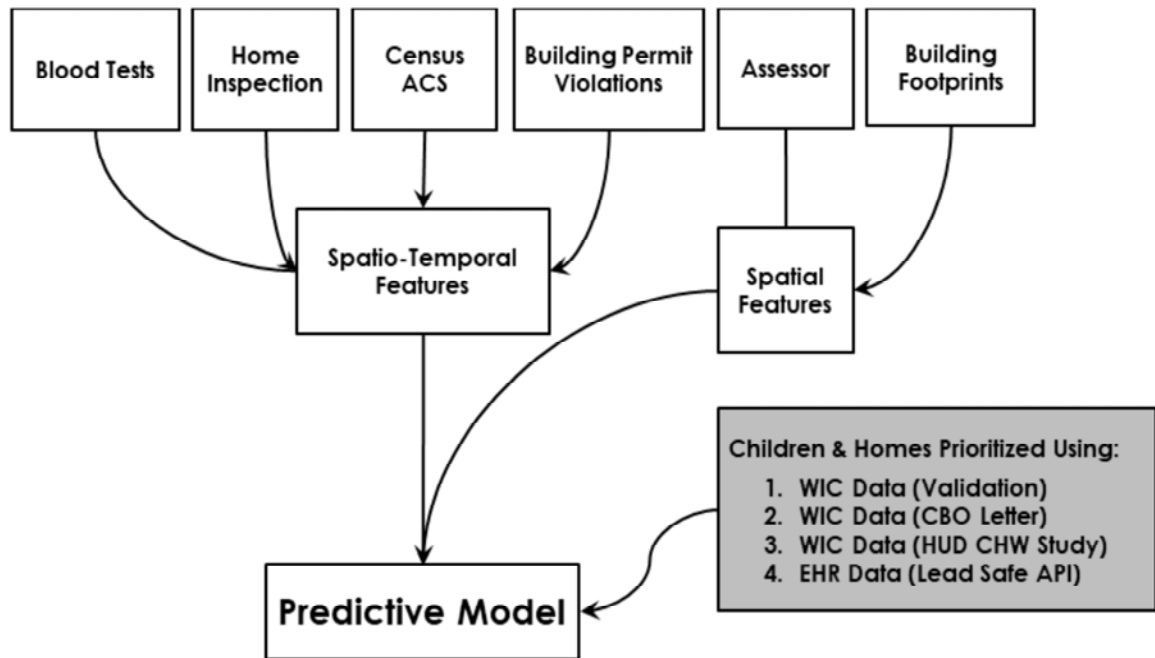
Support for this project is provided by the Data Across Sectors of Health grant from the Robert Wood Johnson Foundation

(Raed) In 2016, we won a grant from RWJF to validate and implement the lead model, but we added an upstream intervention: physicians and using an HER (electronic health record) to access our lead predictive model through AllianceChicago's 4 FQHCs.

A Predictive Model

Blood Tests | Home Inspection | Census ACS | Building Permit Violations | Assessor | Building Footprints

Spatio-Temporal Features

Spatial Features

Children & Homes Prioritized Using:
1. WIC Data (Validation)
2. WIC Data (CBO Letter)
3. WIC Data (HUD CHW Study)
4. EHR Data (Lead Safe API)

Predictive Model

(Rayid)
Lead Model and projects using model. See Organizational and Data Readiness Tools at
http://dsapp.uchicago.edu/projects/health/lead-prevention
Definitions:
EHR = electronic health records
API = application program interface

(Raed) One of the first uses of the model was using it to send a behavioral nudge with our partners at ideas42, and two local CBOs to perform the initial visual inspection before a CDPH lead inspector confirmed it. While the letter writing response increased 50%, the total number of responses was still 2%, which is typical of direct marketing, but we knew we were not getting a large impact.

Validating the Model

To compare our performance in predicting lead paint poisoning with a baseline, given by random predictions.

(Rayid)

Electronic Health Record Clinical Decision Support Tool

- Access the predictive model through an EHR Clinical Decision Support Tool (CDST)

- The CDST can alert providers to the risk of lead exposure based on the patient's current address

- CDST provides recommendation for visual home inspections and patient education on lead abatement strategies

(Rayid)

AllianceChicago healthcare providers educated on tool.

Follows the 5 Rights of Clinical Decision Support (by AHiMA):
1. The right information
2. To the right person
3. In the right intervention format
4. Through the right channel
5. At the right time in workflow

Operationalizing the Model

```
┌─────────────────────┐          ┌─────────────────────────┐
│                     │          │ Clinical Data from       │
│  Predictive Model   │ ◄─────── │ Electronic Health Record │
│                     │          │ using an API             │
└──────────┬──────────┘          └─────────────────────────┘
           │
           │
┌──────────┴──────────┐          ┌─────────────────────────┐
│   Risk Score in     │          │ Referral for visual home │
│  Electronic Health  │ ───────► │ inspections and patient  │
│      Record         │          │ education on lead        │
│                     │          │ abatement strategies     │
└─────────────────────┘          └─────────────────────────┘
```
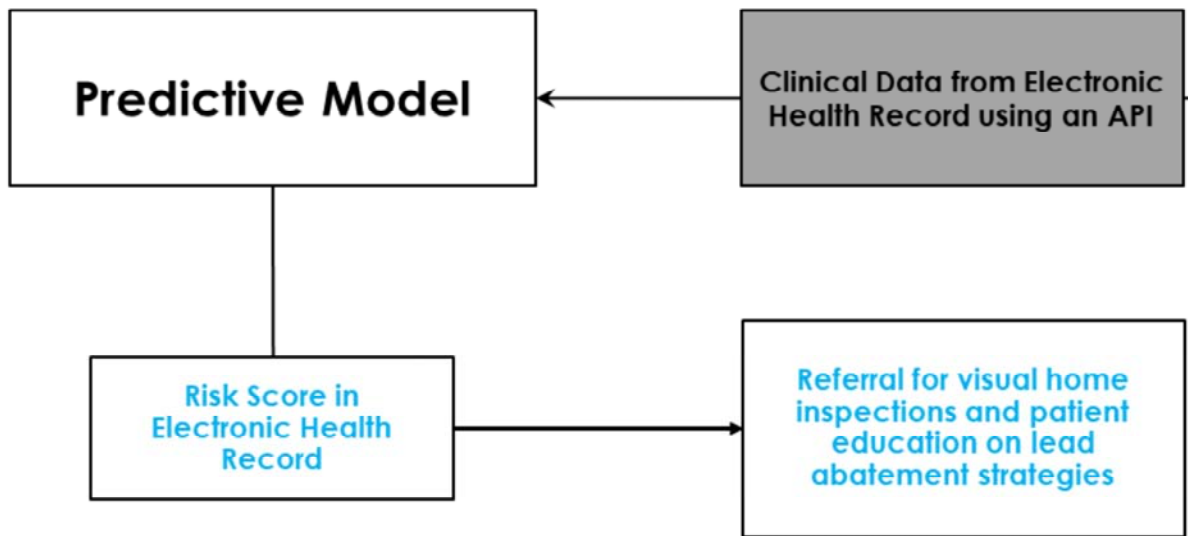
(Rayid)

Lead Safe Standard API Exchange

Submitted to City

- Address / ZIP Code
- Date of Birth
- Expected Due Date
- Gender
- Race
- Ethnicity
- Past visits to doctor's office
- Past lab results

⟷

Returned to Clinic

- Risk Score
- Instructions to doctor

(Rayid)

Challenges Encountered

1. CDPH limited resources for interventions

2. Difficult to access datasets to validate model due to antiquated & siloed IT infrastructures and governance

    - Birth certificates (State)
    - Immunization records (State)
    - WIC enrollment (USDA)

3. Establishing partner with healthcare systems to pilot

    Privacy, interpreting laws, policies and data sharing barriers prevented initial pilot in healthcare settings

4. Compare performance in predicting poisoning with a baseline, given by random predictions

Challenges Encountered

5. Finding appropriate community partners
6. Abiding by clinical decision support tool design standards
7. Time required for data exchange between model & HER (electronic health records)
8. Developing interoperability standards for data exchanges with other EHRs
9. Healthcare has difficulty justifying technology costs on prevention

(Raed)

Community Health Worker Study

Partnering with a health system's community health worker (CHW) to assess:

- Feasibility
- Impact
- Cost effectiveness

of using CHWs to conduct proactive visual inspections for lead-based paint hazards at homes on the south and west sides identified as high risk by our model using WIC enrollment.

Funded by US HUD Office of Lead Hazard Control and Healthy Homes Lead Technical Studies Program

(Janna)

Challenges Overcome

Project identified gaps in Antiquated & Siloed IT Infrastructure
- New mobile Lead Inspector Application
- Data Warehouse

Project identified policies not addressing new methods and data sharing
- Amended authority to include data science and other analytic models in ordinance
- IGA to share data bidirectionally with schools

States agencies can be a rate limiting step
- Work closely to find solutions to HHLPSS automation

EHR Interoperability
- Standardize API

Applicability of model is diverse and still discovering potential uses

(Janna)

Technical Amendments

Municipal Code of Chicago

Title 2 City Government and Administration
Chapter 112 Board of Health
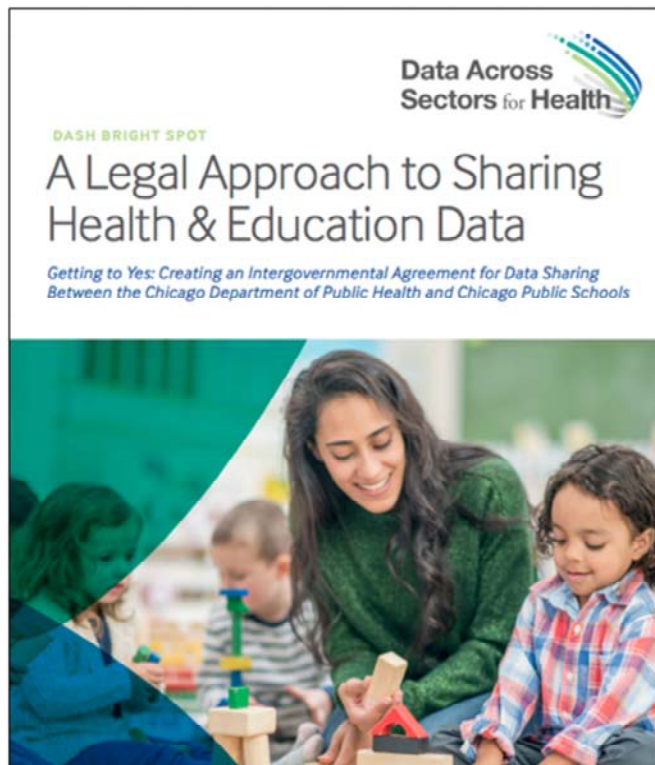160 Commissioner - Additional powers and duties.

The Commissioner of Health shall have the following powers and duties:

(a) Public health related powers and duties:

(7) To request, collect, receive, and maintain confidential information, records, and data, including protected health information consistent with 45 C.F.R. § 164.512(b)(1)(i), for the purpose of preventing or controlling disease, injury, or disability. The confidential information, records, and data may support activities including, but not limited to, the reporting of disease, injury, or vital events such as birth or death; the conducting of public health surveillance, public health investigations, and public health interventions; **the performance of epidemiological studies; and the application of data science methods or other analytic models** that protect and promote public health.

(Raed) In response to not being able to use certain variables within statistical models since it was not outlined within past agreements and no authority to do so, our attorney helped us make an amendment to the Commissioner's authority to add "the performance of epidemiological studies; and the application of data science methods or other analytic model."

Innovative Agreement

DASH BRIGHT SPOT

A Legal Approach to Sharing Health & Education Data

Getting to Yes: Creating an Intergovernmental Agreement for Data Sharing Between the Chicago Department of Public Health and Chicago Public Schools

Data Across Sectors for Health

(Raed) This project helped add lead paint risk values and blood lead level sharing (amongst many other variables) in a unique data-sharing agreement between schools and health; threading the needle between Family Educational Rights and Privacy Act (FERPA) and Health Insurance Portability and Accountability Act (HIPAA) for the benefit of a child's health and education, as we all know, are connected.

https://dashconnect.org/wp-content/uploads/2018/05/DASH-Bright-Spot_Chicago.pdf

Predictive Modeling for Public Health: Preventing Childhood Lead Poisoning

Eric Potash
University of Chicago
epotash@uchicago.edu

Joe Brew
University of Florida
joebrew@ufl.edu

Alexander Loewi
Carnegie Mellon University
aloewi@cmu.edu

Subhabrata Majumdar
University of Minnesota
majum010@umn.edu

Andrew Reece
Harvard University
reece@g.harvard.edu

Joe Walsh
University of Chicago
jwalsh@uchicago.edu

Eric Rozier
University of Cincinnati
eric.rozier@uc.edu

Emile Jorgensen
Chicago Dept of Public Health
Emile.Jorgensen@cityofchicago.org

Raed Mansour
Chicago Dept of Public Health
Raed.Mansour@cityofchicago.org

Rayid Ghani
University of Chicago
rayid@uchicago.edu

ABSTRACT

Lead poisoning is a major public health problem that affects hundreds of thousands of children in the United States every year. A common approach to identifying lead hazards is to test all children for elevated blood lead levels and then investigate and remediate the homes of children with elevated tests. This can prevent exposure to lead of future residents, but only after a child has been poisoned. This paper describes joint work with the Chicago Department of Public Health (CDPH) in which we build a model that predicts the risk of a child to being poisoned so that an intervention can take place before that happens. Using two decades of blood lead level tests, home lead inspections, property value assessments, and census data, our model allows inspectors to prioritize houses on an intractably long list of potential hazards and identify children who are at the highest risk. This work has been described by CDPH as pioneering in the use of machine learning and predictive analytics in public health and has the potential to have a significant impact on both health and economic outcomes for communities across the US.

Categories and Subject Descriptors

J.3 [Life and Medical Sciences]: Health; K.4.1 [Public Policy Issues]: Human Safety

General Terms

Machine Learning, Social Good, Lead Poisoning, Public Health, Public Policy

1. INTRODUCTION

Lead poisoning is a major public health issue, imposing lifelong health and economic costs on hundreds of thousands of children every year in the United States. Although European states banned lead paint as early as 1909 [19], political forces and vested business interests delayed bans on leaded consumer products in the United States until the late 1970s [21]. Throughout most of the 20th century, cars ran on leaded gas, houses were coated with leaded paint, and industry emitted leaded waste products directly into the environment. To this day, lead in paint remains a significant hazard. In Chicago, almost 90% of the housing stock was built before the ban [13].

Exposure to lead has been found to be associated with premature birth and early neurological development issues such as edema, herniation, atrophy, and white-matter degeneration [12, 10]. Lead can cause vomiting; convulsions; paralysis; and, in high concentrations, death [14]. Elevated blood lead levels are associated with lower IQs in children. A retrospective study by Mazumdar et al [20] shows that, on average, a 1 µg/dL increase in blood-lead level is associated with a decrease of 1 IQ point among six-month-olds and 2 IQ points among 10 year olds.
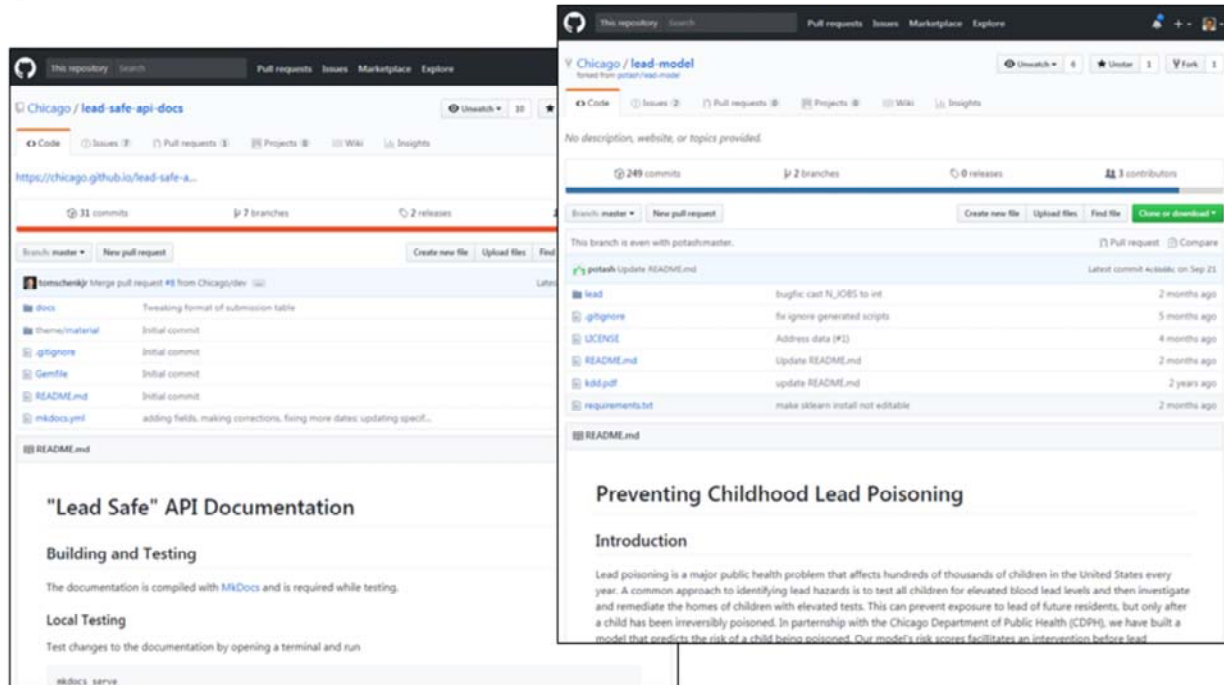
Because of the permanent damage it can inflict, lead poisoning imposes significant indirect costs on society. Based on its well-documented effects on IQ and contributions to neuropsychiatric disorders such as ADHD, lead poisoning has been estimated to significantly lower lifetime earnings for individuals and greatly increase the costs of crime prevention and special-education programs for the government. Lead-related child health issues conservatively cost over $40 billion annually [18]. Completely eliminating lead in the United States could indirectly save $200 billion dollars per year [22], ten times more than needed for removal.

(Raed) available online at http://www.dssgfellowship.org/wp-content/uploads/2016/04/cdph-lead-kdd2015.pdf, and another paper will be out next year.

(Raed) In 2018 and 2019, the team and model were recognized for Innovation and Digital Transformation.

Open Source Code

(Rayid)

https://github.com/Chicago/lead-safe-api-docs

Thank You!