COS-D407. Scientific Modeling and Model Validation

Lecturer: Christina Bohk-Ewald

Week 5

University of Helsinki, Finland
01.11.2021–15.12.2021

Fourth week's class:

Scientific modeling & model validation in practice

- Q&A: recap of material of previous session.
- Present your findings of previous lab session.
- Model validation continued. Case without true realizations.
  - Sensitivity analysis.
  - Analyze how robust the demographic scaling model's COVID-19 infection estimates are with respect to taking input data from different sources.

Fifth week's class in the lab: Robustness of demographic scaling model's results.

- Test how robust the demographic scaling model's results are with respect to taking reference infection fatality rates from different sources.
  - Original and scaled IFRs based on Chinese data provided by Verity and colleagues (2020).
  - Original and scaled IFRs based on French data provided by Salje and colleagues (2020).
  - IFRs based on universal exponential model provided by Levin and colleagues (2020).
- Critically think about the plausibility of these IFR estimates and their implications for the results of the demographic scaling model.

$\rightarrow$ Present and discuss your findings in class at the beginning of the next session on Monday.

Recap material of previous session in a brief Q & A:

- What sources are there for getting the numbers of COVID-19 infections in a country of interest?
- What are the main methodological features of the demographic scaling model? How does it estimate the total numbers of COVID-19 infections in a country of interest?
- What features are supposed to make the demographic scaling model broadly applicable?

$\rightarrow$ Open questions?

Present your findings of previous lab session:

- How large are the COVID-19 infection estimates for Finland produced with the demographic scaling model? And how large are they for another country of interest?
- What are the two key assumptions of the demographic scaling model?
- How likely is it that these two key assumption hold in real-world applications? How does this might impair the validity of the estimation results of the demographic scaling model?

$\longrightarrow$ Open issues?

Different setting or conditions in first and second year of the coronavirus pandemic

Keep in mind that the demographic scaling model estimates infections as deaths over IFRs.

1. In the first year of the pandemic, the IFRs appear to be plausible, so that the infection estimates are larger than confirmed cases.

2. In the second year of the pandemic, the IFRs appear to be too high considering the spread of less deadly COVID-19 variants and vaccination. Consequently, infection estimates are smaller than confirmed cases. $\Rightarrow$ IFR is an important input parameter that needs to be adjusted over time.

Some more thoughts on this

How accurate are the COVID-19 infection estimates
of
the demographic scaling model?

Key assumptions of demographic scaling model

1. COVID-19-related death counts are fairly accurately recorded.

2. Infection fatality rates from reference country are fairly accurately recorded and
become applicable in a country of interest through proper scaling based on remaining life expectancy.

First key assumption of demographic scaling model

COVID-19-related death counts are fairly accurately recorded.

First key assumption of demographic scaling model

Potential bias in COVID-19-related death counts:

- Reporting delays may amount to several days.
- Inconsistent practices for defining and testing COVID-19 deaths within and between countries over time. For example,
  - Only deceased individuals who (1) have been hospitalized or (2) have died from COVID-19 as primary cause of death may be counted.
  - Test coverage may be low so that not all people who have died from COVID-19 are detected and reported.
- ...

First key assumption of demographic scaling model

- If COVID-19 deaths were underreported, infection estimates would be too low, and *vice versa*.
- Correct for potential bias, B, in COVID-19-related death counts using, e.g., emerging studies on excess mortality

  $$I^B = B \cdot \sum_x \frac{D_x}{IFR_x} = B \cdot I^T$$

  in order to increase accuracy of COVID-19 infection estimates, I.

Second key assumption of demographic scaling model

Infection fatality rates from reference country are fairly accurately recorded and become applicable in a country of interest through proper scaling based on remaining life expectancy.

Second key assumption of demographic scaling model

Potential bias in IFRs in RC and scaled IFRs in COI:

- IFR in RC. Misreporting COVID-19 deaths and cases.
- IFR in RC. Misspecification in statistical model to estimate IFRs.
- IFR in COI. Scaling of IFRs from RC onto COI could be impaired if RC and COI differed considerably:
  - In the taken control measures and their acceptance in each population.
  - In their structure and distribution of major diseases that impact both vulnerability to COVID-19 and remaining lifetime.
  - In the occupancy rate of medical services.

Second key assumption of demographic scaling model

- If scaled IFRs were overestimated, infection estimates would be too low, and *vice versa*.
- Correct for potential bias, B, in scaled infection fatality rates

  $$I^B = B \cdot \sum_x \frac{D_x}{IFR_x} = B \cdot I^T$$

  in order to increase accuracy of COVID-19 infection estimates, I.

Key assumptions of demographic scaling model

The two key assumptions may only partially hold at the moment.

However, as soon as better input data will become available,
the demographic scaling model can account for them,
and its COVID-19 infection estimates
are likely to become more accurate.

Topic today

Validity & **robustness**

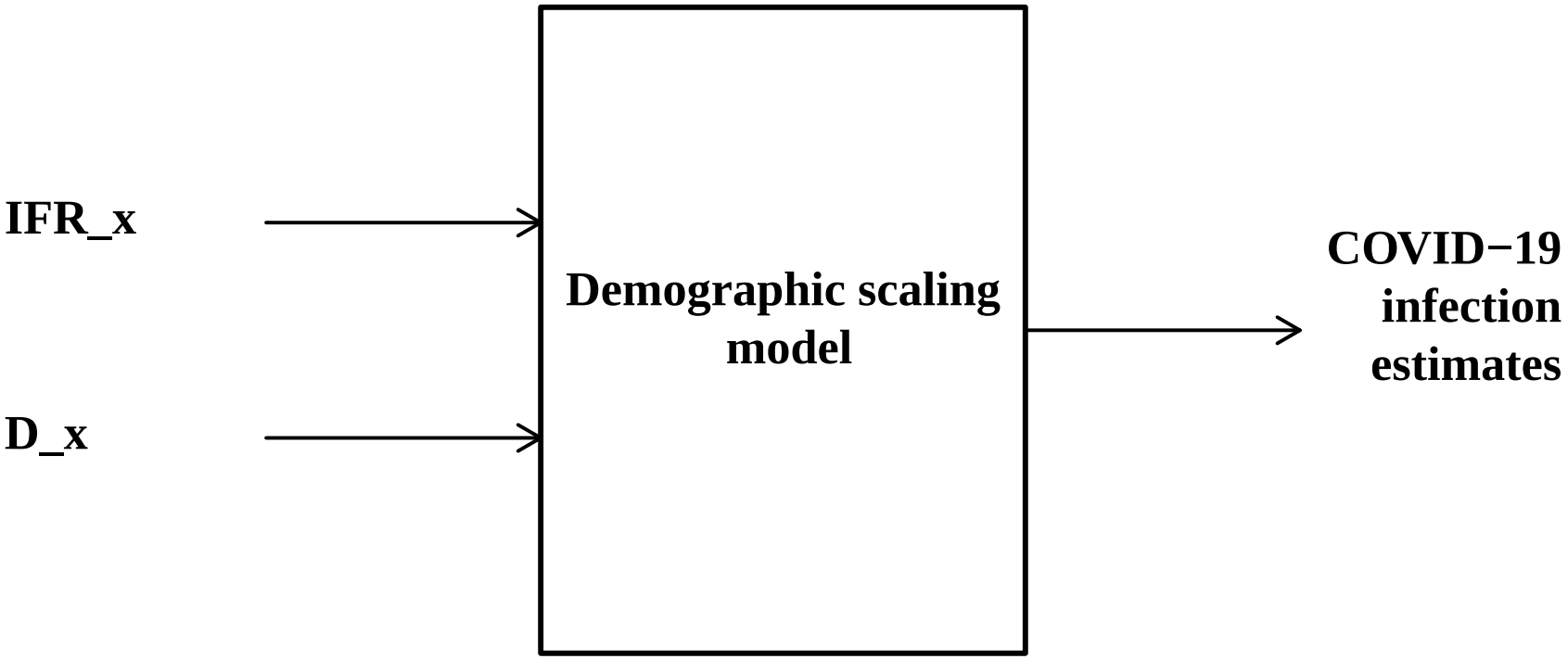
of demographic scaling model's

COVID-19 infection estimates.

$\longrightarrow$ Particularity: no *true* values have been observed

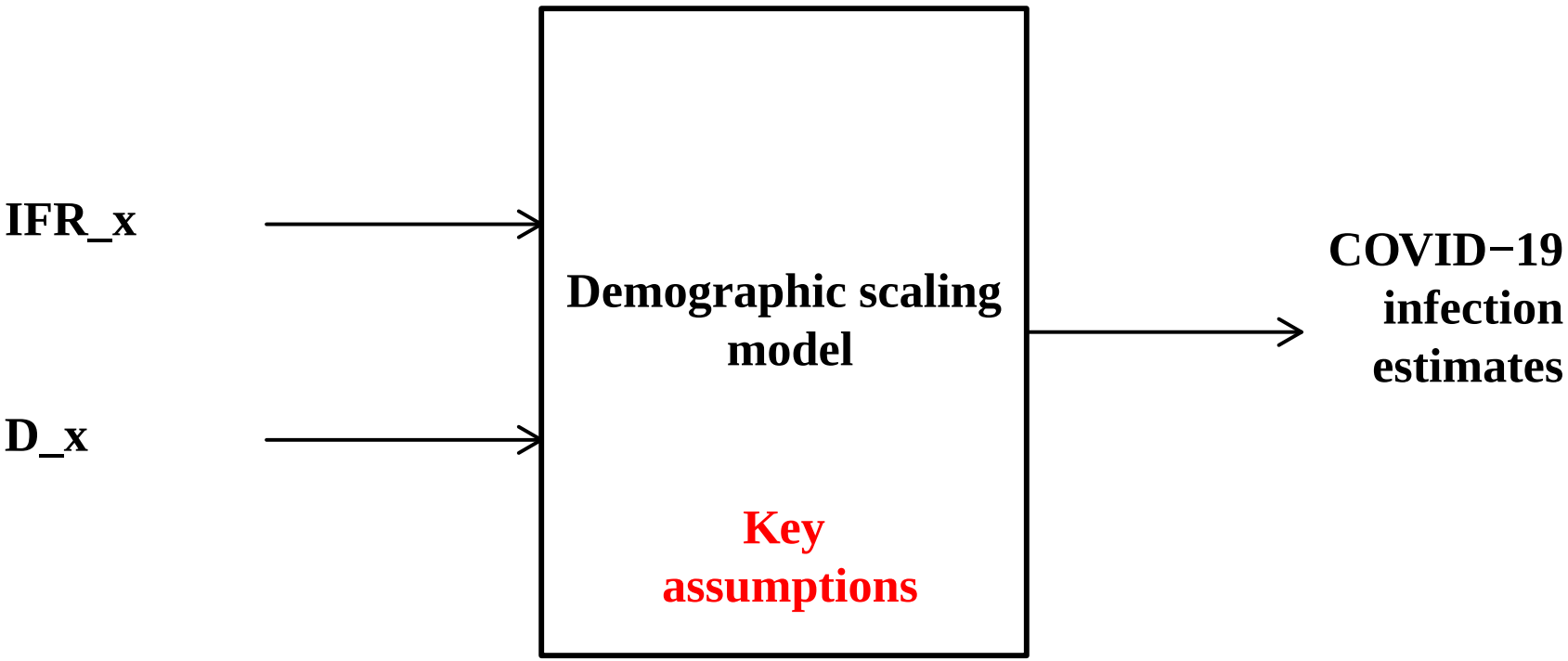
that we could compare with our model infection estimates

The model that needs to be evaluated



→ So, what could we do to evaluate this model, even though there are no *true* values available to compare model infection estimates with?

What we have evaluated already

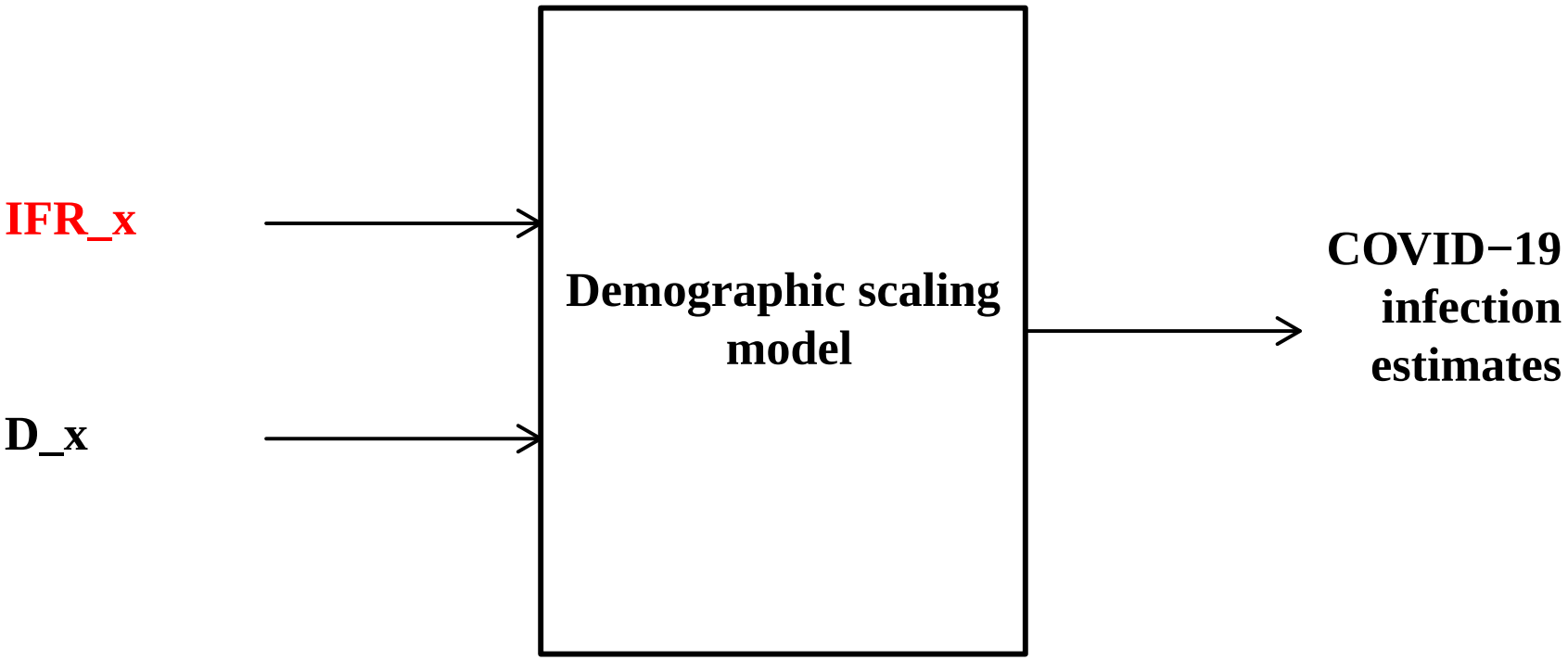


→ Think creatively and critically about the model's key assumptions and how they might impact the model's infection estimates.

→ Think of situations when a model should work well and when it should not work well; how plausible / likely are each of these situations?

Robustness of demographic scaling model's COVID-19 infection estimates

How robust are model infection estimates wrt IFR_x?



$\rightarrow$ Think creatively and critically about the model's key assumptions and how they might impact the model's infection estimates.

Robustness wrt IFR

Robustness of demographic scaling model's COVID-19 infection estimates **with respect to** IFR_x

Robustness wrt IFR

- Aim: test by how much the COVID-19 infection estimates change if we take infection fatality rates from different sources.
- New sources for IFRs become available with ongoing time:
  - Original IFRs based on Chinese data provided by Verity et al. (2020).
  - Original IFRs based on French data provided by Salje et al. (2020).
  - "Universal" IFRs from exponential model of Levin et al. (2020).
  - Scale original IFR estimates of Verity et al. (2020) and Salje et al. (2020) based on remaining life expectancy as proposed by the demographic scaling model (in order to account for cross-country differences in age structure, preconditions, and medical services)
  - ...

Robustness wrt IFR

What do you think?

What to possibly think creatively and critically about here?

Robustness wrt IFR

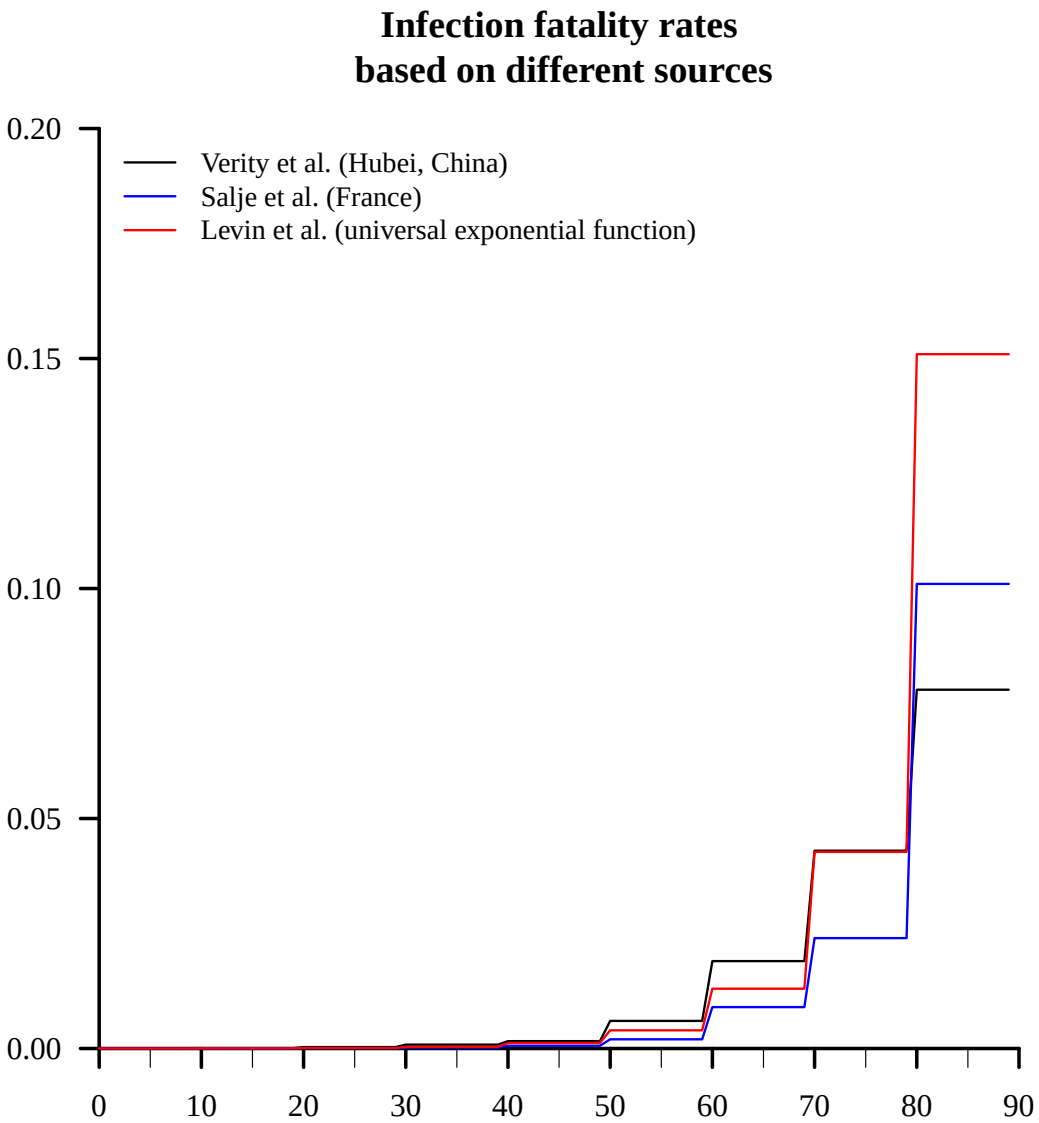
- Aim: test by how much the COVID-19 infection estimates change if we take infection fatality rates from different sources.
- New sources for IFRs become available with ongoing time.
- Think creatively and critically about the plausibility of these IFR estimates and their potential implications for the results of the demographic scaling model.
  - How much do these IFR estimates differ within and between countries?
  - How much do these IFR estimates apply to each country?
  - How much do the COVID-19 infection estimates differ with respect to these IFR estimates?
  - How much does this robustness with respect to these IFR estimates change by country?
  - ...

Robustness wrt IFR

Let us start

and have a look at these different IFR estimates

Original IFR estimates of different sources



Original IFR estimates of different sources

- Infection fatality rate increases with age
- IFR in old age are largest according to estimates of Levin et al.
- Crossover in IFR estimates of Verity et al. and Salje et al.:
  - At younger ages: Chinese IFR larger than French IFR
  - At older ages: Chinese IFR smaller than French IFR
- Similar crossover pattern for IFR of Verity et al. and Levin et al.
- IFR of Levin et al. is consistently larger than IFR of Salje et al.

IFR estimates of Verity et al.

- Based on data of Wuhan city, Hubei province, mainland China and the cruise ship Diamond Princess until early February / March 2020
- Bayesian model based on, e.g., proportion of deaths of passengers tested positive for COVID-19 and time-to-death distribution derived from cases and deaths related to COVID-19 of mainland China
- Conducted several robustness checks and accounted for demographic age structure
- Verity et al. (2020)
  Estimates of the severity of coronavirus disease 2019: a model-based analysis. Lancet Infect Dis. DOI:10.1016/S1473-3099(20)30243-7

IFR estimates of Salje et al.

- Based on data of France and the cruise ship Diamond Princess until early May 2020.
- Models based on *passive surveillance*: real time hospital admissions and deaths related to COVID-19 in France as well as *active surveillance*: as performed for the passengers of the cruise ship Diamond Princess.
- Salje et al. (2020)
  Estimating the burden of SARS-CoV-2 in France.
  Science(eabc3517). DOI: 10.1126/science.abc3517

IFR estimates of Levin et al.

- Based on data of 28 locations
  - Representative samples (England, Ireland, Italy, Netherlands, Portugal, Spain, Geneva, Atlanta, Indiana, New York, Salt Lake City)
  - Convenience samples (Belgium, France, Sweden, Connecticut, Louisiana, Miami, Minneapolis, Missouri, Philadelphia, San Francisco, Seattl)
  - Comprehensive tracing programs (Australia, Iceland, Korea, Lithuania, New Zealand)

- Exponential relationship between age and the IFR in %:
  $IFR = e^{-7.53+0.119 \cdot age}$

- Levin et al. (2020)
  Assessing the Age Specificity of Infection Fatality Rates for COVID-19: Systematic Review, Meta-Analysis, and Public Policy Implications.
  medRxiv 2020; published online July 24
  `https://doi.org/10.1101/2020.07.23.20160895`

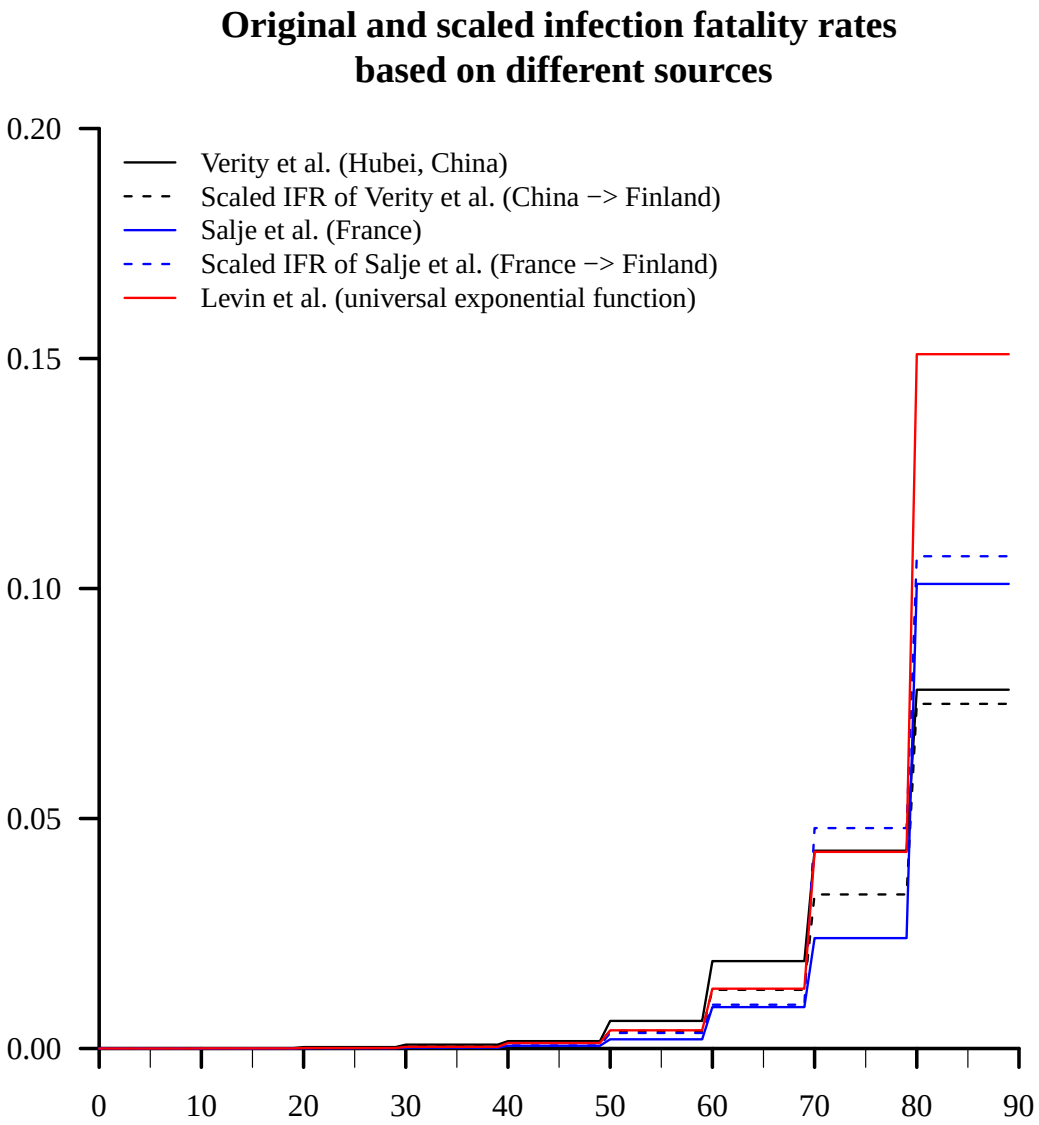
Additional IFR estimates

For the sake of completeness:

- John Ioannidis 2020:
  The infection fatality rate of COVID-19 inferred from seroprevalence data
  medRxiv 2020; publishd online May 19
  https://doi.org/10.1101/2020.05.13.20101253

- Gideon Meyerowitz-Katz and Lea Merone 2020:
  A systematic review and meta-analysis of published research data on COVID-19 infection-fatality rates
  medRxiv 2020; published online May 6
  https://doi.org/10.1101/2020.05.03.20089854

Scaled IFR estimates for Finland



Scaled IFR estimates for Finland

- Scaled IFR fo Finland is consistently smaller than the original IFR for China of Verity et al.
- Scaled IFR fo Finland is consistently larger than the original IFR for France of Salje et al.
- Biggest differences between these original and scaled IFRs:
  - France $\longrightarrow$ Finland: age groups $70-79$ and $80+$
  - China $\longrightarrow$ Finland: age groups $60-69$ and $70-79$

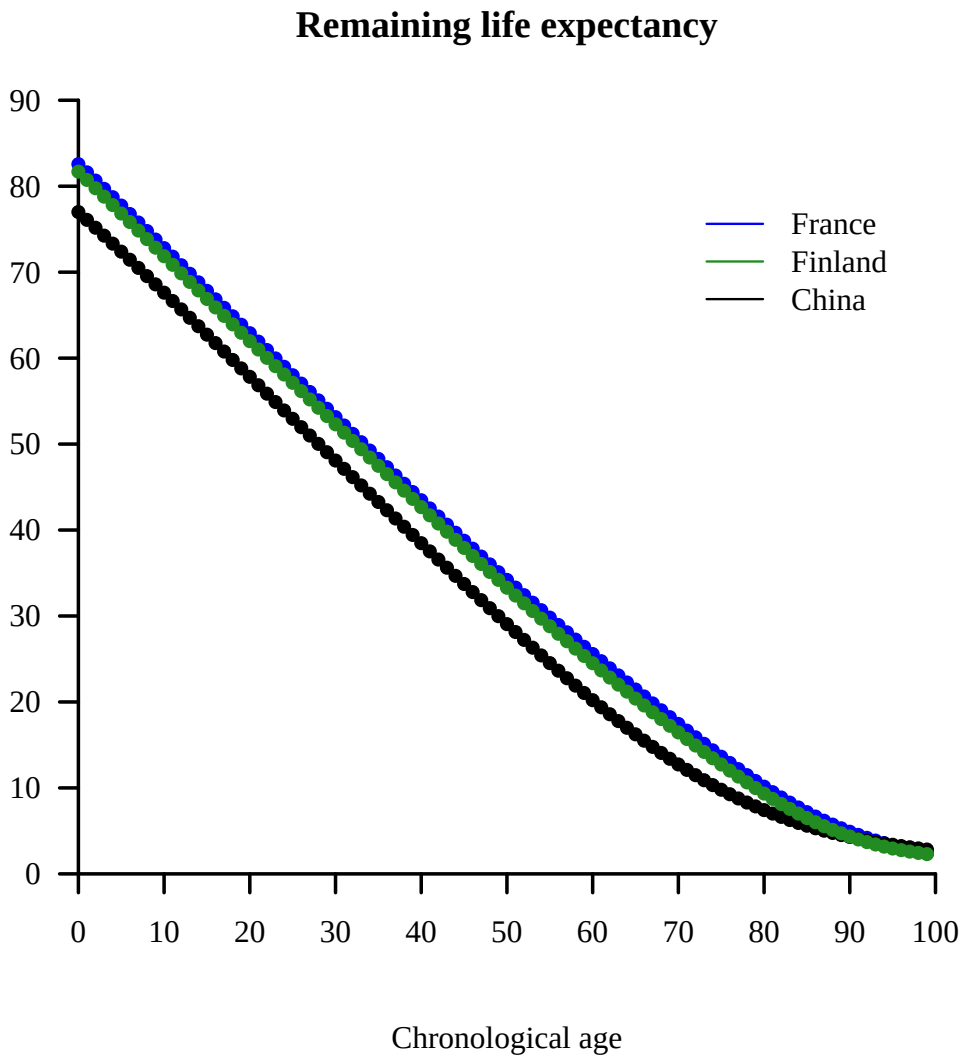
Scaled IFR estimates for Finland

Why are the scaled IFRs for Finland

larger than the original IFRs in France

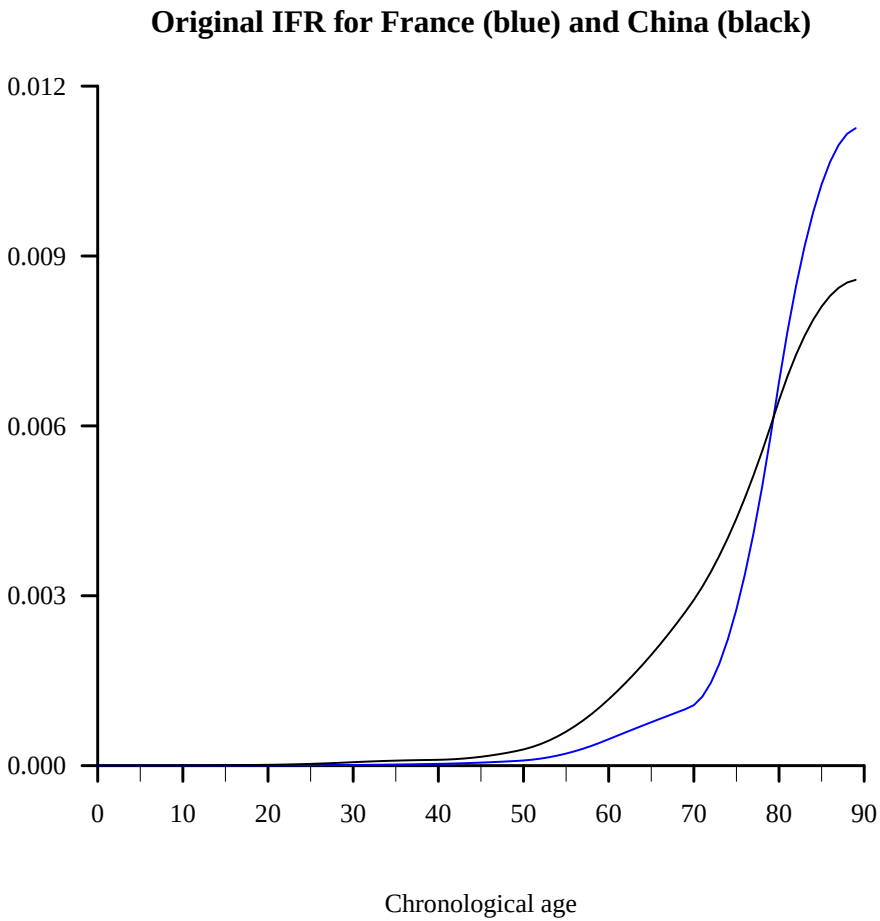
and smaller than the original IFRs in China?

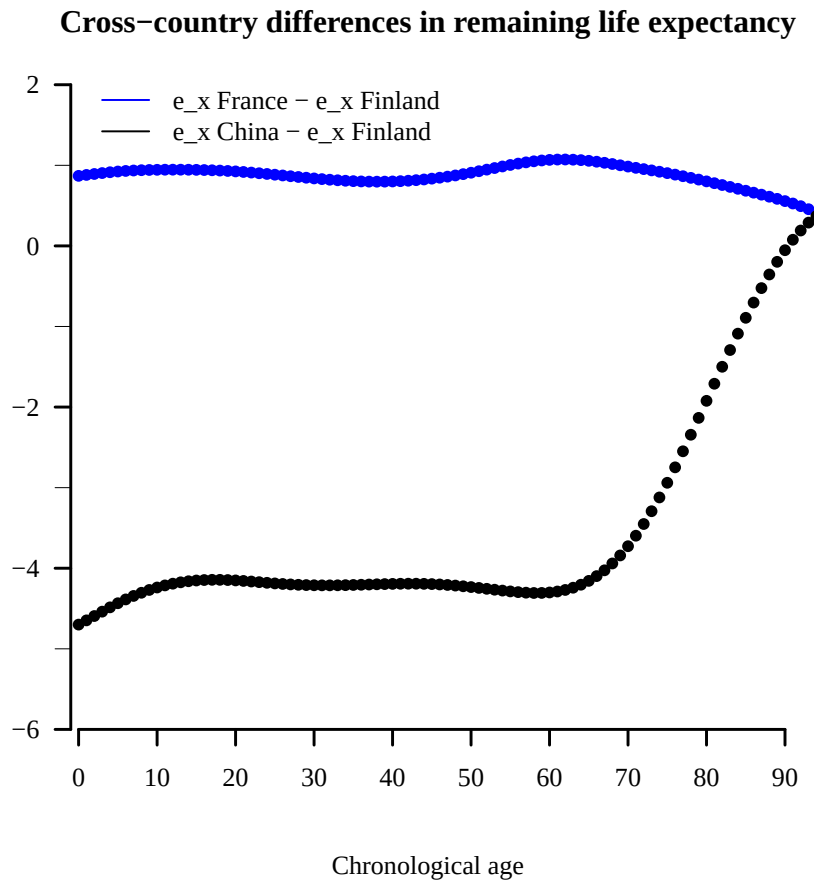
Explain scaled IFR estimates for Finland



$\rightarrow$ Need to go a little deeper in order to understand what is going on

Explain scaled IFR estimates for Finland





Scaled IFR estimates for Finland

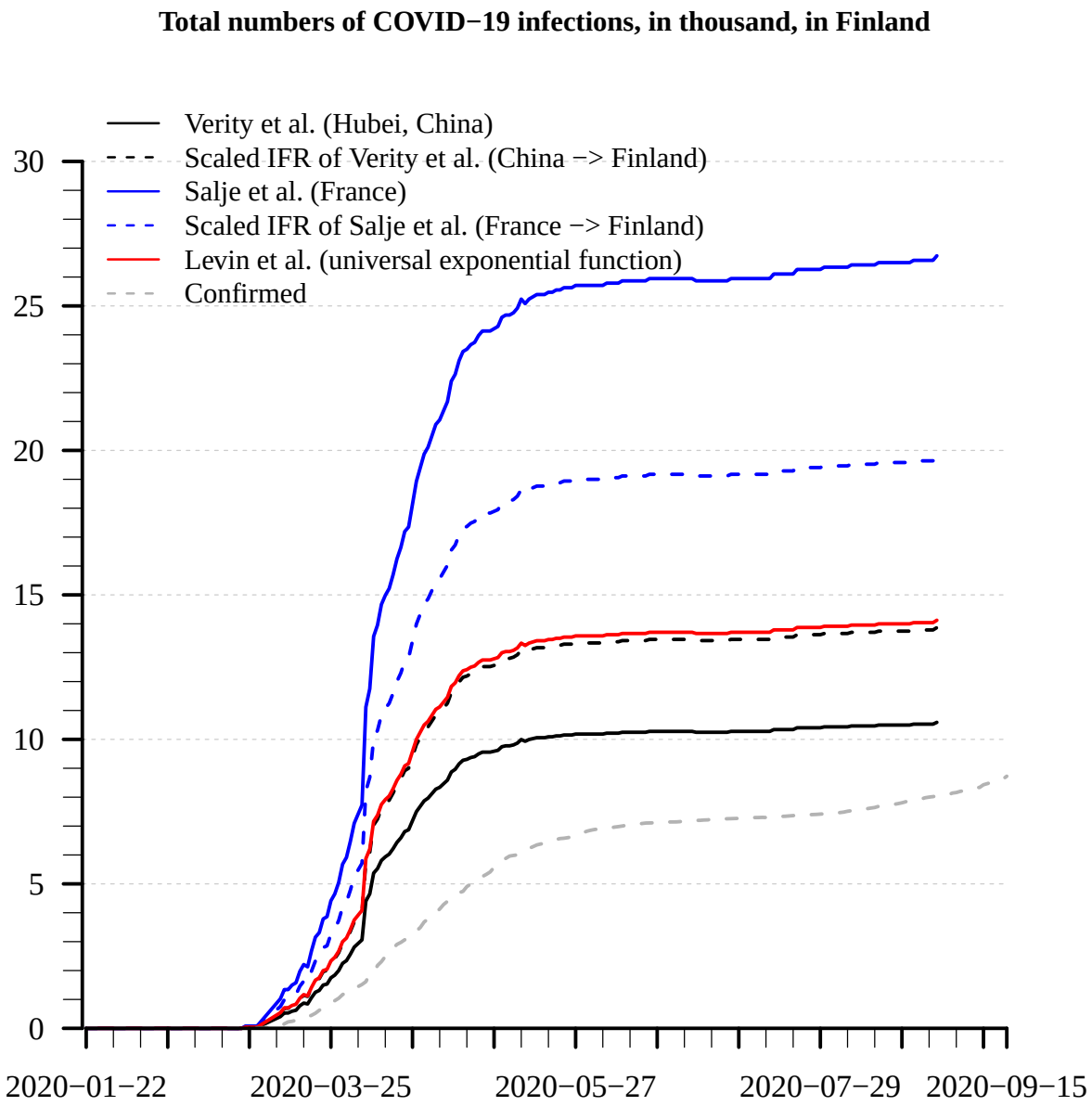
- China: $e_x^{FIN} > e_x^{RC} \rightarrow$ Chinese IFR_x will be assigned to older chronological ages in Finland $\Rightarrow$ Chinese IFR are larger than scaled IFR for Finland
- France: $e_x^{FIN} < e_x^{RC} \rightarrow$ French IFR_x will be assigned to younger chronological ages in Finland $\Rightarrow$ French IFR are smaller than scaled IFR for Finland
- Biggest differences at particular age groups:
  - Original IFR of Verity et al. (2020) increases substantially at ages 60+
  - Original IFR of Salje et al. (2020) increases substantially at ages 70+

Robustness wrt IFR

Let us continue with the impact of these different IFR estimates

on the total number of COVID-19 infection estimates

Robustness wrt IFR



Robustness wrt IFR

The example of Finland from January 2020 until September 2020 shows:

- IFR estimate appears to have big impact on estimations of total numbers of COVID-19 infections of the demographic scaling model
- All COVID-19 infection estimates are consistently larger than the numbers of confirmed cases
- Range of COVID-19 infection estimates narrows down if we consider scaled infection fatality rates for Finland (and the IFR of Levin et al.)
- COVID-19 infection estimates appear to be largest when based on French IFR (original and scaled for Finland)

Robustness wrt IFR

Think creatively and critically ...and extend this analysis:

- How much do the COVID-19 infection estimates differ with respect to this set of IFR estimates over time?
- How much do the COVID-19 infection estimates differ with respect to this set of IFR estimates?
- How much do these IFR estimates apply to each country?
- How much does the robustness of the COVID-19 infection estimates change across countries with respect to this set of IFR estimates?
- How much do these (original and scaled) IFR estimates differ within and between countries?
- ...

What you have learned today about assessing the demographic scaling model

- Explain why the two key assumptions of the demographic scaling model may only partially hold.
- Explain how potential bias in deaths and IFR may bias the COVID-19 infection estimates of the demographic scaling model.
- Describe a way to analyze the robustness of the COVID-19 infection estimates with respect to infection fatality rates.
- Describe different sources of IFR estimates.
- Describe the impact of IFR estimates on the COVID-19 infection estimates for Finland.

Course learning materials

Course learning materials on GitHub:

`https://github.com/christina-bohk-ewald/2021-COS-D407-scientific-modeling-and-model-validation`

Recommended learning material for today's class

- **Verity et al. (2020)**
  Estimates of the severity of coronavirus disease 2019: a model-based analysis.
  Lancet Infect Dis
  DOI:10.1016/S1473-3099(20)30243-7
- **Salje et al. (2020)**
  Estimating the burden of SARS-CoV-2 in France.
  Science(eabc3517)
  DOI: 10.1126/science.abc3517
- **Levin et al. (2020)**
  Assessing the Age Specificity of Infection Fatality Rates for COVID-19: Systematic Review, Meta-Analysis, and Public Policy Implications.
  Published as preprint on medRxiv
  `https://www.medrxiv.org/content/10.1101/2020.07.23.20160895v5`

Thank you for your attention!

christina.bohk-ewald@helsinki.fi

Fifth week's class in the lab: Robustness of demographic scaling model's results.

- Test how robust the demographic scaling model's results for Finland are with respect to taking reference infection fatality rates from different sources.
  - Original and scaled IFRs based on Chinese data provided by Verity and colleagues (2020).
  - Original and scaled IFRs based on French data provided by Salje and colleagues (2020).
  - IFRs based on universal exponential model provided by Levin and colleagues (2020).
- Critically think about the plausibility of these IFR estimates and their implications for the results of the demographic scaling model.

$\rightarrow$ Present and discuss your findings in class at the beginning of the next session on Monday.