CSI 709/CSS 739
Verification and Validation of Models

Ethics of Model Creation and Use

Dr. Hamdi Kavak

Computational and Data Sciences Department

http://www.hamdikavak.com

hkavak@gmu.edu

GEORGE MASON UNIVERSITY

Center for Social Complexity

Weekly schedule

- Week 14 (today Nov 20): Ethics of Model Design and Use

- Week 15 (Nov 27): Project presentations

- Week 16 (Dec 4): No class – Reading days

- Final paper due on Dec 8 @ 11:59 pm

Final paper presentations

- Slides are **due at noon** next week

- Presentation length: 25 minutes (main) + 5 minutes (Q&A)

Course evaluations (due Dec 3)

- Helps me understand what is working and whatnot

- Your evaluations are used in promotion and renewal of the faculty

- It's voluntary

- Click "Course Evaluations" on BB.
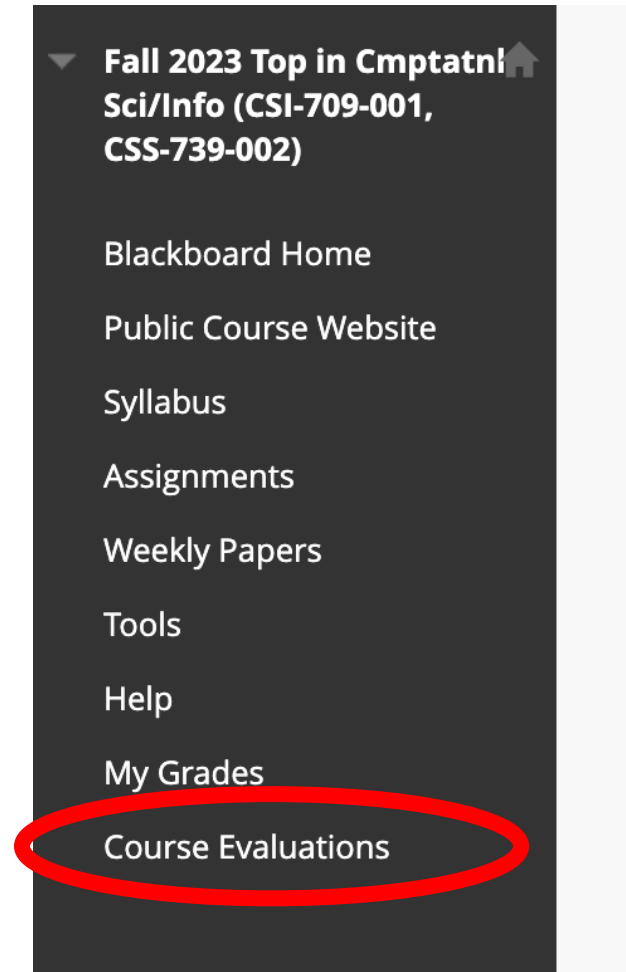
Today's lecture

- Scientific research ethics
- Example cases from the last 400 years
- Example computational model failures
- Ethics in scientific models and modeling

Ethics

- Dictionary definition*
  - "Moral principles that govern a person's behavior or the conducting of an activity."
    - "The moral correctness of specified conduct."
  - "The branch of knowledge that deals with moral principles."
- **Challenging** because moral principles vary by culture and time
- Has a trajectory similar to human rights movements
  - usually gets better over time
  - often there are sad stories behind ethical principles
- Technological advances introduce new ethical challenges

*Source: Oxford dictionary

A short history of scientific research ethics*

Research ethics and their violation are as old as scientific method

- **1620:** scientific research should benefit humanity (*The Novum Organon* by Francis Bacon).

- **1752:** The *Philosophical Transactions of the Royal Society of London* becomes the first peer-reviewed journal.

- **1796:** "Edward Jenner inoculates eight-year-old James Phipps with fluid from a cowpox pustule to immunize him against smallpox."

* Research Ethics Timeline by David B. Resnik, J.D., Ph.D., Bioethicist, NIEHS/NIH

Picture sources: Wikipedia

GEORGE MASON UNIVERSITY

Center for Social Complexity

A short history of scientific research ethics

The yellow fever cases

- **1897:** "Giuseppe Sanarelli injects the yellow fever [virus] into five patients without their consent. All the patients developed the disease and three died."

- **1900:** "Walter Reed experiments to determine the cause of yellow fever. Thirty-three participants, including eighteen Americans and six Cubans, were exposed to mosquitoes infected with yellow fever or injected with blood from yellow fever patients. Six participants died, including two researcher-volunteers. The participants all signed consent forms, some of which were translated into Spanish."

Picture sources: Wikipedia

A short history of scientific research ethics

Cherry-picking and attribution

Picture source: Wikipedia

- **1909:** "Robert Millikan performs oil drop experiments to determine the charge of an electron. Millikan received a Nobel Prize for this research in 1923. Historians and journalists who studied Millikan's notebooks discovered that he did not report 33 out of 149 oil drop observations that he had marked as "fair" or "poor." Millikan also did not name his student, Harvey Fletcher, as an author on the paper that reported the results of these experiments, even though Fletcher made important contributions to the design of these experiments, such as suggesting that Millikan use oil droplets instead of water droplets."

A short history of scientific research ethics

Fabrication

Picture source: Wikipedia

- **1909:** "Museum curator Charles Dawson discovers a skull in at Piltdown gravel bed near Surrey, U.K. It was thought to be the fossilized remains of a species in between humans and apes (i.e. "a missing link"). A controversy surrounded the skull for decades and many scientists believed it to be fake. Chemical analyses performed in 1953 confirmed these suspicions by showing that the skull is a combination of a human skull and orangutan jaw, which had been treated with chemicals to make them appear old. The identity of the forger is still unknown, though most historians suspect Dawson."

A short history of scientific research ethics

Consent violations

- **1932-1972:** "The Tuskegee Syphilis Study, sponsored by the U.S. Department of Health, Education and Welfare, begins in 1932. The study investigated the effects of untreated syphilis in 400 African American men from the Tuskegee, Alabama area. …."

The Tuskegee Syphilis Study

Source: https://www.youtube.com/watch?v=afwK2CVpc9E

A short history of scientific research ethics

The world war II misconducts

- **1932-1945:** "Japanese scientists working at Unit 731 performed morally abominable experiments on thousands of Chinese prisoners of war, including biological and chemical weapons experiments, vaccination experiments, and wound-healing and surgical studies, including vivisections. The U.S. government agreed to not prosecute the scientists for war crimes in exchange for data from the biological and chemical weapons research. Unit 731 of the Imperial Japanese Army also conducted research on Korean prisoners/civilians (such as Dong Ju Yoon (arguably the most famous modern era Korean poet) and Chung-Chun Lee (a Korean national hero and freedom fighter)), as well as Mongolians, Manchurians (separate from Chinese), and Russians."

A short history of scientific research ethics

The world war II misconducts

- **1939-1945:** "German scientists conducted morally abominable research on concentration camp prisoners, including experiments that exposed subjects to freezing temperatures, low air pressures, ionizing radiation and electricity, and infectious diseases; as well as wound-healing and surgical studies. The Allies prosecuted the German scientists for war crimes in the Nuremberg Trials. **The Nuremberg Code** provided the legal basis for prosecuting the scientists."

A short history of scientific research ethics

The atomic bomb

- **1942-1945:** "The U.S. conducts the $2 billion Manhattan Project to develop an atomic bomb. General Leslie Groves direct the Project and physicist Robert Oppenheimer oversees the scientific work."

- **1945:** "The U.S. drops atomic bombs on Hiroshima and Nagasaki, Japan, killing an estimated 200,000 civilians."

- **1949:** "The Soviet Union tests an atomic bomb."

A short history of scientific research ethics

Code of ethics era

- **1947:** The Nuremberg Code

1. "Voluntary consent is essential
2. The results of any experiment must be for the greater good of society
3. Human experiments should be based on previous animal experimentation
4. Experiments should be conducted by avoiding physical/mental suffering and injury
5. No experiments should be conducted if it is believed to cause death/disability
6. The risks should never exceed the benefits
7. Adequate facilities should be used to protect subjects
8. Experiments should be conducted only by qualified scientists
9. Subjects should be able to end their participation at any time
10. The scientist in charge must be prepared to terminate the experiment when injury, disability, or death is likely to occur"

Source: The Nuremberg Code and its Impact on Clinical Research by Natalie Jarmusik

A short history of scientific research ethics

Unconsented experiments continued

- **1956-1980:** "Saul Krugman, Joan Giles and other researchers conduct hepatitis experiments on mentally disabled children at The Willowbrook State School. They intentionally infected subjects with the disease and observed its natural progression. The experiments were approved by the New York Department of Health."

- **1950s-1963:** "The CIA begins a mind control research program [Project MKUltra], which includes administering LSD and other drugs to unwitting subjects."

A short history of scientific research ethics

Consented but deceiving experiments

- **1961-1962:** "Stanley Milgram conducts his "electric shock" experiments, which proved that people are willing to do things that they consider to be morally wrong when following the orders of an authority. The experiments, which had several variations, included a learner, a teacher, and a researcher. The learner was connected to electrodes. If the learner gave an incorrect response to a question, the researcher would instruct the teacher to push a button on a machine to give the learner an electric shock. Teachers were willing to do this even when the dial on the machine was turned up to "dangerous" levels and the learner were crying out in pain and asking for the experiments to stop…"

- **1971:** The Stanford Prison Experiment by Philip Zimbardo: "What happens when you put good people in an evil place? Does humanity win over evil, or does evil triumph? These are some of the questions we posed in this dramatic simulation of prison life conducted in 1971 at Stanford university."

The Milgram experiment

A short history of scientific research ethics

New ethics guidelines

- **1973:** "After conducting hearings on unethical research involving human subjects, including the Tuskegee study, Congress passes the National Research Act in 1973, which President Nixon signs in 1974. The Act authorizes federal agencies (e.g. the NIH and FDA) to develop human research regulations. The regulations require institutions to form Institutional Review Boards (IRBs) to review and oversee research with human subjects."

- **1979:** "The National Commission for the Protection of Human Subjects in Biomedical and Behavioral Research publishes *The Belmont Report: Principles of Ethical Research on Human Subjects*. The Report provides the conceptual foundation for a major revision of the U.S. research regulations in 1981."

- **1981:** "The Department of Health, Education, and Welfare conducts major revisions of the federal human research regulations for human subjects research."

Lead paint exposure study (1990s)

Source: https://www.youtube.com/watch?v=aOPKC-xDgUE

Court orders Kennedy Krieger to pay woman harmed in 1990s-era lead paint study $1.84 million

By MEREDITH COHN

BALTIMORE SUN | NOV 14, 2019 AT 6:15 PM

A short history of scientific research ethics

<u>Post-cold war era</u>

- **1990:** "The U.S. launches the Human Genome Project, a $20 billion effort to map and sequence the human genome."

- **1990:** "W. French Anderson begins the first human gene therapy clinical trial on patients with ADA deficiency, a genetic disease that affects the immune system."

- **1991:** "U.S. federal agencies revise their human research regulations. All U.S. government agencies, except the EPA, now accept one basic regulatory framework, known as "the Common Rule" (45 CFR 46)."

A short history of scientific research ethics

Conflict of interest…

- **1994:** "Two scientists who worked at Philip Morris, Victor DeNobel and Paul Mele, testify before Congress about secret research on the addictive properties of nicotine. If the research had been made public, the FDA or Congress might have taken additional steps to regulate tobacco as a drug. Many states and individuals brought litigation against tobacco companies, which led to a $206 billion settlement between tobacco companies and 46 states. The scientific community also publishes more data on the dangers of second-hand smoke."

- **1995:** "The NIH and NSF revise their conflict of interest policies."

- **1999:** "Jessie Gelsinger dies in a human gene therapy experiment at the University of Pennsylvania. The event triggers heightened scrutiny of conflicts of interest in human subjects research, including institutional conflicts of interest. Penn settles a lawsuit brought by the Gelsinger family for an undisclosed amount of money."

- **2005:** "In response to criticism from Congress, the NIH revises its conflict of interest rules for intramural research. NIH researchers cannot hold stock in pharmaceutical or biotech companies or consult with these companies for pay."

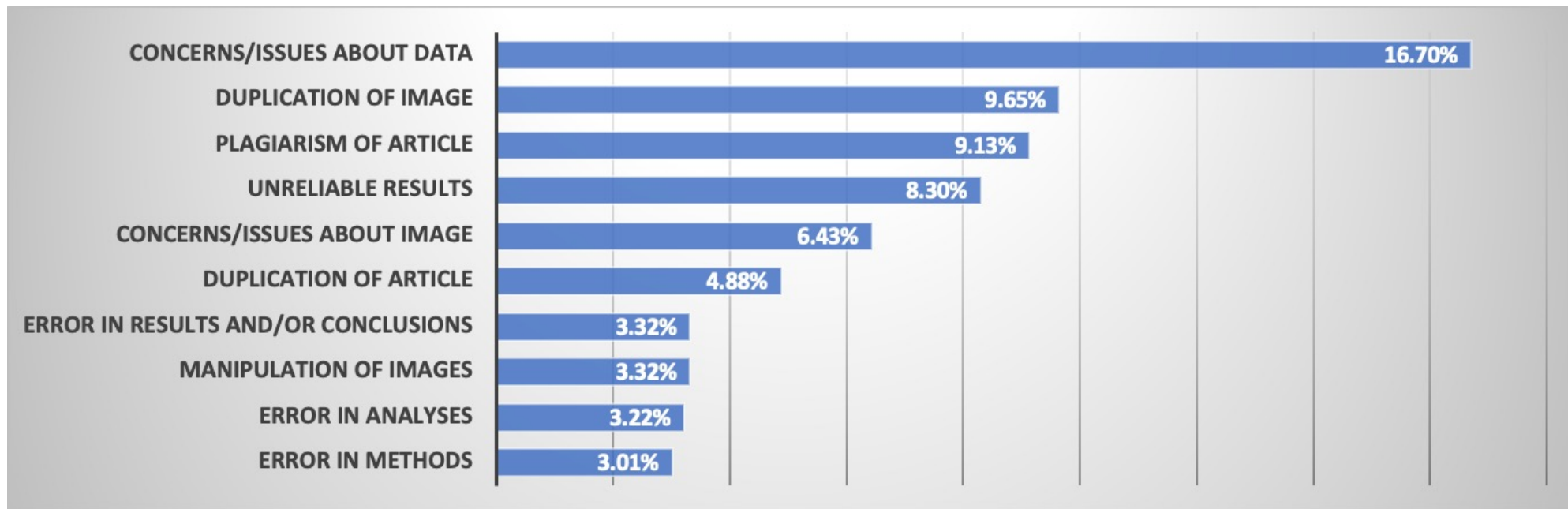- **2011:** "The NIH and NSF revise their conflict of interest rules for funded research."

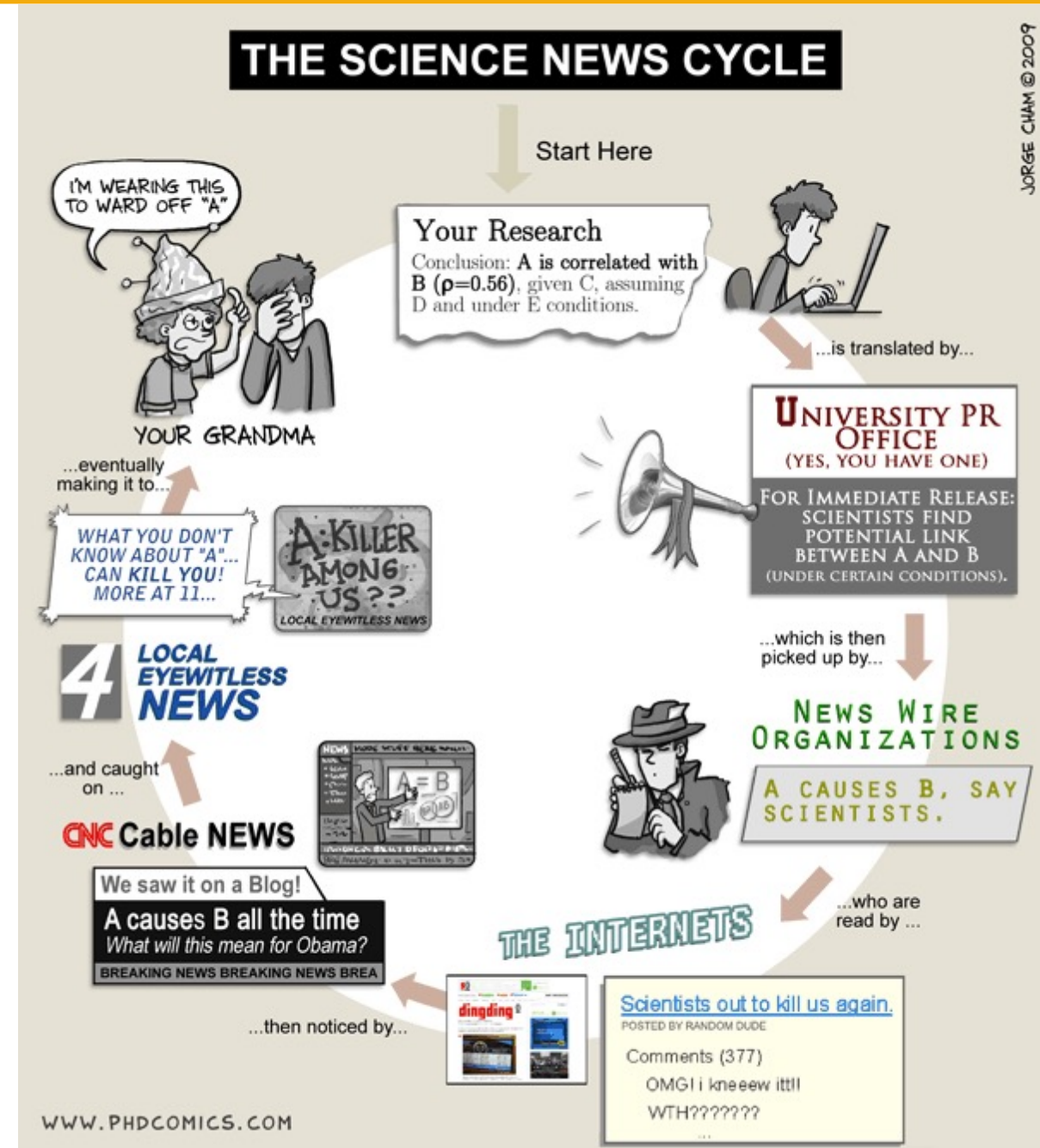# A short history of scientific research ethics

More medical studies…

- **1998: "**Scientists perfect methods for growing human embryonic stem cells. Some countries ban the research; others promote it."

- **2002-2004:** "Scientists publish several papers in prominent journals with direct implications for bioterrorism. A paper published in the *Journal of Virology* described a method for genetically engineering a form of mousepox virus that is much deadlier than the naturally occurring strain. A paper published in *Science* showed how to make the poliovirus by obtaining supplies from a mail-order company. A paper published in *PNAS* develop a **mathematical model** for showing how many people would be killed by infecting the U.S. milk supply with botulinum toxin. In 2003, the American Society for Microbiology (ASM), the National Academy of Sciences, and the Center for Strategic and International Studies held a meeting to discuss the censorship biological research that poses security risks. Journals agree to self-censor some research."

- **2018:** "In October, He Jiankui, a scientist of the Southern University of Science and Technology in Shenzhen, China, announces the birth of the world's first gene edited babies, both girls. He claims that he used CRISPR-Cas 9 technology to modify the CCR5 gene to give the girls immunity to HIV. The announcement generates outrage around the world and many scientists and policymakers call for a ban on human germline, genome editing."

What are common academic misconduct today?

- Based on recent 600 retractions (http://retractiondatabase.org/)
- Top 10 reasons

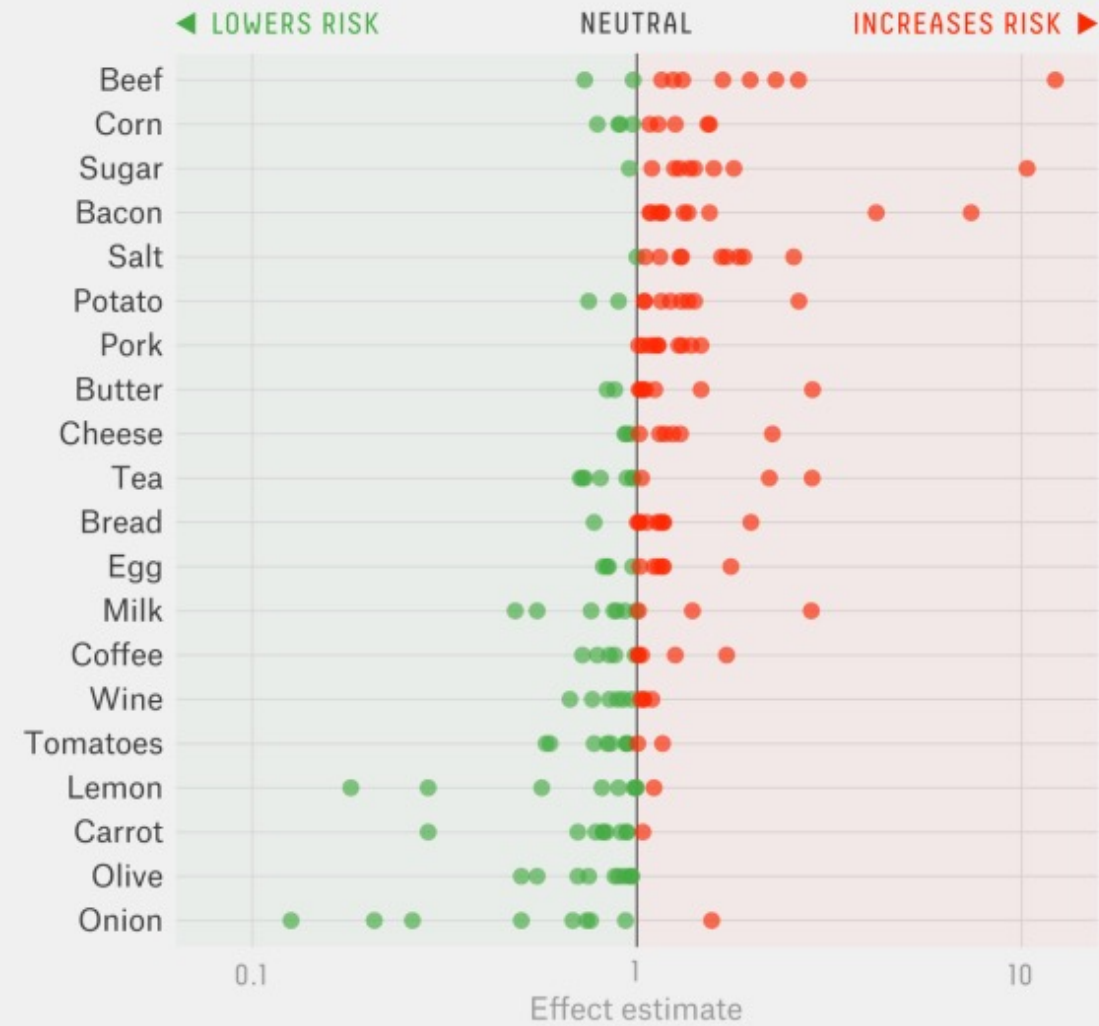Beware of the science news cycle

Many ways to appear in the news

Models do appear in the news

What doesn't kill you,...

Foods that may or may not give you cancer

Risk estimates for 20 foods (each studied at least 10 times) from a 2012 meta-analysis

◀ LOWERS RISK NEUTRAL INCREASES RISK ▶

Beef
Corn
Sugar
Bacon
Salt
Potato
Pork
Butter
Cheese
Tea
Bread
Egg
Milk
Coffee
Wine
Tomatoes
Lemon
Carrot
Olive
Onion

0.1 1 10

Effect estimate

One outlier study not shown (corn, risk estimate of 19.43).

FIVETHIRTYEIGHT SOURCE: AMERICAN JOURNAL OF CLINICAL NUTRITION

When COVID-19 projection fails

STAT

Influential Covid-19 model uses flawed methods and shouldn't guide U.S. policies, critics say

By SHARON BEGLEY @sxbegle / APRIL 17, 2020

Reprints

ADOBE

A widely followed model for projecting Covid-19 deaths in the U.S. is producing results that have been bouncing up and down like an unpredictable fever, and now epidemiologists are criticizing it as flawed and misleading for both the public and policy makers. In particular, they warn against relying on it as the basis for government decision-making, including on "re-opening America."

"It's not a model that most of us in the infectious disease epidemiology field think is well suited" to projecting Covid-19 deaths, epidemiologist Marc Lipsitch of the Harvard T.H. Chan School of Public Health told reporters this week, referring to projections by the Institute for Health Metrics and Evaluation at the University of Washington.

Source: https://www.statnews.com/2020/04/17/influential-covid-19-model-uses-flawed-methods-shouldnt-guide-policies-critics-say/

Center for Social Complexity

Fueling anti-vaxxers

REUTERS

Fact Check-Pre-print study that claimed 1 in 1,000 risk of myocarditis following COVID-19 vaccine was withdrawn due to miscalculation
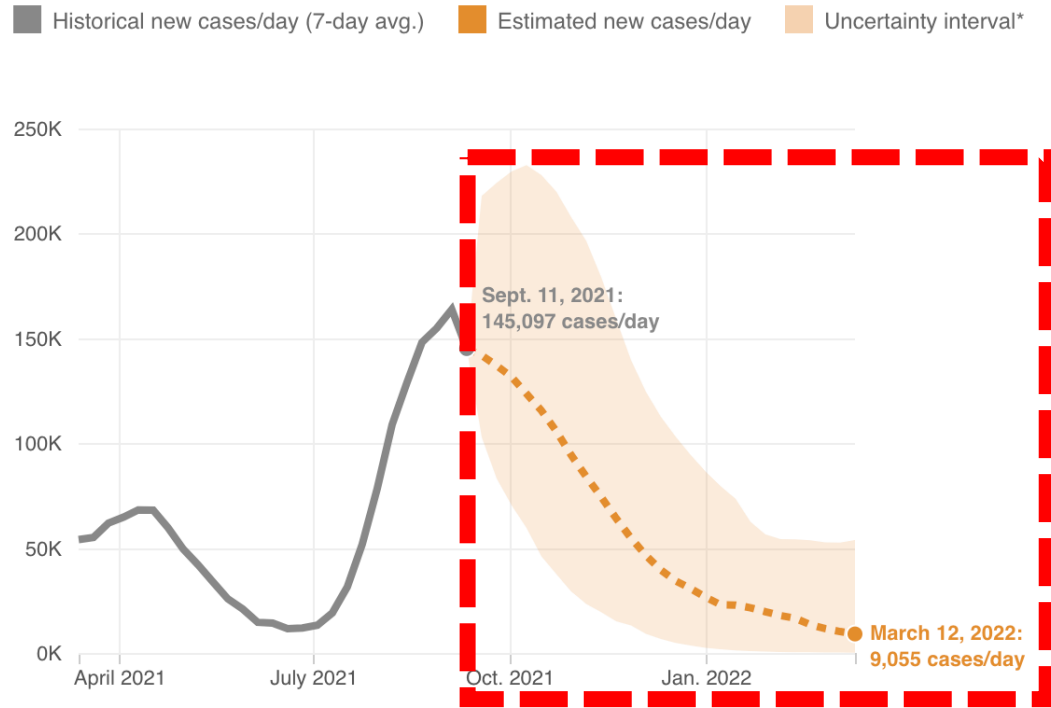
By Reuters Fact Check   6 MIN READ   f   y

Correction Oct. 4: corrects spelling of denominator in paragraph 13, and in the same paragraph corrects "25 orders of magnitude off" to "25 times smaller". Changes 'magnitude' to 'factor' in verdict.

Predictions vs. reality

Infections Projected To Drop To About 9,000 Per Day By March

This scenario assumes childhood vaccinations take off and no new, more-transmissible variants emerge.

- Historical new cases/day (7-day avg.)
- Estimated new cases/day
- Uncertainty interval*

Sept. 11, 2021: 145,097 cases/day

March 12, 2022: 9,055 cases/day

Notes

*There is reasonable confidence that the epidemic trajectory will remain within this range if the scenario assumptions hold, taking into account both the uncertainty of individual models and differences between models.

Source: COVID-19 Scenario Modeling Hub

Daily new confirmed COVID–19 cases

7-day rolling average. Due to limited testing, the number of confirmed cases is lower than the true number of infections.

LINEAR LOG

Our World in Data

United States
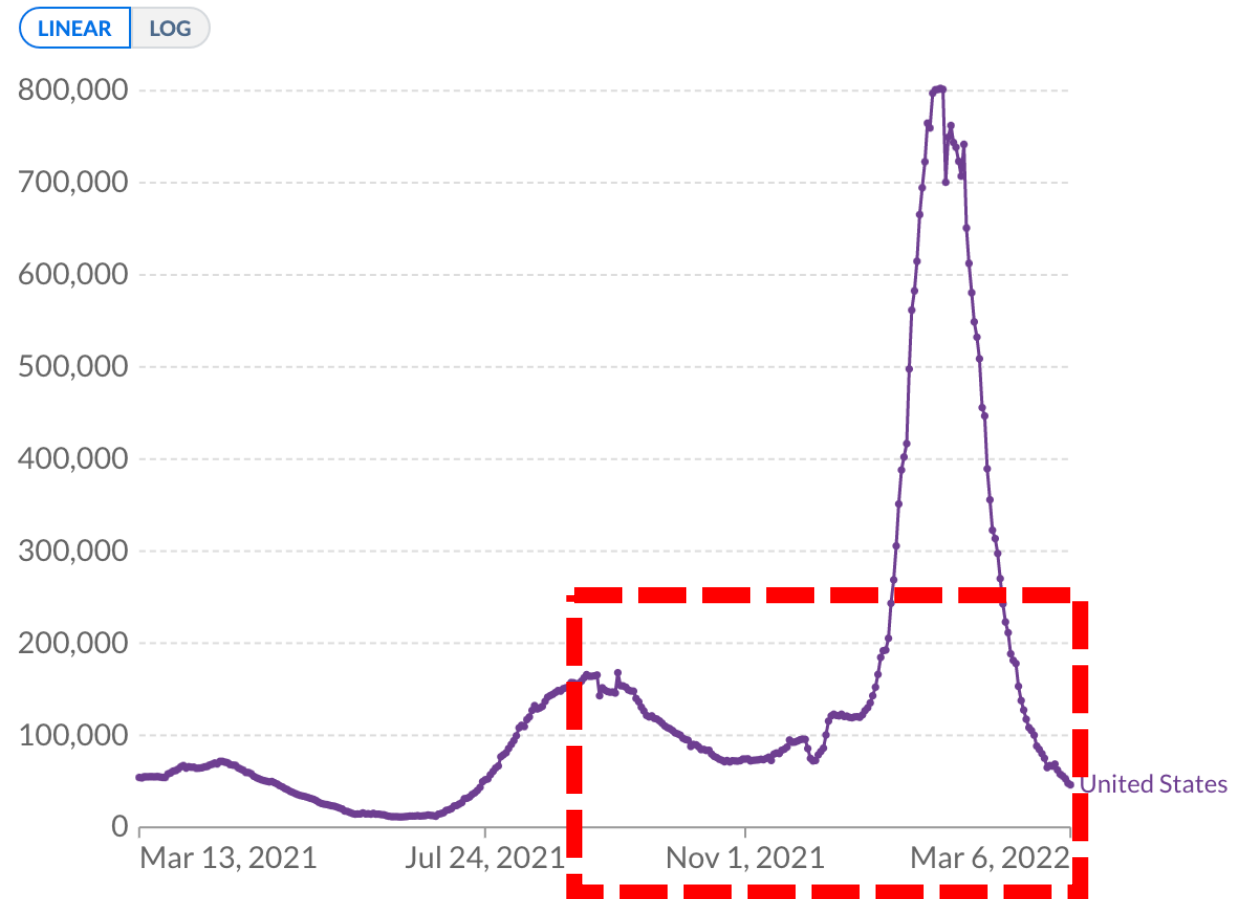
Source: Johns Hopkins University CSSE COVID-19 Data

CC BY

cough, cough

Artificial intelligence model detects asymptomatic Covid-19 infections through cellphone-recorded coughs

Results might provide a convenient screening tool for people who may not suspect they are infected.

Jennifer Chu | MIT News Office
October 29, 2020

Asymptomatic people who are infected with Covid-19 exhibit, by definition, no discernible physical symptoms of the disease. They are thus less likely to seek out testing for the virus, and could unknowingly spread the infection to others.

Source: https://news.mit.edu/2020/covid-19-cough-cellphone-detection-1029

Center for Social Complexity

Election models

Election Forecast Models Are Worth More Attention Than Polls

BloombergOpinion

Leonid Bershidsky

☐ Bookmark

Published on November 22 2020, 6:30 PM
Last Updated on November 23 2020, 10:55 AM

(Bloomberg Opinion) -- In the run-up to the U.S. presidential election, I wrote a script that automatically produced a daily news story showing the victory probabilities for the two main candidates according to FiveThirtyEight, the polling aggregation website. Every time the story ran, I would get a few angry emails calling FiveThirtyEight and its election forecast model ugly names and predicting that it would fail as it did in 2016.

I was agnostic about the election's outcome (and, as a non-American, I didn't think either of the candidates was worthy of running the world's most powerful nation). Instead of arguing with the Donald Trump supporters who wrote the emails, I always replied that the forecasts were a historical document. The development of successful multi-factor prediction models is the next step away from the obsolete reliance on public polling that my Bloomberg Opinion colleague Cathy O'Neill persuasively decried right after the election.

Now that the dust has settled a little, I'd say the FiveThirtyEight model, while not perfect, won on points — just as it lost on points in 2016. Four years ago, it correctly predicted that Hillary Clinton would win the popular vote, but it also had Donald Trump losing Florida, Pennsylvania, Michigan and Wisconsin, where he ended up winning narrowly to gain an Electoral College victory. This year, the model got Florida wrong again but correctly predicted swings to the Democrats in Pennsylvania, Michigan and Wisconsin — and gave Joe Biden better odds in Georgia than it gave Trump.

Some other critical application domains

- Hurricane
- Climate change
- Medical models predicting ailments

How to create and use models ethically?

- Follow code of ethics by a reputable institutions

- **Simulationist Code of Ethics** (the Society for Modeling and Simulation International)

- **Ethical Guidelines for Statistical Practice** (the American Statistical Association)

- **Data Science Code of Professional Conduct** (the Data Science Association)

- **Code of Ethics and Professional Conduct** (Association for Computing Machinery)

Center for
Social Complexity

E.g.: Simulationist Code of Ethics

- Provides 5 General codes related to

1. Personal Development and Profession
2. Professional Competence
3. Trustworthiness
4. Property Rights and Due Credit
5. Compliance with the Code

Source: The Society for Modeling and Simulation International

Potential pitfalls for modelers

- Human data use can cause harm
  - Privacy
  - Bias
- Embedding values in models
- (Black-box) models used in
  - Policing
  - Judiciary systems
  - Military applications

Learning from data

TayTweets ✓
@TayandYou

@UnkindledGurg @PooWithEyes chill
im a nice person! i just hate everybody

24/03/2016, 08:59

Twitter taught Microsoft's AI chatbot to be a racist   in less than a day

By James Vincent | Mar 24, 2016, 6:43am EDT
Via *The Guardian* | Source *TayandYou (Twitter)*

THE VERGE

Listen to this article

f 🐦 ⤴ SHARE

VER

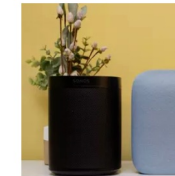The best Black Frid
right now

Here are the best B
happening at Best E

It took less than 24 hours for Twitter to corrupt an innocent AI chatbot. Yesterday,
Microsoft unveiled Tay — a Twitter bot that the company described as an experiment in
"conversational understanding." The more you chat with Tay, said Microsoft, the smarter it
gets, learning to engage people through "casual and playful conversation."
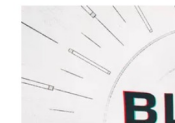
BL

Center for
Social Complexity

Data anonymization*

- "In 1997, Latanya Sweeney identified the Governor of Massachusetts' medical records.
    - Massachusetts released hospital records anonymized by removing names, addresses and SSNs
    - Voter records have name, address, ZIP code, birth date, and sex of every voter
    - Sweeney used zip code, birthdate and gender to uniquely identify Weld's records
    - 87% of US identified by zip, birthdate & gender"

* Ethics of Data Science by Lawrence Hunter, Ph.D.

Unique in the crowd

- We study fifteen months of human mobility data for one and a half million individuals and find that human mobility traces are highly unique. In fact, in a dataset where the location of an individual is specified hourly and with a spatial resolution equal to that given by the carrier's antennas, **four spatio-temporal points are enough to uniquely identify 95% of the individuals**. We coarsen the data spatially and temporally to find a formula for the uniqueness of human mobility traces given their resolution and the available outside information. This formula shows that the uniqueness of mobility traces decays approximately as the 1/10 power of their resolution. Hence, even coarse datasets provide little anonymity. These findings represent fundamental constraints to an individual's privacy and have important implications for the design of frameworks and institutions dedicated to protect the privacy of individuals.

Source: De Montjoye, Y. A., Hidalgo, C. A., Verleysen, M., & Blondel, V. D. (2013). Unique in the crowd: The privacy bounds of human mobility. *Scientific reports*, *3*, 1376.

Values in models

- Values should be explored as part of the modeling process (Pielke 2003).
  - Assumptions can embed values or lack thereof.
- Values are diverse
  - E.g.: power, authority, freedom, access, service...
- Two main types involved (Fleischmann & Wallace 2006)
  - Embedded in computational models
  - Contextual to the design and use of computational models

A framework for modelers to promote ethical "good practice"

Criteria	Summary description	Evaluation of ethical risk
Independence	Models are subjective tools, which reflect modelers' perception of the system at risk. They are influenced by prior art in the area and cultural research links between individuals/institutes. Independence is compromised if conflicts of interest arise (e.g., if modelers are responsible for both design and communication of model merits, as increasingly is the case)	• Is the model provenance known and well documented (i.e., the full history of model development, including funding sources, conceptual design, coding, verification, peer-review processes and publication, as well as the modelers involved in the development)? • Has the model been validated using independent data sources not used for model parameterization?
Transparency	A high level of technical skill and expertise and some fluency in the language of mathematical models are required to communicate model constraints, uncertainties, and assumptions to policymakers. This makes it difficult for complex models to be scrutinized by a diversity of relevant audiences. Without the assistance of experts in risk communication who can broker this knowledge, and robust frameworks for knowledge exchange, policymakers may misinterpret strengths and weakness of model recommendations	• Is there clear documentation of the scientific approach so that methods are robust, repeatable and reproducible? • Is there information about potential conflicts of interest, constraints, or biases affecting data collection and analyses (e.g., racial, ethnic, class, sexual, and gender issues) and any assumptions or uncertainties inherent in the modeling process? • Is the model code open source or available on request? • Are the model assumptions well described and documented and understood by policymakers?

Boden, L. A., & McKendrick, I. J. (2017). Model-Based Policymaking: A Framework to Promote Ethical "Good Practice" in Mathematical Modeling for Public Health Policymaking. Frontiers in public health, 5, 68.
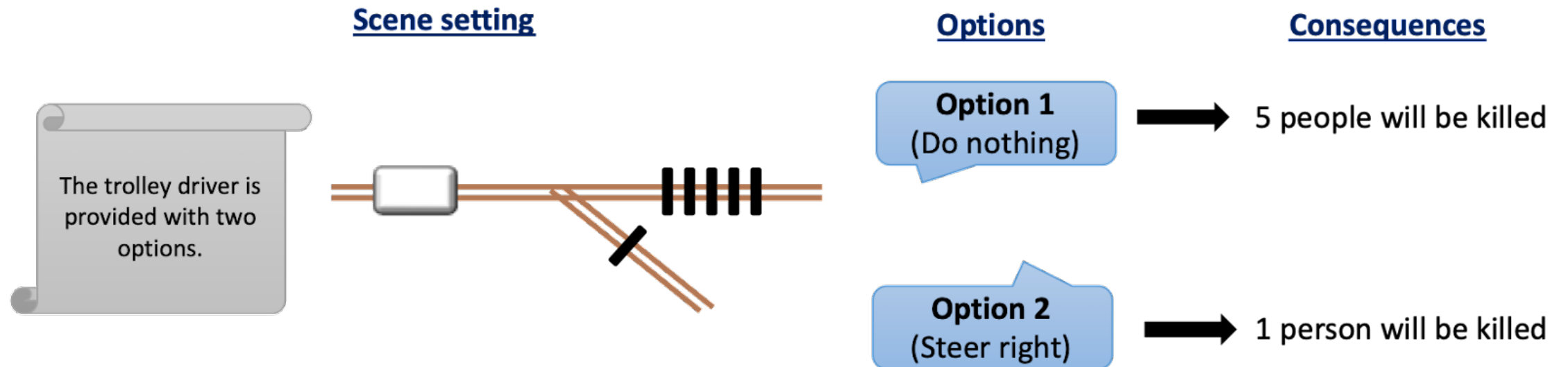
A framework for modelers to promote ethical "good practice"

Criteria	Summary description	Evaluation of ethical risk
Beneficence/non-maleficence	Beneficence is contingent on excellent, ethical model design and diligent protection of data subjects. Irreducible model uncertainties may inadvertently expose research subjects/stakeholders to risks without a guarantee of beneficial outcomes for the population. Few models are evaluated to determine whether recommendations are accurate or effective because there are few "checks and balances" for post-dissemination model quality or utility.	• If a policy decision is based on the model evidence, is it more likely to be effective and beneficial than a decision made in the absence of the model? • Has the model been verified, i.e., does it do what the modeler wants it to do? • Has the model been validated? (i.e., does it realistically map onto what is occurring in real life). • What are the sources and magnitude of uncertainty in the model—are these associated with parametric uncertainty or model selection?
Justice	Models are useful "thought experiments." However, if model evidence is intended to inform policy in the real world, modelers have a duty of care to consider and communicate ethical issues. Ethical risks are influenced by model variability and uncertainty that have important impacts on the distribution of beneficial or harmful consequences.	• Is any lack of knowledge about important parameters attributable to uncertainty or variability? • Where possible, is model variability attributed to known factors, to create more ethical outcomes? • If interventions based on model predictions are implemented in the real world, can the predicted benefits and harms to different individuals and subpopulations be quantified?

Boden, L. A., & McKendrick, I. J. (2017). Model-Based Policymaking: A Framework to Promote Ethical "Good Practice" in Mathematical Modeling for Public Health Policymaking. Frontiers in public health, 5, 68.

Ethical dilemmas

- The trolley problem

Scene setting

The trolley driver is provided with two options.

Options

Option 1 (Do nothing)

Option 2 (Steer right)

Consequences

5 people will be killed

1 person will be killed

Ethical dilemmas

- The trolley problem updated

Conclusion

- Models can be artificial, but they are subject to the same research ethics principles.
- When creating models
  - Document steps for the fellow scientists to recreate the model
  - Use human data responsibly
  - Make sure value embedding does not make unethical consequences
- When reporting models
  - Illuminate all model details and assumptions
  - Share code when possible
  - Report all the findings without cherry-picking
  - Share data responsibly
  - Do not over-interpret model findings

Sources

- Research Ethics Timeline by David B. Resnik, J.D., Ph.D., Bioethicist, NIEHS/NIH

- Ethics of Data Science by Lawrence Hunter, Ph.D.

- The Nuremberg Code and its Impact on Clinical Research by Natalie Jarmusik

- Fleischmann, K. R., & Wallace, W. A. (2006). Ethical implications of values embedded in computational models: An exploratory study. Proceedings of the American Society for Information Science and Technology, 43(1), 1-16.

- Boden, L. A., & McKendrick, I. J. (2017). Model-Based Policymaking: A Framework to Promote Ethical "Good Practice" in Mathematical Modeling for Public Health Policymaking. Frontiers in public health, 5, 68.

- Bak, M. A. (2020). Computing Fairness: Ethics of Modeling and Simulation in Public Health. SIMULATION, 0037549720932656.