COS-R403. Special Research Methods.
Forecasting I: Introduction

Lecturer: Christina Bohk-Ewald

Day 4 of intensive 5-day course

University of Helsinki, Finland
04.05.2020–08.05.2020

Fourth day's class:

- Recap main concepts of last lecture
- Some findings of previous lab session
- The Lee-Carter method, continued
- New directions in mortality forecasting
- COVID-19 pandemic: exceptional situation, unsolved questions, new data. How useful could demographic forecasting be in this context?

Fourth day's class in the lab: Trends of COVID-19 pandemic & Think about how to tackle pressing issues with forecast tools

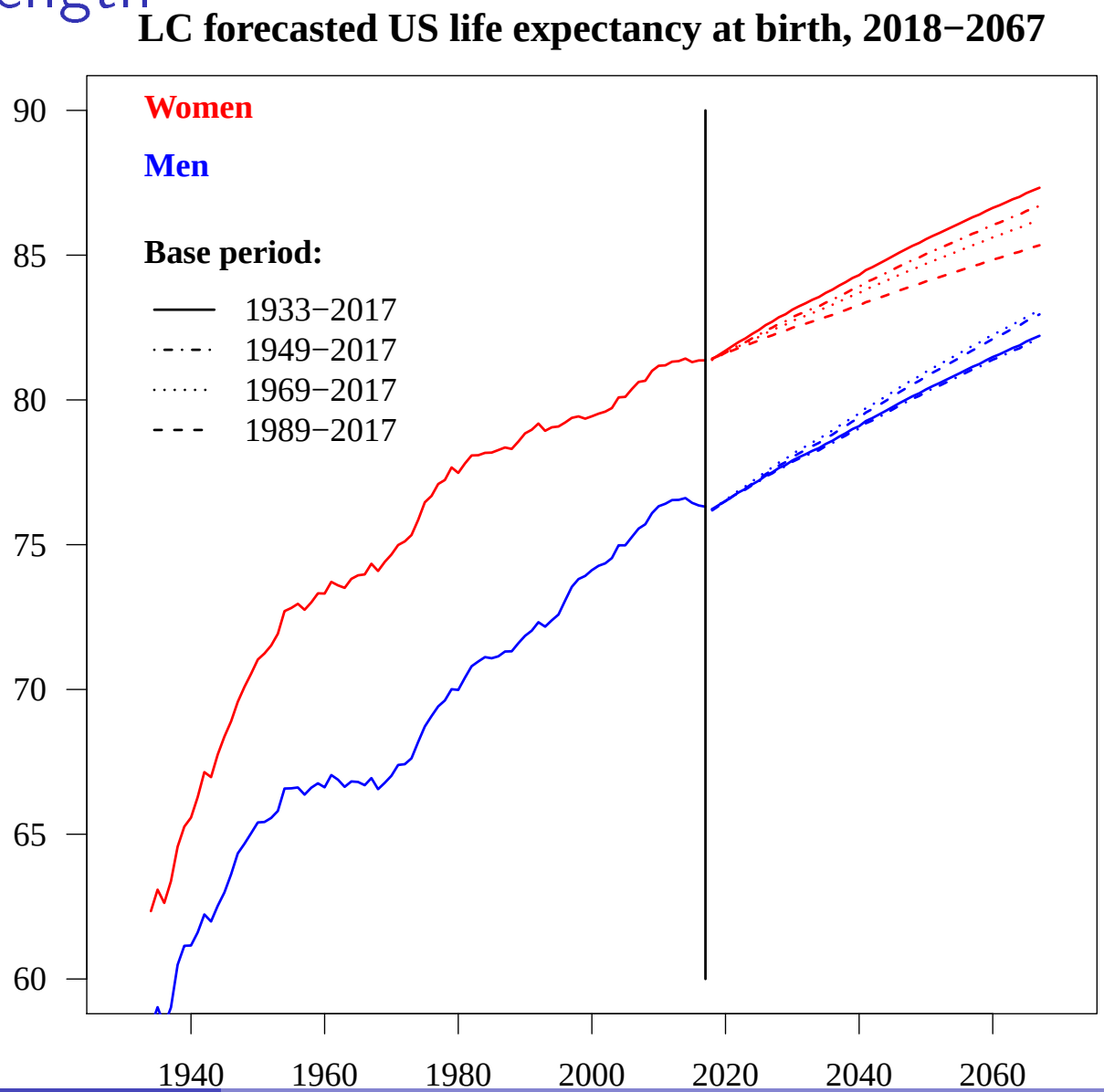
- Analyze trends of COVID-19 pandemic so far
  - Download cases and deaths from JHU CSSE website
  - Explore these data and rank the ten countries with most confirmed cases and reported deaths
- Think about how to tackle pressing issues with forecast tools
  - What are urgent empirical questions during the COVID-19 pandemic?
  - How to use and / or adjust demographic forecast tools for this?
  - What kind of data would you need for this?
  - Example: please read the paper *A demographic scaling model for estimating the total number of COVID-19 infections*, which is vailable at `https://doi.org/10.1101/2020.04.23.20077719`
- Present and discuss your findings. Put this in your report.

Recap main concepts of last lecture

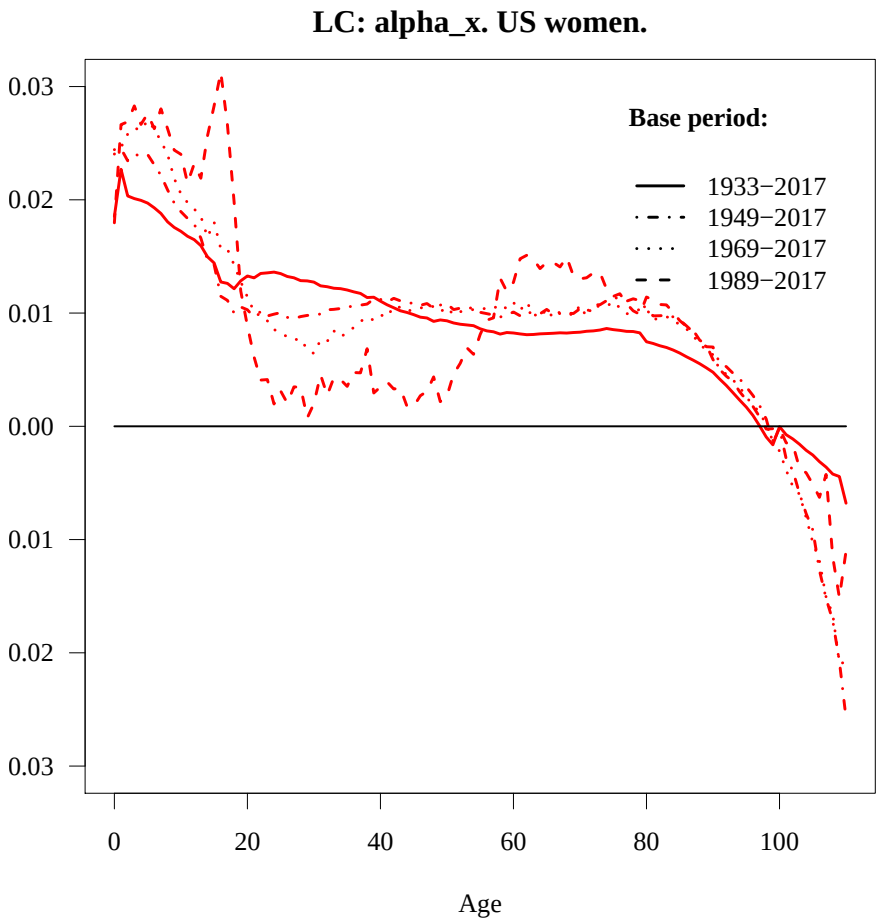
- What types of approaches do we distinguish in mortality forecasting?
- What are the two main steps of the general procedure of the Lee-Carter method?
- What do α_x, β_x, and κ_t mean?
- What are the benefits of the Lee-Carter method?
- What are the drawbacks of the Lee-Carter method?
- ...

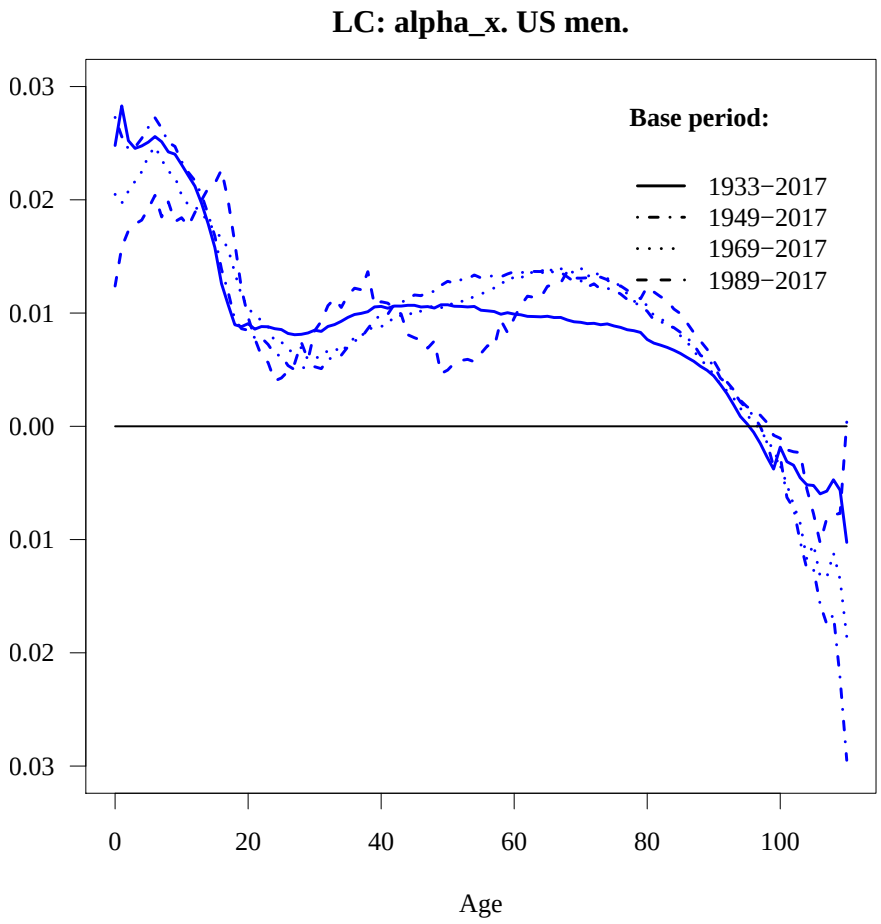
$\Rightarrow$ Questions?

Forecasting US mortality with LC using base periods of different length



Forecasting US mortality with LC using base periods of different length: β_x



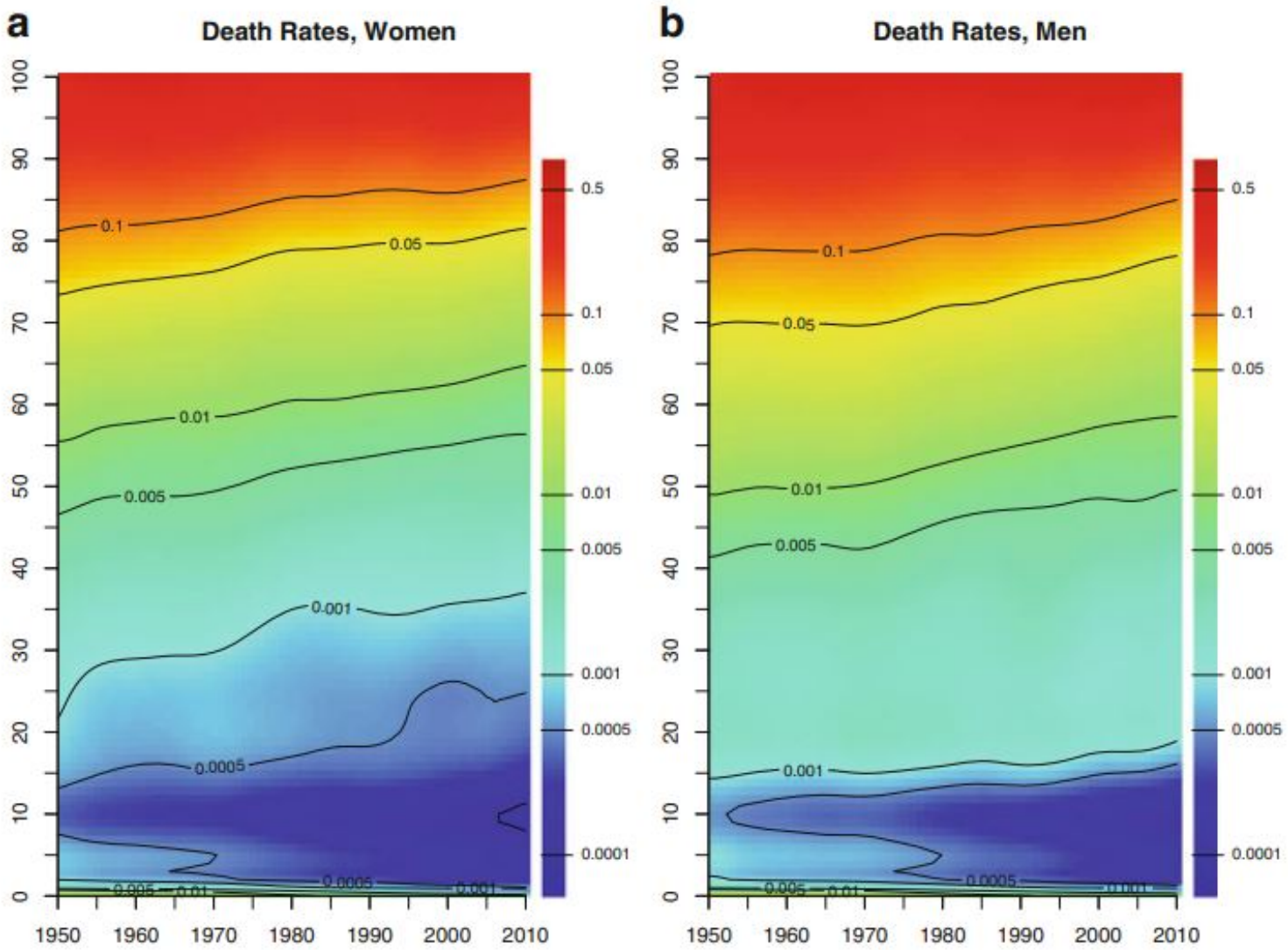


US mortality level and trends

Bohk and Rau (2016) and Rau et al. (2017):

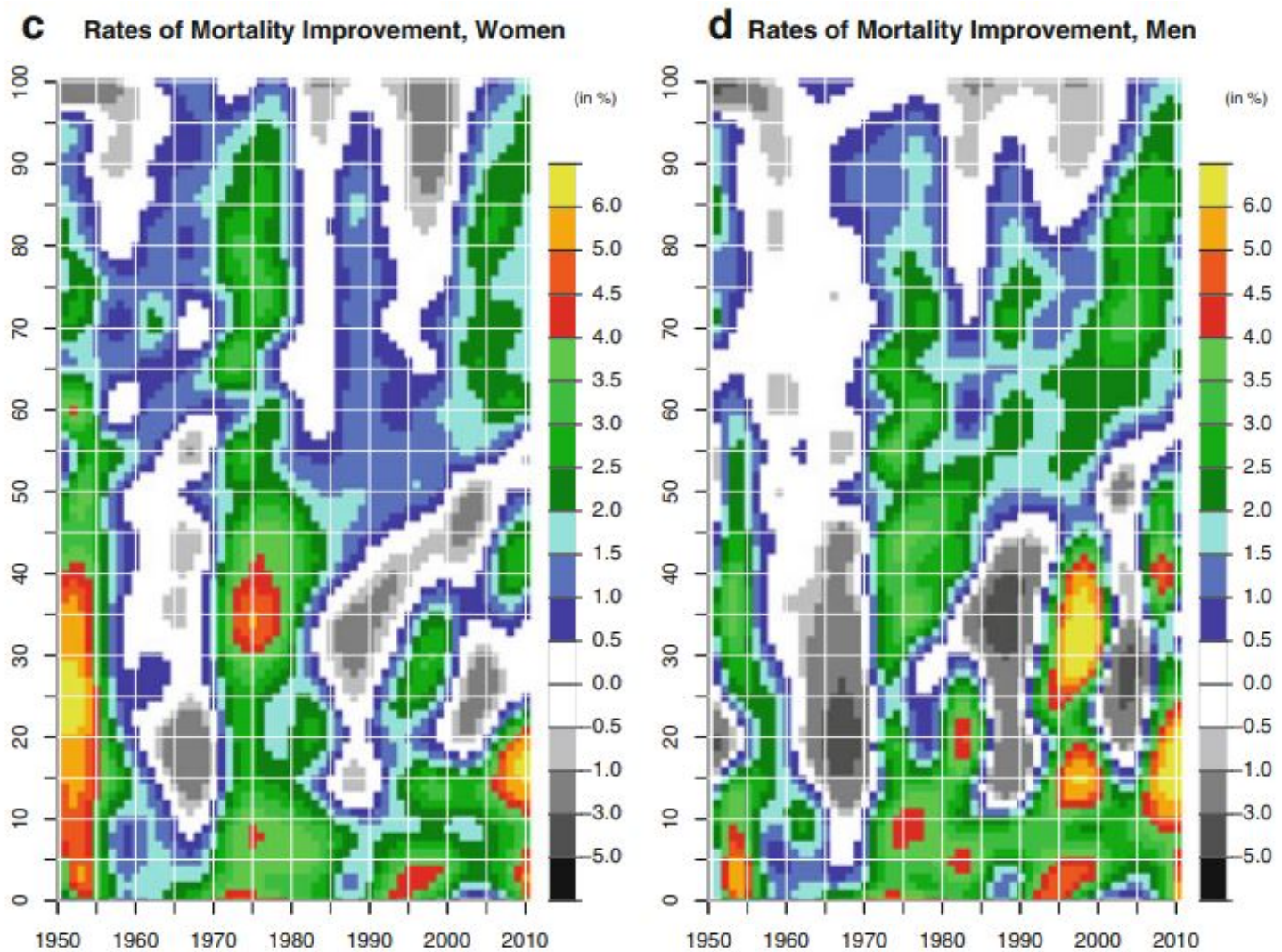
- Rates of mortality improvement
  $\rightarrow \rho_{x,t} = -\left(\frac{m_{x,t+1}}{m_{x,t}} - 1\right)$
- for lung cancer related mortality
  $\rightarrow$ lifestyle risk factor cigarette smoking
- for diabetes related mortality
  $\rightarrow$ link to obesity?
- ...

US mortality level and trends



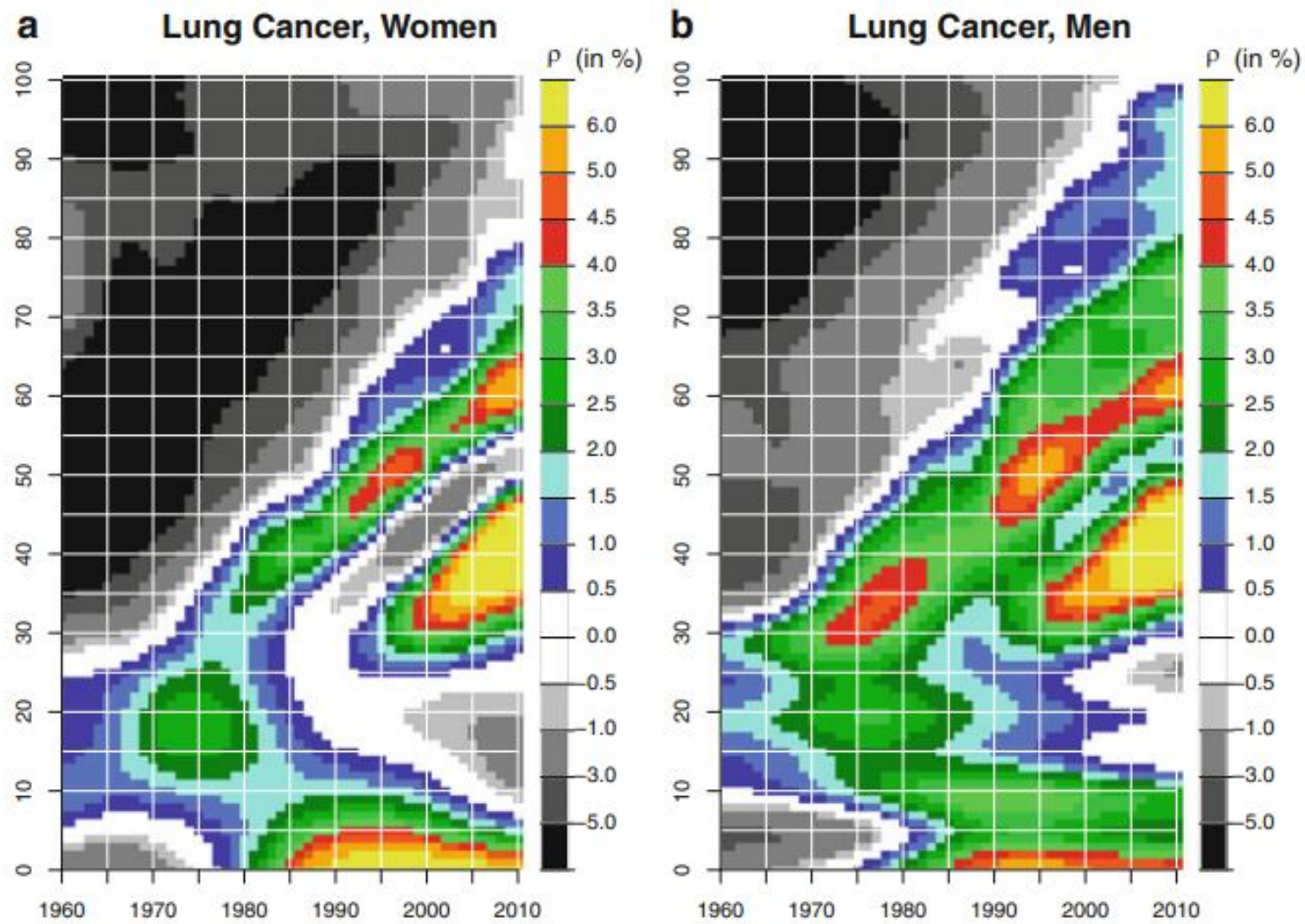
Bohk and Rau (2016, p. 72)

US mortality level and trends



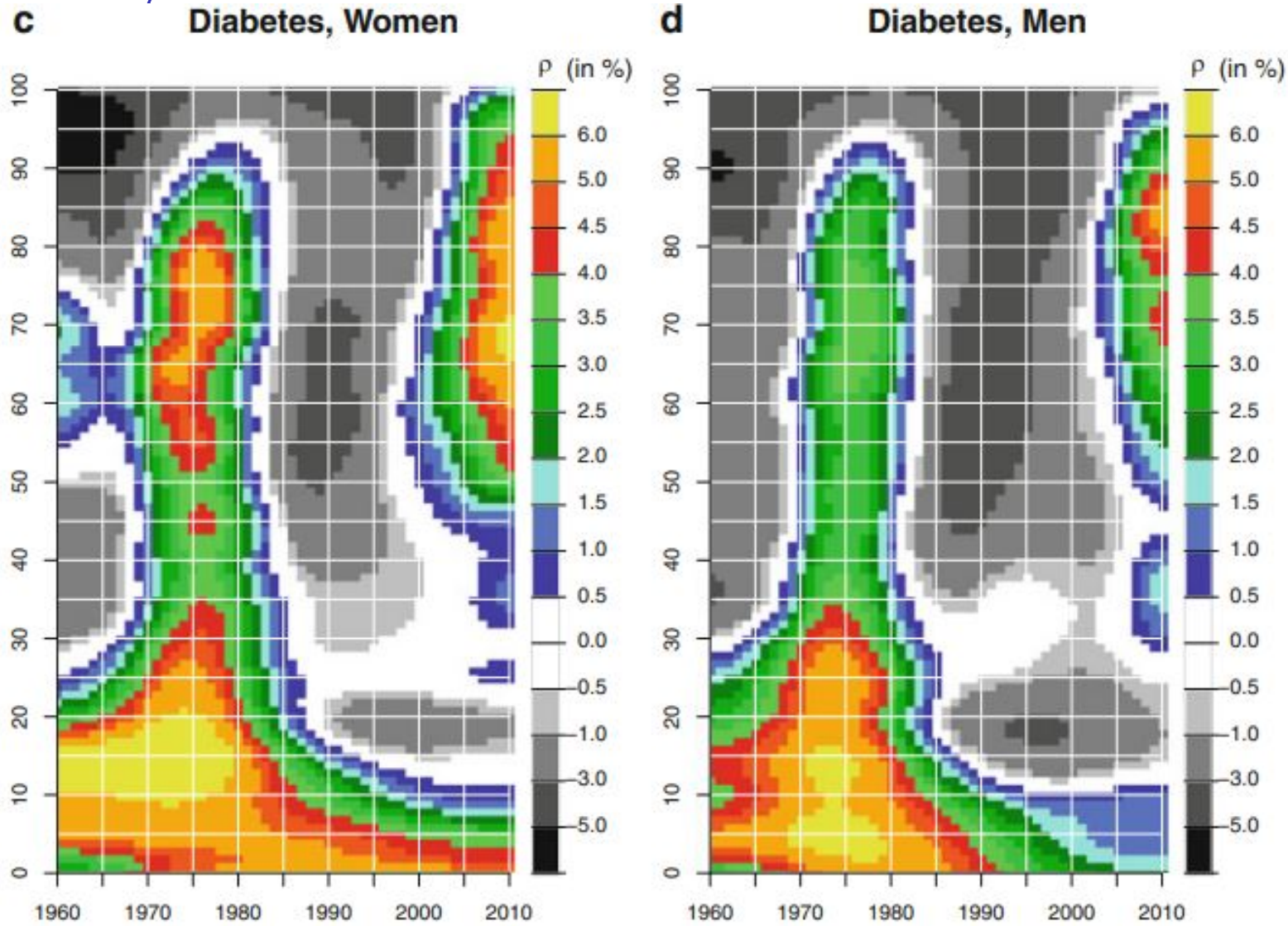
Bohk and Rau (2016, p. 72)

US mortality level and trends



Bohk and Rau (2016, p. 74)

US mortality level and trends



Bohk and Rau (2016, p. 74)

Extensions of the LC method

- Re-fit κ_t in step 1.7 to match
  - Total death counts (e.g. Lee and Carter (1992))
  - Life expectancy at birth (e.g. Lee and Miller (2001))
  - Death counts by age and time (e.g. Booth et al. (2002))
  - Lifespan inequality (e.g. Rabbi and Mazzuco (2020))
  - ...
- Consider mortality of multiple populations (e.g. Li and Lee (2005))
- Consider cohort effects (e.g. Renshaw and Haberman (2006))
- Consider multiple functional principal components to capture non-random patterns (e.g. Hyndman and Ullah (2007))
- Consider time-variant age pattern of mortality change (e.g. Li et al. (2013))
- ...

$\rightarrow$ Overview of some LC extensions in e.g. Booth (2006)

Many R implementations of the Lee-Carter model out there

R-packages and related functions (non-exhaustive):

- `demography: lca(), forecast.lca()`
- `StMoMo:`
  Vignette available at `https://cran.r-project.org/web/packages/StMoMo/vignettes/StMoMoVignette.pdf`
- `MortalityForecast: model.LeeCarter(), predict()`
- `ilc`
- `LifeMetrics`
  available at `http://www.macs.hw.ac.uk/~andrewc/lifemetrics/`

→ helpful to apply LC method but please do not use them as black box

→ to **really understand LC method** it is good to **implement it yourself**!

Other directions in mortality forecasting

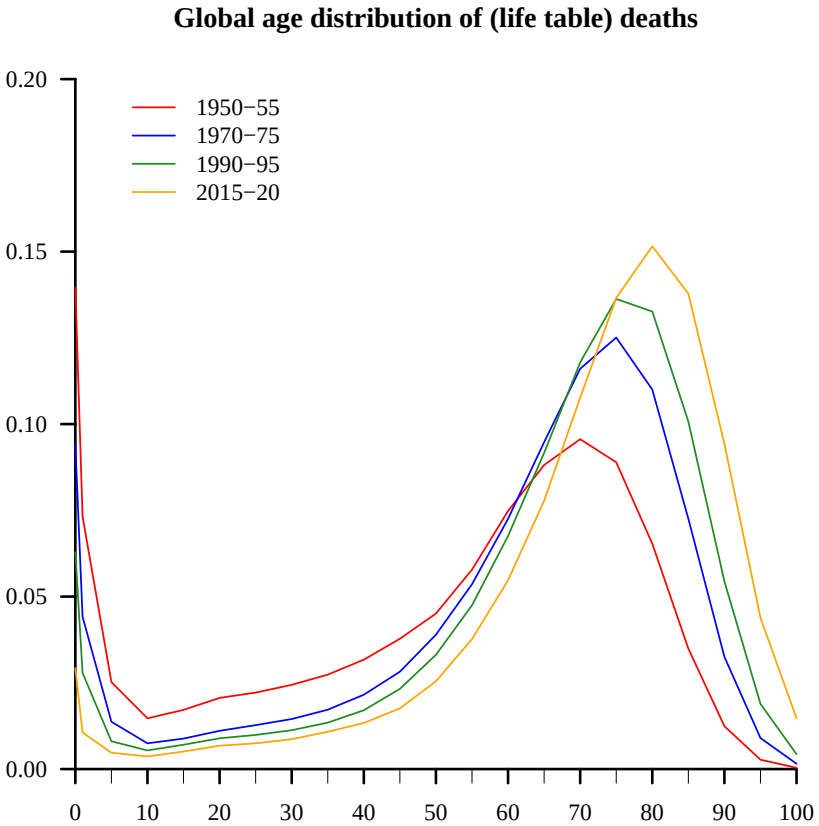
What are the methodological and empirical questions in mortality forecasting of the recent past and at the moment?

Other directions in mortality forecasting

1. Coherent forecasting to consider developments of other populations:
   - Multiple countries (e.g. Li and Lee (2005))
   - Women and men (e.g. Hyndman et al. (2013))
   - Bayesian methodology of recent UNWPP (e.g. Raftery et al. (2013))
   - ...
2. Approaches to consider mortality attributable to health-related behavior:
   - Smoking (e.g. Janssen et al. (2013))
   - ...

Other directions in mortality forecasting

3. Approaches to capture more closely mortality by age over time:
   - Rates of mortality improvement (e.g. Bohk-Ewald and Rau (2017))
   - Distribution of ages at death (e.g. Basellini and Camarda (2019))
   - ...



Data source: UNWPP 2019 (abridged lif tables)

Other directions in mortality forecasting

Other current / future directions:

- Account for climate change (e.g. air quality)
- Account fo COVID-19 pandemic
- ...

Other directions in mortality forecasting

What would you do?

How would you forecast mortality?

$\rightarrow$ ...might also be topic for PhD thesis

COVID-19 pandemic

- COVID-19: **co**rona**vi**rus **d**isease **2019**
- SARS-CoV-2 is the causative agent of COVID-19
- SARS-CoV-2 belongs to coronavirus family,
  as do e.g. SARS-CoV and MERS-CoV
- COVID-19 is
  a mild to severe respiratory illness that is caused by a coronavirus (Severe acute respiratory syndrome coronavirus 2 of the genus Betacoronavirus), is transmitted chiefly by contact with infectious material (such as respiratory droplets), and is characterized especially by fever, cough, and shortness of breath and may progress to pneumonia and respiratory failure

Source of information is the Merriam Webster Dictionary:

`https://www.merriam-webster.com/words-at-play/new-dictionary-words-coronavirus-covid-19`

`https://www.merriam-webster.com/dictionary/COVID-19`

COVID-19 pandemic

COVID-19 compared to other common conditions

SYMPTOM	COVID-19	COMMON COLD	FLU	ALLERGIES
Fever	Common	Rare	Common	Sometimes
Dry cough	Common	Mild	Common	Sometimes
Shortness of breath	Common	No	No	Common
Headaches	Sometimes	Rare	Common	Sometimes
Aches and pains	Sometimes	Common	Common	No
Sore throat	Sometimes	Common	Common	No
Fatigue	Sometimes	Sometimes	Common	Sometimes
Diarrhea	Rare	No	Sometimes*	No
Runny nose	Rare	Common	Sometimes	Common
Sneezing	No	Common	No	Common

*Sometimes for children

Sources: CDC, WHO, American College of Allergy, Asthma and Immunology

BUSINESS INSIDER

https://www.businessinsider.com/coronavirus-symptoms-compared-to-flu-common-cold-and-allergies-2020-3

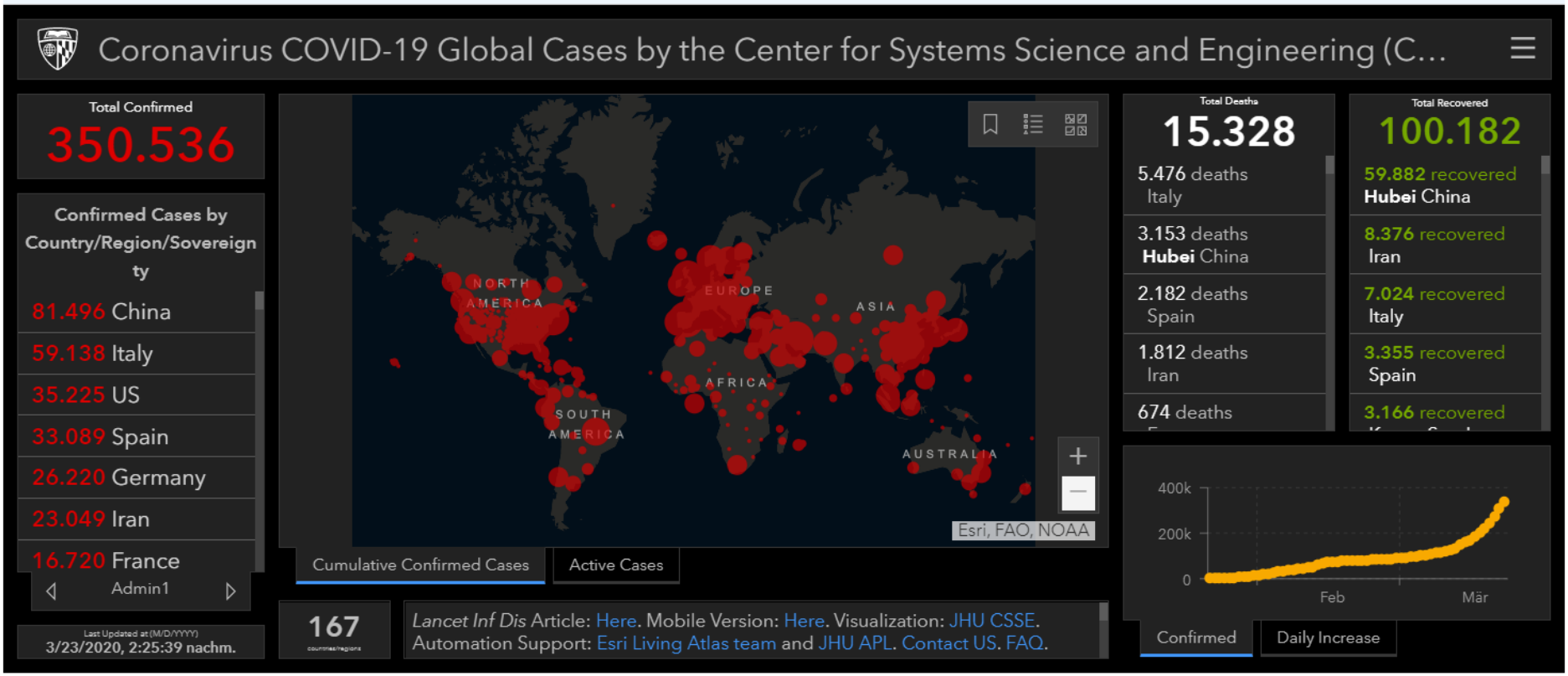
COVID-19 pandemic

Spread of COVID-19 ...

- **Outbreak.** First confirmed cases of COVID-19 in Wuhan, Hubei Province, Mainland China, since November / December 2019
- **Epidemic.** Spread of COVID-19 in Mainland China in December 2019 and January 2020.
- **Pandemic.** COVID-19 spreads to countries across the globe
  - For example, Italy, US, Spain, Germany, Iran, France are countries with many confirmed cases and deaths in March 2019
  - Italy has more confirmed cases and reported deaths than China in March 2019
  - ...

COVID-19 pandemic

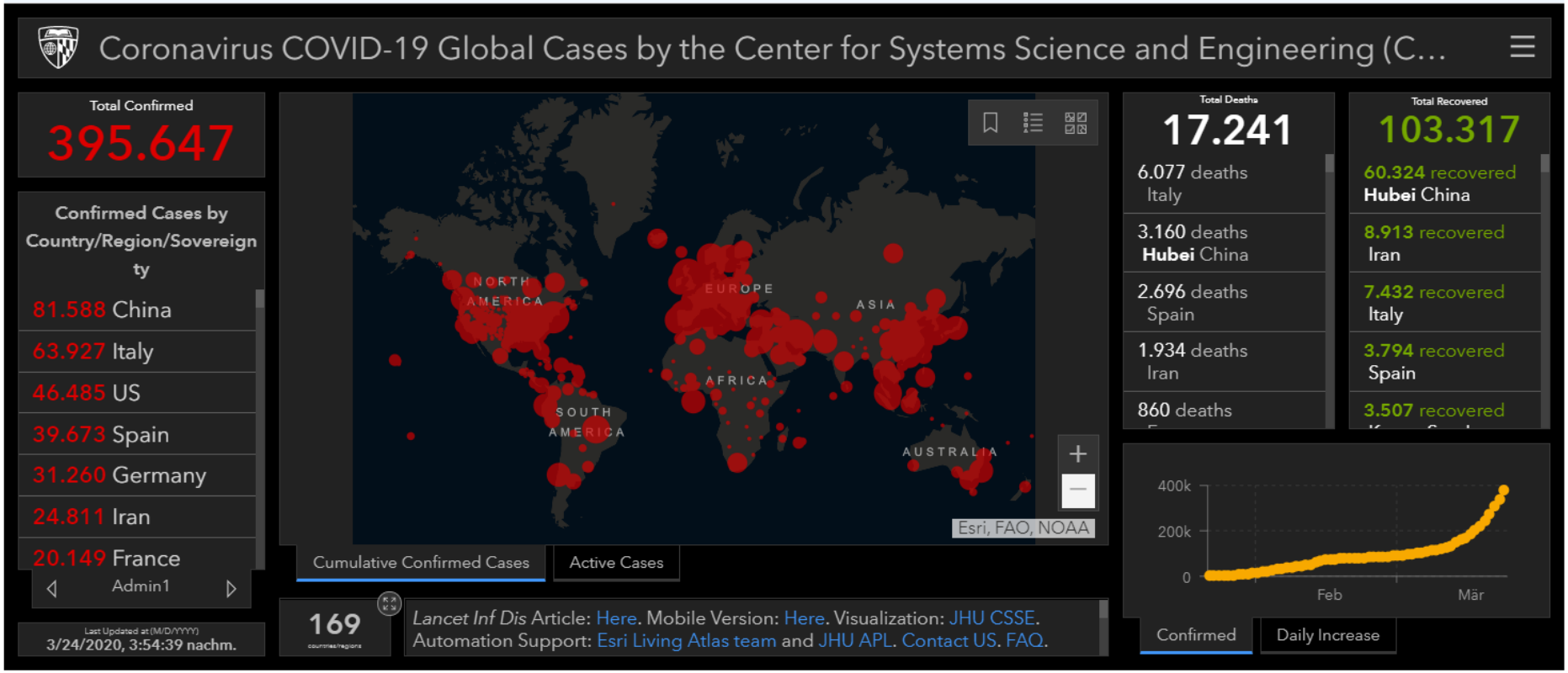
Spread of COVID-19 as of March 23, 2020...



`https://data.humdata.org/dataset/novel-coronavirus-2019-ncov-cases`

COVID-19 pandemic

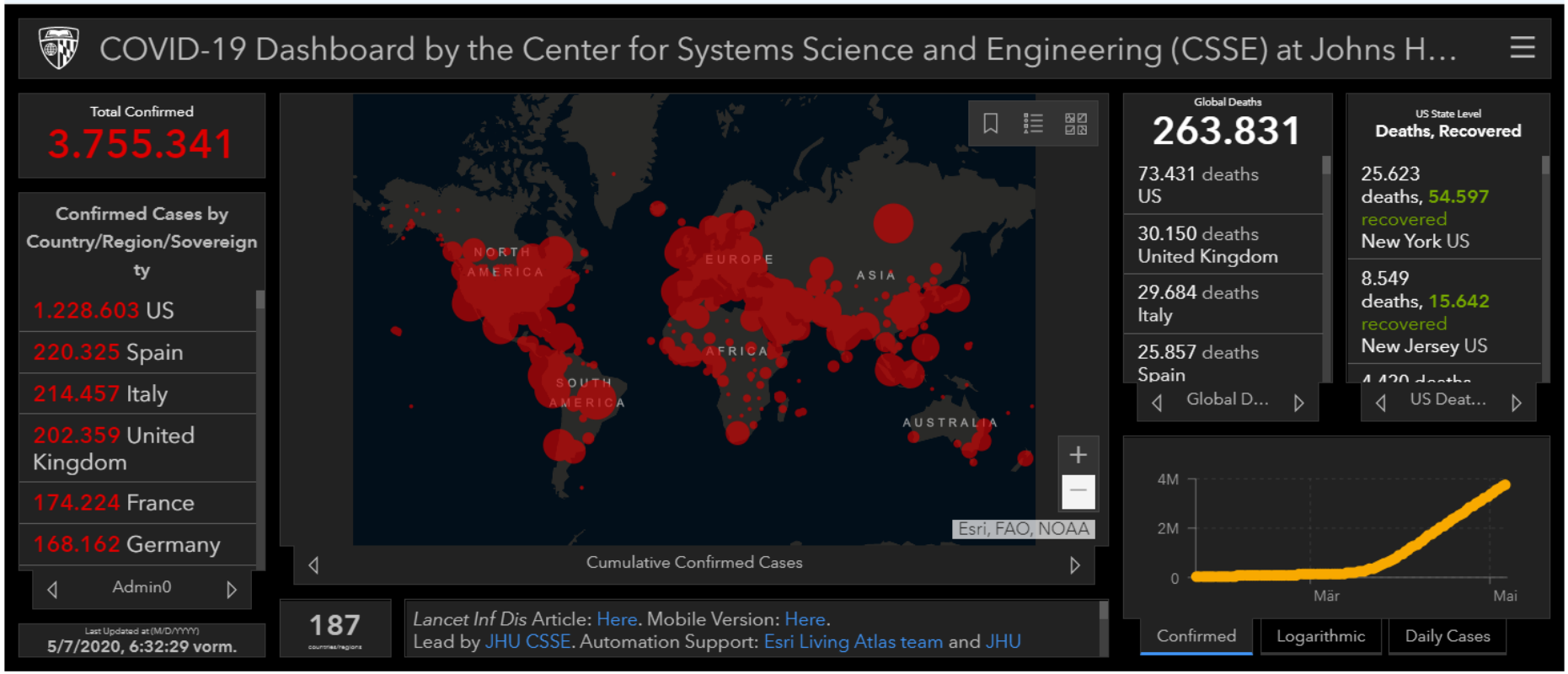
Spread of COVID-19 as of March 24, 2020...



`https://data.humdata.org/dataset/novel-coronavirus-2019-ncov-cases`

COVID-19 pandemic

Spread of COVID-19 as of May 7, 2020...



https://data.humdata.org/dataset/novel-coronavirus-2019-ncov-cases

COVID-19 pandemic

What to think of COVID-19 from demographic-epidemiological view point

- COVID-19 is highly contagious $\longrightarrow$ spreads very fast $\longrightarrow$ requires fast & drastic action of e.g. scientists, medical service & politicians
- Demographic-epidemiological forecasts have high impact. They show:
  - How fast people will get infected
  - How many will be severely ill and will need intensive care
  - How many people are likely to die
  - What actions could help to prevent/slow down further spread of disease
  - ...

$\longrightarrow$ Demographic-epidemiological forecasts can be a game changer in the fight against COVID-19

COVID-19 pandemic

Early studies and projections:

1. **Morbitiy and mortality of COVID-19.** Early paper of *Novel Coronavirus Pneumonia Emergency Response Epidemiology Team* analyzes data of epidemic center in China as of Feb 11, 2020
2. ***Flatten-the-curve*.** For example, Population health analyst Drew A. Harris, Thomas Jefferson University, Philadelphia, USA, twittered about health-care system capacity on Feb 28, 2020
3. **Scenarios projecting COVID-19 cases based on various control measures.** Imperial College London, team of Prof. Neil Ferguson, report 9 published in March 16, 2020
4. ...

COVID-19 pandemic

1. Early paper analyzes data of epidemic center in Mainland China:

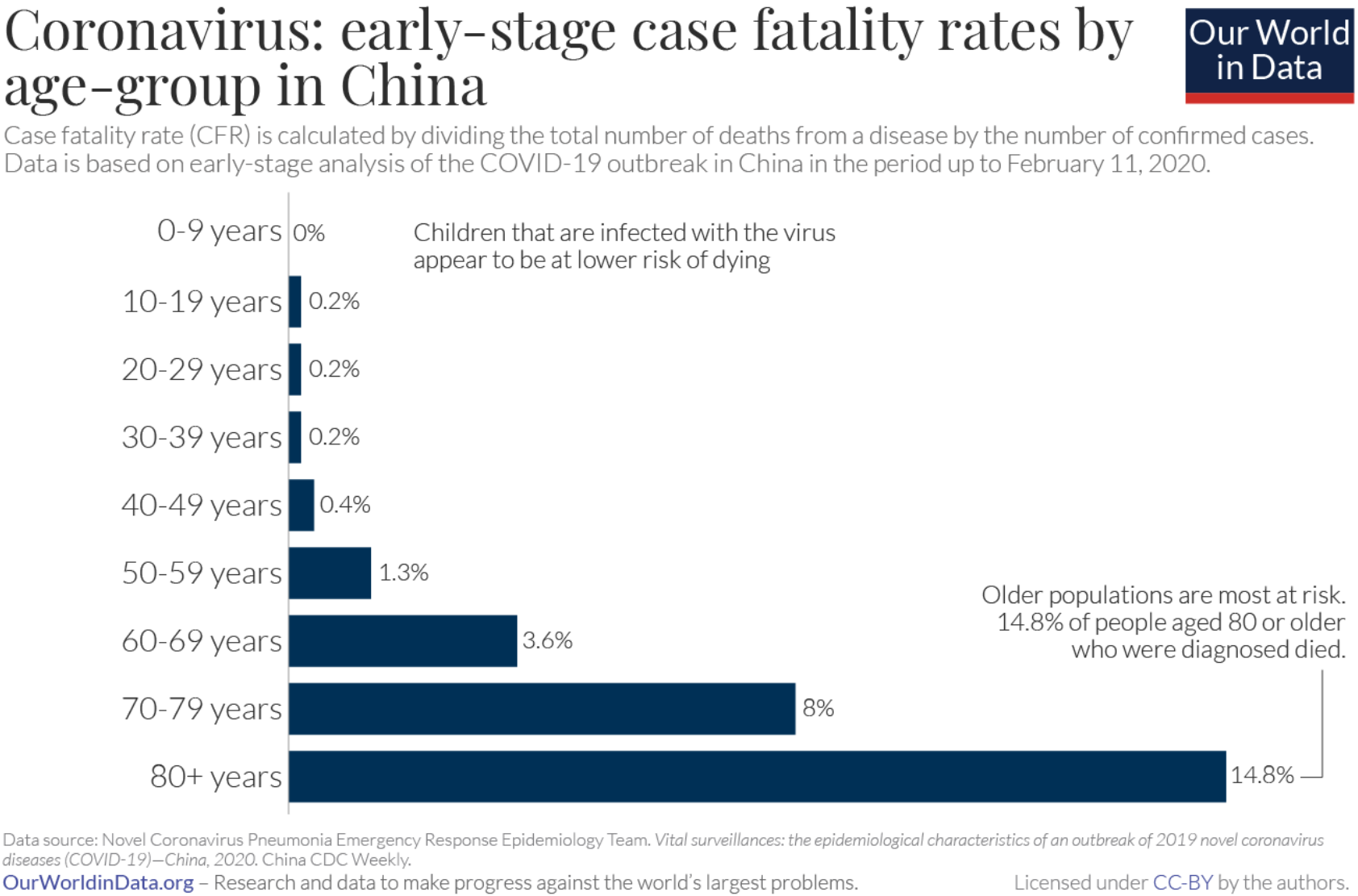
   - Novel Coronavirus Pneumonia Emergency Response Epidemiology Team (2020). Analysis of Epidemiological Characteristics of New Coronavirus Pneumonia. Chinese Journal of Epidemiology, 41 `http://rs.yiigle.com/yufabiao/1181998.htm`
   - Study of 44672 confirmed cases, which were reported in Mainland China as of February 11, 2020
   - 86.6% between 30 and 79 years old
   - 74.7% of Hubei Province

$\rightarrow$ Analyzed confirmed cases by severeness of the course of COVID-19 and its mortality, and how this varies by demographic characteristics

COVID-19 pandemic

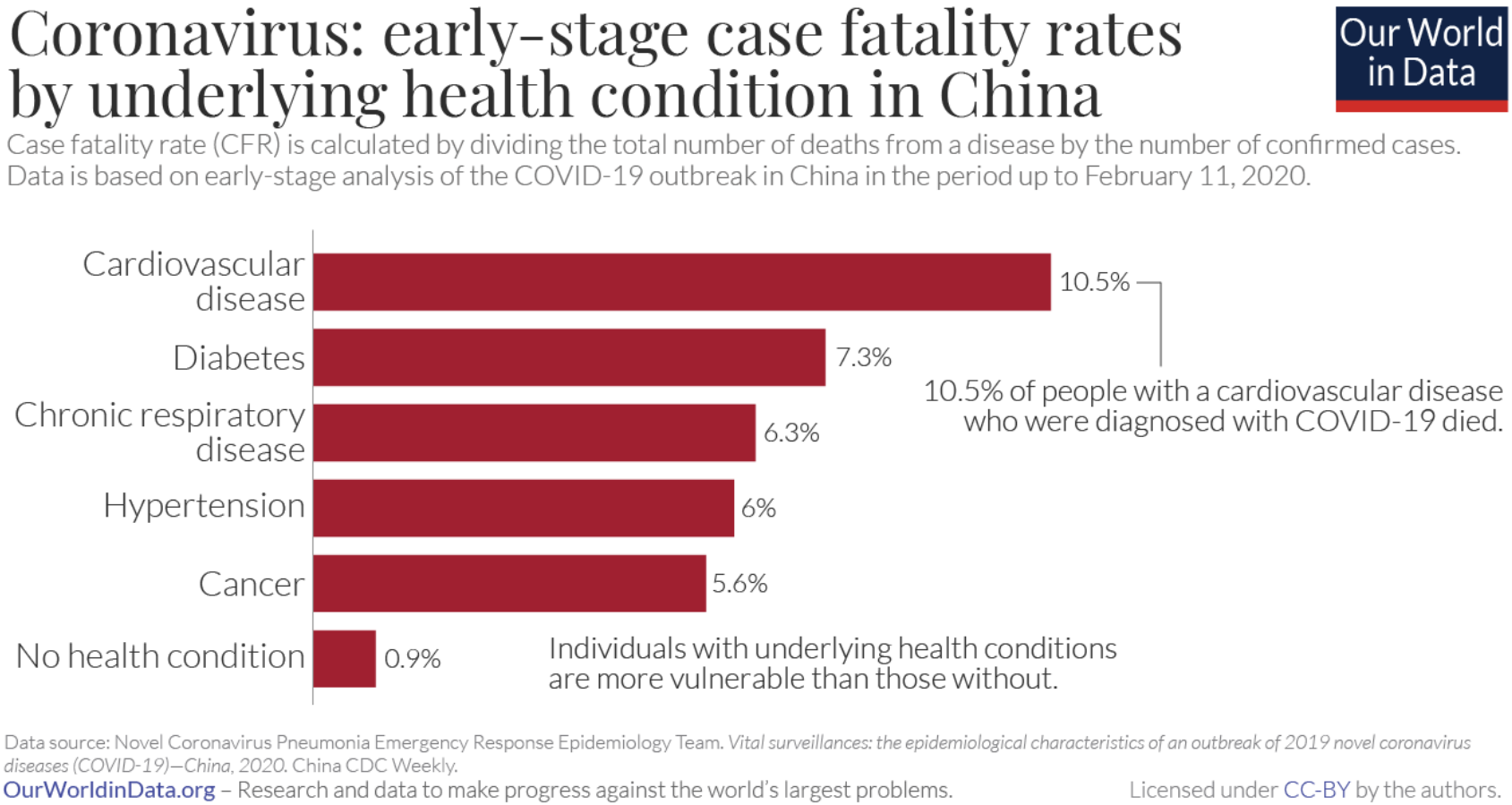
1. Early paper analyzes data of epidemic center in Mainland China:
   - 80.9% mild cases
   - Crude case fatality rate of 2.3% $\rightarrow$ $(\frac{1023 deaths}{44672 confirmed\ \ cases} \cdot 100)$
   - Appears to affect particularly
     - Older people
     - People with preconditions
     - Men
       - ★ Crude case fatality rate for men: 2.8%
       - ★ Crude case fatality rate for women: 1.7%

COVID-19 pandemic



https://ourworldindata.org/uploads/2020/03/Coronavirus-CFR-by-age-in-China-1-800x526.png

COVID-19 pandemic



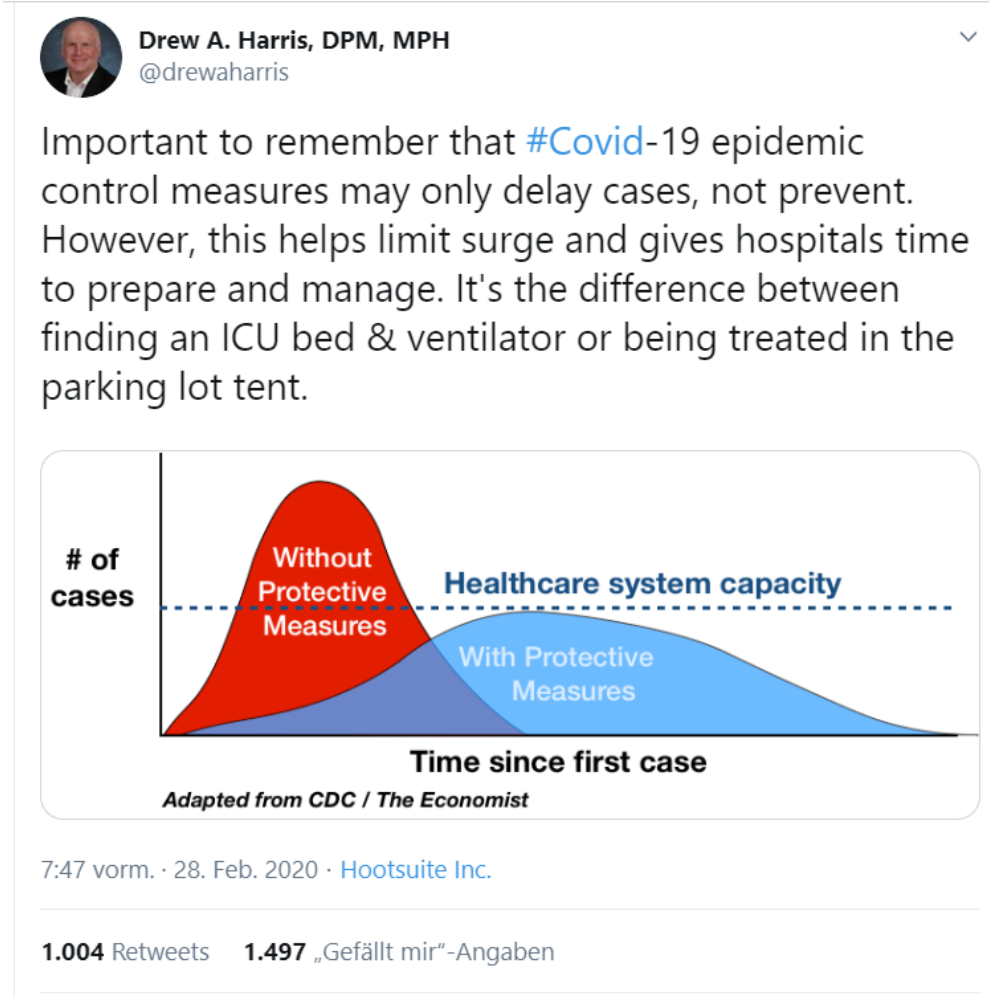
`https://ourworldindata.org/uploads/2020/03/Coronavirus-CFR-by-health-condition-in-China.png`

COVID-19 pandemic

2. Flatten the curve...

- Population health analyst Drew A. Harris, Thomas Jefferson University, Philadelphia, USA, twittered about health-care system capacity on Feb 28, 2020
- Lancet article from March 21, 2019

`https://www.thelancet.com/journals/lancet/article/PIIS0140-6736(20)30567-5/fulltext`



→ `https://twitter.com/drewaharris/status/1233267475036372992`

COVID-19 pandemic

3. Scenarios projecting COVID-19 cases based on various control measures

- Imperial College London, team of Prof. Neil Ferguson, report 9 published in March 16, 2020
- Scenarios for GB and US based on (un)mitigated epidemic

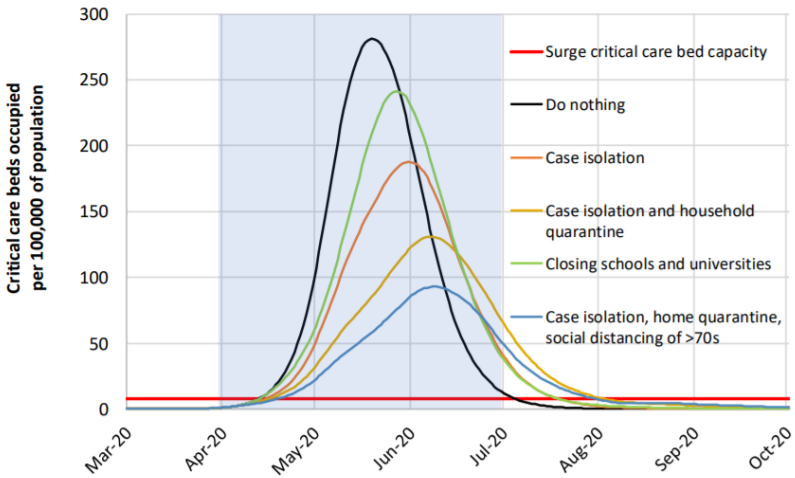


Figure 2: Mitigation strategy scenarios for GB showing critical care (ICU) bed requirements. The bl

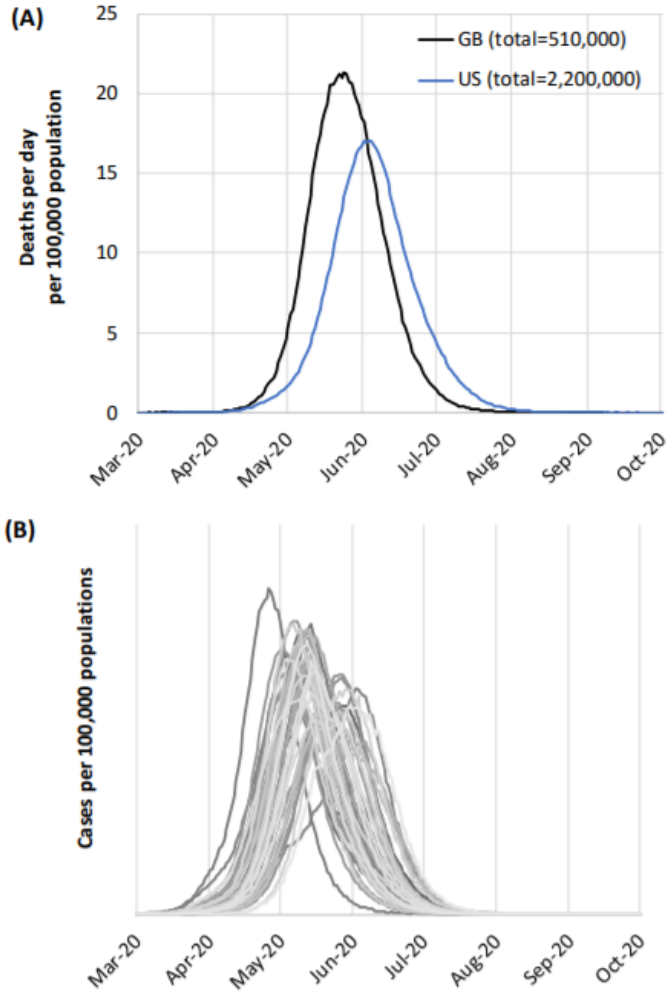


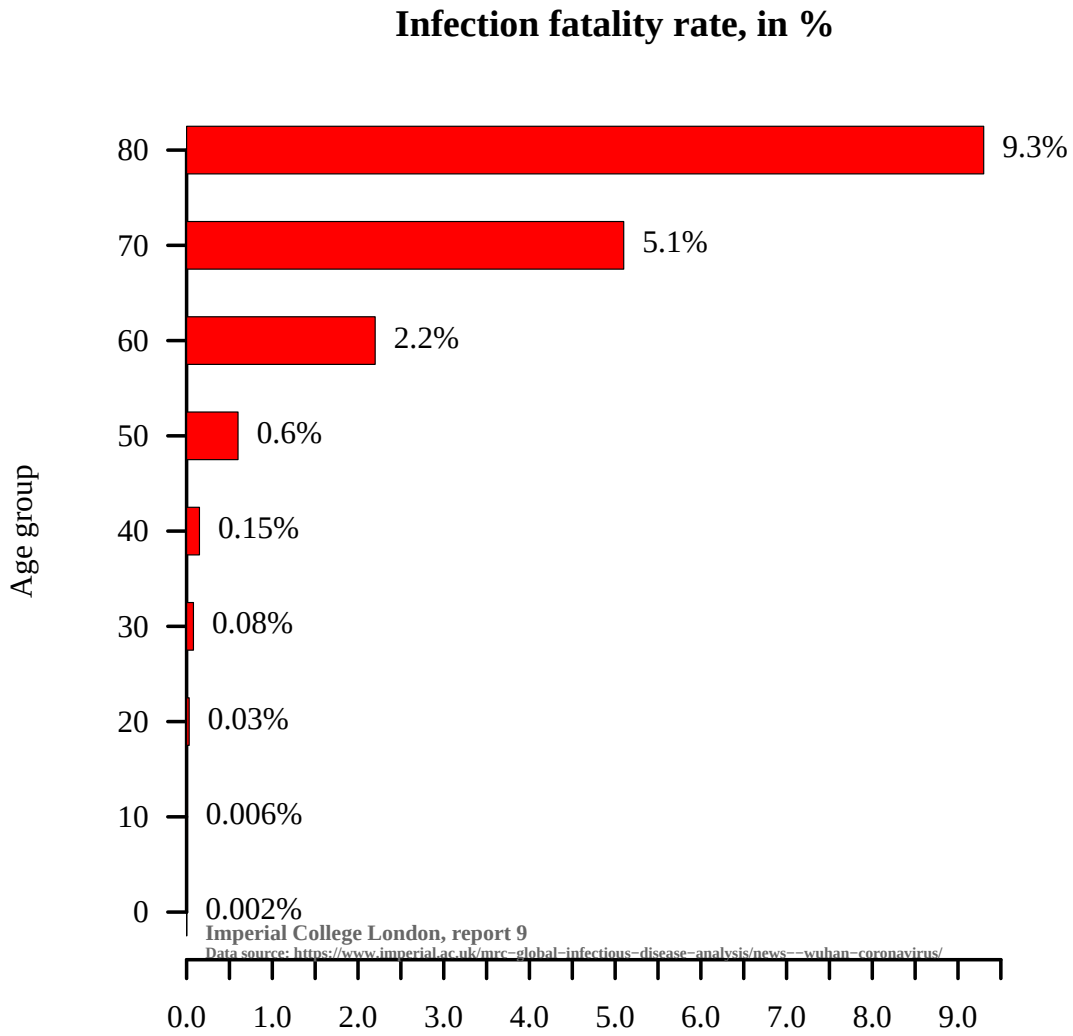
Figure 1: Unmitigated epidemic scenarios for GB and the US. (A) Projected deaths per day p

→ https://www.imperial.ac.uk/mrc-global-infectious-disease-analysis/news--wuhan-coronavirus/

COVID-19 pandemic

3. Scenarios projecting COVID-19 cases based on various control measures

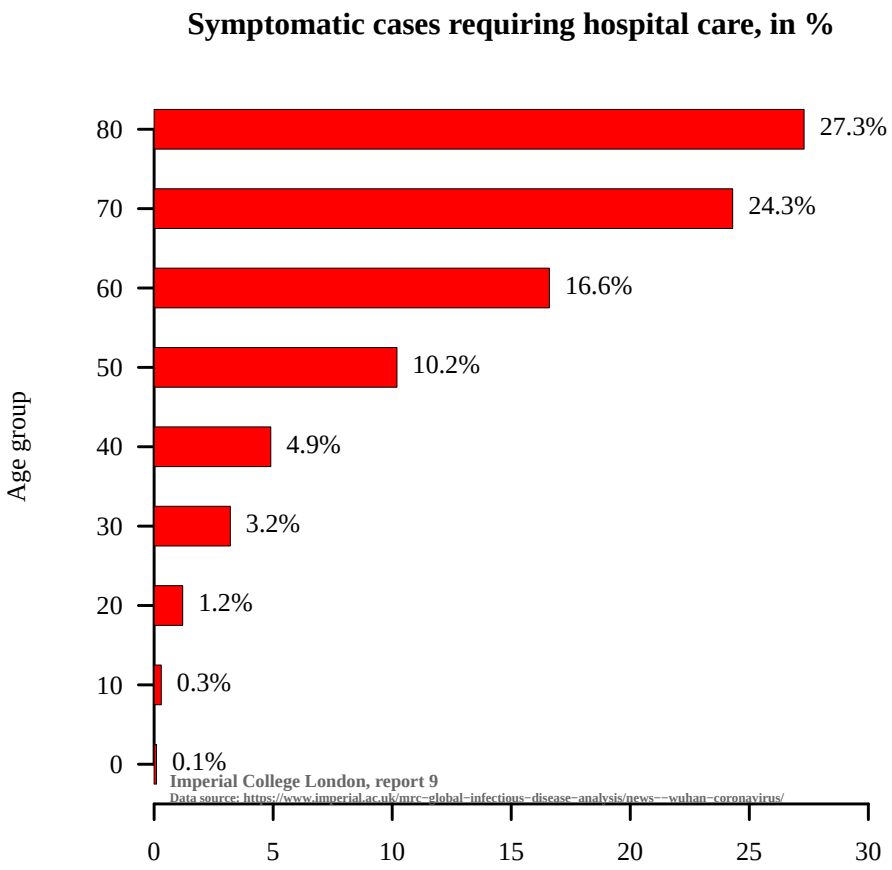
- *Infection fatality rate (IFR)* is the number of deaths divided by the confirmed und undiagnosed cases
- Subset of China, applied to GB:
  - IFR: 0.9%.
  - Hospitalization 4.4%
  - Critical care: 30% of those in hospital

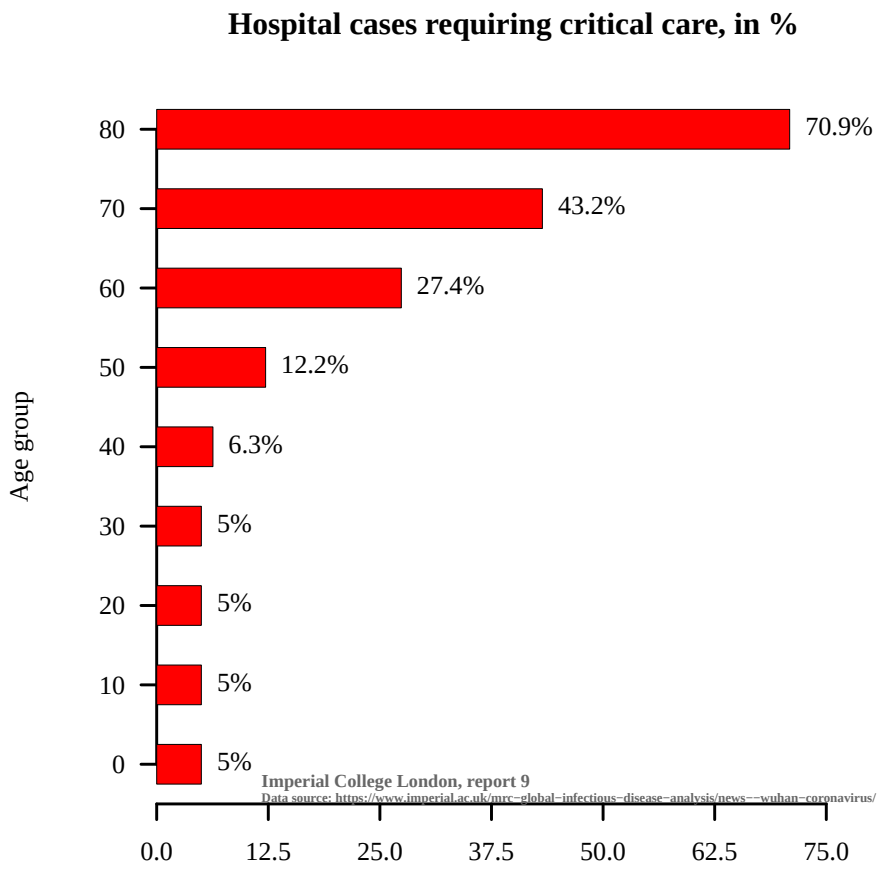


⟶ https://www.imperial.ac.uk/mrc-global-infectious-disease-analysis/news--wuhan-coronavirus/

COVID-19 pandemic

3. Scenarios projecting COVID-19 cases based on various control measures





→ `https://www.imperial.ac.uk/mrc-global-infectious-disease-analysis/news--wuhan-coronavirus/`

COVID-19 pandemic

World Health Organization (WHO)

- First case of unknown pneumonia in Wuhan, China, reported to WHO on Dec 31, 2019
- WHO declared disease *Public Health Emergency of International Concern* on Jan 30, 2020
- WHO anounced name for this new disease: COVID-19 on Feb 11, 2020
- WHO characterizes COVID-19 as a pandemic on March 11, 2020
- WHO suggests measures of mitigation and prevention to governments
- `https://www.who.int/emergencies/diseases/novel-coronavirus-2019/events-as-they-happen`

COVID-19 pandemic

Control measures: governments urge for closing national borders, social distancing, and self-quarantine, ...

- China
  - Jan 23, 2020: Public transport suspended
  - Jan 25, 2020: Lunar New Year celebrations cancelled
  - Isolation, quarantine, social distancing, and community containment
  - More information:

    `https://jamanetwork.com/journals/jama/fullarticle/2762130?mod=article_inline`

`https://twitter.com/ChinaDaily/status/1220052882596286465`

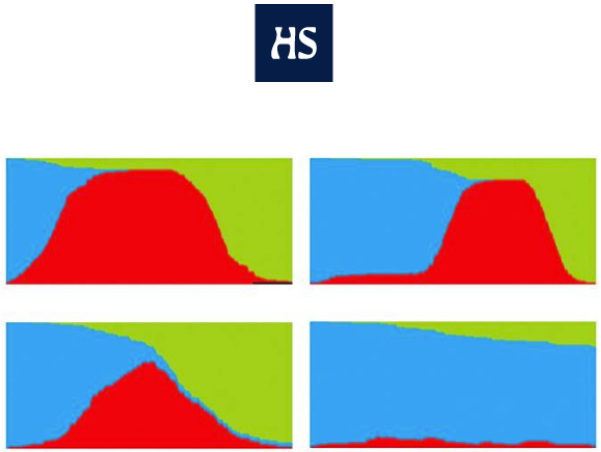
- Europe. Ban on going out in Italy since March 9, 2020, in Spain since March 14, 2020. Other countries, counties followed.
- US and GB reacted later. Ban on going out in California since March 19, 2020. Other states (New York, Illinois, ...) followed.

COVID-19 pandemic

Media, science, politics go hand in hand...

- Huge media attention: The Economist, The Guardian, The NY Times, ... publish and discuss development of COVID-19, measures of control, and their social and economic consequences
- People discuss it on social media platforms
- Example. Finland. Helsingin Sanomat. March 17, 2020:

  `https://dynamic.hs.fi/2020/sosiaalinen-etaisyys/`



- red-infected, blue-not infected, green: recovered
- Top left: unrestricted.
  Top right: leaking quarantine.
- Bottom left: social distaning.
  Bottom right: strict restraint.

COVID-19 pandemic

Still unclear what will eventually have been best strategy...

- **Unrestricted transmission.** Many interactions, many deaths.
- **Flatten the curve.** Restrict interactions but accept infections until epidemic has run its course. This scenario allows for deaths.
- **Trace and track** every case. However, undetected cases might grow fast and make it impossible to maintain this strategy.
- **Stop then restart.** Drastic lockdown. Minimize all unnecessary interaction (public transport, educational institutions, retail) until no new infections are reported. Then slowly restart social and economic activity.
- Vaccine, medical treatment, and / or (herd) immunity is way out
- Source: `https://www.theguardian.com/commentisfree/2020/mar/21/the-case-for-shutting-down-almost-everything-and-restarting-when-coronavirus-is-gone`

COVID-19 pandemic

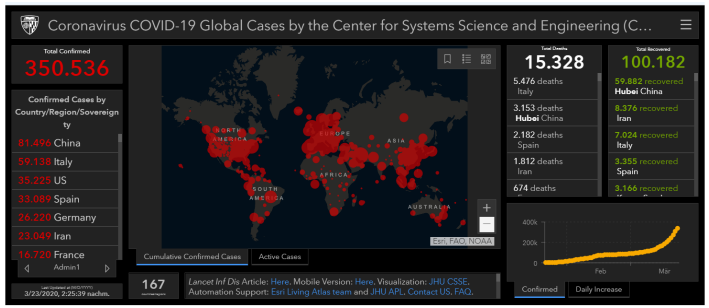
Demographic forecasts

- Demographic forecasts analyze, model, and forecast e.g. mortality and populations, which could be helpful to analyze, model, and forecast mortality from COVID-19 and people having been and being infected with COVID-19
- Requirement: reliable data on confirmed cases, reported deaths, and people recovered from COVID-19 in order to make reliable forecasts and if-then projections
- Experts and decision makers need the correct tools at hand in times of a crisis
- ...

COVID-19 pandemic

Analyze data yourself...

- Johns Hopkins University Center for Systems Science and Engineering (JHU CCSE)
- Epidemiological data since January 22, 2020
- Daily updates
- Compiled data from multiple sources such as WHO, US CDC
- GitHub data repo: `https://github.com/CSSEGISandData/COVID-19`

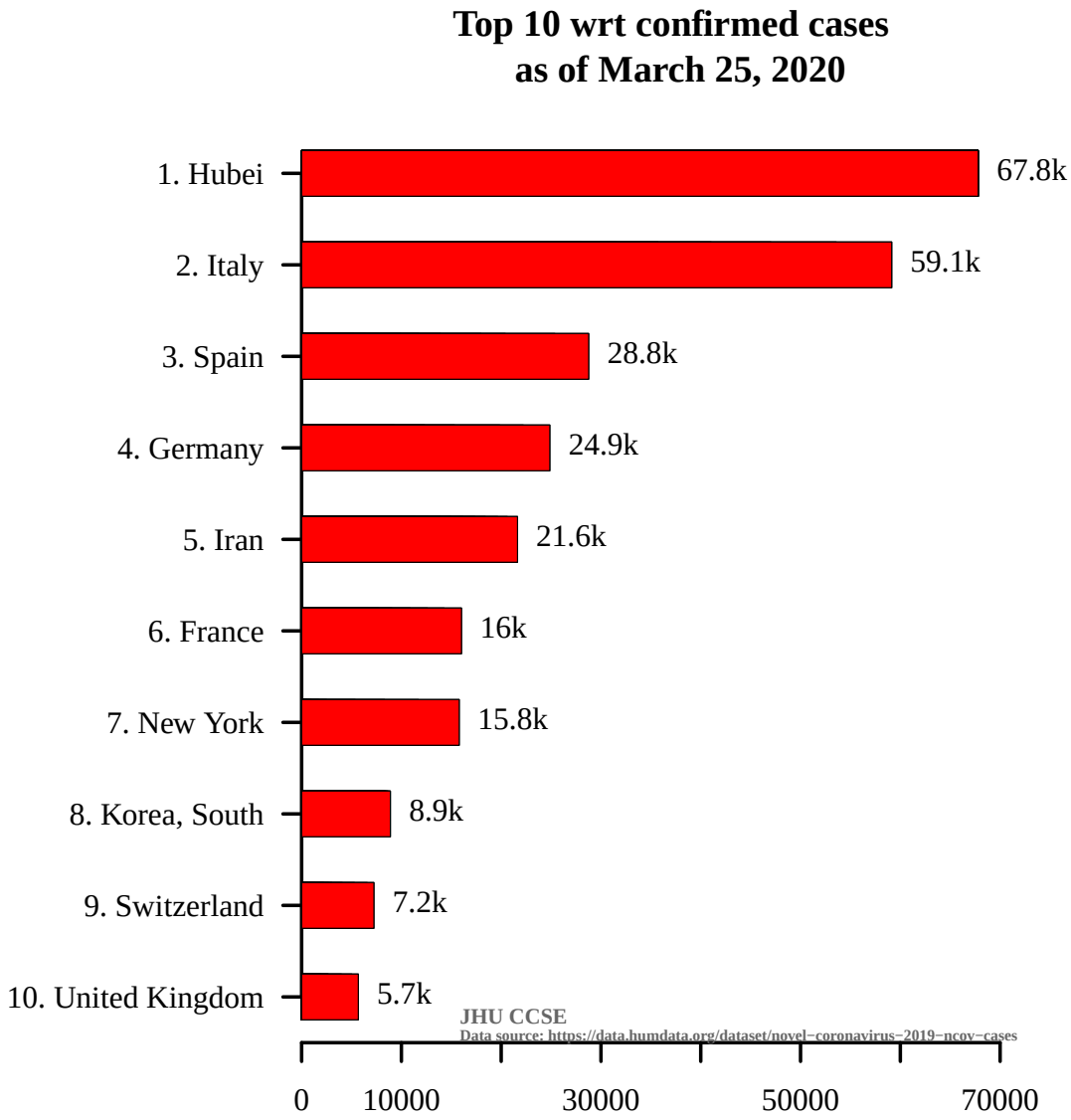


→ What do you think is interesting / important to analyze?

COVID-19 pandemic

Analyze data yourself...

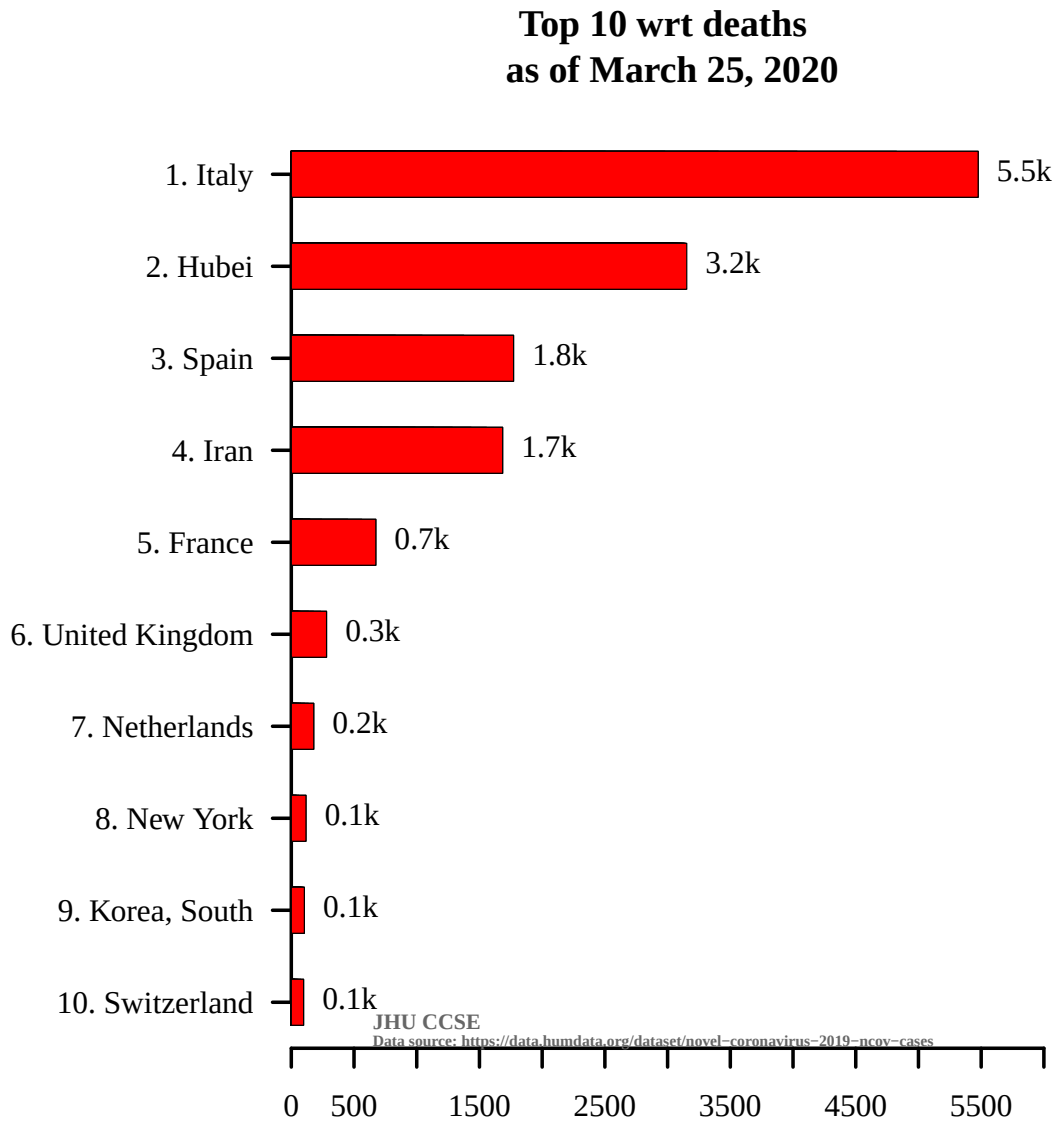
- 10 countries with largest number of confirmed cases as of March 25, 2020
- 3 continents: Asia, Europe, North America



COVID-19 pandemic

Analyze data yourself...

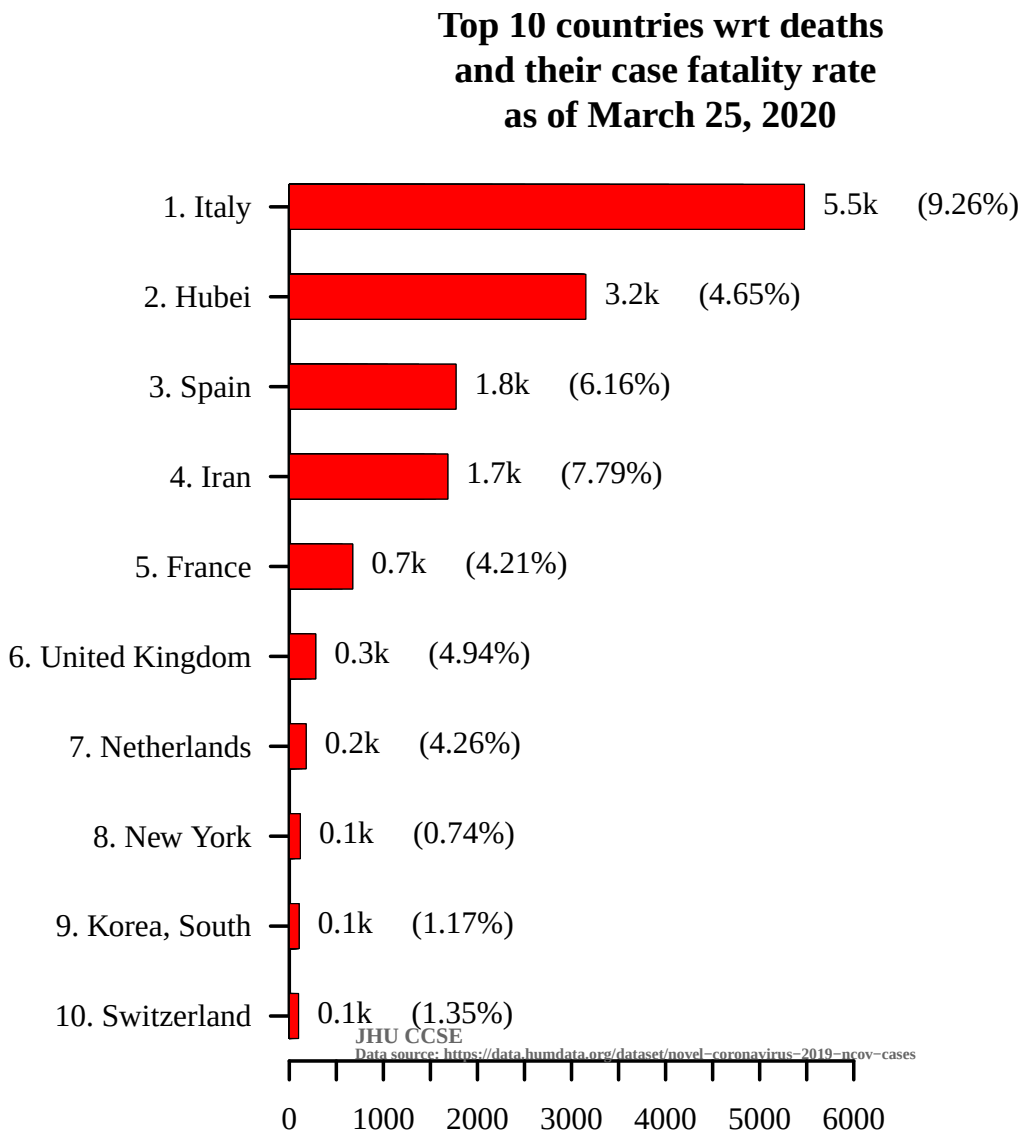
- 10 countries with largest number of reported deaths as of March 25, 2020
- 3 continents: Asia, Europe, North America
- Compared to confirmed cases:
  - Germany out, Netherlands in
  - UK moves up



COVID-19 pandemic

Analyze data yourself...

- 10 countries with largest number of reported deaths as of March 25, 2020
- Italy's CFR almost twice as large as Hubei's CFR. Why?
  - Shut down strategies
  - Health-care capacity
  - Cases may be underreported
  - Age structure of population
  - ...



What you have learned today about demographic forecasting

- Describe extensions that are available for the LC method
- Describe new directions in mortality forecasting
- List different data sources that inform about COVID-19 pandemic
- Describe progress of the COVID-19 pandemic
- Reflect on the potential usefulness of demographic forecast tool for analyzing the COVID-19 pandemic with respect to confirmed cases and reported deaths

Fourth day's class in the lab: Trends of COVID-19 pandemic & Think about how to tackle pressing issues with forecast tools

- Analyze trends of COVID-19 pandemic so far
  - Download cases and deaths from JHU CSSE website
  - Explore these data and rank the ten countries with most confirmed cases and reported deaths
- Think about how to tackle pressing issues with forecast tools
  - What are urgent empirical questions during the COVID-19 pandemic?
  - How to use and / or adjust demographic forecast tools for this?
  - What kind of data would you need for this?
  - Example: please read the paper *A demographic scaling model for estimating the total number of COVID-19 infections*, which is vailable at `https://doi.org/10.1101/2020.04.23.20077719`
- Present and discuss your findings. Put this in your report.

Course learning materials

Course learning materials on GitHub:

`https://github.com/christina-bohk-ewald/2020-course-COS-R403-forecasting-1-introduction`

Recommended learning material for today's class

- **Lee, R. D., & Carter, L. R. (1992)**
  Modeling and forecasting U.S. mortality. Journal of the American Statistical Association, 87(419), 659-671.
- **Booth, H. (2006)**
  Demographic forecasting: 1980 to 2005 in review. International Journal of Forecasting, 22(3), 547-581.
- **Booth, H., & Tickle, L. (2008)**
  Mortality modelling and forecasting: A review of methods. Annals of Actuarial Science, 3(1-2), 3-43.
- **Bohk-Ewald, C., Dudel, C., & Myrskylä, M. (2020)**
  A demographic scaling model for estimating the total number of COVID-19 infections. medRxiv.org 2020. https://doi.org/10.1101/2020.04.23.20077719.

Recommended learning material for today's class

- Johns Hopkins University CSSE. Novel Coronavirus (COVID-19) Cases Data. Published online: https://data.humdata.org/dataset/novel-coronavirus-2019-ncov-cases. Download on April 17, 2020.
- United Nations, Department of Economic and Social Affairs, Population Division. (2019). Population Prospects 2019. Published online: https://population.un.org/wpp/. Download on April 17, 2020.
- Roser M, Ritchie H, Ortiz-Ospina E (2020). Coronavirus Disease (COVID-19) – Statistics and Research. Our World in Data 2020; published online. https://ourworldindata.org/coronavirus.

Recommended learning material for today's class

- Bohk-Ewald C, Dudel C, Myrskylä M. A demographic scaling model for estimating the total number of COVID-19 infections. medRxiv.org 2020; published online April 29. DOI:https://doi.org/10.1101/2020.04.23.20077719.
- Verity R, Okell LC, Dorigatti I, et al. Estimates of the severity of coronavirus disease 2019: a model-based analysis. Lancet Infect Dis 2020; published online March 30. DOI:10.1016/S1473-3099(20)30243-7.
- Epidemiology Group of the New Coronavirus Pneumonia Emergency Response Mechanism of the Chinese Center for Disease Control and Prevention. Epidemiological characteristics of the new coronavirus pneumonia. Chinese Journal of Epidemiology 2020; 41: published online February 17. DOI: 10.3760 / cma.j.issn.0254-6450.2020.02.003.

Thank you for your attention!

christina.bohk-ewald@helsinki.fi

Fourth day's class in the lab: Trends of COVID-19 pandemic & Think about how to tackle pressing issues with forecast tools

- Analyze trends of COVID-19 pandemic so far
  - Download cases and deaths from JHU CSSE website
  - Explore these data and rank the ten countries with most confirmed cases and reported deaths
- Think about how to tackle pressing issues with forecast tools
  - What are urgent empirical questions during the COVID-19 pandemic?
  - How to use and / or adjust demographic forecast tools for this?
  - What kind of data would you need for this?
  - Example: please read the paper *A demographic scaling model for estimating the total number of COVID-19 infections*, which is vailable at `https://doi.org/10.1101/2020.04.23.20077719`
- Present and discuss your findings. Put this in your report.