Institut für Informatik
Praktische Informatik und Bioinformatik
Prof. Dr. Ralf Zimmer
Dr. Evi Berchtold, Markus Joppich, Felix Offensperger, Jewgenija Zhdanovich

# Corona/Online Winter Term 2020/21
# Computational Systems Biology

## Assignments 2020-1add
## Modeling the COVID-19 Pandemics

**Working period**: Three weeks (24.11.-15.12.2020)
Hand-in anytime or in any exercise class
Please hand-in only reproducible results, answers, figures, tables, simulations, ...

This project analyses the current COVID-19[1] pandemics, a world-wide crisis caused by a new corona virus, the SARS-CoV-2[2]. It investigates some basic dynamical systems models and the employed methods. It also introduces some technical issues and frameworks for visualizing results and making models available for parametrization and simulations.

The goal of this first class of the summer term 2020 is to understand as much as possible about an ongoing and pressing disease from a bioinformatics/system biology perspective. The class should also enable to better understand the information, myths, and fakes about the disease. It also should set the stage for a already quite broad range of techniques (from text mining, knowledge extraction, to network reconstruction, high-throughput analysis, and systems simulation) in current bioinformatics/systems biology research, which will be of use in other research projects. Thereby, the COVID-19 analysis introduces already quite some concepts which we will discuss in more detail during the course. We also setup a platform for implementing and communicating current and forthcoming programming tasks.

May 5, 2020 and Nov 1, 2020 updates on the Corona Pandemics. Meanwhile you can review the predictions of papers and models post factum. But, before being to critical, note that the second wave is just under way and, again, predictions (for the future) are being made, which might have drastical consequences for further lockdown with major individual restrictions and econonomic impacts.

### Task 1 The Heinsberg study

Today, finally, the final results from the Heinsberg study (Streeck, et al, 5.5.2020, Preprint) has been published (on a Preprint Server). The Heinsberg study has been the basis of political decisions in

[1] COVID-19 = COronaVIrus Disease 2019

[2] SARS-CoV-2 = severe acute respiratory syndrome coronavirus 2, "Schweres akutes Atemwegssyndrom Coronavirus"; vormals 2019-nCoV, 2019-novel Corona virus, neuartiges Coronavirus 2019 sowie Wuhan-Coronavirus

NRW and target of quite some critics. Now the study is available. Its main findings are a much higher ratio of infected (and recovered), i.e. probably immune, people (about 15%) and, therefore, a much lower fatality rate of 0.37%.

(a) Summarize and discuss the study and its findings. Can you comment on the methodology (study setup and statistics)? Can the results be generalized to NRW or to Germany or to the worldwide situation?

(b) Several similar studies have been made available, e.g. for Brooklyn or Manhattan, which also report quite high ratios of immune peoples. But there are other studies, which reports and estimates much lower ratios. Compare these and identify reasons, why the results are that different!

(c) Does the study help with estimating, whether the capacity of the health system (i.e. number of ICUs) will be exceeded?

Another type of study have been performed by systems biologist Prof. Edda Klipp of Humboldt University, Berlin. Review the paper (Manuscript by Goldenburg et al & Edda Klipp, Geospatial precision simulations of community confined human interactions during SARS-CoV-2 transmission reveals bimodal intervention outcomes, available on medRxiv (May 27, 2020), https://doi.org/10.1101/2020.05.03.20089235) and explain the differences (in the approach perfomed) and the results obtained? What are the major findings of that paper?

**Task 2 (Tests)**

(a) Explain the difference between"Corona-Tests"! What are the major types and what do they measure?

(b) there are a number of immunity studies in Germany, e.g. in Fulda, Munich and also a Germany-wide study. Give an overview of these studies, preliminary results and when we can expect reliable results from these studies.

(c) The used tests are not perfect both with respect to sensitivity and specificity: review the quality of the various tests and include the new one just announced yesterday by Roche, Penzberg.

(d) Compare with Antibody- and PCR-based tests with respect to their respective sensitivities and specificities.

(e) What about test capacities? How many tests might be needed for treasonable coverage? How is the relation (over time) of actual tests and test capacity?

**Task 3 (Reproduction number $R$, $R_0$)**

(a) Define the reproduction number(s) and explain the importance for the pandemic.

(b) how are these numbers estimated? How large are difference between various estimates? What are reasons for these differences?

(c) Reproduction numbers evolve over time. What is the correlation of the r(t) estimates between different sources? How large are the errors and the deviations? Are these errors relevant for the models and their predictions?

**Task 4 (CDC models of COVID-19 epidemics in the US)**

On Apr 28, 2020, the New York Times (NYT) reports on an CDC (Centers of Disease Control) study which predicts a drastic increase of infections and deaths in the US during May and June 2020 ("U.S. Coronavirus Death Toll Is Far Higher Than Reported, C.D.C. Data Suggests").

(a) What is this study based on? What are the assumptions behind these predictions? Are the predictions realistic and is there reason to panic?

(b) The NYT (and other media) also provide a comparative assessment of several epidemiological studies with quite different predictions ("What 5 Coronavirus Models Say the Next Month Will Look Like", NYT, Apr 22, 2020). Assess these studies and the comparison, interpret! What can be learned form the models and their comparison, respectively?