3

CHAPTER 3

Population-based data from the COVID RADAR app as a surveillance tool and predictor of COVID-19 related primary care demand

Willian J. van Dijk, Martijn H.H. Schoot Uiterkamp, Marcel R. Haas, Mattijs E. Numans, Leo G. Visser, Astrid van Hylckama Vlieg, Frits R. Rosendaal, Jessica C. Kiefte-de Jong

Under Review

ABSTRACT

Introduction

Self-reported, population-based data may serve as a valuable surveillance tool for infectious disease outbreaks, especially when testing resources are limited. Using COVID-19 as a case study, we examined whether app-based reports on symptoms and behaviours could predict future primary care workload.

Methods

We analysed data from the COVID RADAR app (2020–2022), where users anonymously reported symptoms and COVID-related behaviours. These data were aggregated daily and linked to COVID-related primary care visits within the Extramural Leiden University Medical Center Academic Network Datawarehouse (≈250,000 patients). We employed recalibrated Poisson regression and Random Forest models, retrained quarterly, and validated performance using Root Mean Square Error (RMSE). Models were compared against those using only confirmed test counts.

Results

Over the study period, 9% of the population contacted primary care for COVID-related reasons. Both models showed moderate predictive accuracy, with the Random Forest model (RMSE ~22) underestimating early demand by a factor 2, and the Poisson model (RMSE ~71) overestimating during peaks by a factor 10. However, both outperformed test-only models (RMSE Poisson ~63,000; Random Forest ~36). The most important predictor was the proportion of app users reporting recent contact with a COVID-positive individual, increasing in importance over time.

Conclusion

These findings suggest that self-reported symptom and behaviour data may enhance surveillance and early prediction of healthcare demand during pandemics. Further improvement may be achieved through app refinement, integration of additional predictors, and enhanced modelling of transmission dynamics.

KEY MESSAGES

- In pandemics, surveillance is important for planning of healthcare demand
- In primary care, as the first line within the healthcare system, care demands are less predictable
- Population-based data from an app may offer valuable surveillance for prediction of primary care demand, particularly when testing capacity is limited.

INTRODUCTION

New infectious disease outbreaks present significant challenges to healthcare, due to limited knowledge and scarce data on disease characteristics, and exponentially increasing patient demand.(1, 2) During the latest pandemic, coronavirus disease (COVID-19), the demand for acute healthcare services (Intensive Care Unit beds) increased substantially, while the number of patients attending elective healthcare (e.g., chronic disease management, cancer screening) decreased.(3-5) Also in primary care this pandemic changed the workload.(6, 7) Although the number of short patient contacts by primary care physicians decreased during the COVID-19 pandemic, the number of contacts with patients in need of more intensive home visits increased in the Netherlands.(8) The limited availability of healthcare resources in general was associated with an increase of the mortality rate of COVID-19.(9) Therefore, careful planning and allocation of these resources is important during a rapidly evolving pandemic. To achieve this, an extensive surveillance system should be in place, i.e., a nationwide test-and-trace policy.

During the COVID-19 pandemic, an upsurge of the number of positive tests usually preceded an increase of hospital healthcare demand,(10, 11) and subsequently (approximately one week later) the number of Intensive Care admissions.(12, 13) In primary healthcare, as the first line within the healthcare system in the Netherlands, care demands are less predictable when based on the number of positive tests, as most patients visit their general practitioner (GP) before they test positive. Therefore, more rapidly available population-based data could support optimal resource distribution and workforce planning of COVID-19 related primary healthcare demand.

A nationwide smartphone app (COVID RADAR), collected anonymous individual data about COVID-19 related symptoms, risk behaviour, and positive test results among community-dwelling participants from April 2020 until February 2022.(14, 15) In prior research, these data were successful in predicting the replication rate of COVID-19.(16)

In this study we will describe the COVID-19 related primary care demand and show the value of population-based data during a pandemic by predicting COVID-19 related workload in primary care in the Leiden and The Hague area in the Netherlands.

METHODS

Study design

We used data from the COVID RADAR app as input to predict the COVID-19 related workload, as expressed by data from the Extramural Leiden University Medical Center Academic Network (ELAN) Datawarehouse. COVID RADAR app users were asked for informed consent upon first use of the app. Patients enlisted with ELAN primary care practice centres could withdraw via an informed opt-out procedure.(17) All data were aggregated by day and covered the region of Leiden and The Hague (with a total of 2 million inhabitants) from January 2020 until February

2022. Ethical approval was granted by the Medical Ethical Board of the LUMC (dossier number N20.070 and 20.080) and the ELAN scientific board (dossier number 895648).

COVID RADAR app

A comprehensive overview of the COVID RADAR app's operations, data collection, and data processing has been previously published.(14) In summary, the COVID RADAR app was a free app that prompted users to anonymously report COVID-19 symptoms and associated behaviours through a brief daily questionnaire. The questions covered symptoms as well as contacts with patients with COVID-19. Users also provided details on SARS-CoV-2 test results and vaccination status; the specific questions are outlined in supplemental table 1. Various national media campaigns, including social media, promoted app usage. The questionnaire design allowed for updates based on changes such as new mitigation measures and evolving scientific insights during the pandemic. No detailed personal information was registered in the app; age was registered in categories of decades and location was registered via the first four digits of the postcode, preserving anonymity of users.

COVID-19 related workload in primary care

The aggregated daily COVID RADAR data were linked to aggregated daily data from the ELAN Datawarehouse. The ELAN Datawarehouse contained electronic health records (EHRs) from approximately 250,000 persons, which was stable during the research period. Data from the ELAN Datawarehouse were extracted from the GP practices each 12 weeks.

We selected EHRs from patients with a respiratory complaint in 2020 - 2022, coded following the International Classification of Primary Care (ICPC). Within these patients we selected records with COVID-19 as the primary reason for contact, from free text and relevant codes (see supplemental document 1 for details about this selection). We included at most one contact per day per person. We excluded records reporting tests results, e-mail or notes, since this is not actual 'patient contact time'. We extracted the following data from the selected patients from the ELAN Datawarehouse: age, sex, type of contact (in-person or not (telephone, etc.), during in- or out-office hours), and if this contact resulted in referral to a hospital.

Descriptive analyses

Firstly, we described the temporal patterns of COVID-19 related workload registered in ELAN during the pandemic expressed in terms of the number and type of contacts, and associations with referral to the hospital.

Secondly, we described the daily temporal patterns of variables from the COVID RADAR about number of reports, symptoms (using the European Centre for Disease Prevention and Control

(ECDC) definition for COVID-19 like symptoms: Cough, fever, shortness of breath or anosmia),(18) behaviour and proportion of users reporting recent contact with a COVID-19 patient.

We compared both data sources (the app and ELAN) with state-reported numbers of positive test results and numbers of new hospitalizations. These data are publicly available and provided by the National Institute for Public Health and the Environment (RIVM).(19)

Predictive analyses

We smoothed the outcome (number of COVID-19 related primary care contacts) using the 3 days prior and 3 days after each point (7-day range) to account for the strong weekly pattern. We used two candidate predictors from the COVID RADAR app: (1) the daily proportion of users who had recent contact with a COVID-19 patient over the past seven days (risk contacts) and (2) the daily proportion of users reporting an ECDC-defined COVID-like symptom over the same period. We divided the study period in blocks of 12 weeks that corresponded with the frequency with which data were extracted from the ELAN Datawarehouse during the pandemic. We designed our model with a training and validation approach in the following way to mimic as much as possible the real world situation in which this model was developed during the pandemic. In a first step, we fitted a regression model on the COVID-19 related healthcare demand data of the first block of 12 weeks using predictors in this block and used this model to make a prediction for the demand of the subsequent block of 12 weeks using the predictors from the first block. Subsequently, this model was recalibrated using data from both the first and second blocks, and we used this new recalibrated model to make a prediction for the third block. This was repeated until February 2022 (end of COVID RADAR data collection), yielding for each of the periods a daily prediction made by a model fitted using data from all preceding periods.

We compared two types of models: a Poisson regression model and a Random forest model (using a grid search with prespecified limits to optimize hyperparameters, details in supplemental document 2). Model fits were compared with the root mean square error (RMSE). We report the coefficients of the Poisson model and the relative importance of the candidate predictors (based on the mean decrease in RMSE) in the Random forest model for each time block.(20)

To compare the added value of the app's data, we also performed the same prediction procedure including all three models with the number of state reported positive tests as a predictor. Statistical analyses were performed using STATA 16.1.

RESULTS

COVID-19 related primary care contacts

Over the research period (January 2020 until February 2022; ~750 days) in approximately 250,000 registered patients (in an area of 2.5 million inhabitants), 47,974 COVID-19 related primary care contacts in 22,338 (9%) unique patients were registered (table 1). The GP referred 13% (n=2,901) of these patients for further workup, which could result in hospital admission. Patients referred to a hospital were more often older than 40 years of age. Contacts that resulted in referral were more often in-person during office hours (see table 1). Most COVID-19 related primary care contacts were in the second half of 2020. From the second half of 2021 onwards the proportion of patients who was referred to hospital increased. The number of COVID-19 related primary care contacts, number of positive tests and hospitalizations showed similar patterns through time (see figure 1).

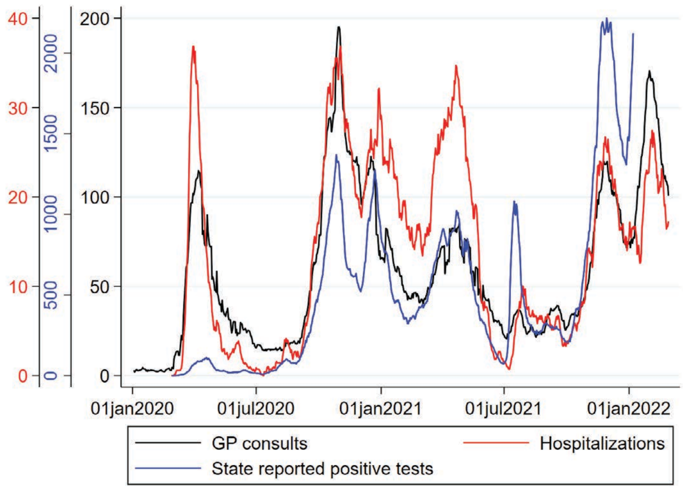

Figure 1: Temporal daily trends of primary care consults, hospitalizations and positive tests due to COVID-19 (all 7 days mean smoothed). Number of positive tests (blue line) stops one month earlier because of a large number of positive tests in February 2022 (35,000) that would have limited the informativeness of the scale.

Risk contact and symptom data

Over the research period, out of approximately 2.5 million inhabitants, 86,037 individuals used the COVID RADAR app in total 2.8 million times (see supplemental table 2). A median number of 7,800 (IQR 4,600; 10,800) unique individuals used the app each week. See for patterns of app usage, reports of positive tests and behaviour the supplementary document 3. Figure 2 shows the daily proportion of users reporting a recent contact with a COVID-19 patient, and figure 3 shows the proportion of users reporting symptoms defined as 'COVID-like symptoms' by the ECDC. The proportion with recent contact with a COVID-19 patient showed higher variability in accordance with the number of COVID-19 related primary care contacts than the proportion of users with symptoms.

Table 1: Patients and their contacts with their GP because of COVID-19, stratified by subsequent referral to the hospital

Referral to hospital		**No**	**Yes**
Number of Patients		19437 (87%)	2901 (13%)
Number of Contacts	1	12702 (97%)	425 (3%)
	2	3669 (88%)	490 (12%)
	3	1456 (74%)	499 (26%)
	>3	1610 (52%)	1487 (48%)
Sex	Female	11508 (87%)	1792 (13%)
Age, mean, SD		43.2 (21.2)	51.0 (19.1)
Age category	<18	2643 (95%)	125 (5%)
	18-39	5832 (90%)	653 (10%)
	40-59	6362 (84%)	1174 (16%)
	>60	4600 (83%)	949 (17%)
Number of contacts		42694 (92%)	3980 (9%)
Contact type	Phone, in office hours	26339 (95%)	1302 (5%)
	Consult, in office hours	12190 (83%)	2549 (17%)
	Phone, out of office hours	1253 (95%)	61 (5%)
	Consult, out of office hours	815 (93%)	60 (7%)
Quarter	2020q1	1732 (97%)	57 (3%)
	2020q2	3877 (93%)	310 (7%)
	2020q3	2261 (92%)	190 (8%)
	2020q4	10865 (95%)	512 (5%)
	2021q1	4475 (91%)	470 (9%)
	2021q2	4329 (87%)	638 (13%)
	2021q3	2280 (84%)	447 (16%)
	2021q4	6168 (89%)	727 (11%)
	2022q1	6655 (91%)	629 (9%)

SD: Standard deviation.

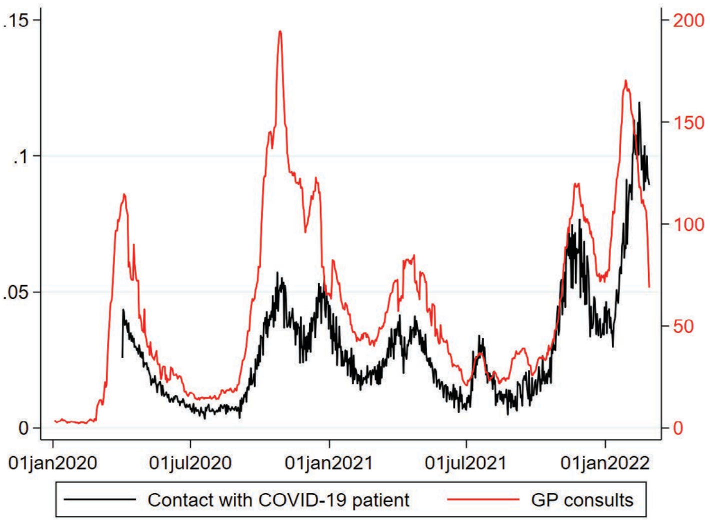

Figure 2: Daily proportion of app users reporting recent contact with COVID-19 patient (<14 days) Red: number of GP consultations because of COVID-19

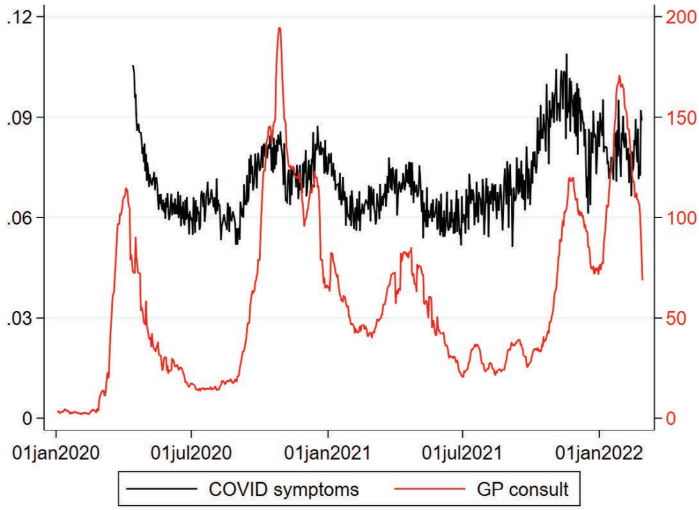

Figure 3: Daily proportion of app users reporting symptoms from ECDC definition of COVID-19 ECDC: (cough, fever, shortness of breath or loss of smell/taste) Red: number of GP consultations because of COVID-19

Prediction

Using a quarter-frequent recalibration, both the Poisson regression and the Random forest model showed patterns in their predictions in line with the healthcare demand outcome (see figures 4 and 5). Their performance, tabulated in table 2, was poor during sudden increases of COVID-19 related primary care demand (Q4 of 2020). Later during the pandemic the Poisson model overestimated the number of COVID-19 related primary care contacts by a factor of 10. The Random forest model underestimated the care outcomes during the first wave (predicting a peak of 100 daily COVID-19-related contacts during the first wave, whereas the actual number was nearly 200) but had more accurate prediction later in the pandemic. Both models anticipated more contacts during the ‘Dansen met Janssen’-wave (a superspreader event in the Netherlands, during which predominantly young people were allowed to ‘go dancing’ directly after a single shot of the Johnson & Johnson vaccine in June 2021). All predictions with COVID RADAR data were superior to predictions using state reported test results (see table 2 and supplemental figure 1; RMSE COVID RADAR between 10 - 252 vs RMSE test between 22 - 441,297).

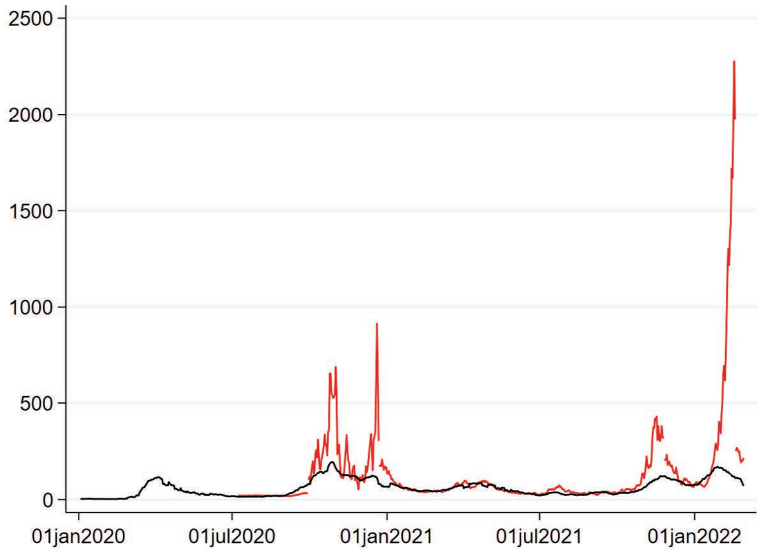

Figure 4: Predicted values for Poisson regression model (red) and actual number of COVID-19 related GP consults (black). Gaps because of recalibration of the model.

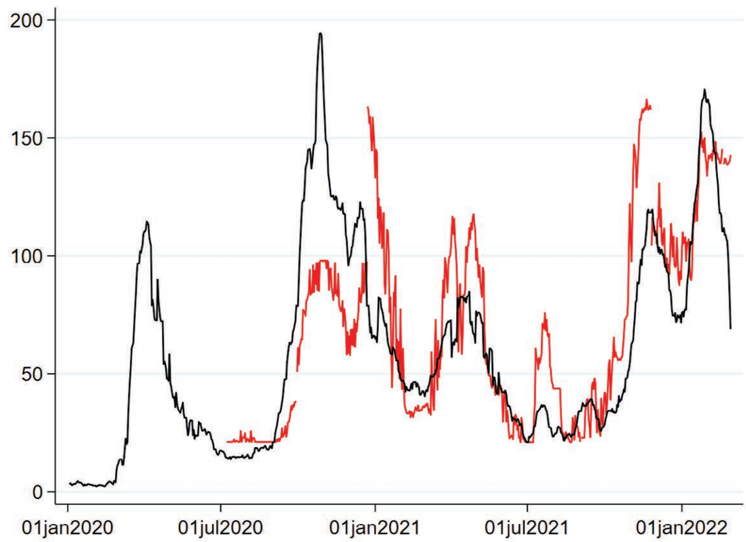

Figure 5: Predicted values for Random forest model (red) against actual number of COVID-19 related GP consults (black). Gaps because of recalibration of the model.

Table 2: Root mean square of the errors of predictions in the quarter after the training period. Quarters are not 'exact' quarters, but periods of 12 weeks, starting in July 2020. Last quarter starts end of November 2021 and ends in February 2022.

RMSE predictions outside training set		**Poisson tests results**	**Random forest tests results**	**Poisson COVID RADAR**	**Random forest COVID RADAR**
Quarter	2020q3	662	31.8	10.1	10.3
	2020q4	2837	67.3	130.6	48.2
	2021q1	25.7	42.1	19.2	23.0
	2021q2	21.9	30.5	10.4	16.5
	2021q3	26.6	31.5	10.8	11.3
	2021q4	52.4	21.9	62.2	24.5
	2022q1	441297	29.5	252	17.2

The variable 'risk contact' had the highest relative contribution in both models (see supplemental table 2). The contributions of the predictors were most stable in the Random Forest model, while the Poisson models exhibited unstable predictor coefficients. During the first 12 weeks symptoms had a slightly higher contribution to the Random forest model than later during the pandemic. No major shifts were seen in the selected days, whereas recent days were the most important.

DISCUSSION

We described COVID-19-related primary care volume in the Netherlands and evaluated whether population-based self-reported data could predict care demand over time. During the pandemic, 9% of the population contacted their GP for COVID-19-related concerns. Using simple models with quarterly recalibration, self-reported symptom and behaviour data moderately predicted care demand and outperformed models based solely on positive test results. While context-specific and in need of refinement, such data offer valuable predictive insights—especially early in a pandemic when testing is limited.

Capacity for testing was limited at the start of the pandemic in the Netherlands, resulting in a underestimation of SARS-CoV-2 infections.(21) Later in the pandemic, care demand declined relative to reported cases, likely due to increased testing, vaccination, new variants, and greater public awareness. Deviations in our predictions may also stem from limitations of models not tailored to capture pandemic complexity. The Poisson model's overestimations may reflect violations of event independence, reduced accuracy with common events, and overdispersion, where variance exceeds the mean.(22, 23) During the pandemic, the Random forest model showed improved performance; however, its performance was poor during the second wave (Q4 2020). Random forests predict within the range of observed training data, as they average target values within decision tree leaves. Thus, they cannot forecast values beyond prior extremes—e.g., the Q4 2020 ceiling of ~100 GP contacts reflected earlier data limits. Predictions near this maximum should be flagged as exceeding historical bounds. Future pandemic surveillance models should adapt to the pandemic phase. Early on, extrapolative models (e.g., linear or Poisson regression) may be more suitable, while later phases could benefit from complex models (e.g., Random Forest) to capture interactions and improve precision. More frequent recalibration could help, though current data availability in quarterly blocks limits this. Incorporating assumptions about delays between risk behaviour, symptom onset, and GP visits—similar to compartmental models—may further enhance predictions.(16)

This study is the first to address prediction of COVID-19 related primary care workload using population-based data from an app. The use of an anonymous and voluntary-to-use app in pandemic surveillance has several advantages, especially during the start of a pandemic. It is easy to use, has privacy by design and operates instantaneously. To improve quality of surveillance and predictive performance in future pandemics, the questionnaire could be adapted, in such a way that the moment of contact with a COVID-19 patient is known more accurately. More complex models could be used, including predictors from other domains (such as the weather), tailored to future pandemics.

Data from a voluntary and anonymous app also has disadvantages. App users were not a random sample of the population and we know that demographics of users were not stable over time.(14) However using periodic recalibration, the impact of this bias was limited. The variable with highest predictive performance was 'risk contact with a COVID-19 patient'. However, knowledge about risk contacts is still partly dependent on testing capacity, which is limited during the start of a pandemic. This may explain why, in the early stages of the pandemic, symptoms carried relatively more weight in the models. Though variables about symptoms do not have this dependency, at the start of a future pandemic it may not yet be known which symptoms are predictive (such as loss of taste and/or smell during COVID-19). A questionnaire based app should be dynamic, allowing updates based on newest insights about disease characteristics. Self-reported data could suffer effects of misspecification of symptoms or underestimation of true contact with COVID-19 patients (because of unknown contacts). Other surveillance apps use Bluetooth or GPS for contact tracing, registering unknown contacts as well; however this approach raises privacy concerns and therefore limited usage uptake.(24-26)

During the latest pandemic, the COVID RADAR app was not used for proactive care planning, therefore its effectivity cannot be determined. Research on effects of digital contact tracing apps and digital surveillance is limited, but available studies show moderate effects.(27, 28) However, these studies focus on the effect of individual tracing on the growth of the pandemic, which is a different aim compared with the current study (pandemic surveillance and prediction of healthcare usage).

Conclusion

Concluding, population-based data from an instantaneously operating app has potential for use as a surveillance tool during a pandemic and prediction of (primary) healthcare demand and outperforms predictions based on the number of positive tests.

Disclosure statement: The authors report no conflict of interest

Data availability statement Raw data of the COVID RADAR app are available in a public, open access repository. Data are accessible on https://doi.org/10.17026/dans-zcd-m9dh

Data deposition Data are accessible on https://doi.org/10.17026/dans-zcd-m9dh

REFERENCES

1. Ni B, Gettler E, Stern R, Munro HM, Steinwandel M, Aldrich MC, et al. Disruption of medical care among individuals in the southeastern United States during the COVID-19 pandemic. Journal of Public Health Research. 2022;11(1):jphr. 2021.497.
2. Sahebi A, Nejati-Zarnaqi B, Moayedi S, Yousefi K, Torres M, Golitaleb M. The prevalence of anxiety and depression among healthcare workers during the COVID-19 pandemic: An umbrella review of meta-analyses. Progress in Neuro-Psychopharmacology and Biological Psychiatry. 2021;107:110247.
3. Moynihan R, Sanders S, Michaleff ZA, Scott AM, Clark J, To EJ, et al. Impact of COVID-19 pandemic on utilisation of healthcare services: a systematic review. BMJ open. 2021;11(3):e045343.
4. Bakouny Z, Hawley JE, Choueiri TK, Peters S, Rini BI, Warner JL, et al. COVID-19 and cancer: current challenges and perspectives. Cancer cell. 2020;38(5):629-46.
5. Müskens JLJM, Hartman TCO, Schers HJ, Akkermans RP, Westert GP, Kool RB, et al. Trends in low-value GP care during the COVID-19 pandemic: a retrospective cohort study. BMC Primary Care. 2024;25(1):73.
6. Kraus M, Stegner C, Reiss M, Riedel M, Børsch AS, Vrangbaek K, et al. The role of primary care during the pandemic: shared experiences from providers in five European countries. Bmc Health Serv Res. 2023;23(1):1054.
7. Matenge S, Sturgiss E, Desborough J, Hall Dykgraaf S, Dut G, Kidd M. Ensuring the continuation of routine primary care during the COVID-19 pandemic: a review of the international literature. Family Practice. 2022;39(4):747-61.
8. Heins M, Hek K, Hooiveld M, Hendriksen J, Korevaar J. Impact coronapandemie op aantal en type huisartscontacten 2020 tot mei 2022. Nivel: Utrecht, The Netherlands. 2022.
9. Xie L, Yang H, Zheng X, Wu Y, Lin X, Shen Z. Medical resources and coronavirus disease (COVID-19) mortality rate: Evidence and implications from Hubei province in China. Plos One. 2021;16(1):e0244867.
10. Vytla V, Ramakuri SK, Peddi A, Srinivas KK, Ragav NN, editors. Mathematical models for predicting COVID-19 pandemic: a review. Journal of Physics: Conference Series; 2021: IOP Publishing.
11. Campillo-Funollet E, Van Yperen J, Allman P, Bell M, Beresford W, Clay J, et al. Predicting and forecasting the impact of local outbreaks of COVID-19: use of SEIR-D quantitative epidemiological modelling for healthcare demand and capacity. International journal of epidemiology. 2021;50(4):1103-13.
12. Fort D, Seoane L, Unis GD, Price-Haywood EG. Locally informed modeling to predict hospital and intensive care unit capacity during the COVID-19 epidemic. Ochsner Journal. 2020;20(3):285-92.
13. Gitto S, Di Mauro C, Ancarani A, Mancuso P. Forecasting national and regional level intensive care unit bed demand during COVID-19: The case of Italy. Plos One. 2021;16(2):e0247726.
14. van Dijk WJ, Saadah NH, Numans ME, Aardoom JJ, Bonten TN, Brandjes M, et al. COVID RADAR app: Description and validation of population surveillance of symptoms and behavior in relation to COVID-19. Plos One. 2021;16(6):e0253566.
15. Splinter B, Saadah NH, Chavannes NH, Kiefte-de Jong JC, Aardoom JJ. Optimizing the Acceptability, Adherence, and Inclusiveness of the COVID Radar Surveillance App: Qualitative Study Using Focus Groups, Thematic Content Analysis, and Usability Testing. JMIR Formative Research. 2022;6(9):e36003.
16. Uiterkamp MH, van Dijk WJ, Heesterbeek H, van der Hofstad R, Jong JC, Litvak N. Value of risk-contact data from digital contact monitoring apps in infectious disease modeling. arXiv preprint arXiv:250321228. 2025.
17. Kist JM, Vos HMM, Vos RC, Mairuhu ATA, Struijs JN, Vermeiren RRJM, et al. Data Resource Profile: Extramural Leiden University Medical Center Academic Network (ELAN). International Journal of Epidemiology. 2024;53(4).
18. ECDC. Case definition for coronavirus disease 2019 (COVID-19) 202 [Available from: https://www.ecdc.europa.eu/en/covid-19/surveillance/case-definition.

19. RIVM NIfPHatE. RIVMdata 2024 [Available from: https://data.rivm.nl/covid-19/.

20. Schonlau M, Zou RY. The random forest algorithm for statistical learning. The Stata Journal. 2020;20(1):3-29.

21. Government D. Coronavirus tijdlijn 2022 [Available from: https://www.rijksoverheid.nl/onderwerpen/coronavirus-tijdlijn.

22. Coxe S, West SG, Aiken LS. The analysis of count data: A gentle introduction to Poisson regression and its alternatives. Journal of personality assessment. 2009;91(2):121-36.

23. Weaver CG, Ravani P, Oliver MJ, Austin PC, Quinn RR. Analyzing hospitalization data: potential limitations of Poisson regression. Nephrology Dialysis Transplantation. 2015;30(8):1244-9.

24. van Brakel R, Kudina O, Fonio C, Boersma K. Bridging values: Finding a balance between privacy and control. The case of Corona apps in Belgium and the Netherlands. Journal of Contingencies and Crisis Management. 2022;30(1):50-8.

25. Hoffman AS, Jacobs B, van Gastel B, Schraffenberger H, Sharon T, Pas B. Towards a seamful ethics of Covid-19 contact tracing apps? Ethics and Information Technology. 2021;23(Suppl 1):105-15.

26. Alanzi T. A review of mobile applications available in the app and google play stores used during the COVID-19 outbreak. Journal of multidisciplinary healthcare. 2021:45-57.

27. Pozo-Martin F, Beltran Sanchez MA, Müller SA, Diaconu V, Weil K, El Bcheraoui C. Comparative effectiveness of contact tracing interventions in the context of the COVID-19 pandemic: a systematic review. European Journal of Epidemiology. 2023;38(3):243-66.

28. Bannister-Tyrrell M, Chen M, Choi V, Miglietta A, Galea G. Systematic scoping review of the implementation, adoption, use, and effectiveness of digital contact tracing interventions for COVID-19 in the Western Pacific Region. The Lancet Regional Health–Western Pacific. 2023;34.

SUPPLEMENTAL MATERIAL

See: