Using Survey Data to Estimate the Impact of the Omicron Variant on Vaccine Efficacy against COVID-19 Infection

Jesús Rufino
IMDEA Networks Institute https://orcid.org/0000-0002-5863-575X
Carlos Baquero
U. Porto & INESC TEC
Davide Frey
Univ Rennes, IRISA, CNRS, Inria
Christin Glorioso
Massachusetts Institute of Technology
Antonio Ortega
U. Southern California
Nina Reščič
Jozef Stefan Institute, Department of Intelligent Systems
Julian Roberts
Gearu LTD
Rosa Lillo
U. Carlos III de Madrid
Raquel Menezes
Centre of Mathematics of U. Minho
Jaya Champati
IMDEA Networks Institute
Antonio Fernandez Anta (antonio.fernandez@imdea.org)
IMDEA Networks Institute https://orcid.org/0000-0001-6501-2377

Analysis

Keywords:

Posted Date: February 15th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1356083/v1

Using Survey Data to Estimate the Impact of the Omicron Variant on Vaccine Efficacy against COVID-19 Infection

Jesús Rufino, MSc[2,1], Carlos Baquero, PhD[3,1], Davide Frey, PhD[4,1], Christin A. Glorioso, PhD[5,1], Antonio Ortega, PhD[6,1], Nina Reščič, MSc[7,1], Julian Charles Roberts, PhD[8,1], Rosa E. Lillo, PhD[9,1], Raquel Menezes, PhD[10,1], Jaya Prakash Champati, PhD[2,1], and Antonio Fernández Anta, PhD[2,1]

[1]CoronaSurveys Team, Spain
[2]IMDEA Networks Institute, Spain
[3]U. Porto & INESC TEC, Portugal
[4]Univ Rennes, IRISA, CNRS, Inria, 35042 Rennes, France
[5]Academics for the Future of Science, Inc. & U. of California San Francisco, USA
[6]U. Southern California, USA
[7]Jožef Stefan Institute, Department of Intelligent Systems, Ljubljana, Slovenia
[8]Gearu LTD, UK
[9]U. Carlos III de Madrid, Spain
[10]Centre of Mathematics of U. Minho, Portugal

Corresponding author: Antonio Fernández Anta

Contact information:
IMDEA Networks Institute
Av. Mar Mediterráneo 22
28918 Leganés, Madrid, Spain
+34 608844802
antonio.fernandez@imdea.org

Abstract

Data collected in the Global COVID-19 Trends and Impact Surveys (UMD Global CTIS), and data on variants sequencing from GISAID, are used to evaluate the impact of the Omicron variant (in South Africa and other countries) on the prevalence of COVID-19 among unvaccinated and vaccinated population, in general and discriminating by the number of doses. In South Africa, we observe that the prevalence of COVID-19 in December (with strong presence of Omicron) among the unvaccinated population is comparable to the prevalence during the previous wave (in August-September), in which Delta was the variant with the largest presence. However, among vaccinated, the prevalence of COVID-19 in December is much higher than in the previous wave. In fact, a significant reduction of the vaccine efficacy is observed from August-September to December. For instance, the efficacy drops from 0.81 to 0.30 for those vaccinated with 2 doses, and from 0.51 to 0.09 for those vaccinated with one dose. The study is then extended to other countries in which Omicron has been detected, comparing the situation in October (before Omicron) with that of December. While the reduction measured is smaller than in South Africa, we still found, for instance, an average drop in vaccine efficacy from 0.53 to 0.45 among those vaccinated with two doses. Moreover, we found a significant negative (Pearson) correlation of around -0.6 between the measured prevalence of Omicron and the vaccine efficacy.

1 Introduction

The Omicron variant of SARS-CoV-2 has seen an expressive increase since its initial classification in November 2021 [16]. In South Africa it appears to have out-competed the Delta variant [7] and has rapidly spread into Europe and other regions. Preliminary observations also indicate that it might spread faster and might have higher immune evasiveness than previous variants [9]. While vaccination still provides a level of protection against a serious disease [20], recent results [18, 15, 11, 13] point towards a reduced level of protection against infection, especially from 15 weeks post the second dose [3], and it is likely that the number of breakthrough infections (i.e., infections among vaccinated people) will rise with the spread of Omicron. It is also possible that the rapid spread of Omicron is not only a consequence of high transmissibility but also of immune evasiveness [13]. Some of the preliminary models [22] showed that high transmissibility in combination with high immune evasiveness could lead to a concerning health system overload [12].

Since the spring of 2020, the University of Maryland in collaboration with Facebook has collected extensive survey data on self-reported symptoms, infection, testing, behavior and, more recently, vaccination status (UMD Global CTIS) [4, 24]. In mid December 2021, researchers used data from this survey concerning the Gauteng province in South Africa to define different combinations of symptoms that are associated with COVID-19 infection, and combined those with self-reported vaccination status to compare vaccine efficacy changes from a Delta dominant period to the current Omicron dominant period [25]. Their findings showed a measurable drop of efficacy towards infection for those vaccinated with two doses.

In this study we use self-reported confirmation of COVID-19 infection, from a subset of the UMD Global CTIS survey responses, to derive an improved proxy for COVID-19 active cases (using a Random Forest classifier) that tracks more closely the evolution of confirmed cases. We use this improved proxy for analysing prevalence and vaccine efficacy changes in South Africa as a whole, and in the Gauteng province, among those unvaccinated, partially vaccinated, and fully vaccinated. We also compute results in other countries that are currently experiencing a rise of Omicron cases, which show a significant negative correlation between the prevalence of Omicron and the vaccine efficacy.

2 Methods

2.1 Self-reported Survey Data

We use the responses to the UMD Global CTIS, which collects more than 100,000 responses daily across the world[1]. We have access to the responses collected by agreement with UMD and Facebook (see Supplement G). All the participants in the CTIS have declared to be at least 18 years of age. The first step is removing abnormal responses, as proposed in Alvarez et al. [2] (see Supplement A).

In order to classify the responses as positive or negative, several criteria have been proposed in the literature. In particular, we consider the following symptom-based COVID-like illness classifiers (see Supplement C for the list of symptoms collected in the survey):

- UMD CLI [6, 2]: A response is considered to be positive if it declares fever (symptom B1_1), along with cough (symptom B1_2), or shortness of breath / difficulty breathing (symptom B1_3). Otherwise, it is negative.
- Stringent CLI [25]: A response is positive if it declares anosmia (symptom B1_10), combined with fever (B1_1), muscle pain (B1_6), or cough (B1_2). Otherwise, it is negative.
- Classic CLI [25]: A response is positive if it declares cough (B1_2), combined with fever (B1_1), muscle pain (B1_6), or anosmia (B1_10). Otherwise, it is negative.
- Broad CLI [25]: A response is positive if it declares muscle pain (B1_6), combined with fever (B1_1), cough (B1_2), or anosmia (B1_10). Otherwise, it is negative.

These methods for classifying cases as positive or negative have two main limitations. First, they do not take into account diagnostic uncertainty, e.g., the same set of symptoms might be associated with some other condition. Second, these criteria are not adaptive to possible changes in the symptoms experienced as conditions change, e.g., as vaccination rates increase or new virus variants emerge. Thus, in this work, we introduce a new machine-learning-based classifier (described in Supplement B). We build a *ground-truth set* with the responses to the survey that report having been tested in the latest 14 days and know the outcome of the test. The responses in the ground-truth set are used to train a model, which is then used to determine the status of users outside that set (users who do not report test information). We use the random-forest technique to design this classifier and the corresponding results are labeled Random Forest in what follows.

[1] Except in the US, where the survey is run by CMU [21].

We refer to the values obtained with each of these five classifiers (namely, Random Forest, UMD CLI, Stringent CLI, Classic CLI, and Broad CLI) as *proxy estimates* (or proxy for short). We compare each proxy estimate with the estimate of active cases obtained from the official number of cases as described by Alvarez et al. [2], where each new case is assumed to remain active for 10 days. These last estimates are called Confirmed.

2.2 Prevalence and Efficacy Estimation

The prevalence of COVID-19 estimated by a given classifier is the ratio between the number of positive cases over the total number of responses. Then, we consider four subsets of responses:

- Unvaccinated: Participants that respond negatively to the question "V1: Have you had a COVID-19 vaccination?"
- Vaccinated: Participants that respond positively to Question V1.
- Vaccinated with 1 dose: Participants that respond positively to Question V1 and declare having received 1 dose in Question "V2: How many COVID-19 vaccinations have you received?"
- Vaccinated with 2 doses: Participants that respond positively to Question V1 and declare having received 2 doses in Question V2.

Unfortunately, from the questions in the UMD Global CTIS it is not possible to know whether those with one dose are fully vaccinated, i.e., they have received a one-dose vaccine, or they simply received only the first dose of a two-dose vaccination. Similarly, it is not possible to know whether a survey respondent received a booster shot.

For each of these subsets, the prevalence of COVID-19 is computed as the fraction of responses classified as positive among the responses that report a given vaccination status. For each proxy we also estimate the *vaccine efficacy* (V_E) against illness as in [25], based on the estimates of prevalence among unvaccinated (P_U) and vaccinated (P_V):

$$V_E = 1 - P_V/P_U.$$

The confidence intervals of this metric are obtained using the Katz-log Method [1]. Since we have three subsets of vaccinated participants, we compute the vaccine efficacy for the subsets Vaccinated, Vaccinated with 1 dose, and Vaccinated with 2 doses.

2.3 Countries and Time Periods

2.3.1 South Africa

The main objective of this work is to evaluate the change in vaccine efficacy due to the Omicron variant. To this end, we evaluate the decrease in vaccine efficacy in South Africa and the Gauteng province, which is among the most affected, from mid-June 2021 until the end of 2021. Moreover, to ensure that we have sufficient data for our estimates, we concentrate on three recent time periods, each lasting about a month, where more data is available. During two of these time periods the Delta variant is dominant: i) June 18 to July 18, 2021, the period considered in [25] with low vaccination level (see eFigure 1), and ii) August 9 to September 6, 2021; while in the last time period, December 1st to 31st, 2021[2] Omicron is dominant (see eTable 1).

2.3.2 World

Beyond South Africa, we study the 50 countries for which the UMD Global CTIS has the largest amount of data. For all of them we compute the vaccine efficacy in the month of October (in which Omicron was still not present) and in the month of December (in which Omicron was present). A computed efficacy value is only considered if i) it is non-negative, ii) both prevalences P_V and P_U are at least 0.01, and iii) the number of samples used to compute them is at least 1000. We only consider further those countries for which these three conditions hold for the efficacy value in December of at least one among the vaccination status cases we consider.

We have observed that the information on prevalence of Omicron becomes available [17] with a significant delay. Hence, most countries do not report relevant presence of Omicron until the second half of December 2021. For that reason, we consider the prevalence of Omicron reported from December 15th, 2021 to January 7th, 2022. Furthermore, among the countries mentioned above, in order to have a reasonable estimate of the prevalence of the Omicron variant, we consider only countries whose data is based on sequencing at least 30 virus samples. We say that these are the countries with *presence of Omicron*. For all countries with presence of Omicron, we compare the estimated vaccination efficacy using Random Forest among all three vaccination groups and for both periods. For this, we adopt simple statistical methods, such as correlation analysis.

[2]The information on variant presence is obtained from [17], which extracts it from [5] via [7].

3 Results

3.1 Prevalence and Vaccination Efficacy in South Africa

Figures 1a and 1b show the prevalence of COVID-19 in South Africa in the period June 18th to December 31st, 2021, with the different proxies. The direct approach of Figure 1a shows a gap between the estimate Confirmed derived from the official number of cases and the other proxies. This gap can be explained in part by under-detection in the official number of cases (in South Africa the test-positivity rate is above 15%, as seen in eTable 8). More generally (in South Africa and elsewhere) symptom-based proxies can overestimate the number of cases when respondents report symptoms that are consistent with COVID-19 but are produced by some other condition. Figure 1b shows that if each curve is independently normalized to the unit scale, all proxies closely track the evolution of the official number of cases Confirmed.

In eFigures 2a-2d we show the COVID-19 prevalence in South Africa depending on the vaccination status with the diffferent proxies. We can observe that the UMD CLI and Stringent CLI proxies show a low infection prevalence in July-September and December when compared with the other proxies. On the other hand, Classic CLI and Broad CLI show a high prevalence in the period October-November, when the official data was showing that the number of cases was very low, possibly because of existing symptoms in the population not related to COVID-19.

Focusing on the Random Forest proxy, Figure 1c shows the prevalence in South Africa across all reported vaccination states. We can observe that the magnitude of the two waves (August-September and December) is similar among the Unvaccinated population, while in the vaccinated groups (Vaccinated, Vaccinated with 1 dose and Vaccinated with 2 doses) there is a much higher rate of prevalence in the December wave. This hints at a decrease of vaccine efficacy towards infection with the introduction of Omicron, as we will show next. We also observe that, as expected, subjects vaccinated with two doses show higher protection that those reporting only one dose (with Vaccinated somewhere in between since it combines both groups).

As for vaccination efficacy, Figure 1d shows the estimates for South Africa, again with Random Forest. While the data in October-November has lower quality due to the reduced number of cases in that country, we can clearly observe the reduction of vaccine efficacy, towards infection, when contrasting the August-September period to the December period when Omicron dominates. Table 1 quantifies the estimated efficacy for the three periods of interest and for the five classifiers, for South Africa and for the Gauteng province.

3.2 Prevalence and Vaccination Efficacy in the World

From the 50 countries with the largest amount of data in the CTIS and having *presence of Omicron*, we select those with an acceptable estimated efficacy value (where estimates are accepted if they follow the three rules listed in Section 2.3.2), resulting in a set of 24 countries. As a reference, in eTable 2 (in Supplement F) we show the level of vaccination in these countries[3]. Then, Tables 2 and 3, present the estimates of virus prevalence in the same countries in the periods of October and December, and also estimates of vaccination efficacy towards infection.

Both prevalence estimates and the derived efficacy estimates are obtained by the Random Forest classifier and shown with 95% confidence intervals. While Table 2 focuses on the data from individuals that declared their overall vaccination status (using groups Vaccinated, Unvaccinated), Table 3 makes a more detailed characterization by considering the number of doses declared (groups Vaccinated with 1 dose, Vaccinated with 2 doses, Unvaccinated). We also observe that there is less data on individuals with only one dose, since this is a transient state in the vaccination sequence. The full information on sample sizes can be consulted in eTables 3 and 4.

Figure 2a shows three pairs of box plots. Each pair allows comparing vaccine efficacy in October and December when considering data from the selected countries. eTable 5 presents the average corresponding to each boxplot, with the 95% confidence interval. We observe that although results are inconclusive for Vaccinated with 1 dose, there is a clear decrease of overall efficacy when considering Vaccinated and Vaccinated with 2 doses.

Figures 2b-2d allow us to see a clear trend when plotting efficacy against the most recent relative level of Omicron presence in each selected country. For each case, we present a smoothed line (Loess fitting curve, in blue), depicting a clear decreasing trend. eTable 6 presents estimates for the correlation coefficient (using Pearson correlation) together with the corresponding p-value, which confirms its statistical significance for the usual $\alpha = 5\%$.

4 Discussion

After its surge in South Africa, the Omicron variant is increasing in prevalence in other countries. Although it is still unclear if this variant is associated to a milder disease [10] several studies have raised concerns over

[3]Vaccination data is obtained from [17, 14].

the decrease of vaccine effectiveness against infection [18, 15, 11, 13] and this can lead to a wider spread of the virus even in countries with a high vaccination uptake.

Daily participatory symptom surveillance has the potential to offer a new instrument for assessing both global and local trends in health status. While limited in assessing the ground truth, due to the smaller control over the sample design and the need to preserve anonymity, we believe that the vast number of daily survey responses can compensate some of these factors. In this study, we developed a method to adapt and calibrate against the reported SARS-CoV-2 infection status the selection of symptoms, and other covariates from the survey, along different time periods and locations. As compared to methods that only use the presence or absence of symptoms reported by survey respondents [25], our proposed method was shown to provide a better proxy for assessing the trend in infections, more closely tracking the official reported cases, in particular in those countries that had a strong surveillance and consistent test positivity rates.

Using this improved classifier we complemented earlier results [25] that used traditional fixed combinations of symptoms, and updated the analysis for South Africa showing the observed decrease in vaccine efficacy when contrasting a Delta-dominated period (August-September 2021) with the recent Omicron-dominated period (December 2021). We confirmed the presence of a measurable drop in vaccine efficacy from 0.62 (with 95% confidence interval $[0.58, 0.65]$) in the Delta period to 0.24 (95% CI $[0.17, 0.30]$) in the Omicron period in the whole country ($0.62[0.54, 0.69]$ to $0.30[0.18, 0.40]$ in the Gauteng province). In addition, we confirmed that having two doses of vaccine confers better protection than one dose, both in Delta ($0.81[0.78, 0.84]$ versus $0.51[0.46, 0.55]$) and Omicron ($0.30[0.23, 0.36]$ versus $0.09[0.00, 0.18]$) dominated periods. However, we have no data on the status of respondents with regard to a possible booster dose.

By January 7th, 2022, there was a small number of candidate countries exhibiting both a high prevalence of Omicron and a high level of sequencing data supporting it. Nevertheless, we extend our analysis to these countries and show the observed changes in efficacy when comparing the months of October (pre-Omicron) with December (with partial presence of Omicron). Although these results should be confirmed once the level of Omicron becomes more dominant in many countries, we have observed a significant level of correlation of around and beyond -0.6 between vaccine efficacy (with either one or two doses) and the prevalence of Omicron. We must also make it clear that our results show a reduction in efficacy in terms of protection against infection, but this does not imply a reduction of vaccine efficacy in protection against serious disease, hospitalization and death.

There are several assumptions that frame our analysis. We assume that UMD Global CTIS answers provide a sample of the population that is interchangeable among the Delta and Omicron dominated periods. Additionally, we did not take into account possible effects from waning immunity and vaccine boost shots. However, within the countries we consider we have a mix of different vaccination timings, so that our observations appear to be valid under different scenarios. We leave for future work a further analysis where vaccination timing is taken into account.

References

[1] Ken Aho and R Terry Bowyer. Confidence intervals for ratios of proportions: implications for selection ratios. *Methods in Ecology and Evolution*, 6(2):121–132, 2015.

[2] Javier Álvarez, Carlos Baquero, Elisa Cabana, Jaya Prakash Champati, Antonio Fernández Anta, Davide Frey, Augusto Garcia-Agundez, Chryssis Georgiou, Mathieu Goessens, Harold Hernández, Rosa Lillo, Raquel Menezes, Raúl Moreno, Nicolas Nicolaou, Oluwasegun Ojo, Antonio Ortega, Estrella Rausell, Jesús Rufino, Efstathios Stavrakis, Govind Jeevan, and Christin Glorioso. Estimating Active Cases of COVID-19. *medRxiv*, 2021.

[3] Nick Andrews, Julia Stowe, Freja Kirsebom, Samuel Toffa, Tim Rickeard, Eileen Gallagher, Charlotte Gower, Meaghan Kall, Natalie Groves, Anne-Marie O'Connell, et al. Effectiveness of covid-19 vaccines against the omicron (b. 1.1. 529) variant of concern. *medRxiv*, 2021.

[4] Christina M Astley, Gaurav Tuli, Kimberly A Mc Cord, Emily L Cohn, Benjamin Rader, Tanner J Varrelman, Samantha L Chiu, Xiaoyi Deng, Kathleen Stewart, Tamer H Farag, et al. Global monitoring of the impact of the COVID-19 pandemic through online surveys sampled from the Facebook user base. *Proceedings of the National Academy of Sciences*, 118(51), 2021.

[5] Stefan Elbe and Gemma Buckland-Merrett. Data, disease and diplomacy: GISAID's innovative contribution to global health. *Global challenges*, 1(1):33–46, 2017.

[6] Junchuan Fan, Yao Li, Kathleen Stewart, Anil R. Kommareddy, Adrianne Bradford, Samantha Chiu, Frauke Kreuter, Neta Barkay, Alyssa Bilinski, Brian Kim, Roee Eliat, Tal Galili, Daniel Haimovich, Sarah LaRocca, Stanley Presser, Katherine Morris, Joshua A Salomon, Elizabeth A. Stuart, Ryan Tibshirani, Tali Alterman Barash, Curtiss Cobb, Andres Garcia, Andi Gros, Ahmed Isa, Alex Kaess, Faisal Karim, Ofir Eretz Kedosha, Shelly Matskel, Roee Melamed, Amey Patankar, Irit Rutenberg, Tal Salmona, and David Vannette. Covid-19 world symptom survey data api. https://covidmap.umd.edu/api.html, 2020.

[7] Emma B. Hodcroft. CoVariants: SARS-CoV-2 Mutations and Variants of Interest. https://covariants.org/, 2021. Accessed: 2022-01-10.

[8] Johns Hopkins University & Medicine. Johns Hopkins Coronavirus Resource Center. https://coronavirus.jhu.edu, 2020. Accessed: 2021-06-02.

[9] Salim S Abdool Karim and Quarraisha Abdool Karim. Omicron SARS-CoV-2 variant: a new chapter in the COVID-19 pandemic. *The Lancet*, 398(10317):2126–2128, 2021.

[10] Matt J. Keeling, Ellen Brooks-Pollock, Rob Challen, Leon Danon, Louise Dyson, Julia R. Gog, Laura Guzmán Rincón, Edward M. Hill, Lorenzo Pellis, Jonathan M. Read, and Michael J. Tildesley. Short-term projections based on early omicron variant dynamics in england. *medRxiv*, 2021.

[11] David S Khoury, Megan Steain, James Triccas, Alex Sigal, Miles Philip Davenport, and Deborah Cromer. Analysis: A meta-analysis of early results to predict vaccine efficacy against omicron. *medRxiv*, 2021.

[12] Epke A Le Rutte, Andrew J Shattock, Nakul Chitnis, Sherrie L Kelly, and Melissa A Penny. Assessing impact of omicron on sars-cov-2 dynamics and public health burden. *medRxiv*, 2021.

[13] Frederik Plesner Lyngse, Laust Hvas Mortensen, Matthew J. Denwood, Lasse Engbo Christiansen, Camilla Holten Møller, Robert Leo Skov, Katja Spiess, Anders Fomsgaard, Ria Lassauniere, Morten Rasmussen, Marc Stegger, Claus Nielsen, Raphael Niklaus Sieber, Arieh Sierra Cohen, Frederik Trier Møller, Maria Overvad, Kåre Mølbak, Tyra Grove Krause, and Carsten Thure Kirkeby. Sars-cov-2 omicron voc transmission in danish households. *medRxiv*, 2021.

[14] Edouard Mathieu, Hannah Ritchie, Esteban Ortiz-Ospina, Max Roser, Joe Hasell, Cameron Appel, Charlie Giattino, and Lucas Rodés-Guirao. A global database of COVID-19 vaccinations. *Nature human behaviour*, pages 1–7, 2021.

[15] Ital Nemet, Limor Kliker, Yaniv Lustig, Neta S Zuckerman, Oran Erster, Carmit Cohen, Yitshak Kreiss, Sharon Alroy-Preis, Gili Regev-Yochay, Ella Mendelson, et al. Third bnt162b2 vaccination neutralization of sars-cov-2 omicron infection. *medRxiv*, 2021.

[16] World Health Organization and 26 November 2021 others. Classification of omicron (b.1.1.529): Sars-cov-2 variant of concern. https://www.who.int/news/item/26-11-2021-classification-of-omicron-(b.1.1.529)-sars-cov-2-variant-of-concern, 2021.

[17] Our World in Data. Data on COVID-19 (coronavirus) by Our World in Data. https://covid.ourworldindata.org/, 2021. Accessed: 2022-01-07.

[18] Juliet RC Pulliam, Cari van Schalkwyk, Nevashan Govender, Anne von Gottberg, Cheryl Cohen, Michelle J Groome, Jonathan Dushoff, Koleka Mlisana, and Harry Moultrie. Increased risk of sars-cov-2 reinfection associated with emergence of the omicron variant in south africa. *MedRxiv*, 2021.

[19] Hannah Ritchie, Edouard Mathieu, Lucas Rodés-Guirao, Cameron Appel, Charlie Giattino, Esteban Ortiz-Ospina, Joe Hasell, Bobbie Macdonald, Diana Beltekian, and Max Roser. Coronavirus Pandemic (COVID-19). *Our World in Data*, 2020. https://ourworldindata.org/coronavirus.

[20] Victoria Rotshild, Bruria Hirsh-Raccah, Ian Miskin, Mordechai Muszkat, and Ilan Matok. Comparing the clinical efficacy of covid-19 vaccines: a systematic review and network meta-analysis. *Scientific Reports*, 11, 2021.

[21] Joshua A Salomon, Alex Reinhart, Alyssa Bilinski, Eu Jing Chua, Wichada La Motte-Kerr, Minttu M Rönn, Marissa B Reitsma, Katherine A Morris, Sarah LaRocca, Tamer H Farag, et al. The us covid-19 trends and impact survey: Continuous real-time measurement of covid-19 symptoms, risks, protective behaviors, testing, and vaccination. *Proceedings of the National Academy of Sciences*, 118(51), 2021.

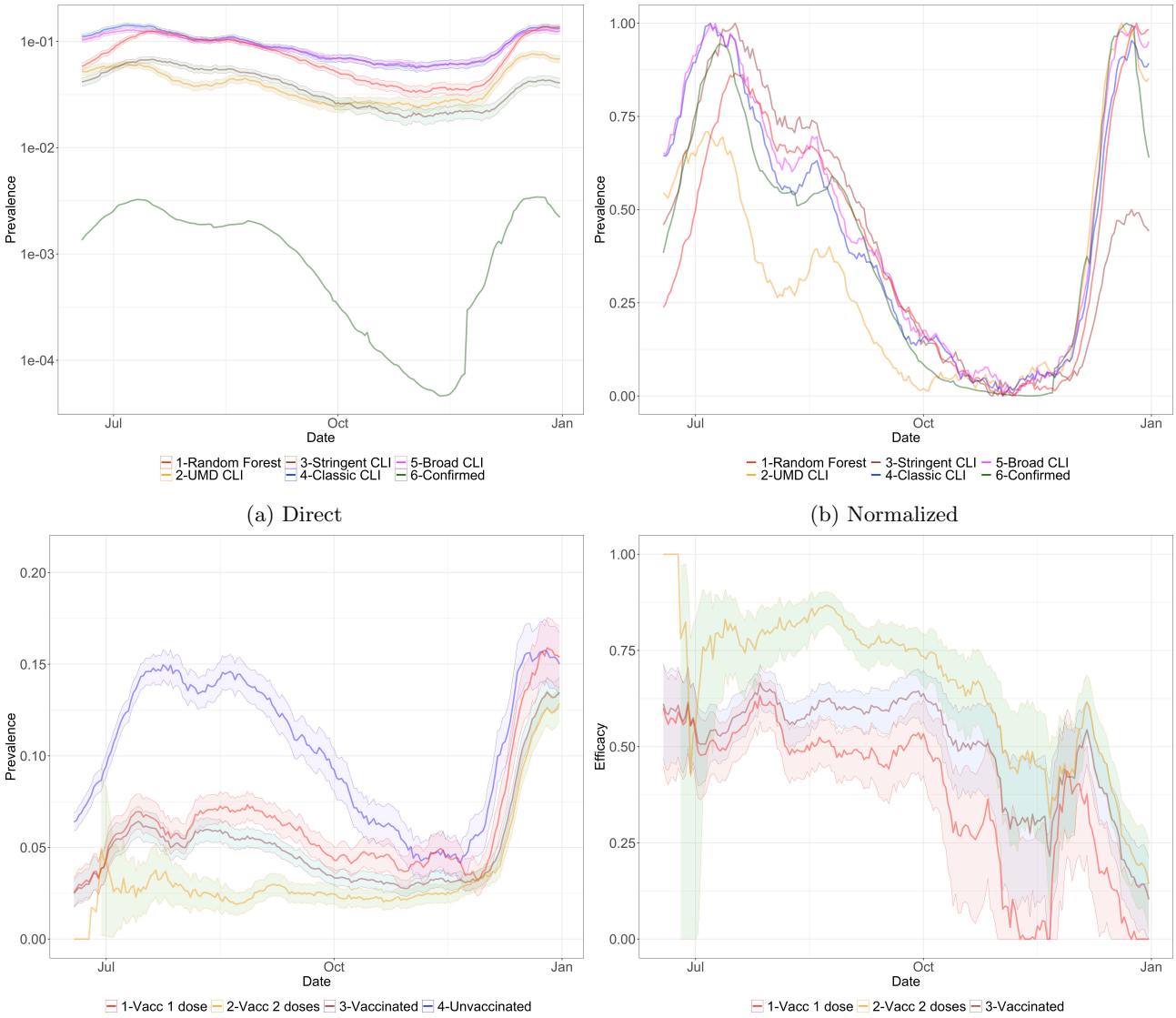
[22] Andrew J. Shattock, Epke A. Le Rutte, Robert P. Dünner, Swapnoleena Sen, Sherrie L. Kelly, Nakul Chitnis, and Melissa A. Penny. Impact of vaccination and non-pharmaceutical interventions on sars-cov-2 dynamics in switzerland. *Epidemics*, 38:100535, 2022.

[23] The University of Maryland Social Data Science Center. COVID19_symptom_survey_intl_V11_ noneu. https://covidmap.umd.edu/document/COVID19_symptom_survey_intl_V11_0723.pdf, 2021. Accessed: 2022-01-10.

[24] The University of Maryland Social Data Science Center. The University of Maryland Social Data Science Center Global COVID-19 Trends and Impact Survey in partnership with Facebook. https://covidmap.umd.edu/, 2021. Accessed: 2022-01-10.

[25] Tanner J Varrelman, Benjamin M Rader, Christina M Astley, and John S Brownstein. Syndromic surveillance-based estimates of vaccine efficacy against covid-like illness from emerging omicron and covid-19 variants. *medRxiv*, 2021.

[26] World Health Organization et al. Public health criteria to adjust public health and social measures in the context of covid-19: annex to considerations in adjusting public health and social measures in the context of covid-19, 12 may 2020. Technical report, World Health Organization, 2020.



(a) Direct

(b) Normalized

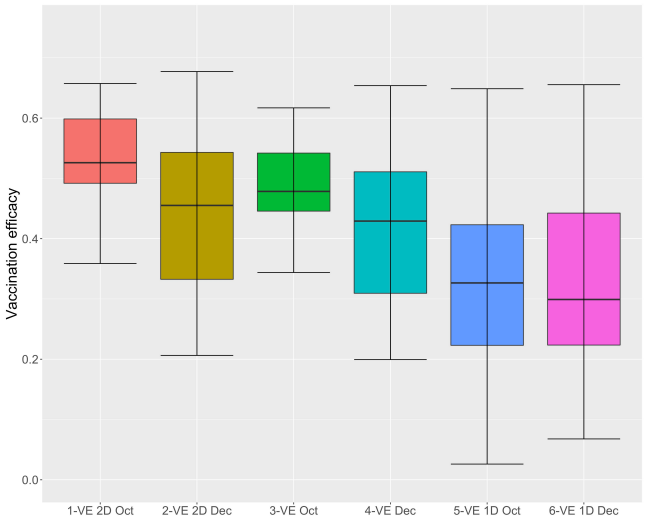
(c) Prevalence

(d) Vaccination efficacy

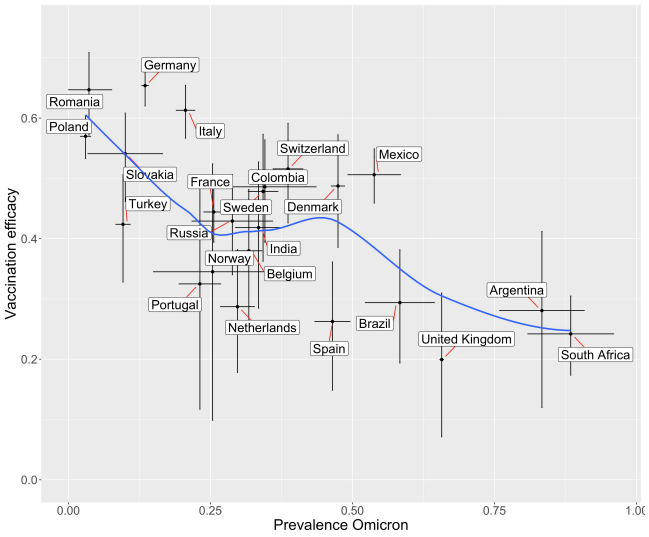
Figure 1: (a)-(b) Prevalence in South Africa obtained with the different proxies, smoothed with a rolling average of 14 days from June 18th to December 31st, 2021. In plot (a) we have the actual ratio (note that the y axis is in logarithmic scale). In plot (b) all curves are normalized so the smallest value is 0 and the largest value is 1. (c) Prevalence and (d) vaccination efficacy in South Africa among people with different levels of vaccination, estimated with Random Forest.

Method	Jun-Jul Efficacy [95%CI]	Aug-Sep Efficacy [95%CI]	Dec Efficacy [95%CI]
South Africa			
Vaccinated			
Random Forest	0.54 [0.48,0.59]	0.62 [0.58,0.65]	0.24 [0.17,0.30]
UMD CLI	0.60 [0.53,0.66]	0.66 [0.61,0.70]	0.46 [0.39,0.51]
Stringent CLI	0.69 [0.63,0.74]	0.70 [0.66,0.73]	0.48 [0.40,0.55]
Classic CLI	0.55 [0.50,0.59]	0.56 [0.52,0.59]	0.38 [0.33,0.43]
Broad CLI	0.50 [0.44,0.54]	0.49 [0.44,0.52]	0.36 [0.30,0.41]
Vaccinated with one dose			
Random Forest	0.50 [0.44,0.56]	0.51 [0.46,0.55]	0.09 [0.00,0.18]
UMD CLI	0.61 [0.54,0.68]	0.56 [0.50,0.62]	0.21 [0.09,0.31]
Stringent CLI	0.67 [0.61,0.73]	0.60 [0.54,0.65]	0.23 [0.07,0.36]
Classic CLI	0.53 [0.47,0.57]	0.47 [0.42,0.51]	0.21 [0.13,0.28]
Broad CLI	0.46 [0.40,0.52]	0.39 [0.34,0.44]	0.18 [0.09,0.26]
Vaccinated with two doses			
Random Forest	0.76 [0.64,0.84]	0.81 [0.78,0.84]	0.30 [0.23,0.36]
UMD CLI	0.75 [0.57,0.86]	0.85 [0.79,0.88]	0.56 [0.50,0.61]
Stringent CLI	0.82 [0.66,0.90]	0.88 [0.84,0.91]	0.59 [0.51,0.65]
Classic CLI	0.77 [0.66,0.84]	0.71 [0.67,0.75]	0.45 [0.40,0.49]
Broad CLI	0.75 [0.63,0.83]	0.66 [0.61,0.71]	0.43 [0.37,0.48]
Gauteng			
Vaccinated			
Random Forest	0.43 [0.33,0.51]	0.62 [0.54,0.69]	0.30 [0.18,0.40]
UMD CLI	0.58 [0.44,0.68]	0.63 [0.51,0.73]	0.52 [0.41,0.61]
Stringent CLI	0.64 [0.53,0.72]	0.70 [0.61,0.78]	0.57 [0.43,0.67]
Classic CLI	0.50 [0.42,0.58]	0.51 [0.42,0.59]	0.48 [0.39,0.55]
Broad CLI	0.49 [0.39,0.57]	0.41 [0.31,0.50]	0.45 [0.35,0.53]
Vaccinated with one dose			
Random Forest	0.40 [0.28,0.49]	0.54 [0.44,0.63]	0.14 [0.00,0.30]
UMD CLI	0.60 [0.46,0.71]	0.58 [0.42,0.70]	0.38 [0.18,0.53]
Stringent CLI	0.62 [0.49,0.71]	0.61 [0.47,0.71]	0.39 [0.13,0.57]
Classic CLI	0.47 [0.37,0.56]	0.47 [0.36,0.56]	0.35 [0.20,0.46]
Broad CLI	0.44 [0.33,0.53]	0.34 [0.20,0.45]	0.29 [0.14,0.42]
Vaccinated with two doses			
Random Forest	0.62 [0.36,0.78]	0.77 [0.67,0.85]	0.36 [0.24,0.46]
UMD CLI	0.69 [0.27,0.87]	0.73 [0.54,0.84]	0.57 [0.45,0.66]
Stringent CLI	0.85 [0.55,0.95]	0.88 [0.76,0.94]	0.65 [0.51,0.74]
Classic CLI	0.79 [0.59,0.90]	0.58 [0.44,0.68]	0.53 [0.44,0.60]
Broad CLI	0.80 [0.59,0.91]	0.54 [0.39,0.65]	0.50 [0.41,0.58]

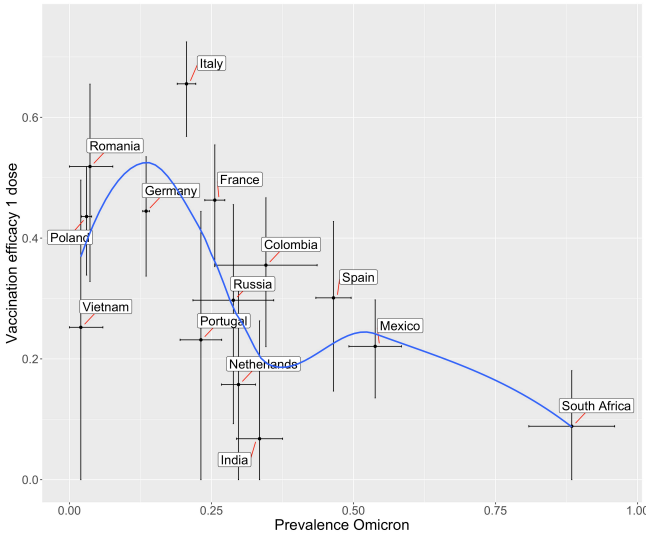
Table 1: Vaccine efficacy in South Africa and the Gauteng province, calculated for three time periods: June 18th to July 18th (Jun-Jul), August 9th to September 6th (Aug-Sep), and December 1st to 31st (Dec).



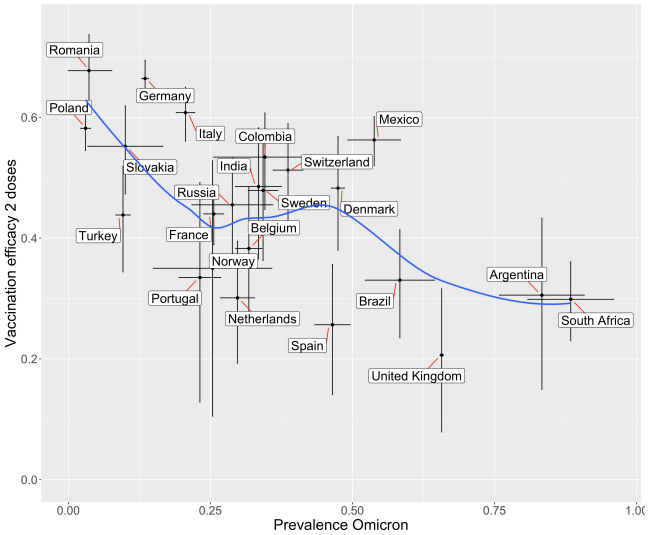
(a) Vaccine Efficacy (VE) distributions



(b) Vaccine efficacy in December



(c) Efficacy 1 dose in December



(d) Efficacy 2 doses in December

Figure 2: Analysis of vaccine efficacy towards preventing infection: Sub-figure (a) shows distributions of efficacy in October and December, for the countries with presence of Omicron (as defined in Section 2.3.2); Sub-figures (b,c,d) show vaccination efficacy versus Omicron prevalence in the same set of countries, depending on vaccination status. For each country the 95% confidence intervals of the two values are shown as black lines. The blue line is the Loess curve fitting of the data.

Country	% Prevalence Omicron	Prevalence Oct	Prevalence Dec	Vac efficacy Oct	Vac efficacy Dec
Argentina	0.83 [0.76,0.91]	0.02 [0.01,0.02]	0.03 [0.03,0.03]	0.48 [0.35,0.58]	0.28 [0.12,0.41]
Belgium	0.32 [0.29,0.34]	0.02 [0.02,0.02]	0.05 [0.05,0.05]	0.53 [0.39,0.64]	0.38 [0.26,0.48]
Brazil	0.58 [0.52,0.64]	0.03 [0.03,0.03]	0.03 [0.02,0.03]	0.43 [0.37,0.49]	0.29 [0.19,0.38]
Colombia	0.35 [0.26,0.44]	0.03 [0.03,0.03]	0.03 [0.03,0.03]	0.55 [0.49,0.61]	0.49 [0.39,0.56]
Denmark	0.47 [0.46,0.49]	0.01 [0.01,0.01]	0.05 [0.05,0.05]	–	0.49 [0.39,0.57]
France	0.26 [0.24,0.27]	0.01 [0.01,0.01]	0.03 [0.03,0.03]	–	0.44 [0.39,0.49]
Germany	0.13 [0.13,0.14]	0.01 [0.01,0.01]	0.02 [0.02,0.02]	–	0.65 [0.62,0.68]
India	0.33 [0.29,0.38]	0.04 [0.04,0.04]	0.03 [0.03,0.03]	0.44 [0.35,0.52]	0.42 [0.28,0.53]
Italy	0.21 [0.19,0.22]	0.01 [0.01,0.01]	0.02 [0.02,0.02]	–	0.61 [0.57,0.65]
Mexico	0.54 [0.49,0.58]	0.05 [0.05,0.05]	0.04 [0.04,0.04]	0.57 [0.54,0.59]	0.51 [0.46,0.55]
Netherlands	0.30 [0.27,0.33]	0.02 [0.02,0.02]	0.05 [0.04,0.05]	0.36 [0.20,0.49]	0.29 [0.18,0.38]
Norway	0.25 [0.15,0.36]	0.01 [0.01,0.01]	0.03 [0.02,0.03]	–	0.35 [0.10,0.52]
Poland	0.03 [0.02,0.04]	0.03 [0.03,0.04]	0.07 [0.06,0.07]	0.50 [0.42,0.56]	0.57 [0.53,0.60]
Portugal	0.23 [0.19,0.27]	0.01 [0.01,0.01]	0.03 [0.03,0.03]	–	0.32 [0.12,0.48]
Romania	0.04 [0.00,0.08]	0.06 [0.06,0.06]	0.02 [0.02,0.02]	0.59 [0.56,0.62]	0.65 [0.57,0.71]
Russia	0.29 [0.22,0.36]	0.04 [0.04,0.05]	0.03 [0.02,0.03]	0.45 [0.39,0.50]	0.43 [0.34,0.51]
Slovakia	0.10 [0.03,0.17]	0.03 [0.03,0.03]	0.06 [0.05,0.06]	0.47 [0.32,0.59]	0.54 [0.46,0.61]
South Africa	0.88 [0.81,0.96]	0.04 [0.04,0.04]	0.12 [0.12,0.13]	0.50 [0.41,0.57]	0.24 [0.17,0.30]
Spain	0.46 [0.43,0.50]	0.01 [0.01,0.02]	0.05 [0.05,0.06]	0.62 [0.50,0.70]	0.26 [0.15,0.36]
Sweden	0.34 [0.32,0.37]	0.01 [0.00,0.01]	0.02 [0.02,0.02]	–	0.48 [0.36,0.57]
Switzerland	0.39 [0.36,0.41]	0.01 [0.01,0.01]	0.04 [0.04,0.04]	–	0.52 [0.43,0.59]
Turkey	0.10 [0.08,0.11]	0.05 [0.05,0.06]	0.05 [0.05,0.05]	0.45 [0.38,0.51]	0.42 [0.33,0.51]
United Kingdom	0.66 [0.65,0.66]	0.03 [0.03,0.03]	0.05 [0.04,0.05]	0.34 [0.22,0.45]	0.20 [0.07,0.31]
Vietnam	0.02 [0.00,0.06]	0.01 [0.01,0.01]	0.03 [0.03,0.03]	–	–

Table 2: Prevalence of Omicron in COVID-19 and vaccination efficacy in the countries with presence of Omicron (as defined in Section 2.3.2). When data is insufficient to meet the defined selection criteria, it is omitted and replaced by "–".

Country	% Prevalence Omicron	Vac 1 dose efficacy Oct	Vac 1 dose efficacy Dec	Vac 2 doses efficacy Oct	Vac 2 doses efficacy Dec
Argentina	0.83 [0.76,0.91]	0.03 [0.00,0.27]	–	0.53 [0.41,0.62]	0.31 [0.15,0.43]
Belgium	0.32 [0.29,0.34]	–	–	0.55 [0.41,0.65]	0.38 [0.26,0.48]
Brazil	0.58 [0.52,0.64]	0.20 [0.11,0.28]	–	0.50 [0.44,0.55]	0.33 [0.23,0.41]
Colombia	0.35 [0.26,0.44]	0.44 [0.35,0.53]	0.36 [0.22,0.47]	0.61 [0.55,0.67]	0.53 [0.45,0.61]
Denmark	0.47 [0.46,0.49]	–	–	–	0.48 [0.38,0.57]
France	0.26 [0.24,0.27]	–	0.46 [0.35,0.55]	–	0.44 [0.39,0.49]
Germany	0.13 [0.13,0.14]	–	0.44 [0.34,0.53]	–	0.66 [0.63,0.69]
India	0.33 [0.29,0.38]	0.19 [0.05,0.31]	0.07 [0.00,0.26]	0.54 [0.47,0.61]	0.49 [0.37,0.58]
Italy	0.21 [0.19,0.22]	–	0.66 [0.57,0.72]	–	0.61 [0.56,0.65]
Mexico	0.54 [0.49,0.58]	0.36 [0.32,0.40]	0.22 [0.14,0.30]	0.66 [0.63,0.68]	0.56 [0.52,0.60]
Netherlands	0.30 [0.27,0.33]	–	0.16 [0.00,0.33]	0.41 [0.26,0.53]	0.30 [0.19,0.39]
Norway	0.25 [0.15,0.36]	–	–	–	0.35 [0.11,0.53]
Poland	0.03 [0.02,0.04]	0.31 [0.13,0.45]	0.44 [0.34,0.52]	0.52 [0.45,0.58]	0.58 [0.55,0.62]
Portugal	0.23 [0.19,0.27]	–	0.23 [0.00,0.44]	–	0.33 [0.13,0.49]
Romania	0.04 [0.00,0.08]	0.65 [0.59,0.70]	0.52 [0.33,0.65]	0.58 [0.55,0.61]	0.68 [0.60,0.74]
Russia	0.29 [0.22,0.36]	0.55 [0.43,0.64]	0.30 [0.09,0.46]	0.44 [0.38,0.50]	0.46 [0.37,0.53]
Slovakia	0.10 [0.03,0.17]	–	–	0.50 [0.35,0.61]	0.55 [0.47,0.62]
South Africa	0.88 [0.81,0.96]	0.29 [0.15,0.40]	0.09 [0.00,0.18]	0.64 [0.56,0.70]	0.30 [0.23,0.36]
Spain	0.46 [0.43,0.50]	0.34 [0.09,0.52]	0.30 [0.15,0.43]	0.66 [0.55,0.74]	0.26 [0.14,0.36]
Sweden	0.34 [0.32,0.37]	–	–	–	0.48 [0.36,0.57]
Switzerland	0.39 [0.36,0.41]	–	–	–	0.51 [0.42,0.59]
Turkey	0.10 [0.08,0.11]	–	–	0.49 [0.42,0.55]	0.44 [0.34,0.52]
United Kingdom	0.66 [0.65,0.66]	–	–	0.36 [0.24,0.46]	0.21 [0.08,0.32]
Vietnam	0.02 [0.00,0.06]	–	0.25 [0.00,0.50]	–	–

Table 3: Prevalence of Omicron and vaccination efficacy with one and two doses in the countries with presence of Omicron (as defined in Section 2.3.2). When data is insufficient to meet the defined selection criteria, it is omitted and replaced by "–". The prevalence of Omicron is replicated from Table 2 for easy reference.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SuplementaryInformation.pdf