A random population sample and 10 follow-ups: one year of COVID-19 pandemic seen through the CHRIS COVID-19 study

Giulia Barbieri[1*], Martin Gögele[1], Roberto Melotti[1], Christian X. Weichenberger[1], Luisa Foco[1], Daniele Giardiello[1], Daniele Bottigliengo[1], Peter P. Pramstaller[1], Cristian Pattaro[1**]

eurac research

[1] Institute for Biomedicine, Eurac Research, Affiliated Institute of the University of Lübeck, Bolzano, Italy

*giulia.barbieri@eurac.edu **cristian.Pattaro@eurac.edu

1 INTRODUCTION

The **Cooperative Health Research In South Tyrol (CHRIS) COVID-19 study** was established in July 2020 to monitor SARS-CoV-2 infections in the middle and upper Val Venosta (South Tyrol, Italy). In the first stage of the study, a population-representative sample of 1812 individuals, reflecting the age and sex distribution of the adult population in the area, was invited to participate.

Aims

a) To estimate the incidence of SARS-CoV-2 infection during the first pandemic wave
b) To characterize the spread and course of the pandemic in between February 2020 and May 2021.

2 METHODS

Study design

At **baseline**, participants

- were submitted an **online screening questionnaire** including questions on COVID-19 related anamnesis, symptoms and lifestyle;
- were contextually invited to undergo a nasopharyngeal swab for **PCR test** and a **serum antibody test**.

Follow-up: all participants testing negative at the baseline assessment were administered a shorter **online questionnaire every 4 weeks** to update anamnesis and symptomatology. Participants reporting positive tests or close contacts with infected individuals were invited to participate to a serological follow-up.

Statistical analyses

- Incidence was estimated with the **Clopper-Pearson** method for extreme proportions.
- Association between baseline incidence and anamnesis and symptoms was assessed using **Fisher's exact test** (significance level: 0.001).
- The temporal trend of the total number of reported symptoms was assessed by fitting **zero-inflated negative binomial models with random intercepts** on individuals, using the month on which symptoms were reported as a predictor.
- We used **longitudinal cluster analysis** based on the k-means method to model the symptoms' dynamic over time, with the optimal number of clusters defined according to the Calinski & Harabatz criterion.

3 RESULTS

Baseline

Overall, 845 were recruited.

- Estimated **cumulative incidence of SARS-CoV-2 infection** in the study area until Aug 2020: **0.95%** (95% confidence interval, CI, **0.41-1.86%**), calibrated to the age and sex distribution of the population.
- **Positivity** was associated with: (1) having **undergone a nasopharyngeal swab**; (2) **having had a previous positive serological test**; (3) **having been isolated** because of suspected or confirmed SARS-CoV-2 infection; (4) **number of reported symptoms**; and (5) specific symptoms such as **loss of taste**, **loss of smell**, and **dry cough**.

Follow-up

- 699 (84%) of the 836 participants testing negative at the baseline completed 1-to-10 follow-up questionnaires in the **period Sep 2020 – May 2021**.
- Of them, 60.5% (95%CI 57.1-63.8%) underwent 1+ nasopharyngeal swabs and **20.5% (95%CI 17.1-24.3%) reported to have tested positive at least once**.
- The **cumulative number of reported symptoms for each participant peaked in October 2020 (2nd pandemic wave) and in February 2021 (3rd pandemic wave)**, closely resembling the shifted trends of hospitalizations and intensive care admissions reported by the local healthcare system in the same period.
- This was confirmed by the **zero-inflated mixed-model results** (**Table 1**).
- Similar patterns emerged from the **dynamic of individual symptoms** (**Figure 1**)**: cluster analysis** showed that the main peaks were associated with clusters containing the most common COVID-19-related symptoms. Lower intensity clusters reflected flat patterns mainly comprising generic symptoms.
- **Random forest and dynamic correlation analyses** based on individual symptoms confirmed the symptomatic profile associated with COVID-19 reported by previous studies (**Figure 2**).

4 CONCLUSIONS

During the first pandemic phase, the middle and upper Val Venosta had lower incidence of SARS-CoV-2 infections than nearby regions. Subsequently, the trend became more similar to that observed in South Tyrol, with a heavy second wave in autumn 2020 and a further third wave in early 2021.

Data collection was conditioned by public health interventions aimed at countering the pandemic itself, which may have also altered individual behavior. These circumstances may translate into selection bias.

Furthermore, testing capacity and screening guidelines have been following different patterns through the period, altering the probability of having undergone a PCR test. Analyzing the reported symptoms in addition to measures of disease occurrence can add value to monitoring the temporal trend of infections, since not all individuals have equivalent propensity to receive a test, mainly during the acute phases of the pandemic.

Table 1: Temporal trend of the total number of self-reported symptoms, assessed by Zero Inflated Negative binomial mixed model, adjusted by gender and age

	Zero Inflated part[a]		Count part[b]	
	β (S.E.)	p-value	β (S.E.)	p-value
(Intercept)	1.08 (0.42)	0.0098	-0.58 (0.35)	0.0939
August	Reference		Reference	
September	1.06 (0.42)	0.0113	0.22 (0.34)	0.517
October	-0.07 (0.4)	0.8523	0.7 (0.29)	0.0145
November	-1.16 (0.34)	0.0007	0.82 (0.3)	0.0059
December	-1.29 (0.34)	0.0002	0.19 (0.3)	0.5229
January	-0.83 (0.36)	0.0215	0.79 (0.29)	0.0075
February	-1.1 (0.35)	0.0016	0.62 (0.29)	0.032
March	-1.96 (0.35)	< 0.0001	0.39 (0.32)	0.2157
April	-0.73 (0.38)	0.0555	0.56 (0.32)	0.0847
Age	-0.53 (0.38)	0.1722	-0.00 (0.01)	0.8791
Gender-M	0.03 (0.01)	< 0.0001	-0.38 (0.14)	0.0077

[a] *Modelling the probability of having 0 symptoms; betas are on logit-link scale*
[b] *Modelling the probability of having j symptoms with j ≥1; betas are on log-link scale*

Figure 1: Occurrence (%) of single symptoms (a) and mean trajectories of the four clusters (b)

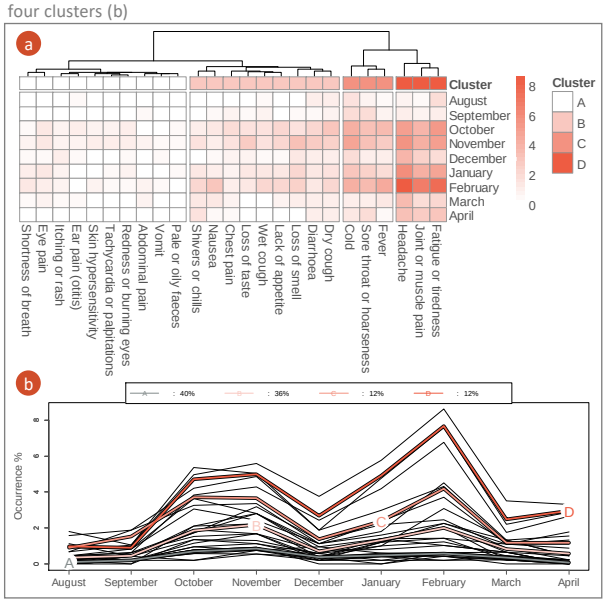
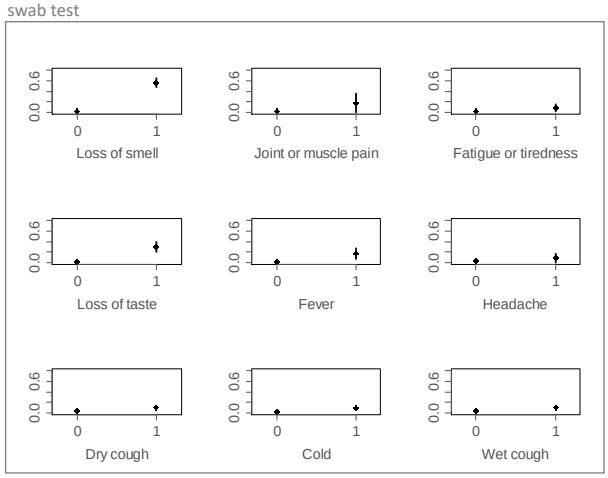


Figure 2: Marginal effects of the main predictors of a self-reported positive swab test



Acknowledgements

The research leading to these results is part of the PACE project funded by the Legge Provinciale 14/2006 of the Provincia Autonoma di Bolzano/Bozen – CUP D52F20000770003