



COL·LEGI
D'ACTUARIS
DE CATALUNYA



CONFERENCIA

MODELO EMPÍRICO PARA LA PREDICCIÓN A CORTO PLAZO DE LA PROPAGACIÓN DE LA COVID-19

Enrique Álvarez Lacalle, Grupo BIOCOSMOS. UPC.

Barcelona, 28 de mayo de 2020

The use of empiric models for the short-term prediction of COVID-19. Risk assessments in epidemiological dynamics.

28th May 2020

Enric Alvarez

Investigador del Grup de recerca de Biologia Computacional i Sistemes Complexos (BIOCOM-SC, UPC)



Tracking epidemics. A collaboration between IGTP and UPC...



CMCiB

Comparative Medicine & Bioimage
Centre of Catalonia



**Computational Biology
and Complex Systems
Research Group**

Martí Català

Pere-Joan Cardona, PhD

*Comparative Medicine and Bioimage Centre of
Catalonia; Institute for Health Science Research
Germans Trias i Pujol*

Clara Prats, PhD

Sergio Alonso, PhD

Enric Álvarez, PhD

Miquel Marchena

David Conesa

Daniel López, PhD

*Computational Biology and Complex Systems;
Universitat Politècnica de Catalunya - BarcelonaTech*

With the collaboration of: Guillem Álvarez, Oriol Bertomeu, Laura Dot, Lavínia Hriscu, Helena Kirchner, Daniel Molinuevo, Pablo Palacios, Sergi Pradas, David Rovira, Xavier Simó, Tomás Urdiales



Outline



1. Context and objectives

2. An empirical model to deal with COVID19. Prediction horizon

3. Estimating real incidence of COVID19

4. Identifying risk

Context and objectives

- Research group BIOCOMSC with previous experience on modelling infectious diseases (tuberculosis, Chagas disease, malaria...) and multi/scale modelling in diseases (subcellular, cellular, organ scale).
- Follow up on COVID19 epidemic in China during January-February.
- Start **applying developed models** when COVID19 reaches **Europe**.
- Clara Prats leads the creation of a COVID19 team from the **European Commission** (DG-CONNECT, DG-SANTE) to carry out a daily analysis and prediction of the situation in EU countries (daily reports → <https://biocomsc.upc.edu/en/covid-19/daily-report>).
- Collaboration with **AQuAS** team that follows up epidemic in **Catalunya** → analysis and prediction at lower scale (health regions, AGAs, hospitals), including hospitalizations, critical points and ICUs. Collaboration extended to other autonomous governments
- Track evolution of the epidemics as mobility increases with the goal to detect hotspots.

Outline



1. Context and objectives

2. An empirical model to deal with COVID19. Prediction horizon

3. Estimating real incidence of COVID19

4. Identifying risk

An empirical model to deal with COVID19

Why an **empirical model**?

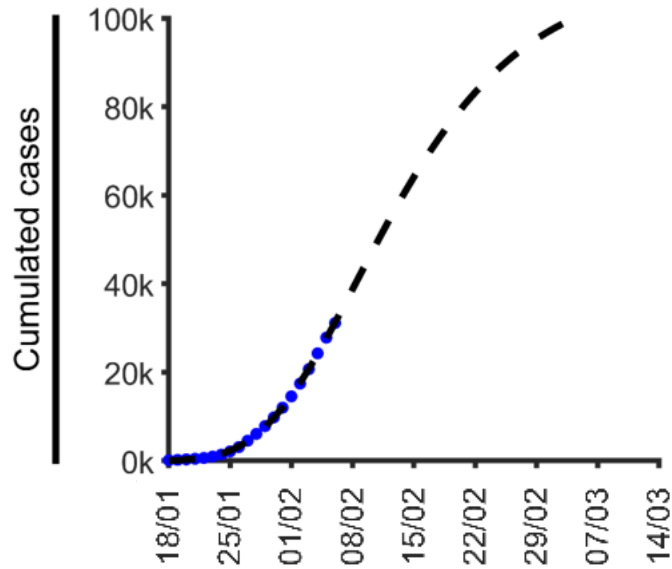
- ✓ February 2020 → Lack of precise knowledge about the dynamics of COVID19: incubation period, infectiousness, disease duration... → SEIR models \approx empirical fittings with many parameters.
- ✓ Lack of herd immunity or vaccine → Dynamics of the disease is **not driven by susceptible population** (not limiting) but by control measures → SEIR models are less appropriate.

Our choice: an **empirical model** that correctly describes the cumulative cases curve and that has only **2 parameters**.

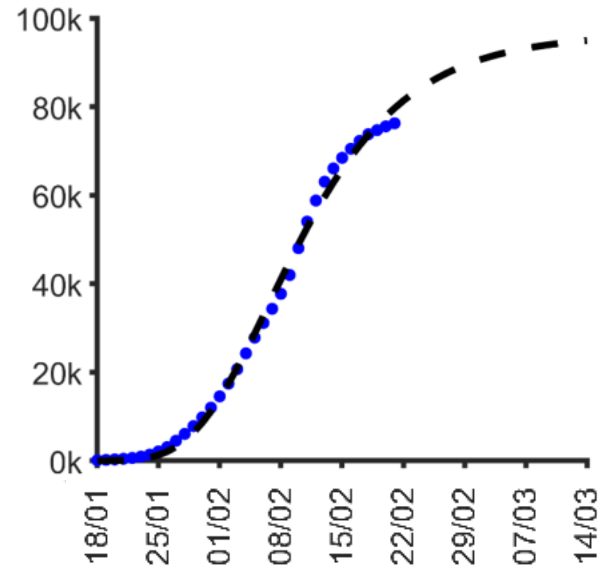
Gompertz model

Parameters of Gompertz model: a and K

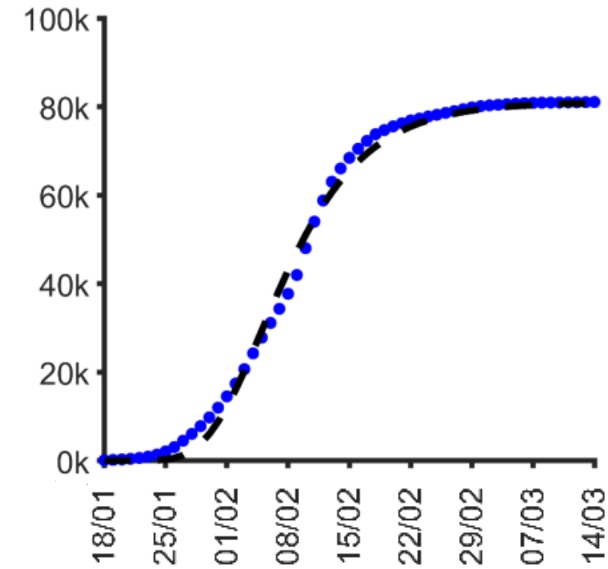
Exponential growth (μ_0)



Growth slows down (a)



Maximum level (K)



Time (days)

$$\mu_0 = a \ln\left(\frac{K}{N_0}\right)$$

$$\frac{d\mu}{dt} = -a\mu$$

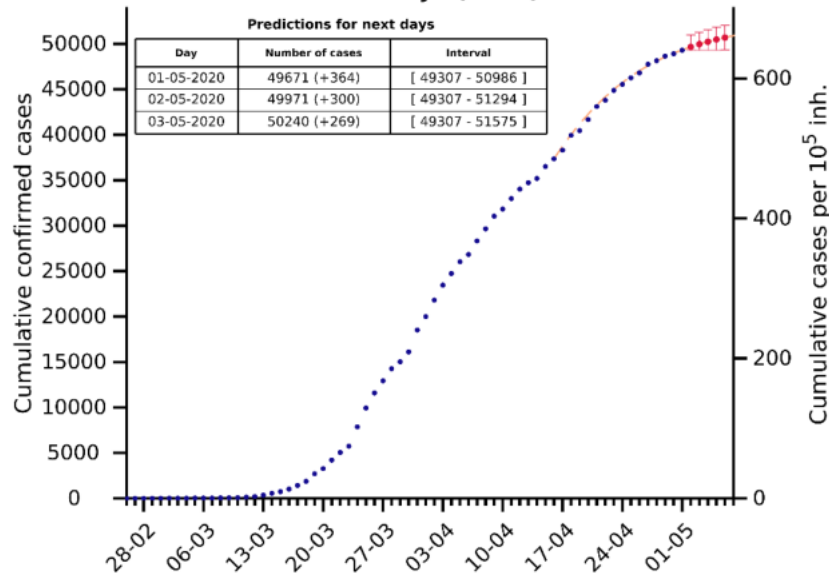
$$N(t) = K e^{-\ln\left(\frac{K}{N_0}\right) e^{-at}}$$

Gompertz model: predictions

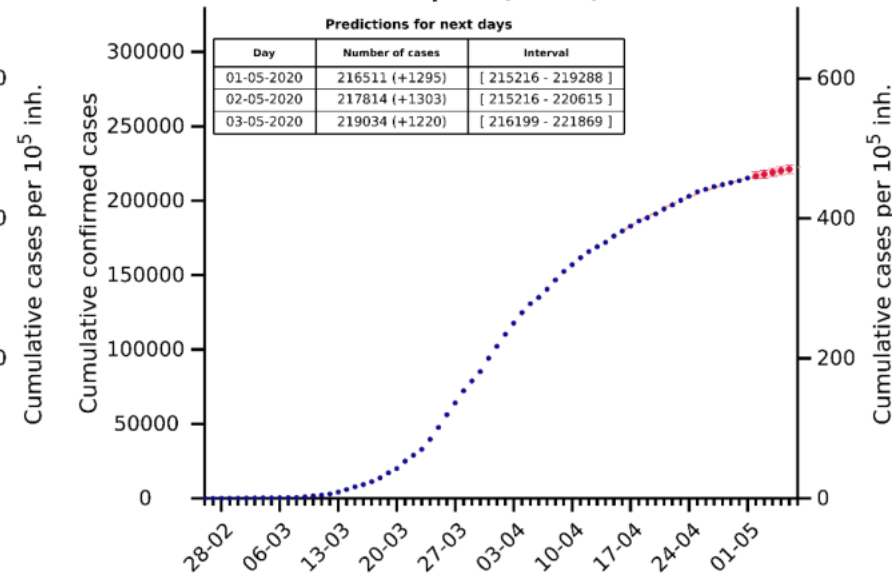
Daily fitting Gompertz model to cumulative cases (countries and regions)

- Predictions at **short term** → expected cases at 1 to 5 days, depending on length of historical series

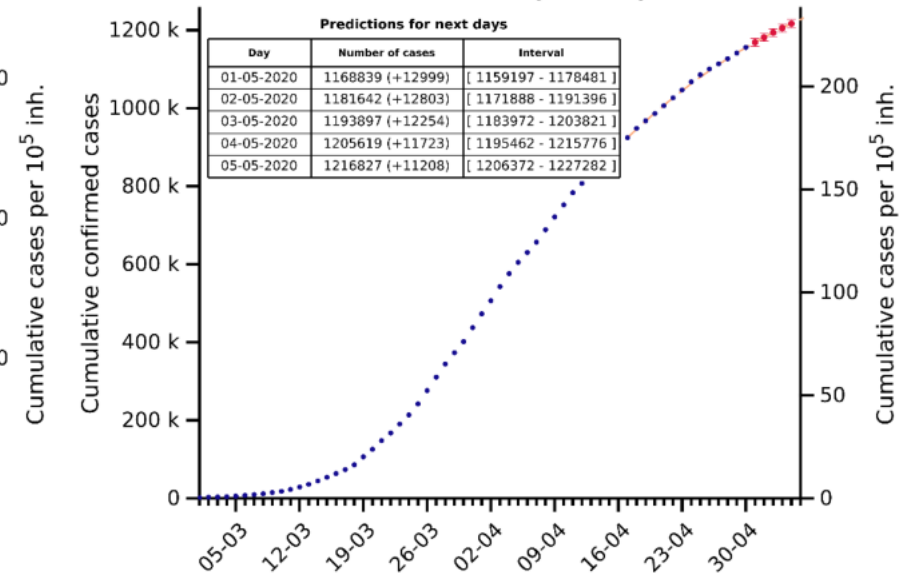
Catalunya (7.7M)



España (47.0M)



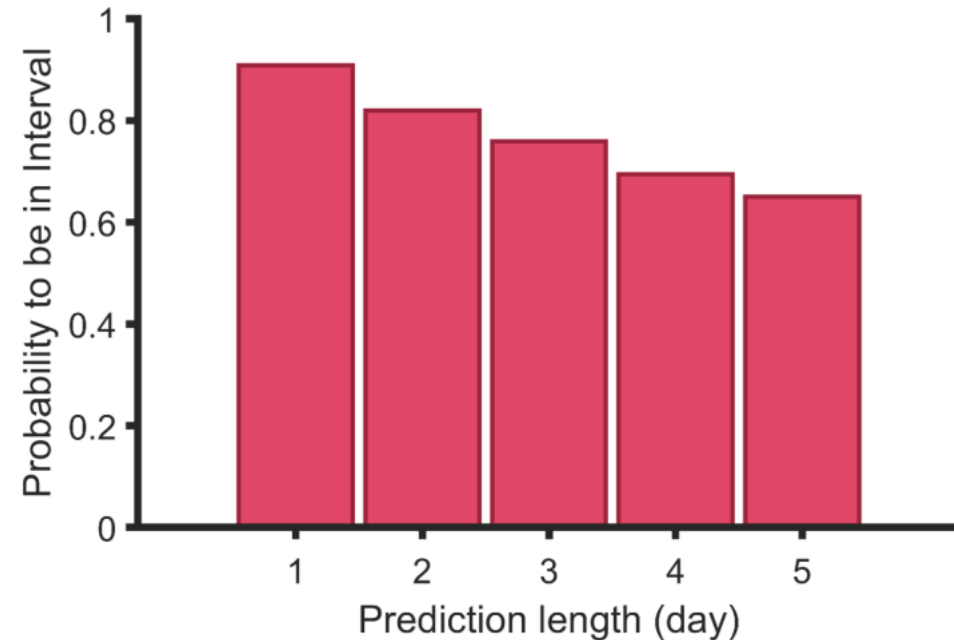
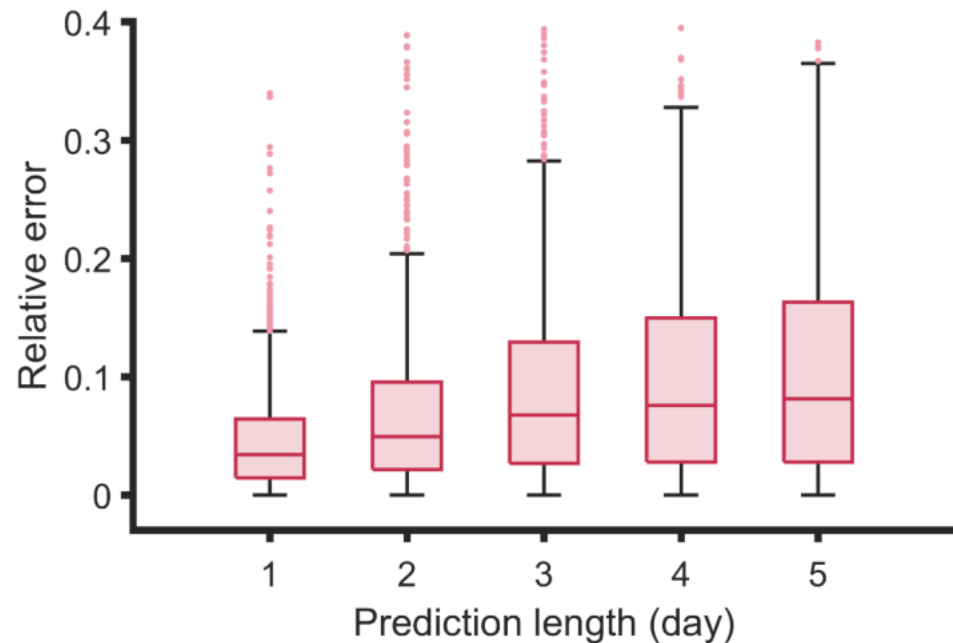
EU+EFTA+UK (527.9M)



Gompertz model: predictions

Daily fitting Gompertz model to cumulative cases (countries and regions)

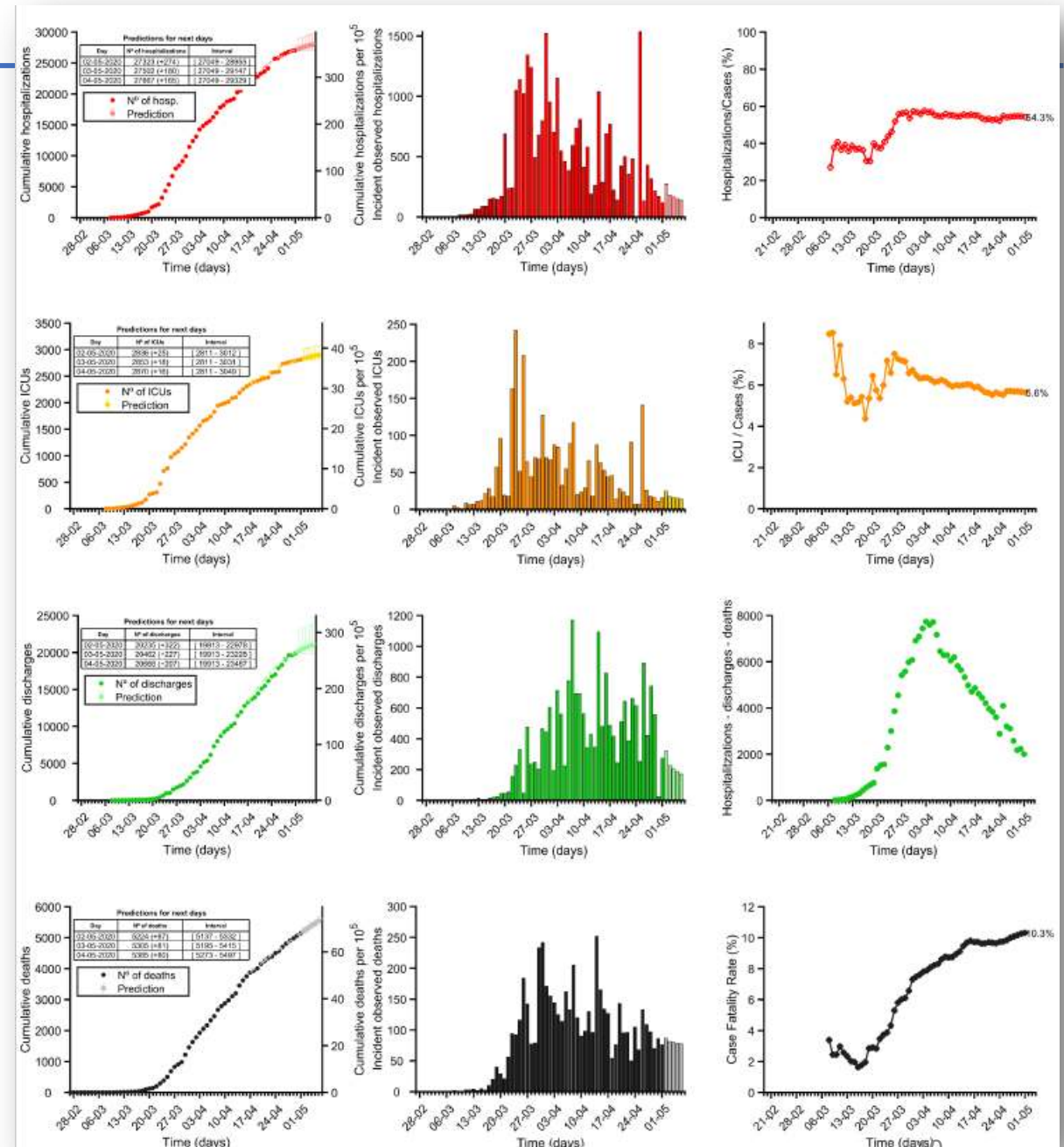
- Predictions at **short term** → expected cases at 1 to 5 days, depending on length of historical series



Gompertz model: Hospital coverage risk

Daily fitting Gompertz model to cumulative data

- Predictions at **short term** → also for ICUs, hospitalizations, discharges, deaths



Outline



1. Context and objectives

2. An empirical model to deal with COVID19. Prediction horizon

3. Estimating real incidence of COVID19

4. Identifying risk

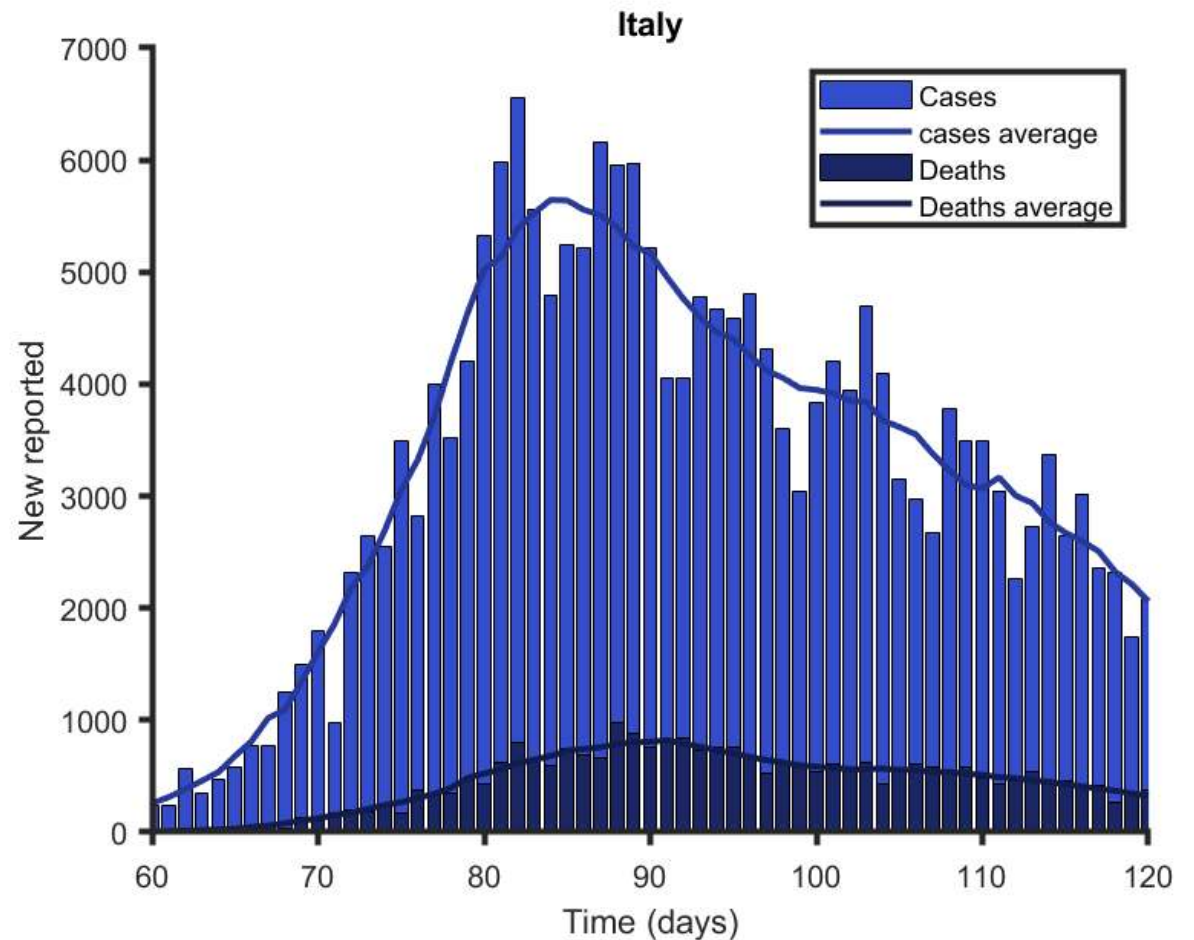
Why estimations of real incidence are needed to evaluate risk

- People with mild symptoms can infect others. Uncertainty about the ability of people with no symptoms (20-40%) to spread the disease
- Most countries do not have the ability to detect people with mild symptoms. Some exceptions are South Korea and Israel.
- It is important to estimate the people that are really active
Correlation between cases and deaths → Diagnostic to Death (DtD)

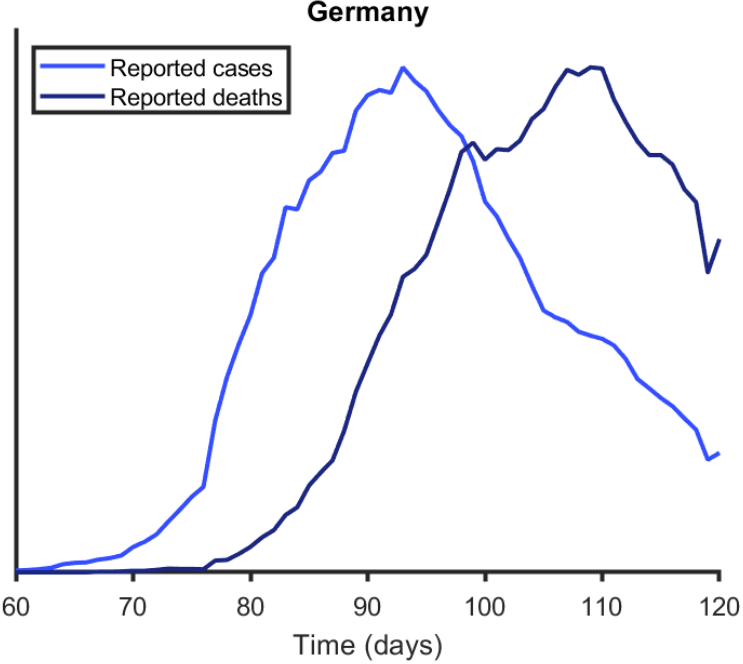
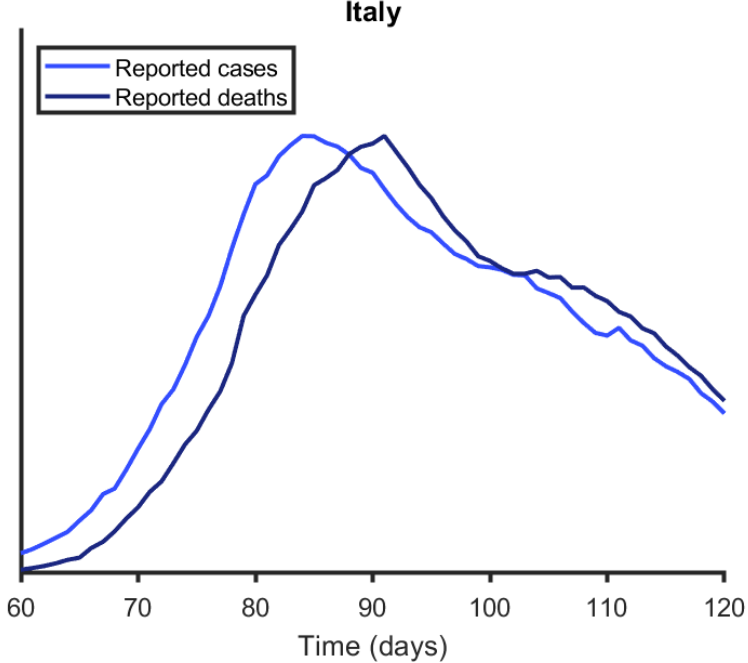
Estimating real incidence

Bibliographical research and new evidence

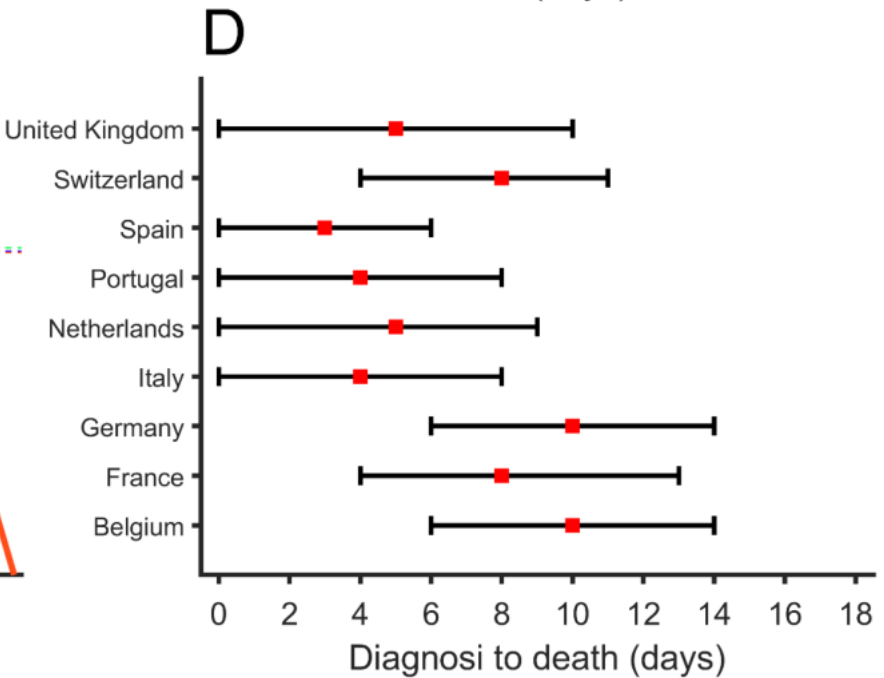
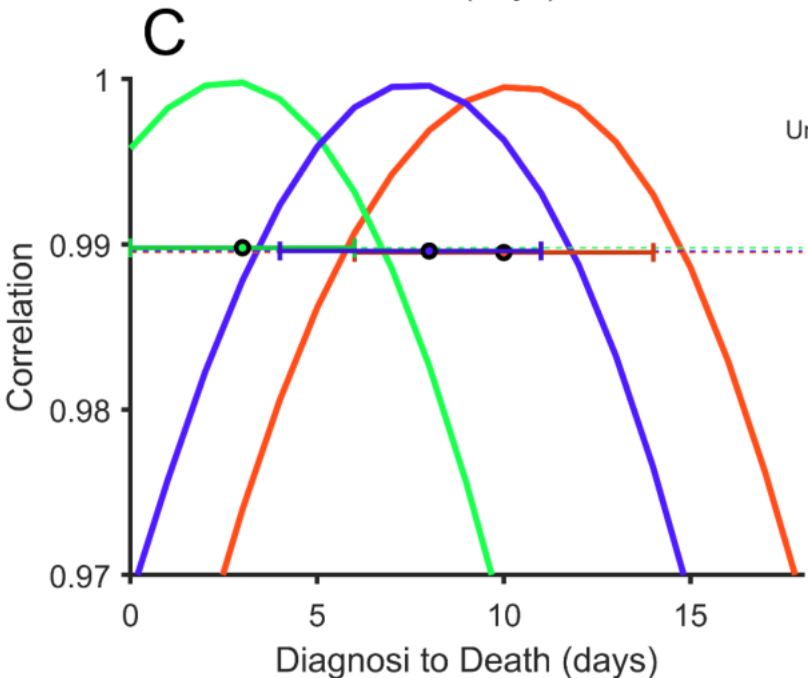
- ✓ 1 % lethality on average
- ✓ 0.5-0.8% lethality for low penetration regions
- ✓ Up to 1.2-1.5 % on average for high penetration regions (residences, nursing homes have very high lethality)
- ✓ Time from onset to death, $TtD \leq 18$ days
- ✓ Correlation between cases and deaths → Diagnostic to Death (DtD)



Estimating real incidence. First rescale.



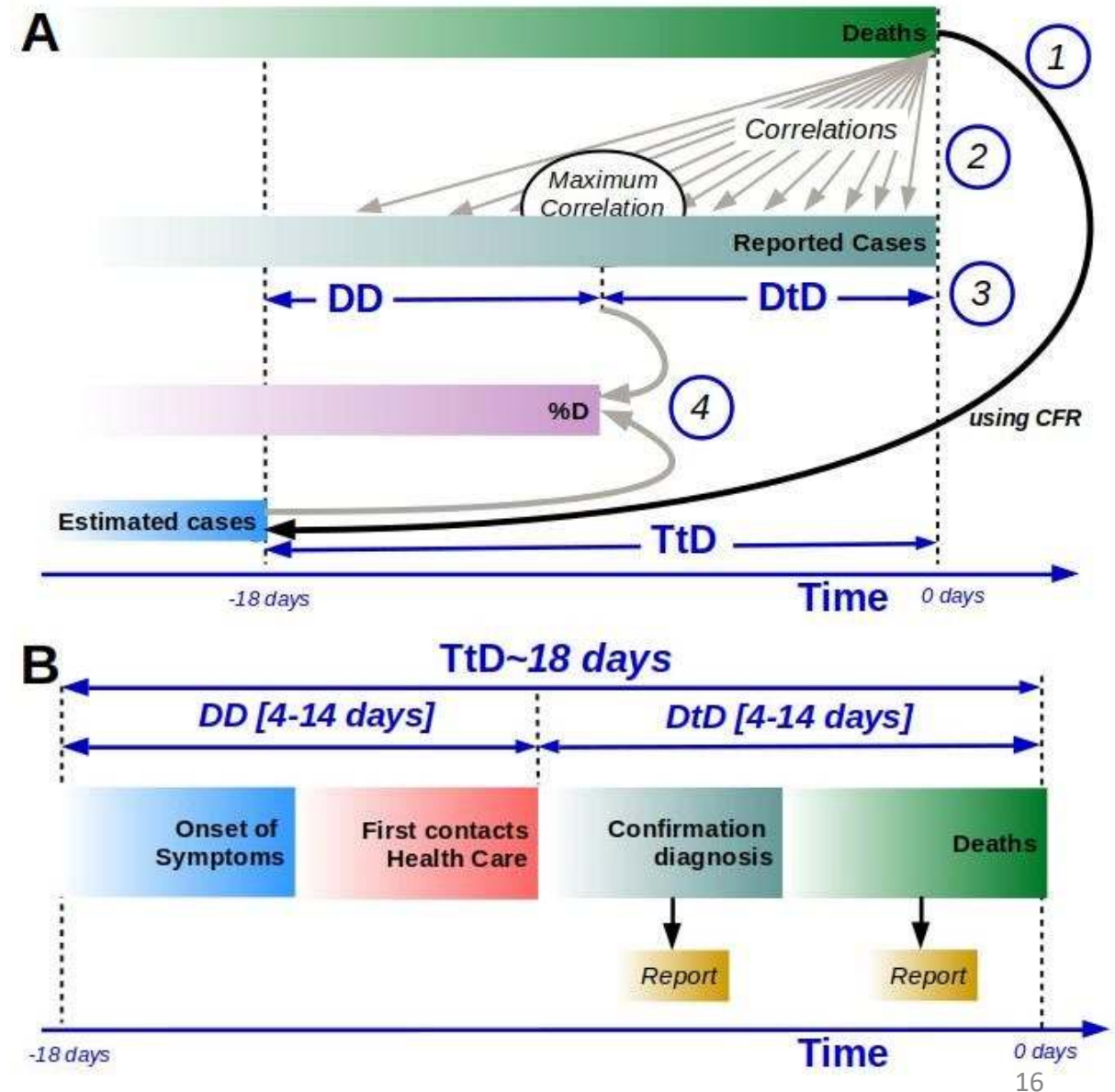
Estimating real incidence. Obtain delay



Estimating real incidence. Now consider all delays

Now take into account delays

- ✓ 1 % lethality
- ✓ Time from onset to death, $TtD \leq 18$ days
- ✓ Correlation between cases and deaths \rightarrow Diagnostic to Death (DtD)
- ✓ Diagnosis Delay $DD = TtD - DtD$
- ✓ DD includes delay in data recording



Estimating real incidence

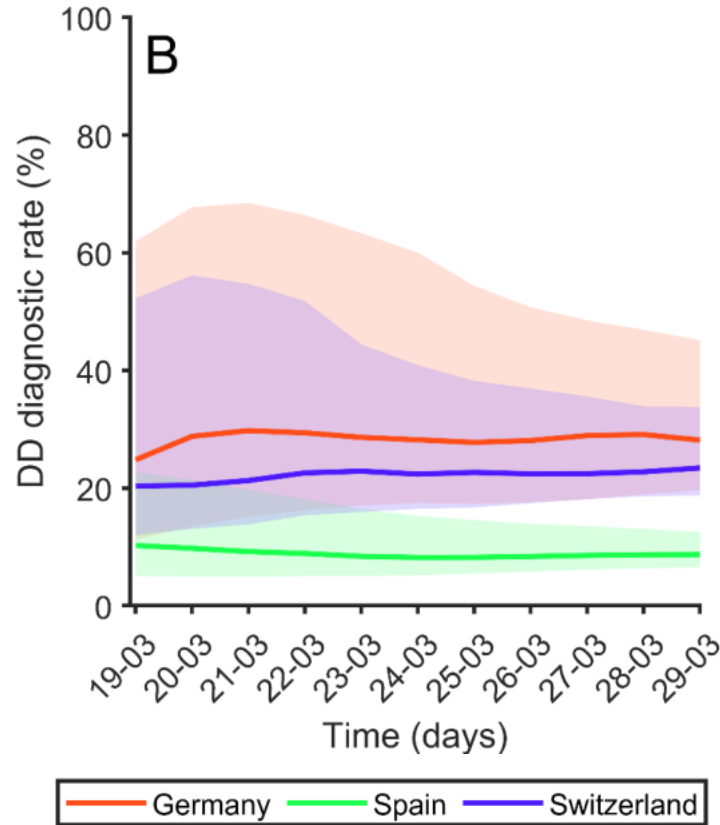
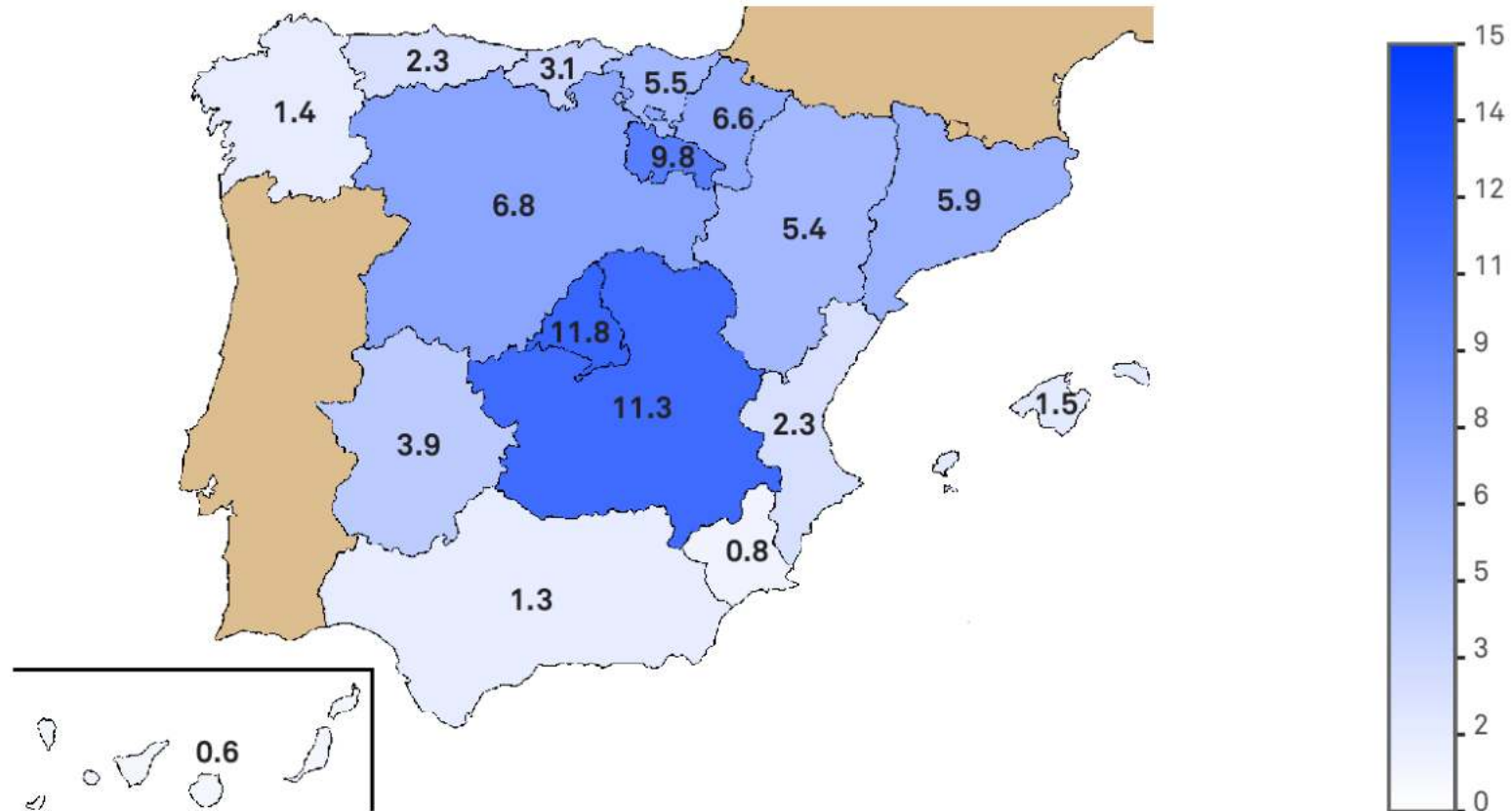


Table from late April

Countries	Detection Delay			Diagnosis Delay Detection Rate (%)			Estimated cumulative cases			Estimated attack rate /10 ⁵ inh.		
	Value	min	max	Value	min	max	Value	min	max	Value	min	max
Belgium	8	4	12	4%	3%	6%	759000	596000	1060000	6553	5140	9143
France	10	5	14	5%	3%	5%	2331000	1956000	3044000	3571	2996	4663
Germany	8	4	12	27%	21%	32%	489000	412000	633000	583	492	755
Italy	14	10	14	7%	6%	7%	2352000	2352000	2631000	3889	3889	4352
Netherlands	13	9	14	7%	6%	8%	382000	362000	471000	2230	2115	2749
Portugal	14	10	14	27%	21%	27%	68000	68000	87000	668	668	852
Spain	14	12	14	9%	8%	9%	2043000	2043000	2171000	4344	4344	4617
Sweden*	9	4	14	6%	4%	8%	197000	142000	289000	1955	1406	2860
Switzerland	10	7	14	23%	21%	25%	116000	104000	127000	1335	1197	1472
United Kingdom	13	8	14	6%	4%	6%	1777000	1625000	2683000	2617	2394	3953

INDICE DE SEROPREVALENCIA POR COMUNIDADES

Predicción: @BIOCOMSC1

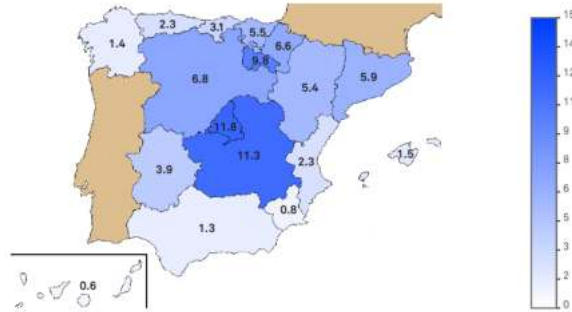


INDICE DE SEROPREVALENCIA POR COMUNIDADES

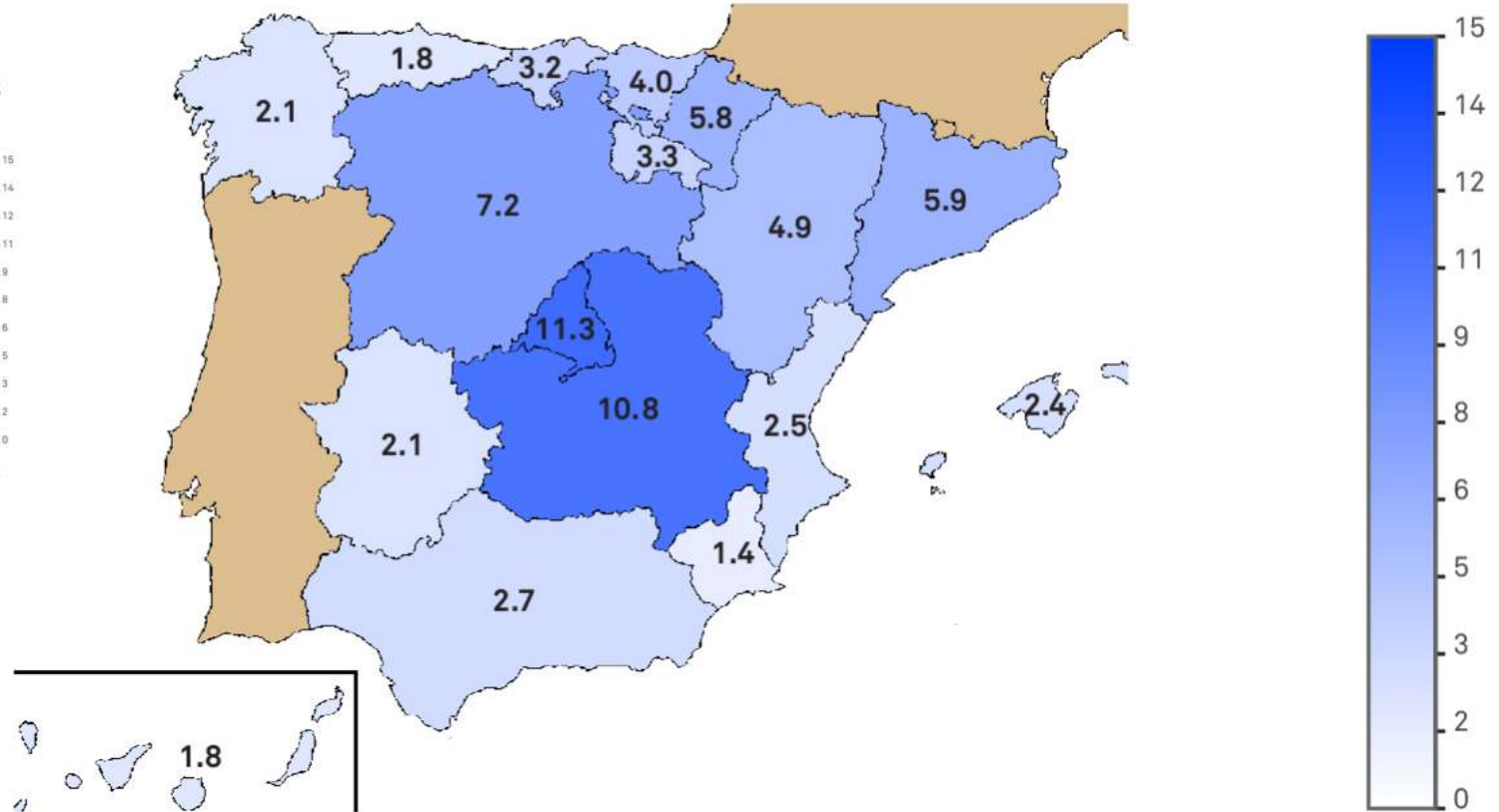
Datos estudio Ministerio Salud: @SALUDISCI3

INDICE DE SEROPREVALENCIA POR COMUNIDADES

Predicción: @BIOCOMSC1



FUENTE: BIODON-SC UPL, OMSB KDP / GRAFICO: INFOGRAFIA.CAT



FUENTE: Instituto de Salud Carlos III (ISCI3) / GRAFICO: INFOGRAFIA.CAT

Assessment of risk in Spain. A look at phase 1

CASOS ACTIVOS POR COMUNIDADES AUTÓNOMAS

Datos a 19 de mayo

	Estimación	% total	
ANDALUCÍA *	3.500	2.82	
ARAGÓN	5.000	4.02	
ASTURIAS	700	0.56	
ISLAS BALEARES	1.000	0.80	
COMUNIDAD VALENCIANA	5.900	4.75	
CANARIAS *	500	0.40	
CANTABRIA	600	0.48	
CASTILLA Y LEÓN	13.200	10.62	
CASTILLA-LA MANCHA	10.800	8.69	
CATALUÑA	34.000	27.35	
EXTREMADURA	2.000	1.61	
GALICIA	2.500	2.01	
LA RIOJA **	600	0.48	
MADRID	34.900	28.08	
MURCIA	600	0.48	
NAVARRA	2.500	2.01	
PAÍS VASCO	5.300	4.26	
ESPAÑA	124.300		

Estimación con errores del $\pm 15\%$

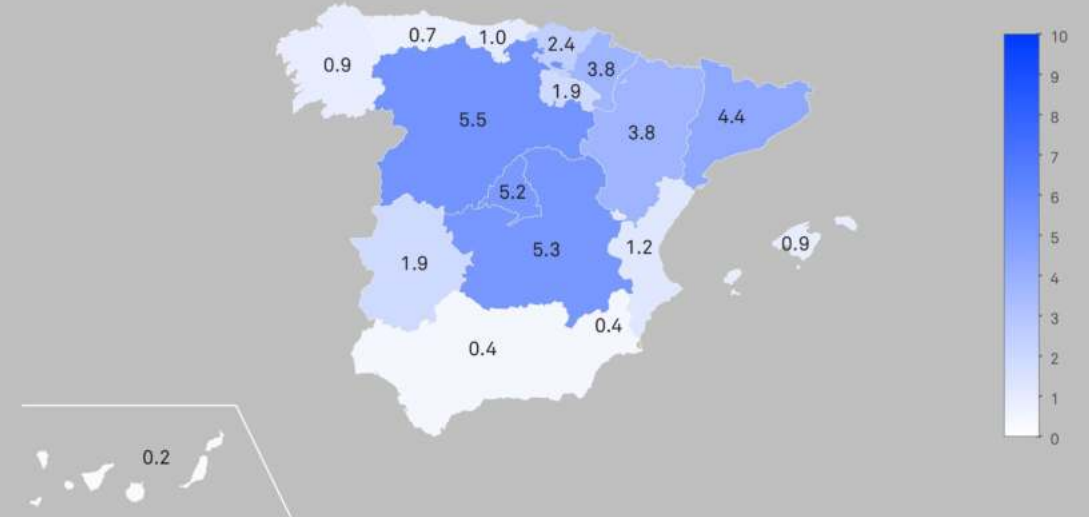
*Andalucía y Canarias pueden presentar un número mayor de casos activos respecto a lo indicado si se confirman los resultados de letalidad en el estudio ENE-COVID19

** La Rioja ha presentado datos de seroprevalencia en ENE-COVID19 estructuralmente muy diferentes del resto de Comunidades sin tener claro su origen.

Nuestra estimación es marcadamente diferente a dichos resultados.

CASOS ACTIVOS POR MIL HABITANTES

Estimación a 19 de mayo



Estimación con errores del $\pm 15\%$

*Andalucía y Canarias pueden presentar un número mayor de casos activos respecto a lo indicado si se confirman los resultados de letalidad en el estudio ENE-COVID19

** La Rioja ha presentado datos de seroprevalencia en ENE-COVID19 estructuralmente muy diferentes del resto de Comunidades sin tener claro su origen.

Nuestra estimación es marcadamente diferente a dichos resultados.

Graphics design

[Ferran Caymel \(@infografia_cat\)](https://twitter.com/infografia_cat)

Outline



1. Context and objectives
2. An empirical model to deal with COVID19. Prediction horizon
3. Estimating real incidence of COVID19
4. Identifying risk
 - Index
 - Following-up deconfinement
 - Early detection

Index 1: Spreading rate

✓ Basic reproduction number: R_0



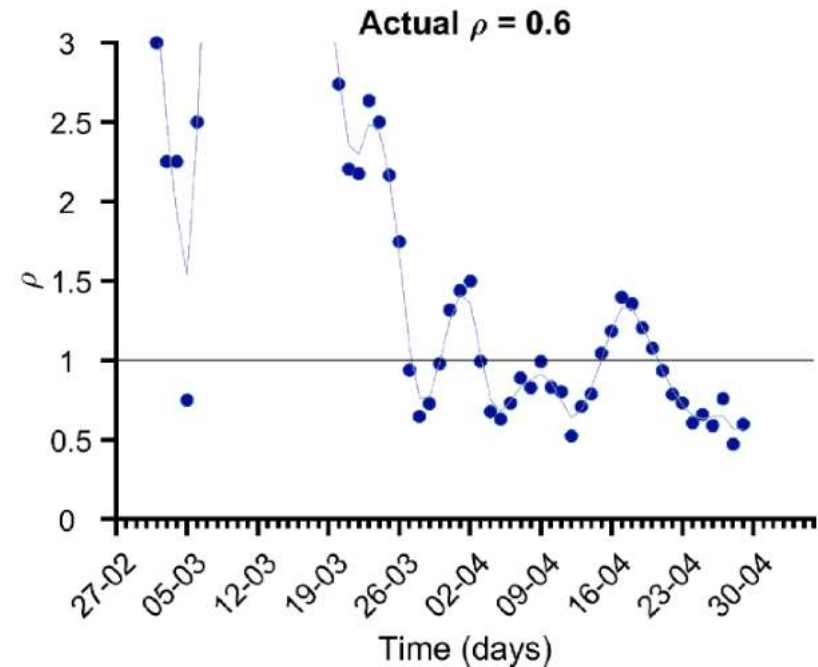
Estimated with SIR models

✓ Effective reproduction number: R_t

Empirical evaluation of spreading rate: ρ_t

$$\rho_t = \frac{N_{new}(t+1) + N_{new}(t) + N_{new}(t-1)}{N_{new}(t-4) + N_{new}(t-5) + N_{new}(t-6)}$$

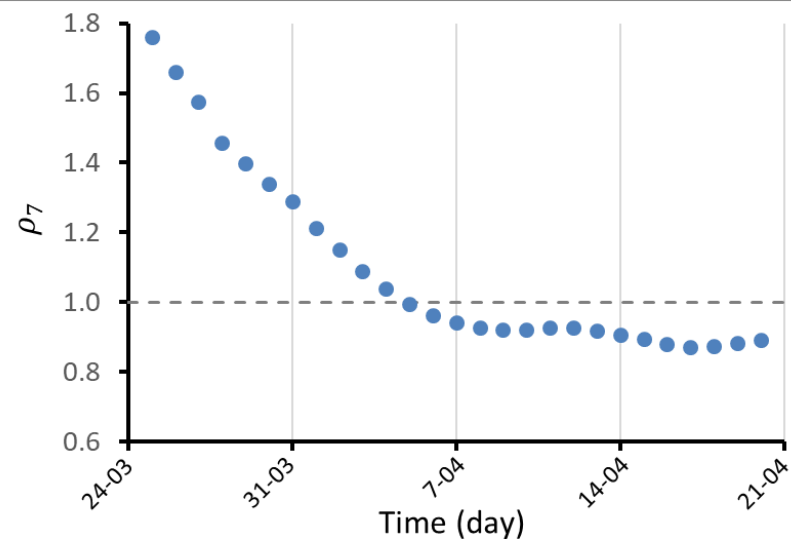
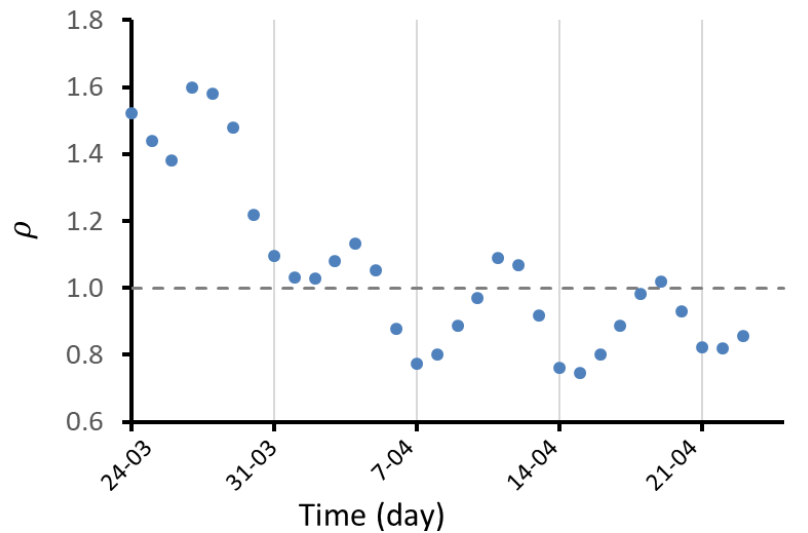
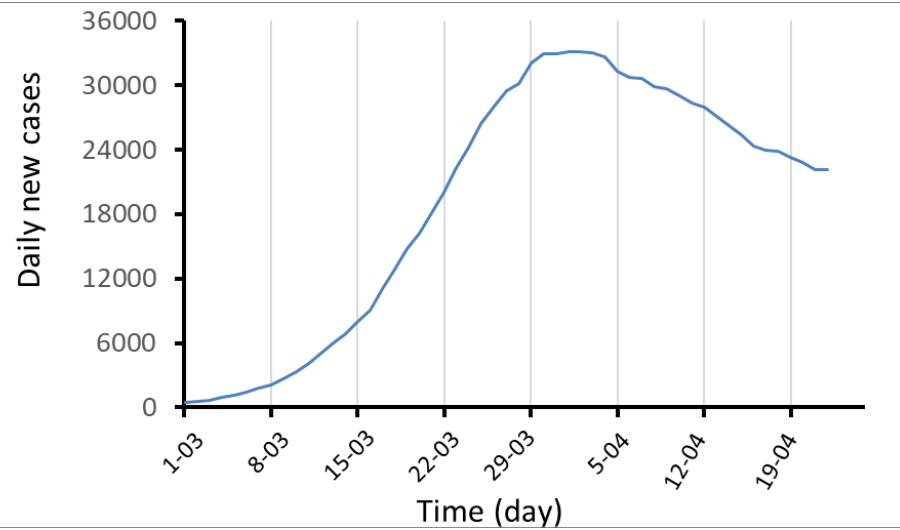
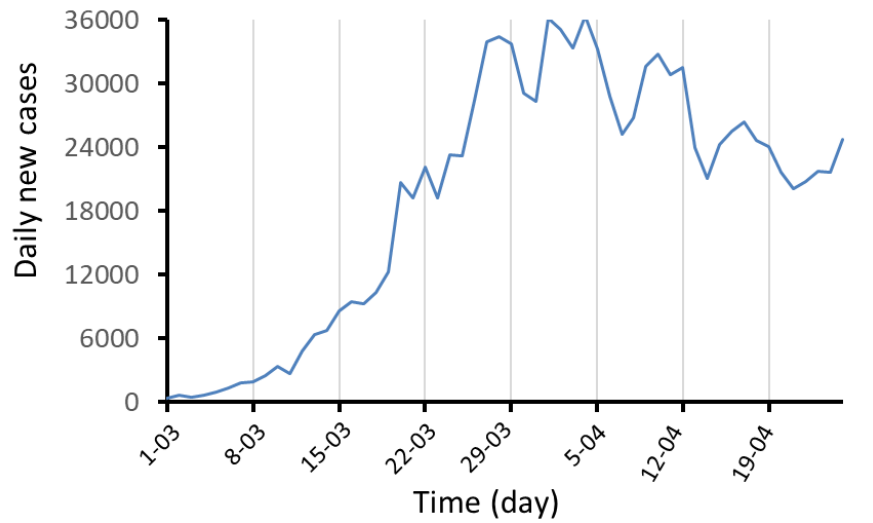
Catalunya, 01-05-2020



Index 1: Spreading rate and weekend effect

EU+EFTA+UK: weekend effect

Daily data

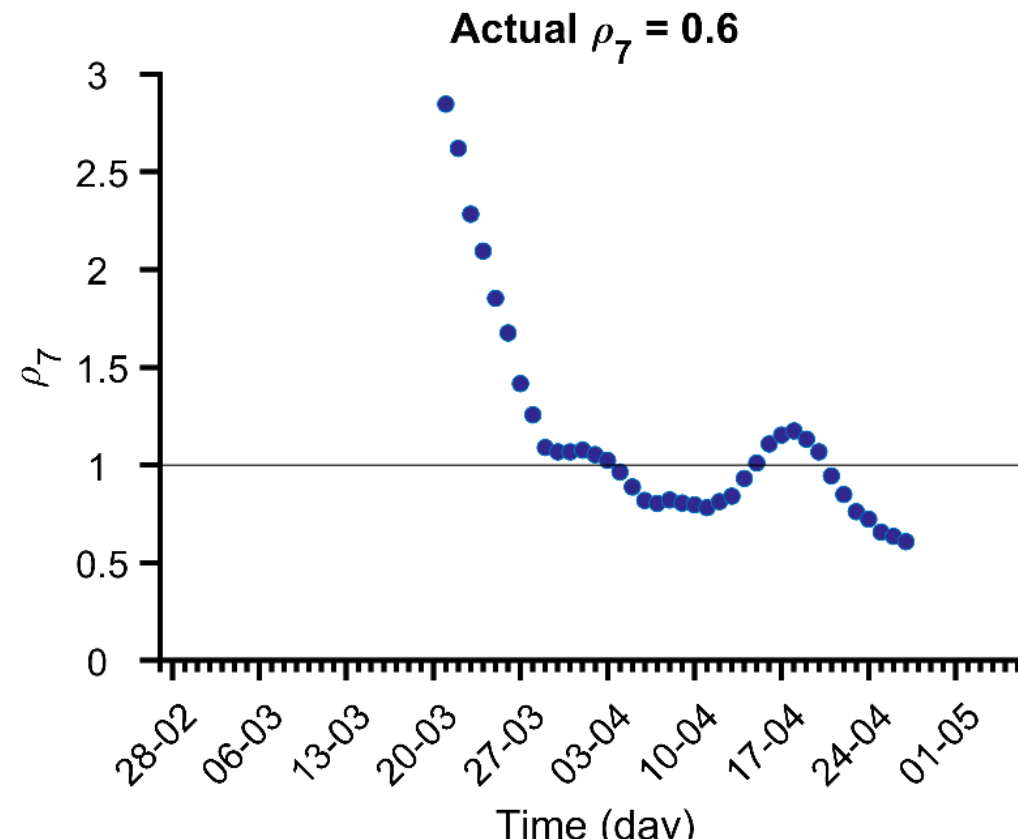


7-day moving average

Index 1: Spreading rate

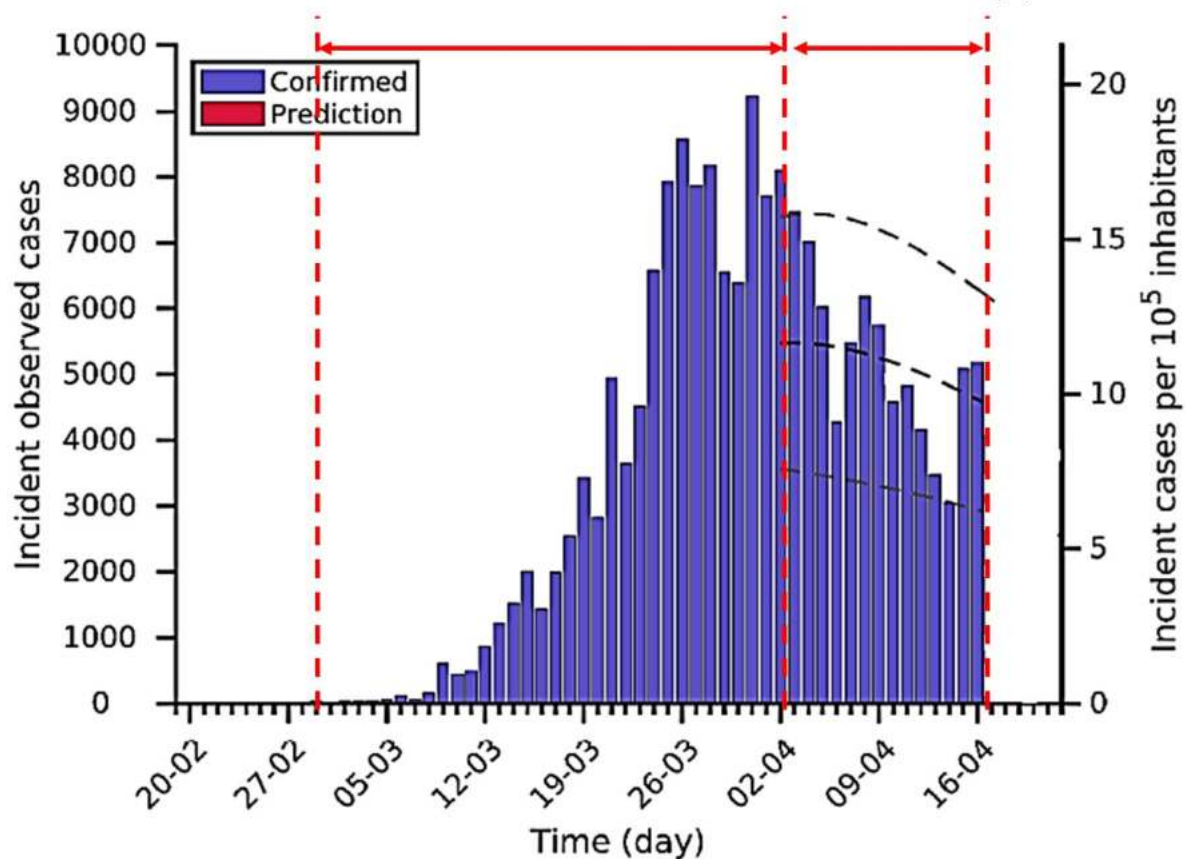
Empirical evaluation of spreading rate: $\rho_t \rightarrow \rho_7$

$$\rho_7 = \frac{\rho(t-3) + \rho(t-2) + \rho(t-1) + \rho(t) + \rho(t+1) + \rho(t+2) + \rho(t+3)}{7}$$



Index 2: 14-day attack rate

- ✓ 14-day cumulative incidence (14-day attack rate, A_{14}) is used as an indicator of **active cases** (ECDC, Ministerio...).



Infectious pool → Contagious potential at spreading rate ρ

Index 3: Effective Potential Growth (EPG)

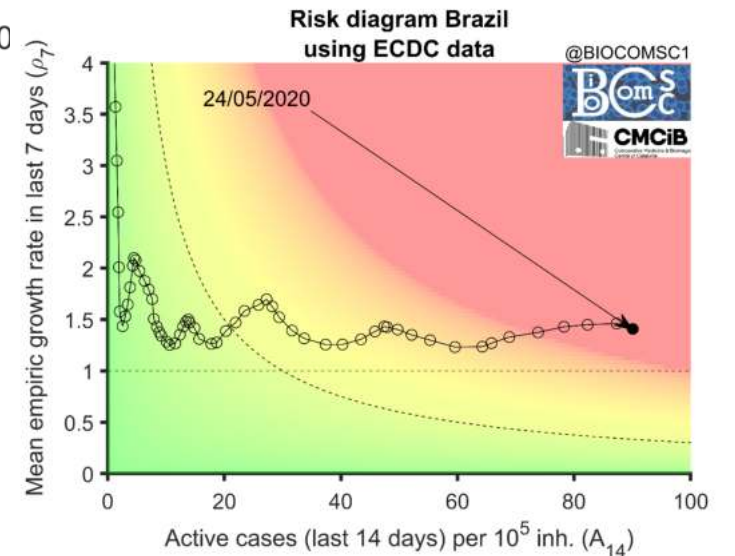
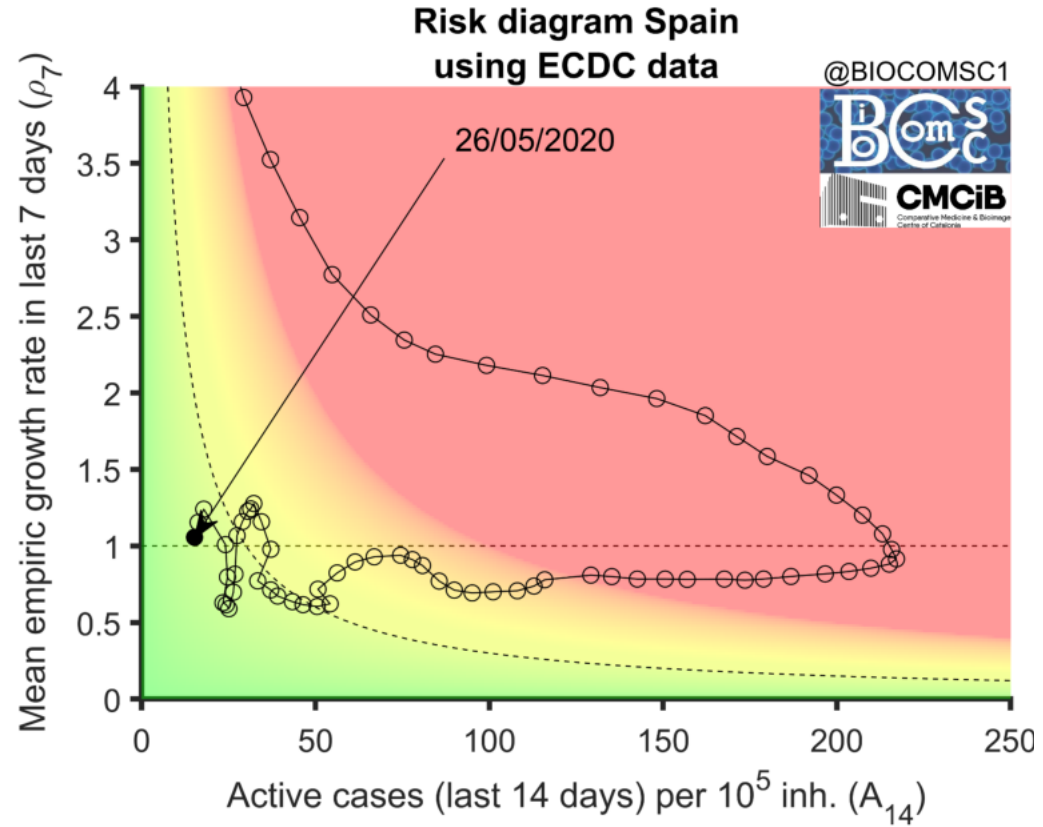
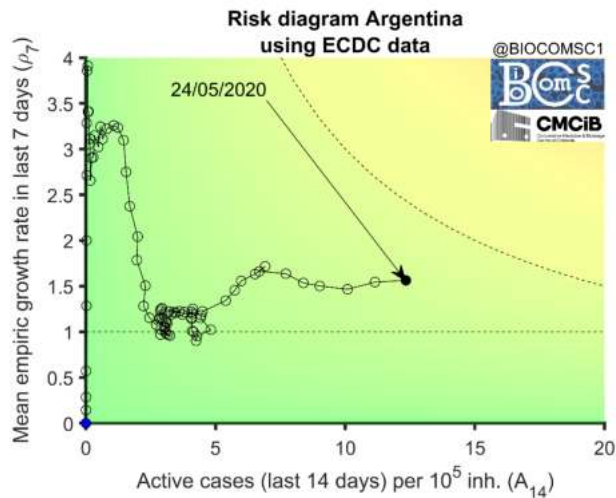
Effective potential growth is given by:

- Spreading rate ρ (ρ_7)
- Active cases (A_{14})

$$\longrightarrow EPG = \rho_7 \cdot A_{14}$$

Country	Reported data						Indexes		
	Cumulative cases	Attack rate /10 ⁵ inh.	Cumulative deaths	Mortality /10 ⁵ inh.	Active cases (last 14 days)	14-day attack rate /10 ⁵ inh.	$\rho_7^{(1)}$	EPG _{REP} ⁽²⁾	EPG _{EST} ⁽³⁾
Spain	216,582	467.3	25,100	54.2	30,150	65.1	0.69	45	652
Italy	209,328	352.2	28,710	48.3	33,403	56.2	0.79	44	623
United Kingdom	182,260	274.3	28,131	42.3	68,043	102.4	1.02	104	1,780
Germany	162,496	198.4	6,649	8.1	22,599	27.6	0.67	18	86
France	130,979	202.4	24,760	38.3	19,158	29.6	0.69	20	410
Belgium	49,517	436.0	7,765	68.4	12,334	108.6	0.65	71	1,182
Netherlands	40,236	236.9	4,987	29.4	8,647	50.9	0.64	33	422
Switzerland	29,734	347.0	1,466	17.1	2,412	28.1	0.75	21	110
Portugal	25,190	242.9	1,023	9.9	5,505	53.1	0.66	35	154
Sweden	22,082	224.5	2,669	27.1	8,260	84.0	0.98	83	1,400
Ireland	21,176	448.1	1,265	26.8	6,418	135.8	0.74	100	707
Austria	15,558	178.6	596	6.8	896	10.3	0.97	10	39
Poland	13,375	35.0	664	1.7	4,633	12.1	0.94	11	69
Romania	12,732	64.4	771	3.9	4,314	21.8	0.94	20	143
Denmark	9,407	164.7	475	8.3	2,165	37.9	0.89	34	182
Norway	7,759	144.5	204	3.8	775	14.4	0.76	11	31
Czech Republic	7,755	73.1	245	2.3	1,101	10.4	0.79	8	28
Finland	5,179	94.1	220	4.0	1,498	27.2	0.87	24	123
Luxembourg	3,812	661.8	92	16.0	275	47.7	0.74	35	ND
Hungary	2,998	30.7	340	3.5	1,082	11.1	0.87	10	136
Greece	2,620	23.4	143	1.3	413	3.7	0.61	2	14
Croatia	2,088	49.6	77	1.8	256	6.1	0.55	3	ND
Iceland	1,798	493.6	10	2.7	38	10.4	0.50	5	ND
Estonia	1,699	129.5	53	4.0	187	14.3	0.79	11	ND
Bulgaria	1,594	22.4	72	1.0	716	10.0	1.10	11	ND
Slovenia	1,439	69.2	94	4.5	122	5.9	0.87	5	ND
Slovakia	1,407	25.8	24	0.4	318	5.8	0.30	2	ND
Lithuania	1,406	48.3	46	1.6	108	3.7	NA	NA	NA
Latvia	871	44.2	16	0.8	159	8.1	0.91	7	ND
Cyprus	864	73.8	20	1.7	103	8.8	1.10	10	ND
Malta	467	108.9	4	0.9	41	9.6	ND	ND	ND
Liechtenstein	83	215.3	1	2.6	1	2.6	ND	ND	ND

Following up on deconfinement: Risk diagrams



Following up on deconfinement: How to define high risk. Red colour

DTL: Daily PCR tests performed per 100.000 people.

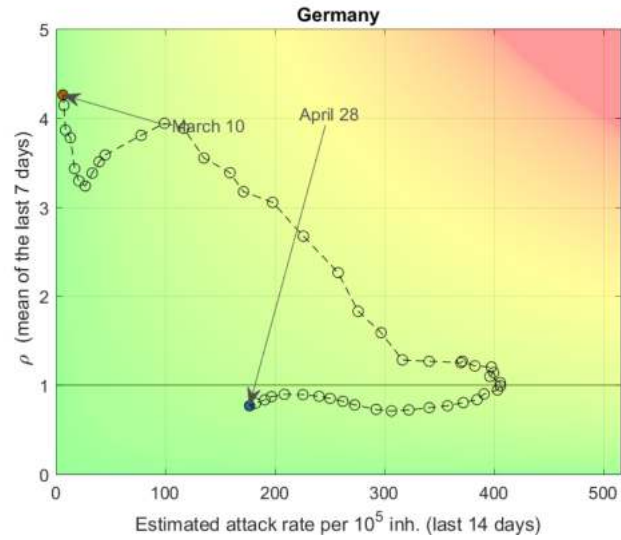
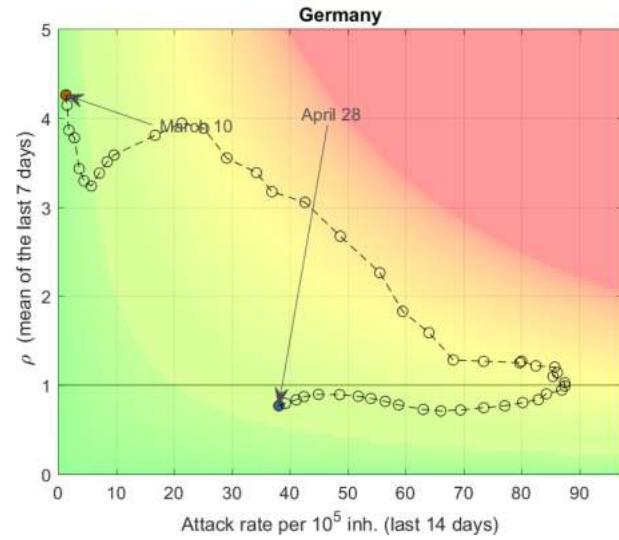
Country	Population (M)	Daily tests*	DTL	DR (%)	IA14 (rep.)	IA14 (est.)	TCF=DTL/IA14e
USA	328	320.000	100	15*	88	590	0.17
Italy	60	60.000	100	7	26	370	0.27
Spain	47	40.000	85	8	23	290	0.29
France	67	24.000	35	5	18	360	0.10
Germany	83	67.000	80	20	15	75	1.01
UK	66.5	65.000	100	7	87	1240	0.08
Denmark	6	15.000	250	20	26	130	1.92
South Korea	51.5	15.000	30	40*	5	12	2.5
Israel	9	15.000	165	60*	4	7	23.6



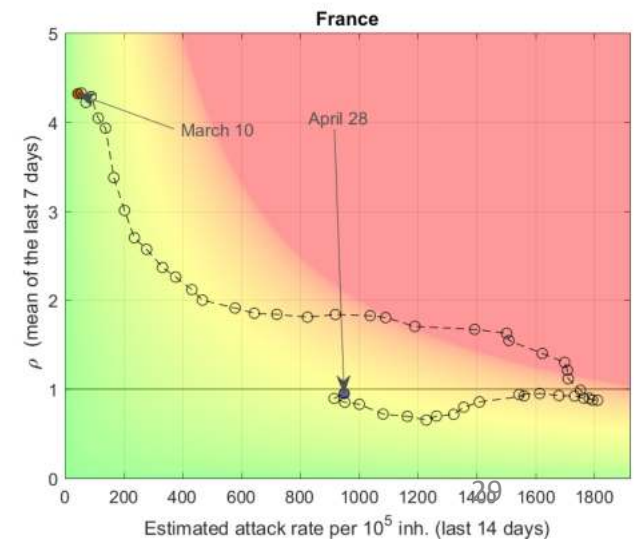
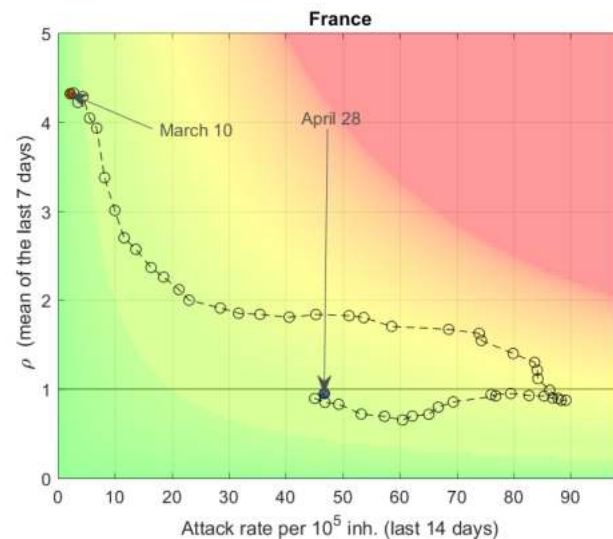
If there are 1000-2000 active cases per 100.000 people with DTL=100 it is impossible to control the epidemics by contact tracing

Following up on deconfinement: Risk diagrams. Reported vs estimated.

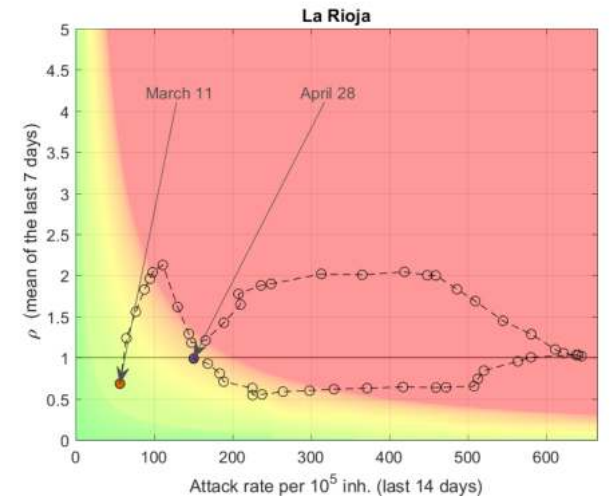
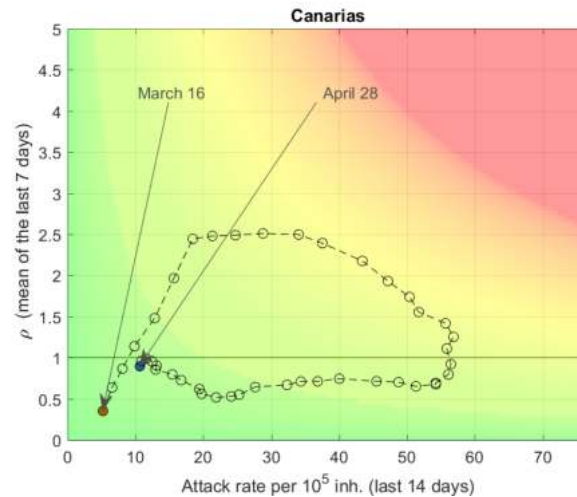
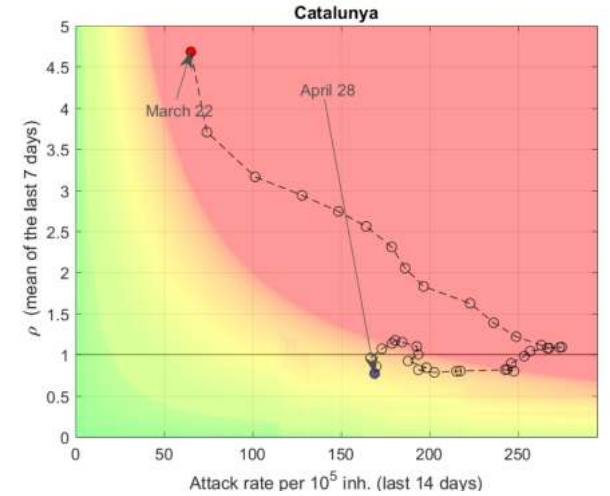
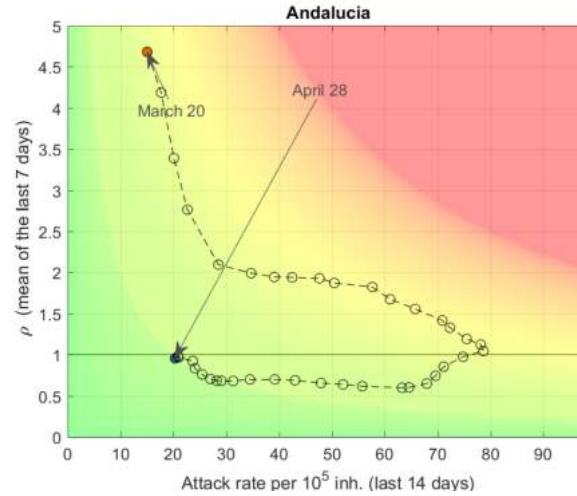
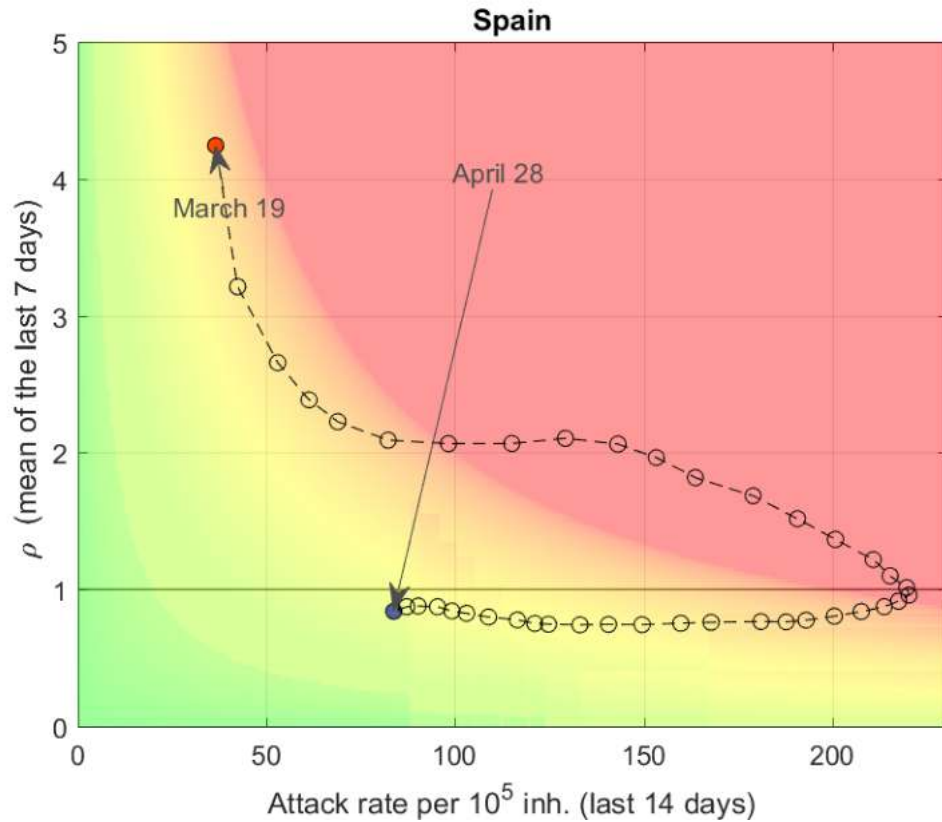
Germany: real situation better than reported



France: real situation worse than reported



Following up on deconfinement: Risk diagrams. The importance of the region.

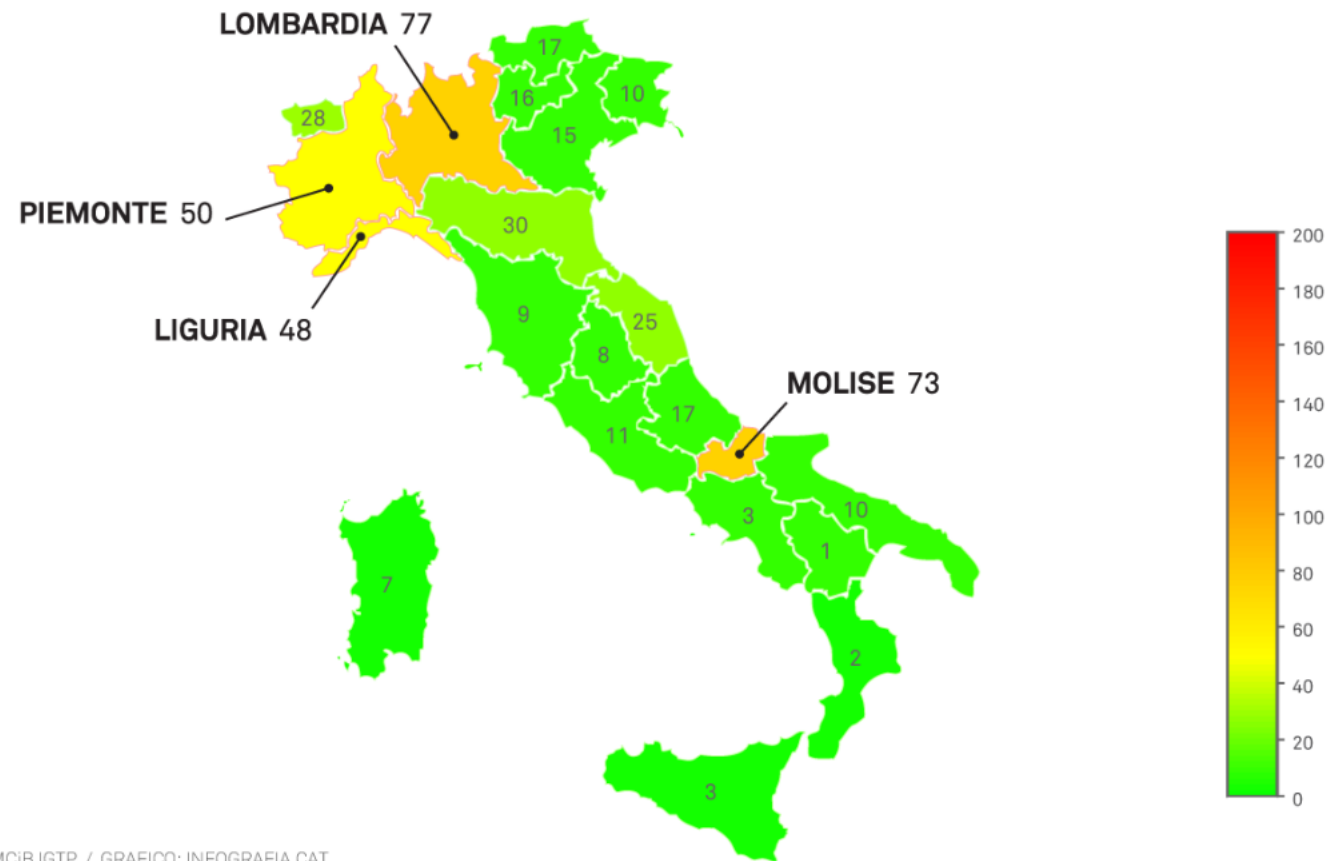


Provinces? Sanitary regions? Lower level?

Successful identification of super-spreader

RISCHIO EPIDEMIOLOGICO PER REGIONI

Dati dell'11 maggio



KEY IDEAS

- ✓ Risk evaluation often requires a short-term predictive model.
- ✓ Risk evaluation requires the proper assessment of the global picture of the situation with the elaboration of estimations that must be validated by surveys.
- ✓ Risk evaluation needs the development of more local or case-specific analysis that need to be set in the proper general perspective.

Thanks for your (virtual) attention!



CMCiB

Comparative Medicine & Bioimage
Centre of Catalonia



COL·LEGI D'ACTUARIS DE CATALUNYA

www.actuaris.org
actuaris@actuaris.org