institute imdea networks

Developing the Science of Networks

Using Indirect Reporting to Monitor the COVID-19 Pandemic

Antonio Fernández Anta

and the CoronaSurveys Team

https://coronasurveys.org/team/

COVID-19, March 13th, 2020

EL PAÍS

www.elpais.com

EL PERIÓDICO GLOBAL

VIERNES 13 DE MARZO DE 2020 | Año XLV | Número 15.581 | EDICIÓN MADRID | Precio: 1,70 euros

ALEMANIA El servicio secreto vigila al ala más radical de la ultraderechista AfD P2

EE UU Detenidos 200 nar cartel heredero de El C

EL AVANCE DEL CORONAVIRUS DESAFÍA AL ESTADO

España, en emergencia

España ya es el segundo país de la UE con más contagiados: 3.142 casos

Sánchez propone un Presupuesto "extrasocial" y Arrimadas ofrece su voto

Casado critica p insuficiente el p apoyará en el C

España ya es el segundo país de la UE con más contagiados: 3.142 casos

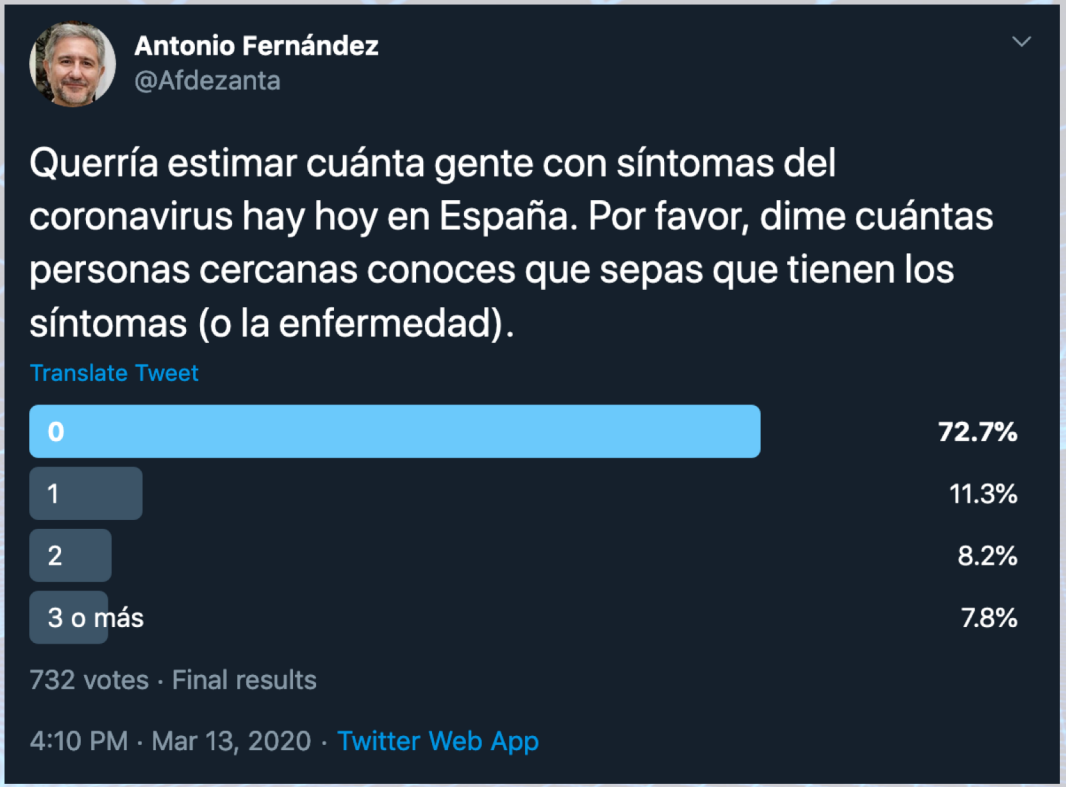
3,142 confirmed cases.
How many actual cases?

Twitter Survey

From confirmed deaths (a posteriori):

- 13 days onset to death
- CFR = 1.38%
- 4,089 deaths by March 27th
- 4,089/1.38% = **296,304 cases**

Contacts (reach)	Cases estimate
150 people (Dunbar)	**160,113 cases**
77 people	**311,909 cases**



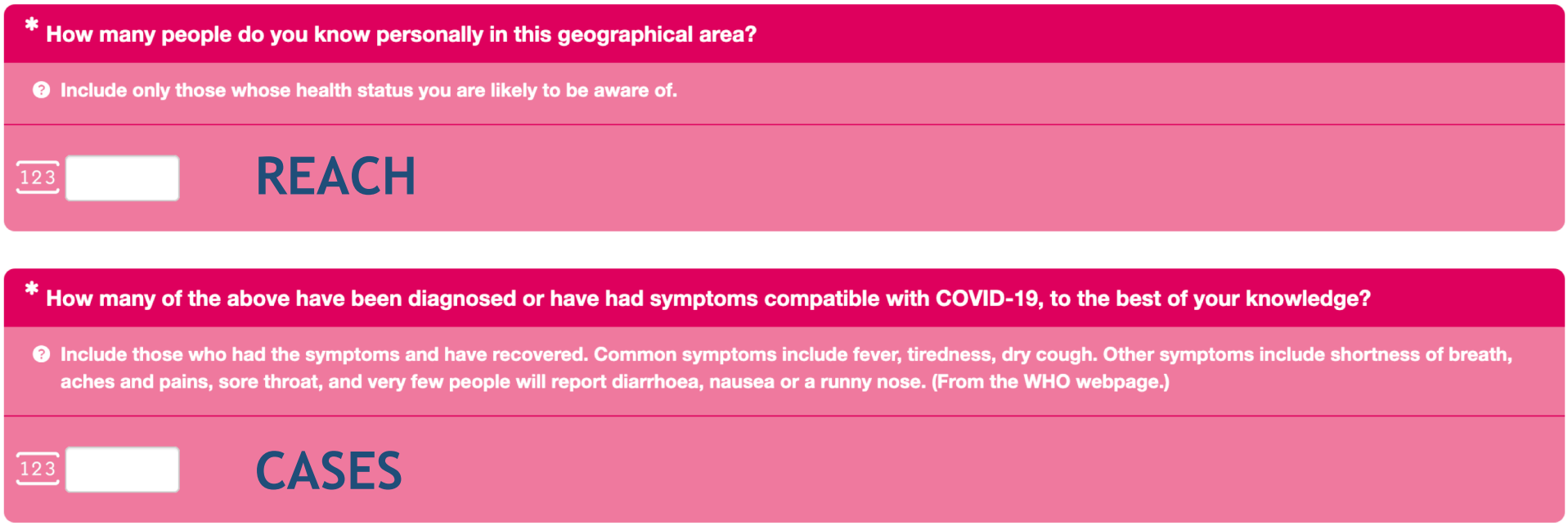
CoronaSurveys Project

coronasurveys.org

@CoronaSurveys: Monitoring the Incidence of COVID-19 via Open Surveys

- Survey with indirect reporting to monitor the evolution of the pandemic
- Available in more than 60 languages for all countries (at regional level in 150)

Indirect Reporting



The incidence is estimated with the Network Scale-up Method (NSUM) [Bernard et al, 1991]:

$$p = \frac{\sum CASES}{\sum REACH}$$

Direct vs Indirect Reporting

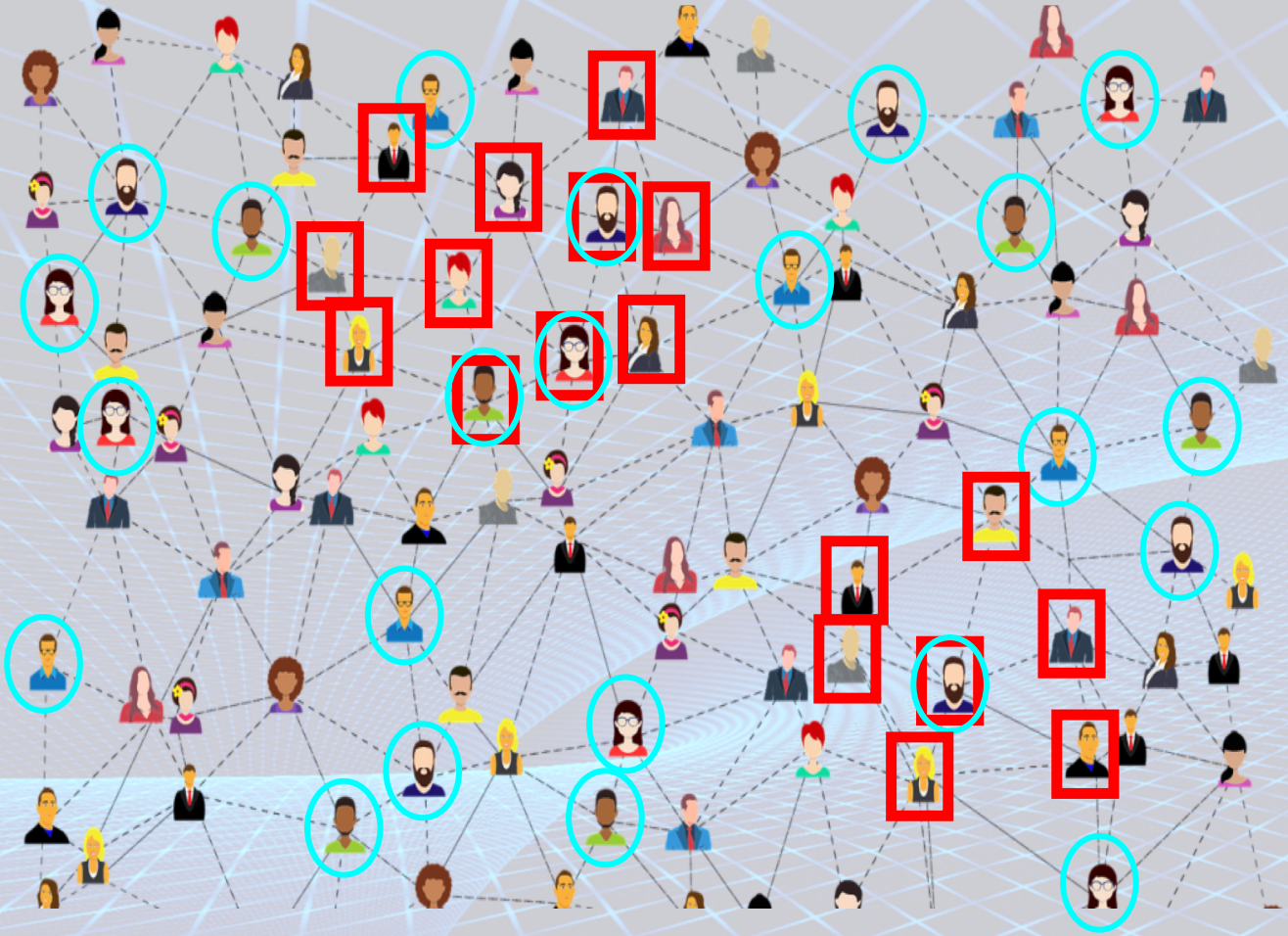
Ground truth:
24/102 = 23.5%

Direct vs Indirect Reporting

Estimate direct:
2/3=66.7%

Ground truth:
24/102 = 23.5%

Direct vs Indirect Reporting



Estimate direct:
2/3=66.7%

Ground truth:
24/102 = 23.5%

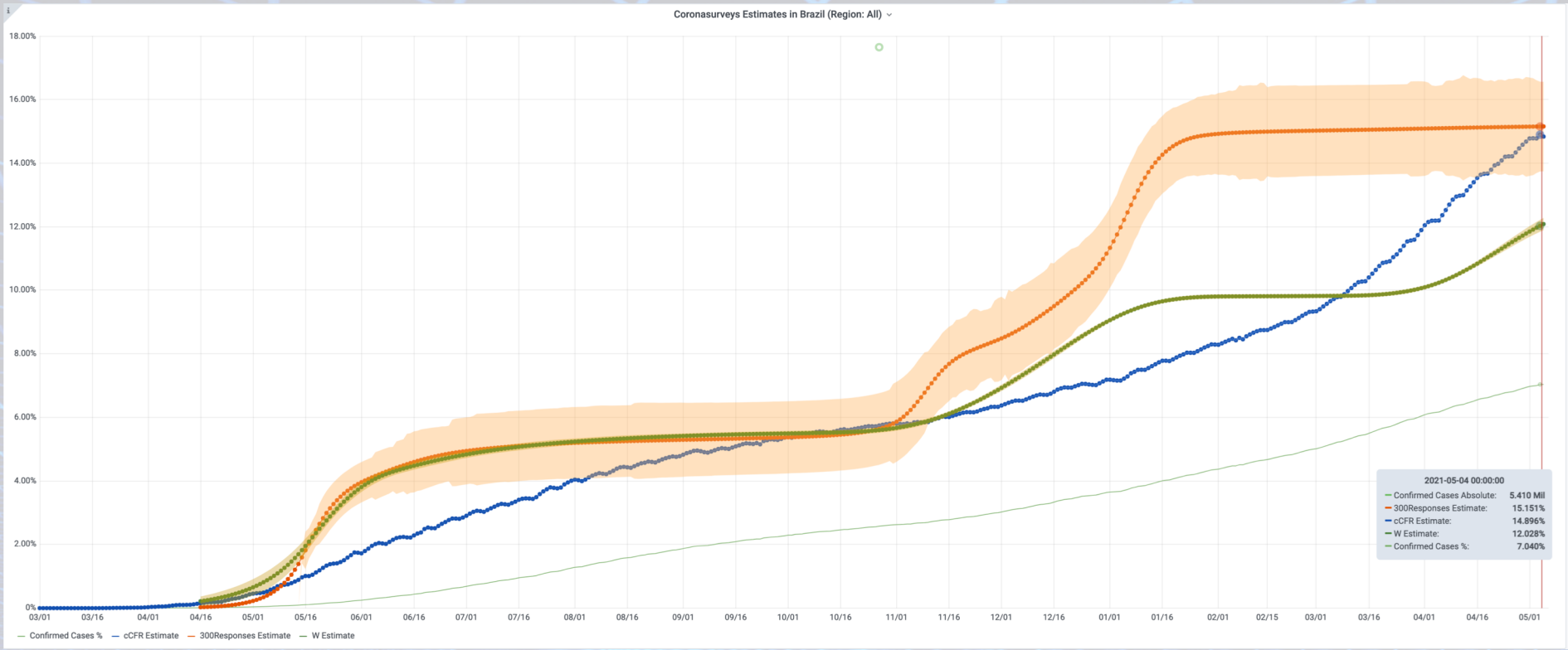
Estimate NSUM:
(2+1+2)/(7+7+6)=
5/20=25%

Known Biases

- Transmission and recall errors: The participant
  - Is not aware of health status of some contacts
  - Does not recall all contacts or some infected contacts
  - Wisdom of the crowd
- Barrier effect: All contacts of a participant are close to her
  - Collect “enough” responses
  - Collect responses at regional level, and aggregate if convenient
- Response bias: Deliberate misreporting
  - Filter out outliers and inconsistent responses

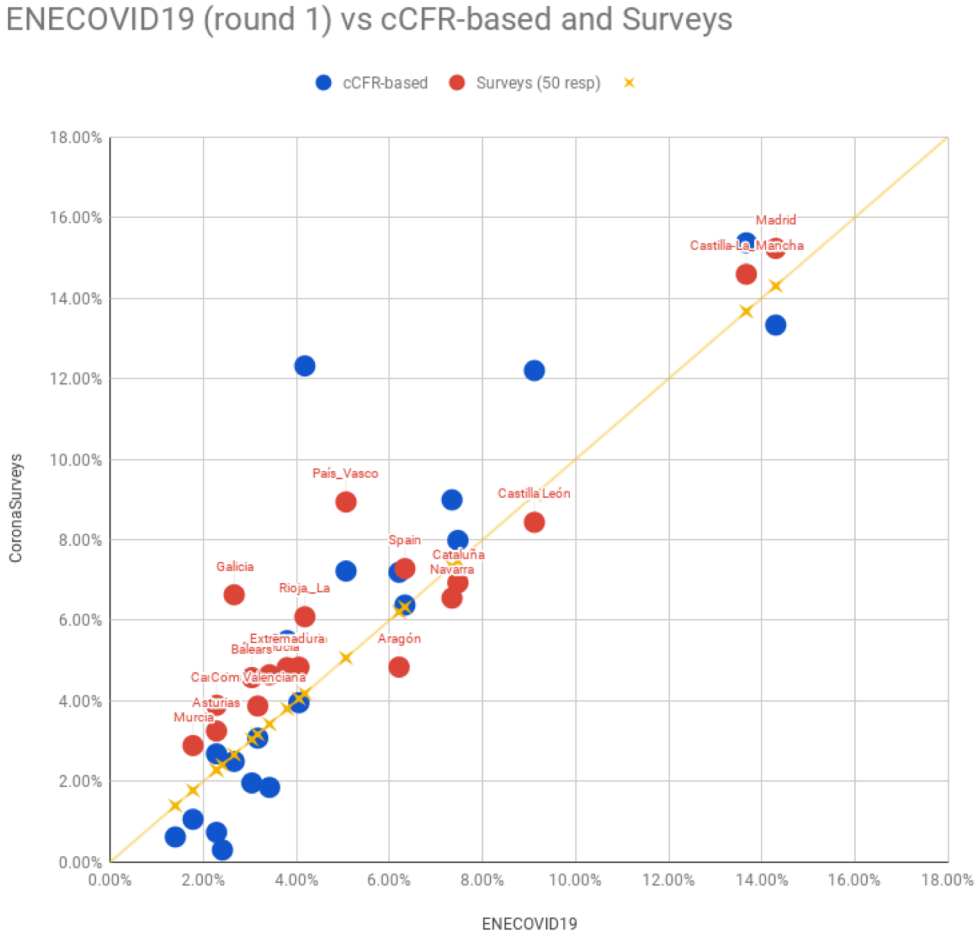
I. Laga, L. Bao, X. Niu. Thirty Years of The Network Scale up Method. arXiv:2011.12516 [stat.ME], 2020

Country Estimates



T Russel et al. Using a delay-adjusted case fatality ratio to estimate under-reporting. Centre for Mathematical Modelling of Infectious Diseases Repository (2020)

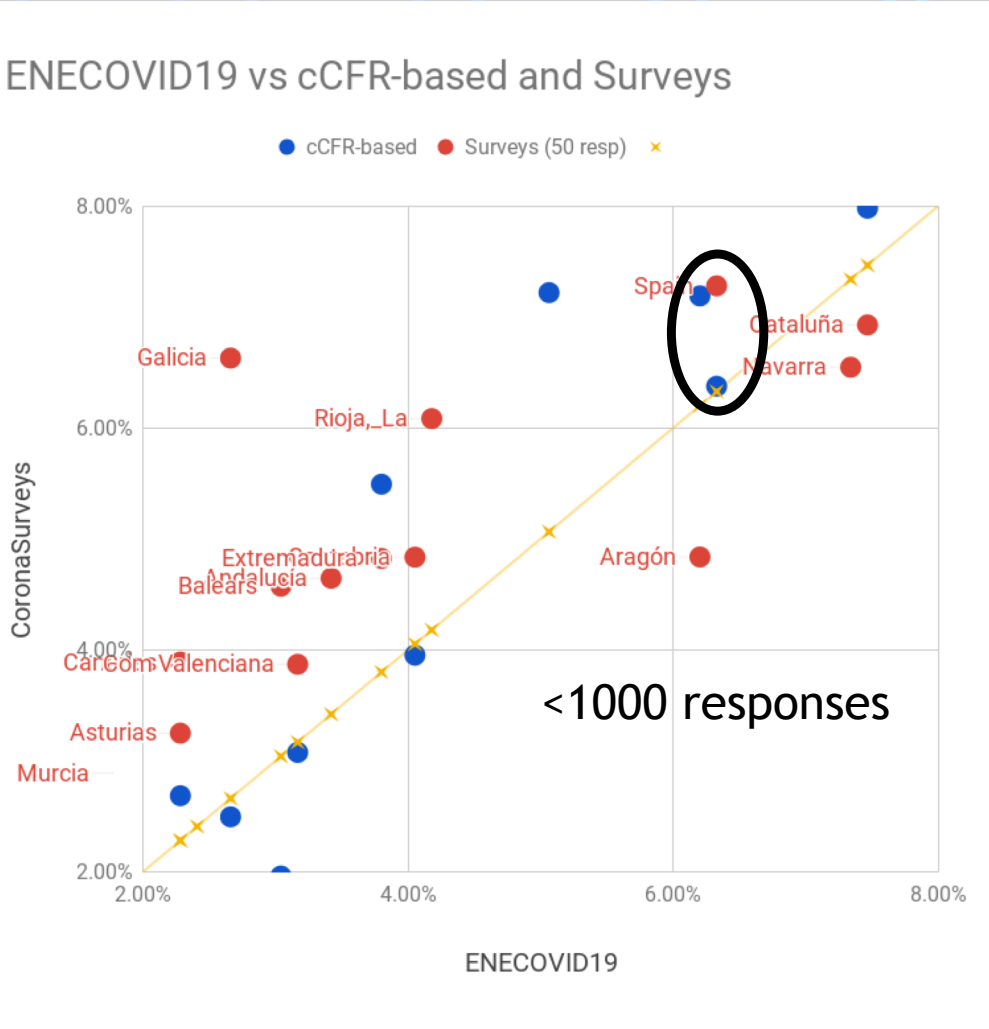
Validation: Serology Study & cCFR



Pollán, Marina, et al. "Prevalence of SARS-CoV-2 in Spain (ENE-COVID): a nationwide, population-based seroepidemiological study." *The Lancet* (2020)

Garcia-Agundez, Augusto, et al. "Estimating the COVID-19 Prevalence in Spain with Indirect Reporting via Open Surveys." *Frontiers in Public Health* 9 (2021)

Validation: Serology Study & cCFR

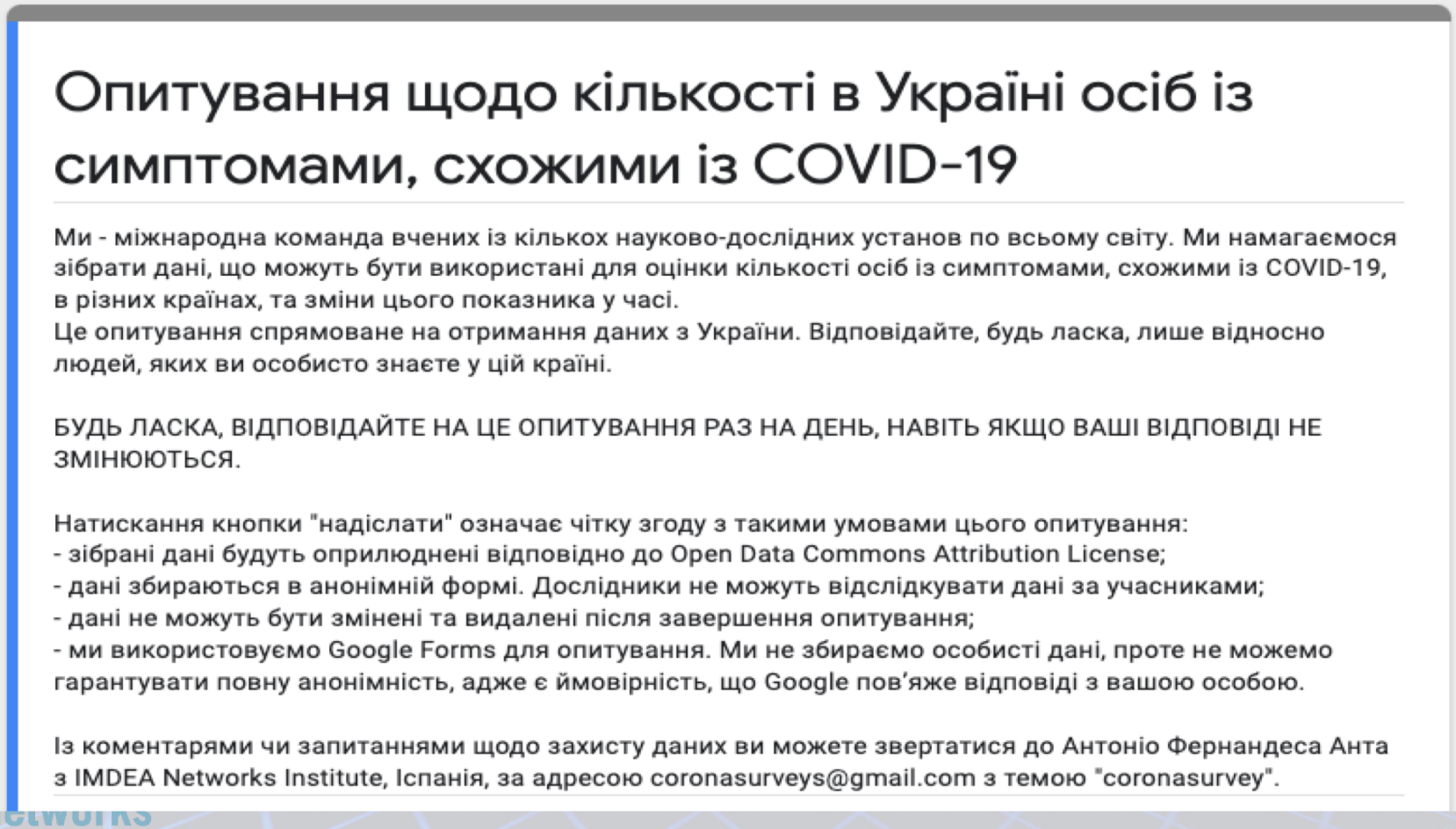


Pollán, Marina, et al. "Prevalence of SARS-CoV-2 in Spain (ENE-COVID): a nationwide, population-based seroepidemiological study." *The Lancet* (2020)

- 60,000 people tested (2 tests in each wave)
- Medical teams all over the country
- 2 weeks of collecting samples
- $$$

Results in Ukraine

- Thanks to Anna Ishchenko and Oleksiy Kebkal a survey was deployed in Ukraine in April 2020



Опитування щодо кількості в Україні осіб із симптомами, схожими із COVID-19

Ми - міжнародна команда вчених із кількох науково-дослідних установ по всьому світу. Ми намагаємося зібрати дані, що можуть бути використані для оцінки кількості осіб із симптомами, схожими із COVID-19, в різних країнах, та зміни цього показника у часі.
Це опитування спрямоване на отримання даних з України. Відповідайте, будь ласка, лише відносно людей, яких ви особисто знаєте у цій країні.

БУДЬ ЛАСКА, ВІДПОВІДАЙТЕ НА ЦЕ ОПИТУВАННЯ РАЗ НА ДЕНЬ, НАВІТЬ ЯКЩО ВАШІ ВІДПОВІДІ НЕ ЗМІНЮЮТЬСЯ.

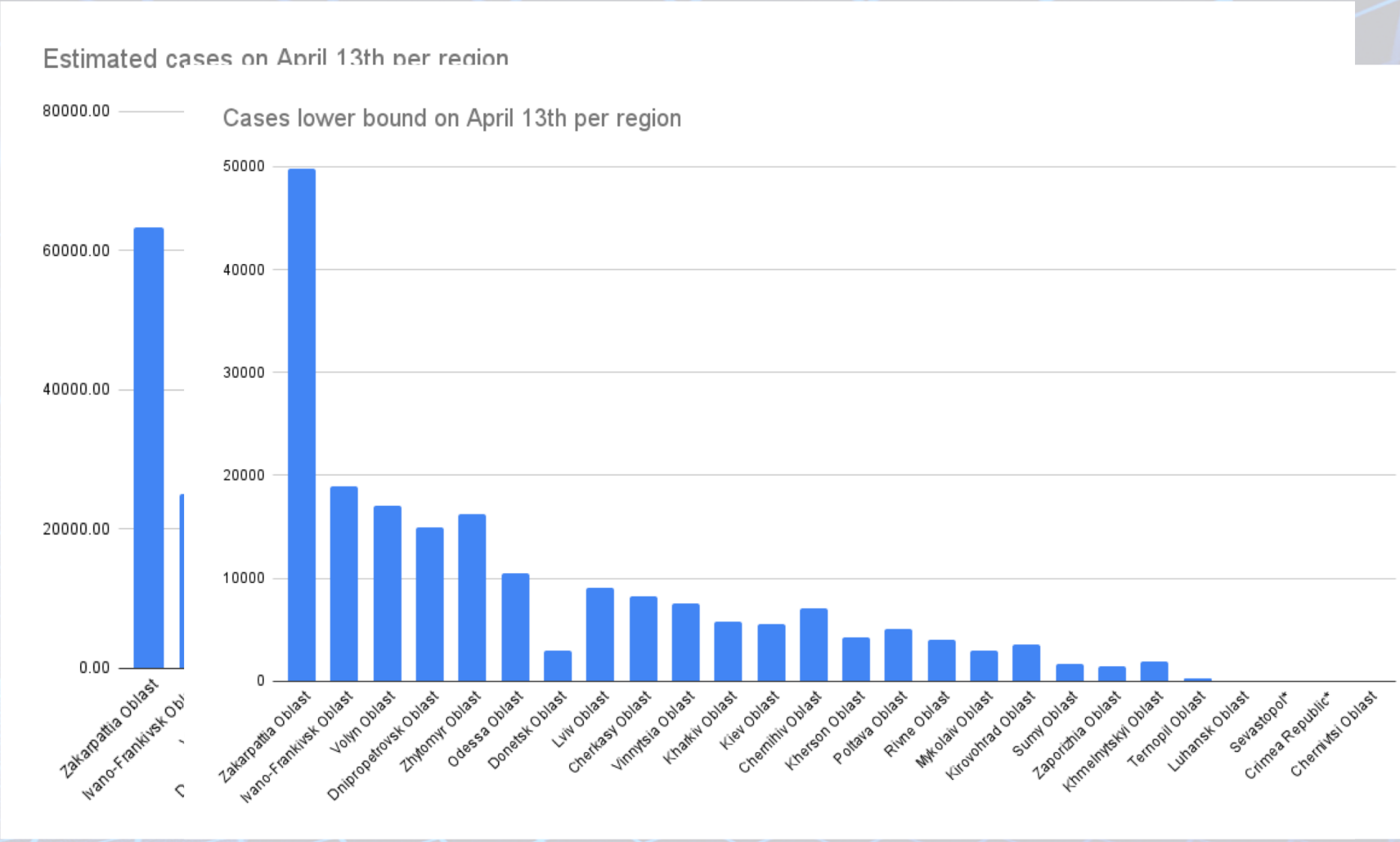
Натискання кнопки "надіслати" означає чітку згоду з такими умовами цього опитування:
- зібрані дані будуть оприлюднені відповідно до Open Data Commons Attribution License;
- дані збираються в анонімній формі. Дослідники не можуть відслідкувати дані за учасниками;
- дані не можуть бути змінені та видалені після завершення опитування;
- ми використовуємо Google Forms для опитування. Ми не збираємо особисті дані, проте не можемо гарантувати повну анонімність, адже є ймовірність, що Google пов'яже відповіді з вашою особою.

Із коментарями чи запитаннями щодо захисту даних ви можете звертатися до Антоніо Фернандеса Анта з IMDEA Networks Institute, Іспанія, за адресою coronasurveys@gmail.com з темою "coronasurvey".

Results in Ukraine

- Thanks to Anna Ishchenko and Oleksiy Kebkal a survey was deployed in Ukraine in April 2020
- In one week by April 13th had 2000+ responses
- Estimates:
  - Average reach: ~90 persons
  - CoronaSurveys cases: ~300,000 (min. 200,000 95%CI)
  - Confirmed cases: 3,102
  - Cases from 209 deaths on April 26th: 15,145
- On July 13th, officially
  - ~55,000 cases
  - Cases from 1624 deaths on July 26th: 117,754

Cases in Ukraine per Region



Beyond Cumulative Incidence

↳ How many are still sick?

123

↳ How many started with symptoms in the latest 7 days?

123

↳ If you get sick, how many of the people you know personally would be aware of it?

123

↳How many of the people you know personally have been given a COVID-19 vaccine?

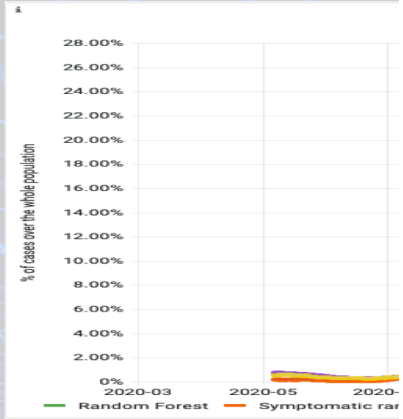
123

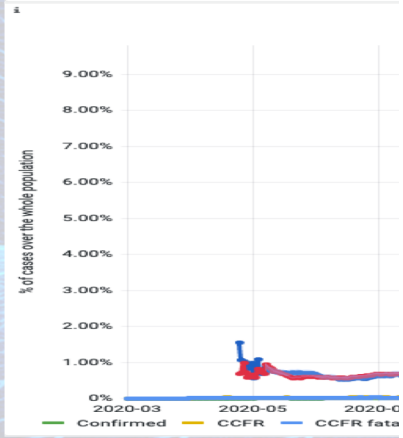
↳How many of the people you know personally refuse to be vaccinated?

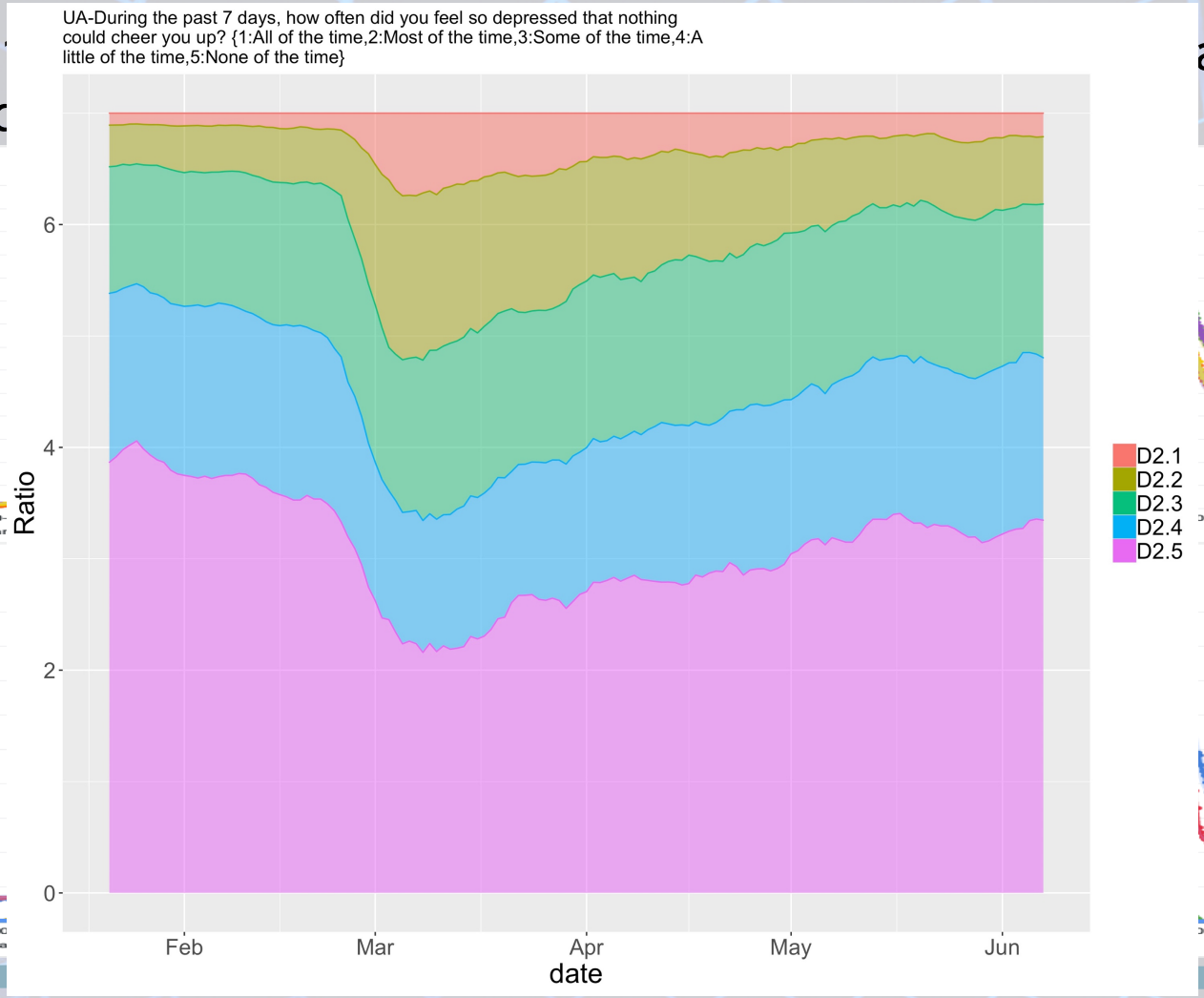
123

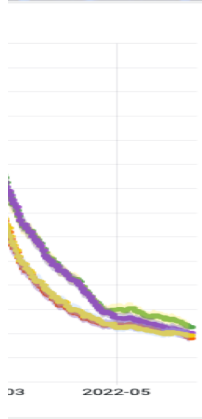
Integration with Symptom Survey

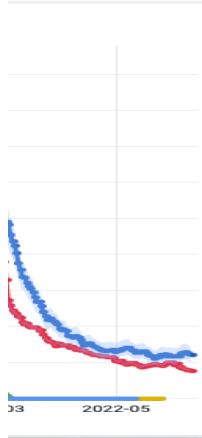
- The COVID-
  has lots of c
  and FB











Conclusions

- Indirect reporting (Aggregated Relational Data) with Network Scale-up Method are a useful tool to track the pandemic evolution
- Easy, fast, and cheap to deploy
  - Possible game changer for LMIC
  - Complements and contrasts the available data
- Good estimates even with limited number of responses

Future Work

- Campaigns to maintain constant flow of responses
- Deploy questions and mechanisms to compensate for the known biases
- Integrating with other data sources
  - CMU/UMD COVID-19 Symptom Survey
  - Oxford data on interventions
  - Google data on mobility
- Forecasting methods with all the data
- Propose interventions

Thank you!
https://coronasurveys.org

- Augusto Garcia-Agundez et al. Estimating the COVID-19 Prevalence in Spain with Indirect Reporting via Open Surveys, Front. Public Health, 09 April 2021
- INESC TEC. Estimating active cases of COVID-19: The unknown matters, Medium.com, July 16th, 2020.
- Oluwasegun Ojo et al. CoronaSurveys: Using Surveys with Indirect Reporting to Estimate the Incidence and Evolution of Epidemics, arXiv:2005.12783 v2 [cs.DC], June 2020. KDD Workshop on Humanitarian Mapping, San Diego, California USA, August 24, 2020.
- Carlos Baquero et al. Measuring Icebergs: Using Different Methods to Estimate the Number of COVID-19 Cases in Portugal and Spain, medRxiv, April 2020.