COVID-19: Estimation of the transmission dynamics in Spain using a stochastic simulator and Black-Box optimization techniques

Marcos Matabuena

CiTIUS (Centro Singular de Investigación en Tecnoloxías Intelixentes)
Universidad de Santiago de Compostela

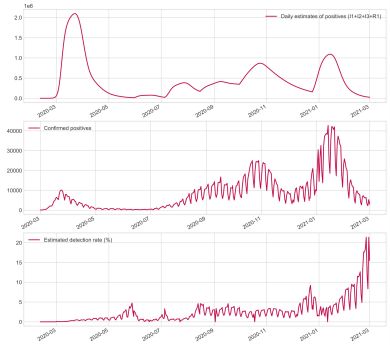
marcos.matabuena@usc.es (@MarcosMatabuena)

University of Erlangen–Nuremberg (01/10/2021)

Talk outline

1. Motivation + summary (of analysis & modeling strategy).
2. General overview of epidemic modeling strategies.
3. Mathematical details of our proposal.
4. Results.
5. Some insights about the new personalized medicine era.

Motivation



Our modeling goals are two-fold:

1 Develop a mathematical tool to estimate and monitoring the current epidemic state, e.g., seroprevalence estimations.

1. Reconstruct the latent dynamic of COVID-19 infections in Spain with the aims:

   2.1 Understanding the past to improve the future.

   2.2 Create a formal tool to guide personalized epidemiological interventions, e.g., local lown-downs.

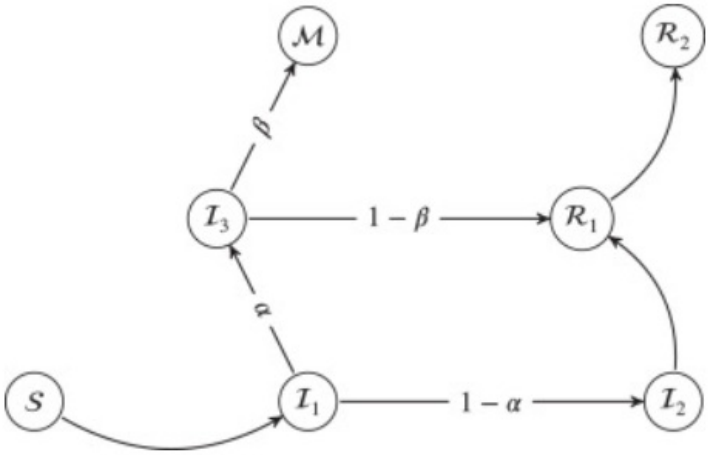
Analysis goals

Answer the following epidemiological questions:

- What was the **spread of the virus in the first wave** in different regions of Spain like?. For example, when did the **peak of infections occur**?. How many **infected people** were there in Spain at the **end of the lockdown policies**?.
- Using a longer time frame, **until March 1, 2021, how were the overall dynamics of SARS-CoV-2 in the Spanish population** as a whole?. For example, the **healthcare situation was critical in October of 2020**, and there were discussions about applying a national lockdown; What could be the **real epidemiological situation at that time**?.
- Given that, from a theoretical point of view, we can reconstruct the dynamics of infections with our model, **how was the actual day-to-day capacity to detect new cases in Spain**?.

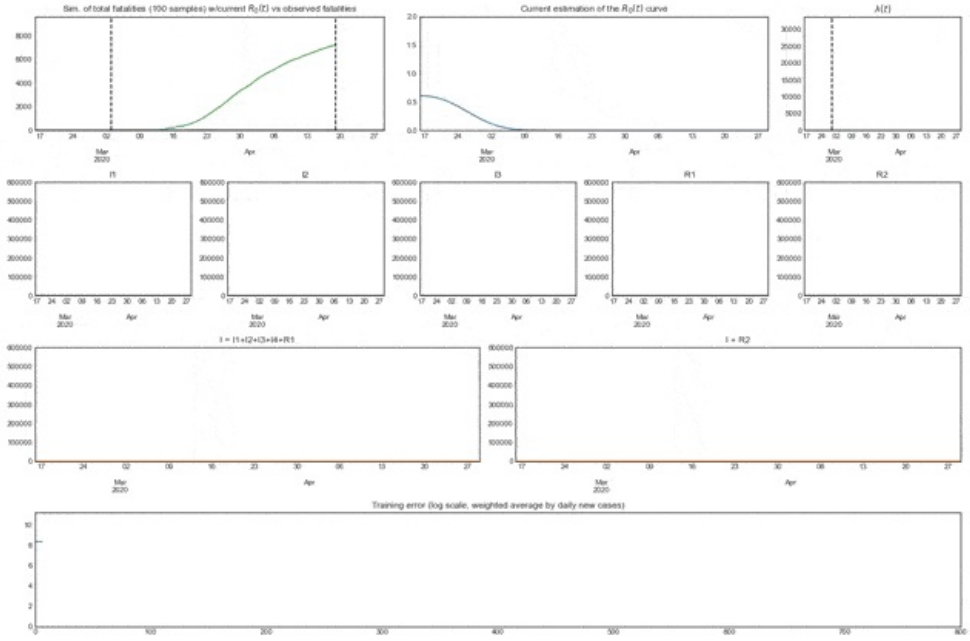
Summary of modeling strategy

- Inverse problem approach using mortality records.
- We assume that the dynamic of infections following the next probabilistic model (Markov chain model that incorporates dynamic non-homogeneous Poisson process)



Each patient belong to states $\mathcal{I}_1, \mathcal{I}_2, \mathcal{I}_3, \mathcal{R}_1$, have an individual probability of infections $\mathcal{R}_i(t) \sim Poisson(m(t))$. $m(t)$ is a parametric function, e.g. Inverse logistic function.

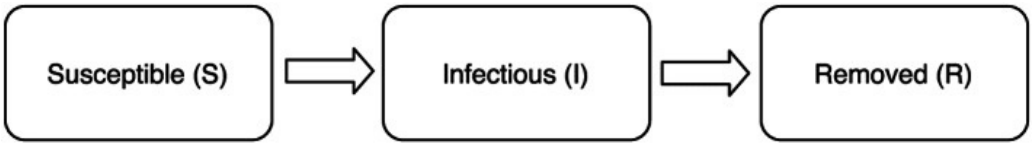
Summary of modelling strategy



General overview of epidemic modeling strategies

- Mechanism models.
  - Compartmental models: Break down the population into a number of discrete “compartments".
  - Agent-based models: individual-based model.
- Phenomenological models:
  - Direct regression models.
  - Time series models.

Mechanism models



2.1. Specification of the Susceptible–Infectious–Removed Model

We use $S(t)$, $I(t)$ and $R(t)$ to denote the time-course subpopulation sizes (i.e. the number of individuals) distributed into each of the three compartments at a given time t, where t is continuous. Clearly, $S(t)+I(t)+R(t)=N, t \geq 0$, where N is the total population size, which is a fixed constant. The starting time is denoted as $t=0$. The rates of change among these subpopulations are represented by a system of ODEs:

$$\begin{aligned} \frac{dS(t)}{dt} &= -\beta \frac{S(t)I(t)}{N}, \\ \frac{dI(t)}{dt} &= \beta \frac{S(t)I(t)}{N} - \gamma I(t), \\ \frac{dR(t)}{dt} &= \gamma I(t), \end{aligned} \tag{1}$$

with $\beta \geq 0$ and $\gamma \geq 0$ and initial conditions $S(0)>0$, $I(0)>0$, $R(0) \geq 0$ and $S(0)+I(0)+R(0)=N$. Because at a given time t, the constraint $S(t)+I(t)+R(t)=N$ implies $dS(t)/dt+dI(t)/dt+dR(t)/dt=0$, which is satisfied by the SIR in Equation (1), these three ODEs define a dynamic system of three deterministic functional trajectories over time, including the susceptible trajectory $S(t)$, the infectious trajectory $I(t)$ and the recovered trajectory $R(t)$ for $t \geq 0$. This SIR dynamic system is well posed in the sense that non-negative initial conditions lead to non-negative solutions of the three functional trajectories. These trajectories collectively demonstrate the evolutionary mechanism of an infectious disease.

Mechanism models

Pyfectious: An individual-level simulator to discover optimal containment polices for epidemic diseases

Arash Mehrjou[*]
Max Planck Institute for Intelligent Systems
Tübingen, Germany &
ETH Zürich, Zürich, Switzerland
amehrjou@ethz.ch

Ashkan Soleymani[*]
Max Planck Institute for Intelligent Systems
Tübingen, Germany
ashkan.soleymani@tuebingen.mpg.de

Amin Abyaneh
Max Planck Institute for Intelligent Systems
Tübingen, Germany
amin.abyaneh@tuebingen.mpg.de

Samir Bhatt
Faculty of Medicine, School of Public Health
Imperial College
London, UK
s.bhatt@imperial.ac.uk

Bernhard Schölkopf
Max Planck Institute for Intelligent Systems
Tübingen, Germany
bs@tuebingen.mpg.de

Stefan Bauer
Max Planck Institute for Intelligent Systems &
CIFAR Azrieli Global Scholar
Tübingen, Germany
stefan.bauer@tuebingen.mpg.de

Abstract

Simulating the spread of infectious diseases in human communities is critical for predicting the trajectory of an epidemic and verifying various policies to control the devastating impacts of the outbreak. Many existing simulators are based on compartment models that divide people into a few subsets and simulate the dynamics among those subsets using hypothesized differential equations. However, these models lack the requisite granularity to study the effect of intelligent policies that influence every individual in a particular way. In this work, we introduce a simulator software capable of modeling a population structure and controlling the disease's propagation at an individualistic level. In order to estimate the confidence of the conclusions drawn from the simulator, we employ a comprehensive probabilistic approach where the entire population is constructed as a hierarchical random variable. This approach makes the inferred conclusions more robust against sampling artifacts and gives confidence bounds for decisions based on the simulation results. To showcase potential applications, the simulator parameters are set based on the formal statistics of the COVID-19 pandemic, and the outcome of a wide range of control measures is investigated. Furthermore, the simulator is used as the environment of a reinforcement learning problem to find the optimal policies to control the pandemic. The obtained experimental results indicate the simulator's adaptability and capacity in making sound predictions and a successful

Assaying Large-scale Testing Models to Interpret COVID-19 Case Numbers.

Michel Besserve[1,2,*], Simon Buchholz[1] and Bernhard Schölkopf[1]

[1] Max Planck Institute for Intelligent Systems, Tübingen, Germany.
[2] Max Planck Institute for Biological Cybernetics, Tübingen, Germany.
[*] Correspondence to: michel.besserve@tuebingen.mpg.de

Abstract

Large-scale testing is considered key to assess the state of the current COVID-19 pandemic. Yet, the link between the reported case numbers and the true state of the pandemic remains elusive. We develop mathematical models based on competing hypotheses regarding this link, thereby providing different prevalence estimates based on case numbers, and validate them by predicting SARS-CoV-2-attributed death rate trajectories. Assuming that individuals were tested based solely on a predefined risk of being infectious implies the absolute case numbers reflect the prevalence, but turned out to be a poor predictor, consistently overestimating growth rates at the beginning of two COVID-19 epidemic waves. In contrast, assuming that testing capacity is fully exploited performs better. This leads to using the percent-positive rate as a more robust indicator of epidemic dynamics, however we find it is subject to a saturation phenomenon that needs to be accounted for as the number of tests becomes larger.

Phenomenological models

Nonmechanistic forecasts of seasonal influenza with iterative one-week-ahead distributions
Logan C. Brooks, David C. Farrow, Sangwon Hyun, Ryan J. Tibshirani, Roni Rosenfeld

Delta density method

Consider the task of estimating the density function $f_{Y_{t+1..T}|Y_{1..t}}$ using an instance-based approach. Kernel density estimation and kernel regression use smoothing kernels to produce flexible estimates of the density of a random variable (e.g., $f_{Y_{t+1..T}}$) and the conditional expectation of one random variable given the value of another (e.g., $\mathbb{E}[Y_{t+1..T} \mid Y_{1..t}]$), respectively; we can combine these two methods to obtain estimates of the conditional density of one random variable given another. One possible approach would be to use the straightforward estimate

$$\hat{f}_{Y_{t+1..T}|Y_{1..t}}(y_{t+1..T} \mid y_{1..t}) = \frac{\sum_{s=1}^{S} I^{1..t}(y_{1..t}, Y^s_{1..t}) O^{t+1..T}(y_{t+1..T}, Y^s_{t+1..T})}{\sum_{s=1}^{S} I^{1..t}(y_{1..t}, Y^s_{1..t})},$$

where $\{1..S\}$ is the set of fully observed historical training seasons, and $I^{1..t}$ and $O^{t+1..T}$ are smoothing kernels describing similarity between "input" trajectories and between "output" trajectories, respectively. However, while basic kernel smoothing methods can excel in low-dimensional settings, their performance scales very poorly with growing dimensionality. During most of the season, neither $Y_{1..t}$ nor $Y_{t+1..T}$ is low-dimensional, and the current season's observations are extremely unlikely to closely match any past $Y^s_{1..t}$ or $Y^s_{t+1..T}$. This, in turn, can lead to kernel density estimates for $Y_{t+1..T}$ based almost entirely on the single season s with the closest $Y^s_{1..t}$ when conditioning on $Y_{1..t}$, and excessively narrow density estimates for $Y_{t+1..T}$ even without conditioning on $Y_{1..t}$. So, instead of applying kernel density estimation directly, we first break the task down into a sequence of low-dimensional sub-tasks. We avoid the high-dimensional output problem by chaining together estimates of $f_{\Delta Y_u|Y_{1..u-1}}$ for each u from $t+1$ to T, where $\Delta Y_u = Y_u - Y_{u-1}$; estimating these single-dimensional densities requires relatively little data. However, this reformulation exacerbates the high-dimensional input problem since we are conditioning on $Y_{1..u-1}$, which can be considerably longer than $Y_{1..t}$. We address the high-dimensional input problem by approximating $f_{\Delta Y_u|Y_{1..u-1}}$ with $f_{\Delta Y_u|\mathbf{R}^u}$, where $\mathbf{R}^u$ is some low-dimensional vector of features derived from $Y_{1..u-1}$. Smoothing kernel methods are used to approximate the conditional density functions using data from past seasons.

We use two sets of choices for the approximate conditional density function and summary features to form two versions of the method.

- **Markovian delta density**: approximates the conditional density of ΔY_u given $Y_{1..u-1}$ with its conditional density given just the previous (real or simulated) observation, Y_u:

$$\begin{aligned}
\hat{f}_{Y_{t+1..T}|Y_{1..t}}(y_{t+1..T} \mid y_{1..t}) &= \prod_{u=t+1}^{T} \hat{f}_{\Delta Y_u|Y_{1..u-1}}(\Delta y_u \mid y_{1..u-1}) \\
&= \prod_{u=t+1}^{T} \hat{f}_{\Delta Y_u|Y_{u-1}}(\Delta y_u \mid y_{u-1}) \\
&= \prod_{u=t+1}^{T} \frac{\sum_s I^u(y_{u-1}, Y^s_{u-1}) \cdot O^u(\Delta y_u, \Delta Y^s_u)}{\sum_s I^u(y_{u-1}, Y^s_{u-1})},
\end{aligned}$$

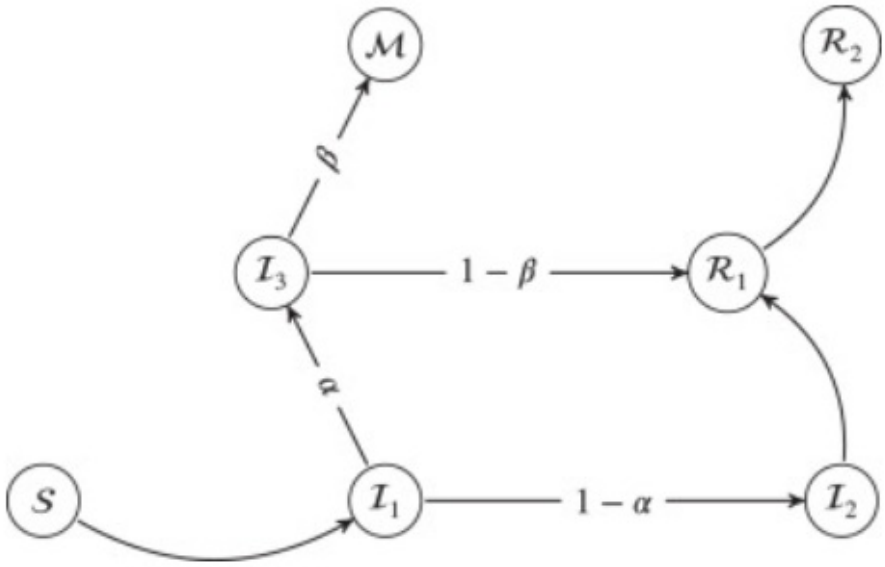
Our probabilist model

2.1. Model elements

Suppose that $\mathscr{D} = \{0, 1, \cdots, n\}$ is the set of days under study. Consider the following random processes whose domain is defined on $\mathscr{D}$.

- $\mathscr{S}(t)$: Number of people susceptible to become infected on day t.
- $\mathscr{I}_1(t)$: Number of infected individuals who are incubating the virus on day t.
- $\mathscr{I}_2(t)$: Number of infected people who have passed the theoretical incubation period and who: i) don't show symptoms or ii) symptoms are mild on day t.
- $\mathscr{I}_3(t)$: Number of infected people who have passed the incubation period and do show moderate or severe symptoms on the day t.
- $\mathscr{R}_1(t)$: Number of recovered cases which are still able to infect on the day t.
- $\mathscr{R}_2(t)$. Number of recovered cases that are not able to infect anymore on the day t.
- $\mathscr{M}(t)$: Number of deaths on day t.

Our probabilist model



Transition	Random variable	Used references
$I_1 \to I_2$	$Gamma(5.807, 0.948)$	Lauer, Grantz, Bi, Jones, Zheng, Meredith, Azman, Reich and Lessler (2020); Abdel-Salam and Mollazehi (2020)
$I_1 \to I_3$	$Gamma(5.807, 0.948)$	Lauer et al. (2020); Abdel-Salam and Mollazehi (2020)
$I_2 \to R_1$	$Uniform(5, 10)$	
$I_3 \to R_1$	$Uniform(9, 14)$	Abdel-Salam and Mollazehi (2020)
$I_3 \to \mathcal{M}$	$Gamma(6.67, 2.55)$	Novel et al. (2020); Salje, Kiem, Lefrancq, Courtejoie, Bosetti, Paireau, Andronico, Hoze, Richet, Dubost et al. (2020) Abdel-Salam and Mollazehi (2020) Verity et al. (2020b)
$R_1 \to R_2$	$Uniform(7, 14)$	Bi, Wu, Mei, Ye, Zou, Zhang, Liu, Wei, Truelove, Zhang et al. (2020) Ehmann et al. (2019)

Table 1
Random variables of the time of each transition

Coefficient	Value	Used references
α	0.8	Day (2020); Nishiura, Kobayashi, Miyama, Suzuki, Jung, Hayashi, Kinoshita, Yang, Yuan, Akhmetzhanov and Linton (2020) Tabata, Imai, Kawano, Ikeda, Kodama, Miyoshi, Obinata, Mimura, Kodera, Kitagaki, Sato, Suzuki, Ito, Uwabe and Tamura (2020); Mizumoto, Kagaya, Zarebski and Chowell (2020)
β	0.06	Rajgor, Lee, Archuleta, Bagdasarian and Quek (2020); Fauci, Lane and Redfield (2020) Verity, Okell et al. (2020a); Wu, Leung, Bushman, Kishore, Niehus, de Salazar, Cowling, Lipsitch and Leung (2020) Mahase (2020); Dudel, Riffe, Acosta, van Raalte and Myrskyla (2020)

Table 2
Probability of each transition

Stochastic Model implementation

2.3. Stochastic Model implementation

Our model does not have a closed-form solution. Therefore, in a real-world setting, it is necessary to use statistical simulation methods to approximate specific population characteristics of the stochastic process as quantile functions. Also, we must fit some parameters of the model to characterize the behaviour of the study population. For this purpose, we use a sample of the deceased patients $\{\mathscr{M}_1, \mathscr{M}_2, \cdots, \mathscr{M}_s\}$ along the set of days $\mathscr{O} = \{1, \cdots, s\}$.

Next, we suppose that our model (M) is dependent on a vector of parameters $\theta = (\theta_1, \theta_2) \in \mathbf{R}^{p_1} \times \mathbf{R}^{p_2}$ (with $p_1 + p_2 = p$), where θ_1 is a vector of dimension p_1, defined in beforehand, and θ_2 must be estimated from the sample. Furthermore, let us assume that the initial state of the system is characterized by

$\mathscr{S} = (\mathscr{S}(0), \mathscr{I}_1(0), \mathscr{I}_2(0), \mathscr{I}_3(0), \mathscr{R}_1(0), \mathscr{R}_2(0), \mathscr{M}(0)) \in \mathbf{N}^7$ and $\mathscr{T} = (\mathscr{T}_1(0), \mathscr{T}_2(0), \mathscr{T}_3(0), \mathscr{T}_4(0), \mathscr{T}_5(0), \mathscr{T}_6(0)) \in \mathbf{N}^m \times \cdots \times \mathbf{N}^m$. $\mathscr{S}$ has the number of elements for each compartment of the model on day 0. $\mathscr{T}$ also contains the amount of remaining days to complete the transition they are in for each individual in the initial state, being m a natural number that represents the maximum number of registered days.

To simplify the notation, for each day $t \in \mathscr{D}$, we denote the average dead trajectory by the function $Mean(\theta_1, \theta_2, \mathscr{S}, \mathscr{T})(t)$.

The next step is to estimate $\hat{\theta}_2$. To do this, we propose to solve the following optimization problem:

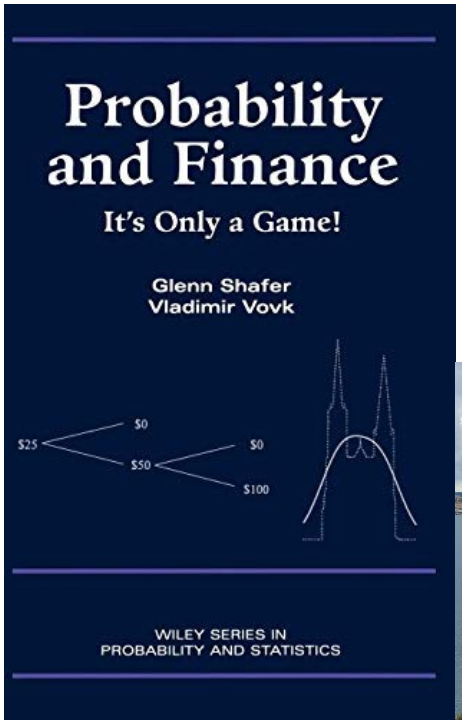
$$\hat{\theta}_2 = \arg \min_{\theta_2 \in S \subset \mathbf{R}^{p_2}} \sum_{i=1}^{s} \omega^i (\mathscr{M}_i - Mean(\theta_1, \theta_2, \mathscr{S}, \mathscr{T})(i))^2, \tag{2}$$

where $\omega = (\omega^1, \cdots, \omega^s)$ is a weighted vector that can help to improve model estimation. Examples of these weights may be:

$\omega^i = \mathscr{M}_i / \sum_{i=1}^{s} \mathscr{M}_i$ or $\omega^i = (1/\mathscr{M}_i) / (\sum_{i=1}^{s} 1/\mathscr{M}_i)$ $(i = 1, \cdots, s)$.

Model uncertainity with Conformal Inference

Vladimir Vovk (Student Andrei Kolmogorov and Professor Aleksei Semenov.)



Conformal Simulation Prediction Intervals

$$Y = m(X) + \epsilon \text{ with } \mathrm{E}(\epsilon \mid X) = 0.$$

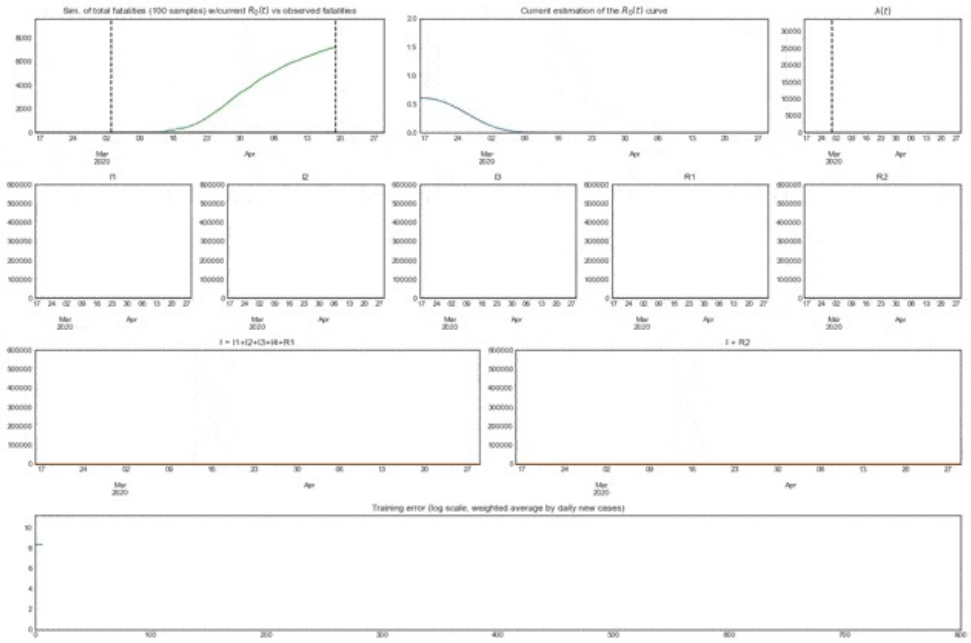
- Symmetrical random error
- Exchangeability

Below, we introduce the specific mathematical detail of the used conformal simulation bands that can handle heterocedastic noise. First, suppose that $\theta = (\theta_1, \hat{\theta}_2)$ is the optimal parameter configuration, where $\hat{\theta}_2$ was estimated according to methodology proposed in the Section 2.3. Then,

1. Perform $B = 10000$ simulation of the model $M(\theta_1, \hat{\theta}_2, \mathcal{S}, \mathcal{T})$ and evaluate the mean square error metric (RSS). $R\hat{S}S^s$ and $M^s(\theta_1, \hat{\theta}_2, \mathcal{S}, \mathcal{T})$ $(s = 1, \ldots, B)$, denote the results for iteration s of mean square error and the estimation of mortaly records in the simulation model respectively.
2. Let $Sel = \{i \in \{1, \ldots, B\} : R\hat{S}S^i \leq R\hat{S}S_{(1000)}\}$, the set of index of simulation with the lesser or equal 1000 value of RSS estimations. $R\hat{S}S_{(1000)}$ denote the element 1000, considering the order sample of $\{R\hat{S}S^s\}_{s=1}^{B}$.
3. Using the subsample of death simulation trajectories $\{M^i(\theta_1, \hat{\theta}_2, \mathcal{S}, \mathcal{T})\}_{i \in Sel}$, estimate pointwise the standart deviation $\hat{\sigma}(t)$, $\forall t \in \mathcal{O}$.
4. Define the conformal score $Score_i = \max_{t \in \mathcal{O}} \frac{|M^i(\theta_1, \hat{\theta}_2, \mathcal{S}, \mathcal{T})(t) - \mathcal{M}_t|}{\hat{\sigma}(t)}$ if $\hat{\sigma}(t) > 0$, $\forall i \in Sel$. Otherwise $Score_i$ is equal to 0.
5. Calculate the quantile $q_\alpha = \arg\min_{t \in \mathbb{R}^+} \{\frac{\sum_{s=1}^{1000} 1\{Score_s \leq t\}}{1000} \geq \alpha\}$, with $\alpha = 0.95$, to guarantee distributional intervals that cover a confidence level of 90%.
6. Define for each $t \in \mathcal{O}$, the confidence interval prediction as $[\overline{M}(\theta_1, \hat{\theta}_2, \mathcal{S}, \mathcal{T})(t) - \hat{\sigma}(t)q_\alpha, \overline{M}(\theta_1, \hat{\theta}_2, \mathcal{S}, \mathcal{T})(t) + \hat{\sigma}(t)q_\alpha]$, where $\overline{M}(\theta_1, \hat{\theta}_2, \mathcal{S}, \mathcal{T})$ denote the simulation mean that correspond with our given mortality estimations.
7. Finally, to build confidence bands for the rest of the stochastic process that makes up our epidemic model, we must select simulation trajectories that lead to the mortality outcome falling within the mortality band calculated in step 6).

Black-Box Optimization techniques

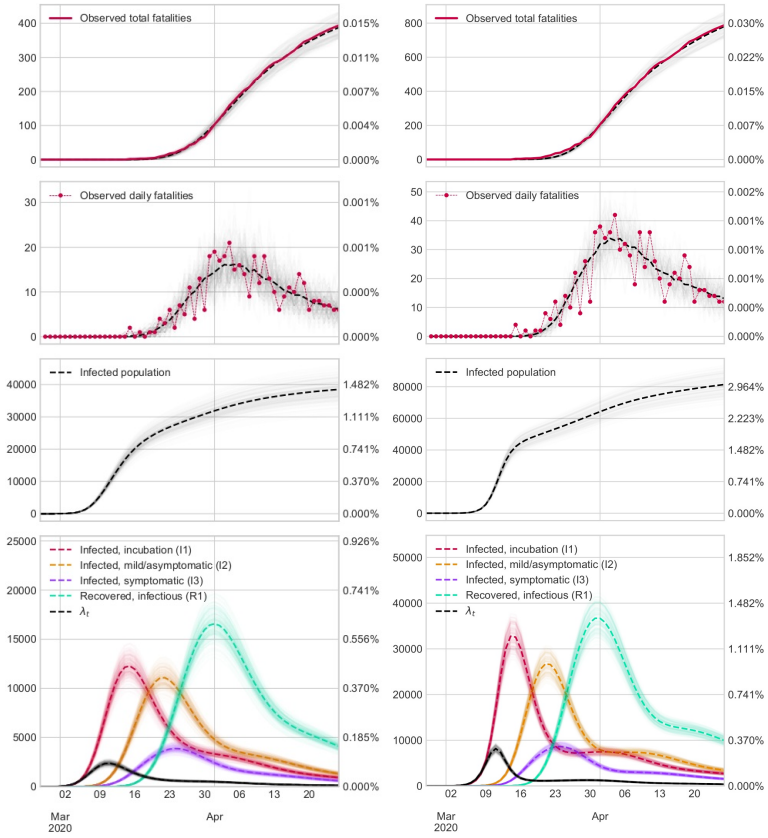
- Obtain the best model parameter only sampling.
- CMAES: model derivative free-optimization solver.
- Probabilist generalization of EM-algoritm.



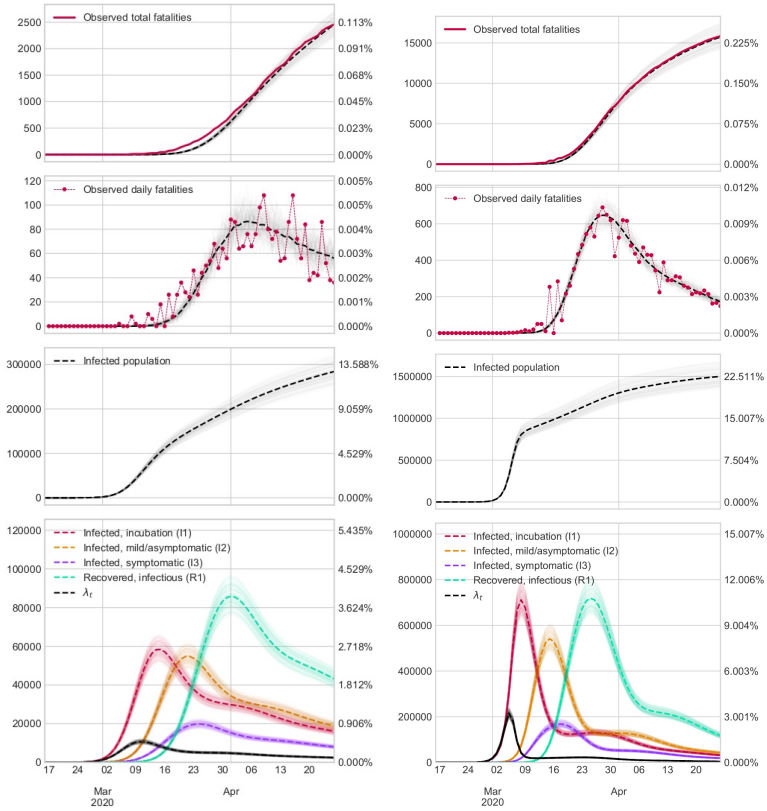
Modeling Strategies

- First wave, no reliable information, two scenarios:
    - Optimistic: Fatalities reported.
    - Pessimistic: Double Fatalities reported.
- National analysis with an excess of mortalities

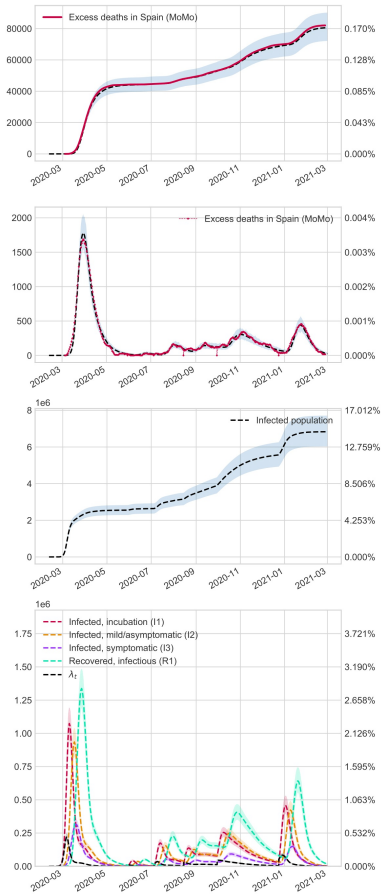
Galicia Results



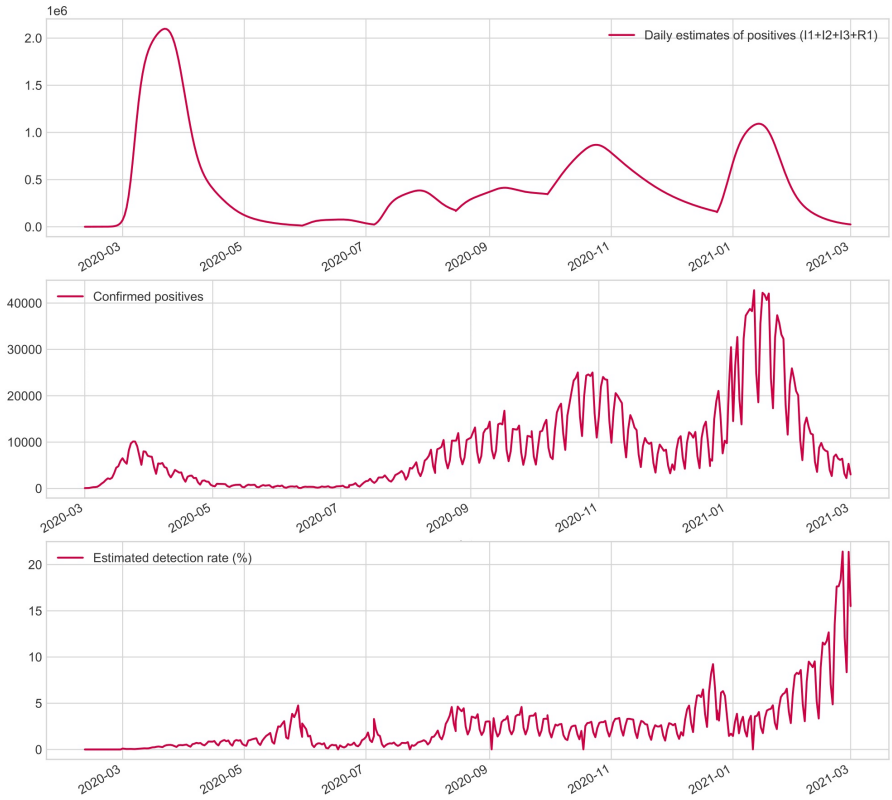
Basque Country and Madrid Results



Spain Results



Evaluate Testing Capacity in Spain



What novelties provide our model?

- Our framework is general and can handle non-parametric distribution or complex structures.
- Our optimization strategy is valid in these cases.
- In more complex model extensions, we use machine learning techniques to accelerate Black-Block Optimization strategies.
- From a modeling point of view, we use daily mortality records as the source of information that increases the difficulty of model identification.

Personalized medicine as motivation of methodological research

Monitoring real epidemic situations with our model and optimize therapeutical strategies with machine learning techniques

ARTICLE

https://doi.org/10.1038/s41467-020-20816-7 OPEN

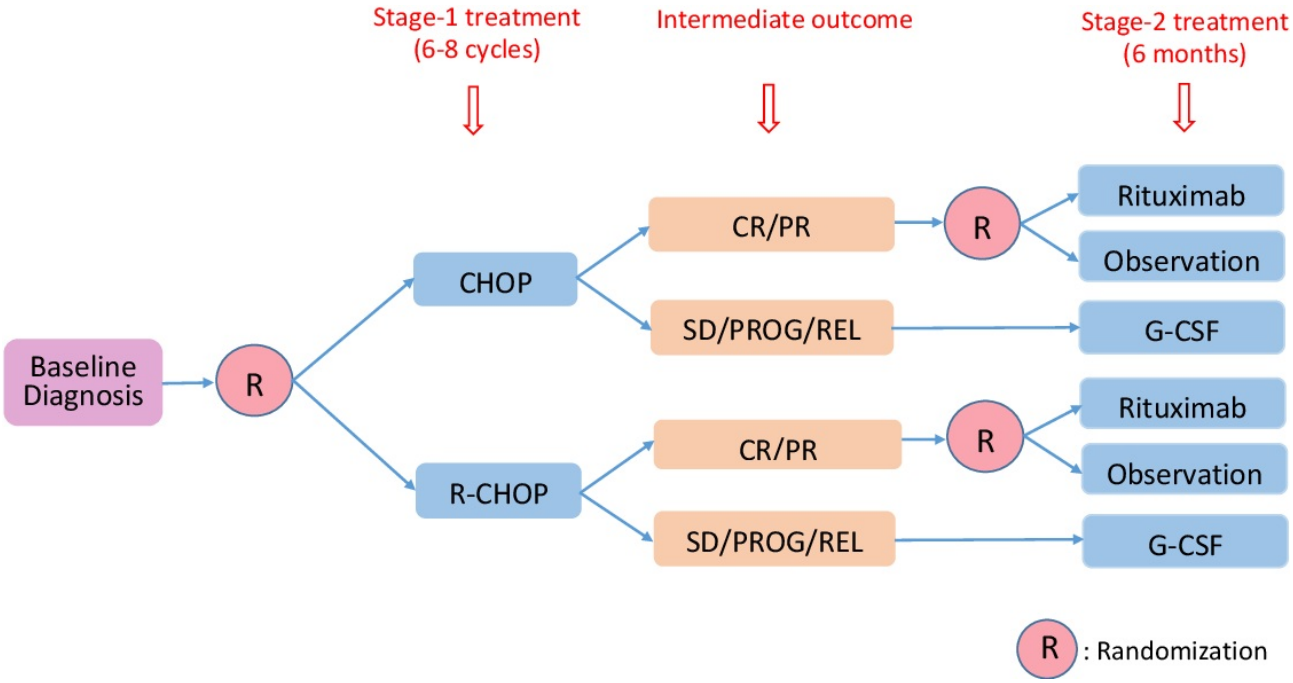
Real-time prediction of COVID-19 related mortality using electronic health records

Patrick Schwab[1✉], Arash Mehrjou[2,3], Sonali Parbhoo[4], Leo Anthony Celi[5,6], Jürgen Hetzel[7,8], Markus Hofer[8], Bernhard Schölkopf[2,3] & Stefan Bauer[2,9]

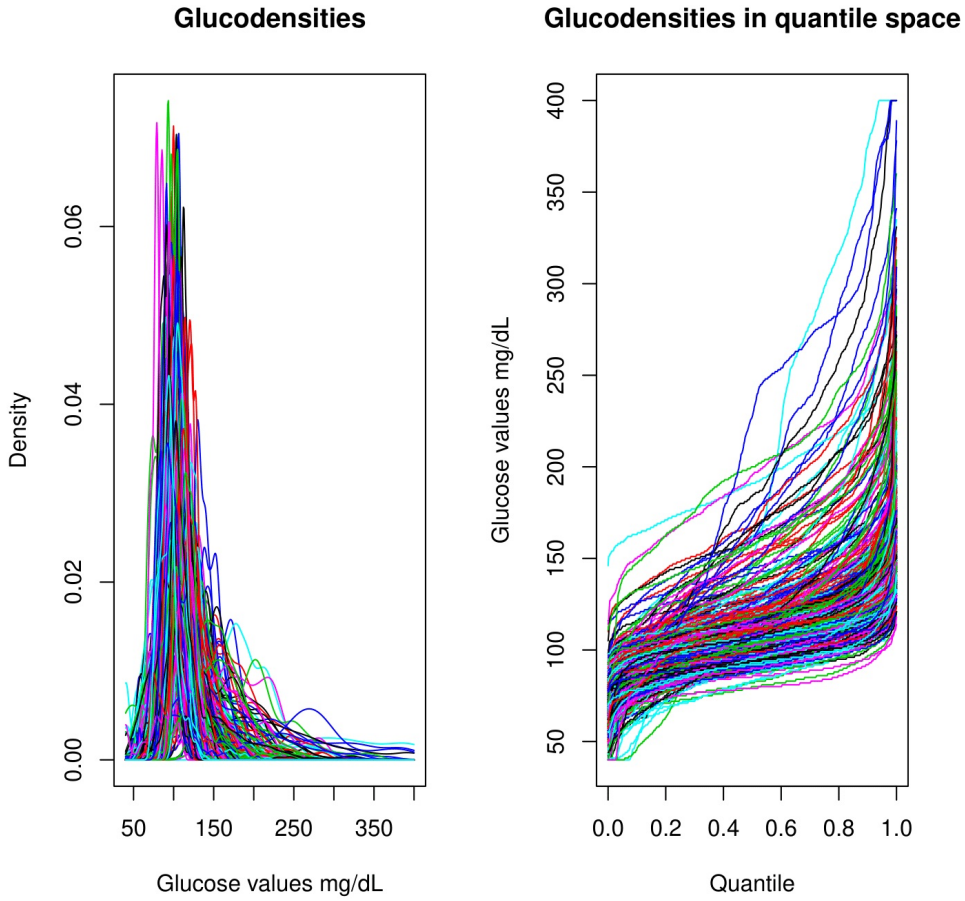
Coronavirus disease 2019 (COVID-19) is a respiratory disease with rapid human-to-human transmission caused by the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). Due to the exponential growth of infections, identifying patients with the highest mortality risk early is critical to enable effective intervention and prioritisation of care. Here, we present the COVID-19 early warning system (CovEWS), a risk scoring system for assessing COVID-19 related mortality risk that we developed using data amounting to a total of over 2863 years of observation time from a cohort of 66 430 patients seen at over 69 healthcare institutions. On an external cohort of 5005 patients, CovEWS predicts mortality from 78.8% (95% confidence interval [CI]: 76.0, 84.7%) to 69.4% (95% CI: 57.6, 75.2%) specificity at sensitivities greater than 95% between, respectively, 1 and 192 h prior to mortality events. CovEWS could enable earlier intervention, and may therefore help in preventing or mitigating COVID-19 related mortality.

The new era of personalized medicine

- Precision medicine seeks to maximize the quality of health care by individualizing the healthcare process to the uniquely evolving health status of each patient
- Precision medicine is formalized as a treatment regime that comprises a sequence of decision rules, one per decision point, which map up-to-date patient information to a recommended action.

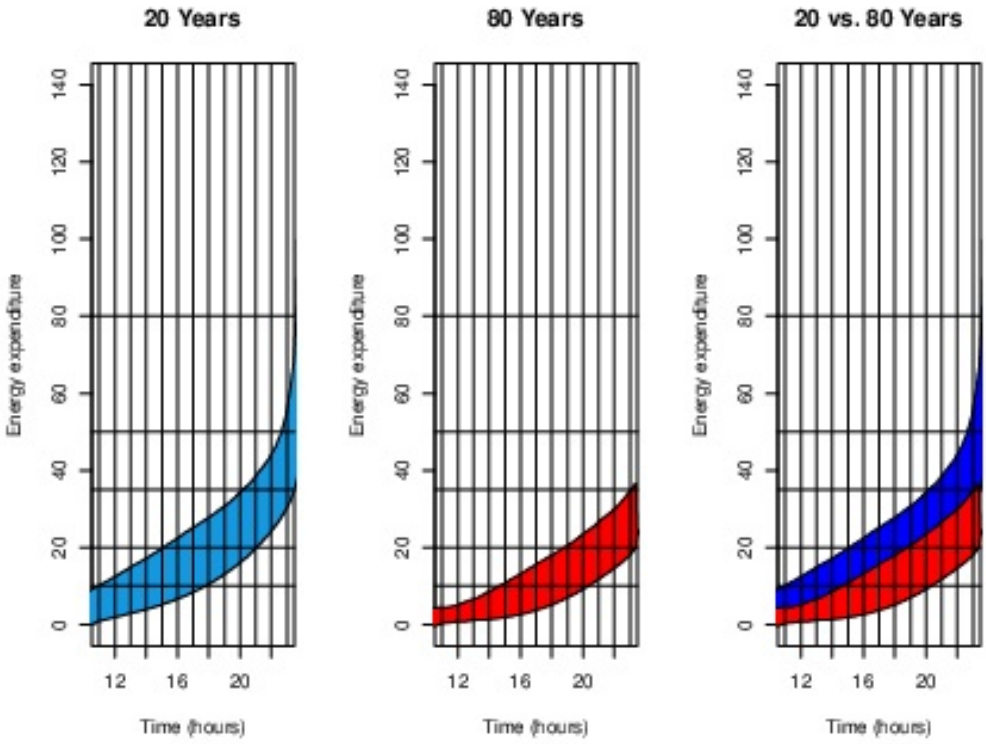


Fréchet analysis in diabetes



-->

Fréchet analysis can help to prescription exercise doses in general populations!



Local structures: Biclustering algoritms for new subtypes of patients

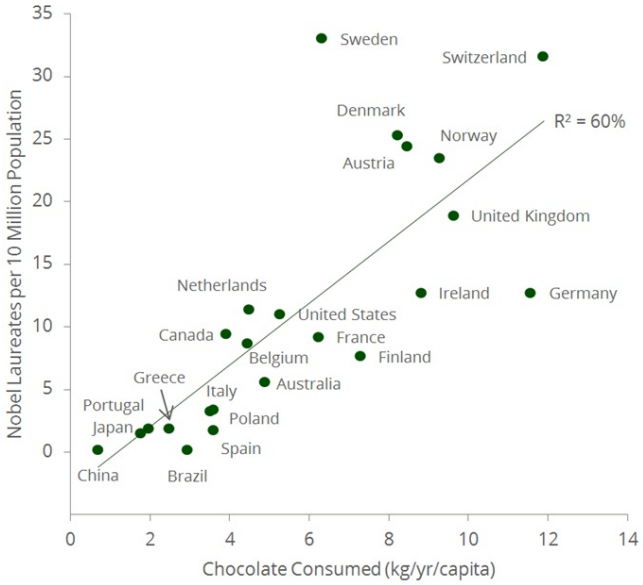
Kernel Biclustering algorithm in Hilbert Spaces

the date of receipt and acceptance should be inserted later

Abstract Biclustering algorithms partition data and covariates simultaneously, providing new insights in several domains, such as analyzing gene expression to discover new biological functions. This paper aims to establish a new model-free biclustering algorithm in abstract spaces using the notions of energy distance (ED) and the maximum mean discrepancy (MMD) –two distances between probability distribution in a separable Hilbert space and capable of handling complex data as curves or graphs. The proposed method can learn more general and complex cluster shapes than most existing literature approaches, usually focused on detecting mean and variance differences or other particular geometries shapes according to specific parametric distributions. Despite, the biclustering configurations of our approach are constrained to create disjoint structures at the datum and covariate levels, results are similar to state-of-the-art methods in their optimal scenarios, assuming a proper Kernel choice, outperforming them when cluster differences are concentrated in higher-order moments. Our approach has been tested in several situations that involve simulated and real-world datasets. Finally, new theoretical consistency results are established using some tools of the theory of optimal transport.

Patient heterogeneity and causality

- Bertrand Russell (in 1926-1927) that went along the lines "Probability is amongst the most important science, not least because no one understands it".
- "If it were not for the great variability among individuals, medicine might as well be a science and not an art." Sir William Osler, The Principles and Practice of Medicine 1892.




Thanks for your attention!