Real-time prediction model of cVDPV2 outbreaks to aid outbreak response vaccination strategies

Arend Voorman[a], Kathleen O'Reilly[b], Hil Lyons[c], Ajay Kumar Goel[d], Kebba Touray[e], Samuel Okiror[e]

Abstract

Background

Circulating vaccine-derived poliovirus outbreaks are spreading more widely than anticipated, which has generated a crisis for the global polio eradication initiative. Effectively responding with vaccination activities requires a rapid risk assessment. This assessment is made difficult by the low case-to-infection ratio of type 2 poliovirus, variable transmissibility, changing population immunity, surveillance delays, and limited vaccine supply from the global stockpile. The geographical extent of responses have been highly variable between countries.

Methods

We develop a statistical spatio-temporal model of short-term, district-level poliovirus spread that incorporates known risk factors, including historical wild poliovirus transmission risk, routine immunization coverage, population immunity, and exposure to the outbreak virus.

Results

We find that proximity to recent cVDPV2 cases is the strongest risk factor for spread of an outbreak, and find significant associations between population immunity, historical risk, routine immunization, and environmental surveillance ($p < 0.05$). We examine the fit of the model to type 2 vaccine derived poliovirus spread since 2016 and find that our model predicts the location of cVDPV2 cases well (AUC = 0.96). We demonstrate use of the model to estimate appropriate scope of outbreak response activities to current outbreaks.

Conclusion

As type 2 immunity continues to decline following the cessation of tOPV in 2016, outbreak responses to new cVDPV2 detections will need to be faster and larger in scope. We provide a framework that can be used to support decisions on the appropriate size of a vaccination response when new detections are identified. While the model does not account for all relevant local factors that must be considered in the overall vaccination response, it enables a quantitative basis for outbreak response size.

Background:

Outbreaks of circulating vaccine-derived poliovirus type 2 (cVDPV2) are spreading widely, putting the entire Global Polio Eradication Initiative (GPEI) at risk, and requiring a new strategy for their control [1,2]. Preventing further transmission of circulating vaccine-derived poliovirus type 2 (cVDPV2) in affected areas of low sanitation can only occur through use of the monovalent oral polio vaccine type 2 (mOPV2)[3]. However, the Sabin 2 virus in mOPV2 can transmit from person-to-person, eventually reverting key attenuating mutations and causing new emergent events and outbreaks of cVDPV2 [2]. Thus, a vaccination response to a cVDPV2 outbreak needs to be large enough to cover populations infected with the outbreak virus, but small enough to avoid unnecessary exposure to mOPV2 which may seed a new outbreak. This is a difficult balance to determine [4]. A novel genetically stable OPV2 (nOPV2) which is designed to retain its attenuation is currently under development, and received approval for emergency use in November 2020. Because its use will be restricted to emergency contexts, care must also be taken to avoid unnecessary exposure until reversion risk can be established and it receives full licensure. Consequently, the challenges of determining scope and geographic extent for cVDPV2 vaccination response with mOPV2 will likely remain through at least 2021, and the challenges with using nOPV2 may be similar.

Defining the at-risk population for further cVDPV2 spread, and outbreak response, is difficult. Surveillance systems for acute flaccid paralysis (AFP) will identify only a small proportion of infected individuals: wild poliovirus type 2 is estimated to cause paralysis in only 1 out of every 2000 individuals infected, and a similar case-to-infection ratio is often assumed for cVDPV2 [5]. Thus, each detection of cVDPV2 likely represents thousands of infected individuals distributed over an unknown geographic area. In addition, a successful vaccination response must cover areas infected at the time of the response, which typically occurs two or more months after paralysis onset of a case, due to laboratory processing times and operational constraints. In the intervening period, the outbreak virus may infect

new populations similarly subject to poor and delayed detectability, which is difficult to account for when making risk assessments.

Prior to withdrawal of OPV2-containing vaccines from routine use in 2016, population immunity to type 2 poliovirus was generally high in most places [6]. Thus, cVDPV2 circulation was limited to small populations in a few countries with poor immunity. Now, however, nearly all areas with poor sanitation are susceptible to transmission of type 2 poliovirus, regardless of routine immunization coverage[7]. As a result, there are fewer barriers to virus spread, and so the extent of spread may be larger and more rapid than would be assumed based on patterns of spread observed in cVDPV2 outbreaks prior to 2016 or in outbreaks of serotypes 1 and 3 poliovirus.

The outbreak response protocol permits a wide range of target populations, from 400 thousand to 4 million for a new detection, and thus additional and rapid quantitative analysis is useful to guide the response strategy[3]. While there have been previous studies informing geographic scope of outbreak response, there are no models currently available which can be used by countries to assess the likely extent of poliovirus spread at the district or second administrative unit level, where outbreak responses are organized. Duintjer-Tebbens and others used a deterministic model to show that larger responses to cVDPV2 were required in areas with lower immunity or higher transmissibility [8]. Spatio-temporal models of poliovirus spread exist, but are often limited to a single country and operate on longer time scales, such as 6 months, than are relevant for cVDPV2 outbreak responses [9,10]. Likewise, global models exist for wild polio virus (WPV1) spread and cVDPV emergence, but since risk is given at the national level, they cannot be used to assess subnational risk necessary for outbreak response [11].

Here we describe a district-level model of cVDPV2 spread that is fitted to historical data and used to forecast short-term spread of outbreaks (1-12 months). This model accounts for polio spread informed from historical observations, the immunity changes induced by withdrawal of tOPV, routine

immunization program performance, and environmental surveillance (ES) data, and can be extended to include additional factors. Since we use a relatively simple regression framework, this model can be quickly updated and be used in near real-time to inform outbreak response activities. As the model was being developed in 2019-2020, the COVID19 pandemic spread globally and resulted in all supplemental immunization activities (SIAs) being suspended due to the contraindication to physical distancing requirements. In preparation for the resumption of SIAs in June 2020, vaccination responses needed to be updated and this model was used to provide guidance to the GPEI.

The rest of the paper is organized as follows. In the Methods section, we list the data used and how it was obtained (Data), and we describe the regression model including how data is used in the model for fitting and validation (Statistical Methods). In the Results section, we describe the risk factors used in the model and provide basic descriptive statistics (Description of the data), as well as the fit of the model to the data (Model results), how well the model performs when forecasting cVDPV2 outbreaks in the subsequent 1-12 months (Forecasting Risk), and how the model is used in an example risk assessment in central Africa (Use in Risk Assessments). Lastly, we conclude in the Discussion with the outlook for cVDPV2 crisis affecting the GPEI, and list some model limitations and future work.

Methods

Data

Our main outcome variable is a new detection of poliovirus in an AFP case or ES sample within a specific district-month. We use polio surveillance data comprising AFP and ES from a database maintained by the WHO for GPEI. While our focus is on estimating cVDPV2 risk, we developed a WVP1 model to capture the shared geographic risks of detection that are common to both epidemics. For our WPV1 model, we used data from 2005-2015 from countries in AFRO and EMRO (from 2015 WPV has been geographically restricted to Pakistan and Afghanistan, with a few exceptions). For our cVDPV2 model we

used data from January 2021 – May 2020. Each isolate of VDPV2 is classified according to an emergence group though genetic clustering described elsewhere [12]. Isolates of cVDPV2 are either cases of poliomyelitis or positive ES samples, ES was not considered in the WPV1 model because ES has not been in widescale use during this time.

To estimate immunity against transmission of serotype 2 poliovirus we used the SIA database maintained by the WHO, which gives the timing and vaccines used for SIAs from 2000-2019. We combined this with estimates of routine immunization coverage from Institute for Health Metrics and Evaluation (IHME) to estimate serotype-specific mucosal immunity for 6- to 36-month-old children for each month at the district level [13].

All data is coded to the district level (m=7125 districts within 69 countries), using the WHO polio geodatabase. Any changes in district boundaries prior to this time are simplified to this specification.

Statistical methods

We modeled poliovirus detection in a new district in a given month, as a function of estimated type-specific immunity, cVDPV2 exposure, historical WPV1 risk, and population size. District-level immunity is estimated based on immunization campaign history using methods described elsewhere[13]. We used a radiation model of population movement which we found to fit the data best, consistent with previous models of measles and polio spread [9,14]. Historical WPV1 risk is estimated with a random effects model measuring the relative risk of WPV1 in a province, compared to a district with similar measured risk factors. To relate these predictors to cVDPV2 detections, we used a Poisson model for cVDPV2 cases and a binomial model for environmental detections, and combined both models into a joint framework. An added advantage of this framework is that the model allows for increases in sensitivity to detections in districts with environmental surveillance. Details of the model equations are given in the technical Appendix.

The model was fitted to data from January 2010 to May 2020 inclusive. To estimate the predictive ability of the model, forecasts were generated starting from each month between January 2010 to May 2020 inclusive, forecasting the subsequent 12 months, and then compared to the observed location of cases where surveillance data was complete (i.e. from February 2010 to May 2020 inclusive). The forecasts and observations were compared visually and by using area under the curve (AUC) diagnostics.

Results

Description of the data

Figure 1 displays the some of the key risk factors for cVDPV2 spread, displayed on a map for the month of June 2020. We see that immunity to Type 2 polio (panel A) is highly restricted to areas where mOPV2 SIAs have been implemented recently (eg. many districts within Angola, Democratic Republic of the Congo, Ethiopia, Nigeria, Niger and Somalia). Examining estimated DPT3 coverage (panel B) shows that the areas with highest Type 2 immunity are often the areas with lowest routine immunization (shown in dark grey). The WPV1 risk score (panel C), which is the estimated relative risk of a WPV1 detection in that district, all other risk factors being equal, and which also coincides with areas with recent mOPV2 campaigns, suggesting common risk factors for cVDPV2 outbreaks and WPV1. Lastly, we display estimated exposure to cVDPV2 (Panel D, where red indicate higher exposure), which decreases rapidly with distance from recent cases. Altogether, one can see stark contrasts of the various risk factors for cVDPV2: in many areas of sub-Saharan Africa we see high and localized immunity due to outbreak response, and at the same time weak routine immunization and high historical risk of WPV1 spread. The relative importance of these risk factors, and thus the likelihood of cVDPV2 spread in areas that are exposed to cVDPV2, will be estimated in the model.

The data used in the model is summarized further quantitatively in Table 1, and contrasts areas with cVDPV2 detections to the general population. There were 7125 districts in AFRO and EMRO available for analysis, with data available for 125 months. There were 510 districts with cVDPV2 detections over this period. Among these detections, 72% were detections from AFP, 30% had detections from ES, and 2% had detections from both AFP and ES. Among districts with cVDPV2 detections, 90% had 1 or fewer cases (max = 4 cVDPV2 cases in one month).

In general, district-months with cVDPV2 detections had lower DPT3 coverage, lower immunity, higher exposure to cVDPV2, higher under-5 population, and higher WPV risk scores.

Model results

Table 2 summarizes the estimated relationship between the model inputs and the risk of cVDPV2 detection in the subsequent month. The variables are associated in the direction and roughly the magnitude one would expect: exposure from cVDPV2 cases is by far the strongest predictor of cVDPV2 spread as measured by the chi-square statistic, and areas with high WPV risk have higher cVDPV2 risk or with increased susceptibility (i.e. lower immunity) are also at higher risk. We also find that districts with environmental sites are around 8 times more likely to detect cVDPV2 than comparable districts without environmental sites. This highlights the importance of environmental surveillance in assessing the geographic extent of cVDPV2 spread. Population size was not significantly ($p<0.05$) associated with cVDPV2 spread, after adjusting for the other risk factors; however we opted to retain it in the model due to its biological plausibility.

Forecasting risk

Overall, the model predicts expanding cVDPV2 outbreaks through 2020 and into 2021 (Figure 2), due to the spread of VDPV2 outside of recent SIA response districts and no further response due to the COVID19 restrictions at the time. The estimated risk of cVDPV2 for February 2010 through May 2021 are

illustrated in Figure 2, based on surveillance data from January 2020 through May 2020 and accounting for SIAs that have been conducted. We estimate the total number of newly infected districts using the most recently available data prior to the month of interest: for February 2010 through May 2020 estimates for a given month are based on the information available in the previous month, while risk estimates for June 2020 through May 2021 are based on forward simulations from the data available through May 2020.

The model achieves high sensitivity and specificity for selecting areas of likely infection by cVDPV2 (Figure 3). The area-under the receiver operating curve, for classifying infection status of districts 1 month in the future is 0.96, and decreases monotonically with time to 0.88 for forecasts 12 months ahead. However, the sparsity of cVDPV2 cases and large geographic scope of the model make it such that broad areas at risk can be identified reliably (ie. regions within a country), but the individual districts in which cases occurs can rarely be predicted with high certainty. For instance, for districts with cases, the average predicted probability of a detection based on the previous month's data is 10%. However, the estimate aggregate number of infected districts Jan 2010 – May 2020 shows that the 95% prediction interval of the number of infected districts includes the observed number in 103 of 124 months (83%).

Use in Risk Assessments.

The model has been used for risk assessment to support outbreak response activities since June 2020. The maps in Figure 4 illustrate an example for Chad, Sudan, South Sudan, and Central African Republic, where at the time of analysis (September 2020) there have been 92 cVDPV2 cases in 2020, from at least 4 different outbreaks, one of which (CHA-NDJ-1) has spread to all four countries. Using the model, we estimate the risk of spread to additional districts. To support vaccination response, districts are ordered by descending VDPV2 risk and the cumulative risk is plotted against cumulative target population size.

These metrics are used to indicate the mOPV2 doses required to vaccinate the highest risk districts, which is then aligned with the mOPV2 supply for that epidemiological block. If no response was carried out (corresponding to a target population of 0), the estimated risk is equivalent to over 35 infected districts. As the target population of the response increases, the remaining risk reduces; as the districts are ordered by decreasing risk, the increasing size of the target population has an initially large effect on reducing risk. The scenarios corresponding to A, B, and C are described in more detail below, but in this example result in a response in excess of 8 million across the four countries.

Criteria for response can vary based on risk tolerance and the perceived cost of the response, and we illustrate several examples here (Figure 4 response A, B and C). One could respond where the risk of cVDPV2 spread outweighs the risk of mOPV2 use[1], as the consequence of either event (additional cVDPV2) can be considered comparable (Figure 4 A). This suggests responses covering all areas not already covered by mOPV2 responses among countries in the risk assessment, and in this example is the largest response. Alternate criteria are also considered, such as choosing a response such that that the expected newly infected districts outside the response zone is less than 1 (Figure 4 B), or less than 0.5 (Figure 4 B), where in this example these criteria result in a moderately smaller response. Still other criteria may be considered, incorporating different models of mOPV2 risk, consequences of cVDPV2 spread, or efficacy of response.

Since this analysis was conducted outbreak response SIAs have been planned covering the areas indicated in Figure 4 A. At the same time, spread of the cVDPV2 outbreaks have been observed prior to the outbreak response, most notably across South Sudan and further into Sudan, and with related

[1] One may estimate the risk of mOPV2 by relating the number of children vaccinated with mOPV2 which could have generated an observed emergence (approximately 150 million at the time of writing) to the number of emergences following tOPV cessation (approximately 50) to estimate a crude risk of 1 outbreak per 3 million children vaccinated.

environmental detections in Egypt. These new detections present additional risk to other areas and will need to be evaluated in light of the anticipated impacts of the planned mOPV2 SIAs.

In practice, the modeling provides one input, which must be complemented with information on additional immunization indicators available at the country and region level, surveillance quality, prior campaign performance, population movements and presence of high-risk groups such as refugees, internally displaced populations, trends of insecurity and instability to come up with proposed scope of response. The overall scope is also invariably influenced by availability of vaccines. In general, we have found that the model suggests larger areas for response than can be approved when also considering mOPV2 emergence risk and vaccine supply.

Discussion

Implications for cVDPV2 outbreak management

This model provides a framework for assessing the risk of cVDPV2 spread using historical epidemiology, for use in risk assessments and planning the scope of outbreak response activities. By using a relatively simple framework, it can be easily updated in near real-time as new data become available. The pairing of the risk model with criteria for response enable a more complete quantitative basis for outbreak response planning. This is particularly important given the relatively poor detectability of cVDPV2 compounded by the need to limit mOPV2 use owing to the potential risk of seeding new cVDPV2 and finite stockpile.

While the primary goal of this tool is ongoing use in risk assessments, it also provides insights into overall considerations for the response to cVDPV2 in the coming years. We find that risk of spread is increased in areas with lower immunity, and conversely, increasing immunity (i.e. outbreak response) is associated with decreased risk. Thus, as population immunity to type 2 poliovirus continues to decline following cessation of tOPV, and following outbreak response campaigns, the number of areas at risk and thus the speed of geographic spread of cVDPV2 are increasing. In order to address this risk, generally larger responses will be required, and our model adds evidence to the overall efficacy of outbreak response activities. Additionally, by forecasting risk over time, we show how the consequences of delayed and inadequately large responses accumulate and place additional areas at risk. This in turn may be used to adjust the scope of response depending on the timing of isolates and the likely time of response, generally recommending increased target populations for delayed campaigns. As shown in the model, the level of risk varies widely based on a range of factors, and thus simple rules of thumb for response are inadequate to guide vaccination programs, and thus quantitative methods as described here will be critical to inform outbreak response until elimination is achieved.

Following this analysis, many geographical areas highlighted as at-risk (Figure 2 C), including Egypt, South Sudan, Senegal, Liberia, and Sierra Leone reported cVDPV2 outbreaks and have required outbreak responses. This spread was not inevitable, but likely a result of delayed and inadequate responses to outbreaks, exacerbated by a pause in activities due to COVID-19.

While mOPV2 is available, outbreak responses will involve its increased use which in turn is likely to seed more outbreaks, though with lesser consequence than allowing current outbreaks to spread if used judiciously, as we describe. With the development and use of the novel OPV vaccine the approaches described here can be readily adapted to inform strategic use against outbreaks, with a reduced risk of emergence.

Looking forward, the prospects of cVDPV2 elimination are not certain, but will depend on the ability of the GPEI to mount high quality and epidemiologically targeted responses. While our model is able to accurately predict how cVDPV2 spread in the past, where the scale and speed of responses was a given, the scale and speed of future responses are not known in advance and so our model is not able to reliably predict the evolution of cVDPV2 outbreaks beyond the near future. However, we suggest that use of near real time modelling will be an essential tool to guide outbreak response teams towards a suitable scale of response that will limit further transmission and cases. This model has been developed with this objective in mind.

Model limitations and future work

Compared to a mechanistic model of polio transmission, such as SEIR models, the regression framework used assumes variables without specifying causal interpretation, but which demonstrate reliable associations with polio epidemiology. While this model is useful for prediction of polio outbreak spread, there are a few notable limitations. Exposure to cVDPV2 is approximated with a radiation model of movement, which will not capture temporal or geographic variation in movement patterns, such as

migratory populations, or movement restrictions such as those implemented due to COVID-19 control measures[15]. Additionally, local factors that affect baseline risk of cVDPV2 are estimated from WPV1 spread. However, this may be inadequate in situations where the local factors have changed, or where WPV1 did not spread over the period considered. Immunity and the impact of SIAs is estimated from a simple model which has its own limitations and does not in general account for local factors such as response quality or variations in vaccine efficacy [13].

There are several reasons the model may underestimate risk. For one, we model observed cases rather than asymptomatic infections, which are necessarily more widespread than observed cases. One could account for this by examining risk in a longer window over which latent infections would be expected to result in an observed case, or by constructing a model that estimates latent infections[16]. However, models with latent infections are computationally intensive and typically make stronger assumptions on disease dynamics [17]. Additionally, the model also relies on clinical surveillance, both currently and in historical outbreaks that are used to inform the model, while in practice surveillance systems may not detect all cases, and while a majority of cases are reported promptly there is a delay between onset and confirmation[18], both leading to an under-assessment of risk. Lastly, the model estimates spread of existing cVDPV2 outbreaks, but does not explicitly estimate the risk of emergence of new outbreaks. However, vaccination responses with mOPV2 are recommended only in response to detections of VDPV2, and therefore a model of spread is sufficient for organizing outbreak response activities.

Further model development will focus on inclusion of IPV immunity into the model and an investigation on how international migration, nomadic and seasonal population movements can further improve model prediction.

References

[1] The Global Polio Eradication Initiative. Strategy for the Response to Type 2 Circulating Vaccine-Derived Poliovirus 2020–2021. 2020.

[2] Macklin GR, O'Reilly KM, Grassly NC, Edmunds WJ, Mach O, Santhana Gopala Krishnan R, et al. Evolving epidemiology of poliovirus serotype 2 following withdrawal of the type 2 oral poliovirus vaccine. Science (80-) 2020. https://doi.org/10.1126/science.aba1238.

[3] World Health Organization. Responding To a Poliovirus Event or Outbreak. 3.1 March. 2020.

[4] Pons-Salort M, Burns CC, Lyons H, Blake IM, Jafari H, Oberste MS, et al. Preventing Vaccine-Derived Poliovirus Emergence during the Polio Endgame. PLoS Pathog 2016. https://doi.org/10.1371/journal.ppat.1005728.

[5] Kalkowska DA, Duintjer Tebbens RJ, Pallansch MA, Cochi SL, Wassilak SGF, Thompson KM. Modeling undetected live poliovirus circulation after apparent interruption of transmission: Implications for surveillance and vaccination. BMC Infect Dis 2015. https://doi.org/10.1186/s12879-015-0791-5.

[6] Pons-Salort M, Molodecky NA, O'Reilly KM, Wadood MZ, Safdar RM, Etsano A, et al. Population Immunity against Serotype-2 Poliomyelitis Leading up to the Global Withdrawal of the Oral Poliovirus Vaccine: Spatio-temporal Modelling of Surveillance Data. PLoS Med 2016. https://doi.org/10.1371/journal.pmed.1002140.

[7] Famulare M, Selinger C, McCarthy KA, Eckhoff PA, Chabot-Couture G. Assessing the stability of polio eradication after the withdrawal of oral polio vaccine. PLoS Biol 2018. https://doi.org/10.1371/journal.pbio.2002468.

[8] Duintjer Tebbens RJ, Pallansch MA, Wassilak SGF, Cochi SL, Thompson KM. Characterization of outbreak response strategies and potential vaccine stockpile needs for the polio endgame. BMC Infect Dis 2016. https://doi.org/10.1186/s12879-016-1465-7.

[9] Mangal TD, Aylward RB, Mwanza M, Gasasira A, Abanida E, Pate MA, et al. Key issues in the persistence of poliomyelitis in Nigeria: a case-control study. Lancet Glob Heal 2014;2:e90-7. https://doi.org/10.1016/S2214-109X(13)70168-2.

[10] Upfill-Brown AM, Lyons HM, Pate M a, Shuaib F, Baig S, Hu H, et al. Predictive spatial risk model of poliovirus to aid prioritization and hasten eradication in Nigeria. BMC Med 2014;12:92. https://doi.org/10.1186/1741-7015-12-92.

[11] O'Reilly KM, Lamoureux C, Molodecky NA, Lyons H, Grassly NC, Tallis G. An assessment of the geographical risks of wild and vaccine-derived poliomyelitis outbreaks in Africa and Asia. BMC Infect Dis 2017;17. https://doi.org/10.1186/s12879-017-2443-4.

[12] Burns CC, Shaw J, Jorba J, Bukbuk D, Adu F, Gumede N, et al. Multiple Independent Emergences of Type 2 Vaccine-Derived Polioviruses during a Large Outbreak in Northern Nigeria. J Virol 2013. https://doi.org/10.1128/JVI.02954-12.

[13] Voorman A, Lyons H, Bennette C, Kovacs S, Kumar J, Vertefeuille J, et al. Analysis of population immunity to poliovirus following cessation of trivalent oral polio vaccine. Vaccine (in Press 2020.

[14] Bjørnstad ON, Grenfell BT, Viboud C, King AA. Comparison of alternative models of human movement and the spread of disease. BioRxiv 2019. https://doi.org/10.1101/2019.12.19.882175.

[15] Bharti N, Tatem AJ, Ferrari MJ, Grais RF, Djibo A, Grenfell BT. Explaining seasonal fluctuations of measles in Niger using nighttime lights imagery. Science (80-) 2011. https://doi.org/10.1126/science.1210554.

[16] Molodecky NA, Jafari H, Safdar RM, Ahmed J, Mahamud A, Bandyopadhyay AS, et al. Modelling the spread of serotype-2 vaccine derived-poliovirus outbreak in Pakistan and Afghanistan to inform outbreak control strategies in the context of the COVID-19 pandemic. Vaccine (in Press 2020.

[17] Martinez-Bakker M, King AA, Rohani P. Unraveling the transmission ecology of polio. PLoS Biol 2015. https://doi.org/10.1371/journal.pbio.1002172.

[18] Blake IM, Chenoweth P, Okayasu H, Donnelly CA, Aylward RB, Grassly NC. Faster detection of poliomyelitis outbreaks to support polio eradication. Emerg Infect Dis 2016. https://doi.org/10.3201/eid2203.151394.

Figures and tables

	All District-months	*District-months with cVDPV2 Detections*
Districts (n)	7125	510
Months (n)	125	110
District-months (n)	890,625	972
Cases (mean, IQR)	0.0012 (0, 0)	1.1 (1, 1)
DPT3 Coverage (mean, IQR)	0.72 (0.57, 0.92)	0.51 (0.3, 0.73)
Type 2 Immunity (mean, IQR)	0.71 (0.6, 0.92)	0.51 (0.005, 0.87)
Exposure (log-10) (mean, IQR)	-6.8 (-7.6, -6)	-2.5 (-3.8, -0.79)
Population (100k) (mean, IQR)	0.38 (0.11, 0.47)	0.81 (0.28, 0.82)
WPV Risk Score (mean, IQR)	1.8 (0.62, 2.7)	3.5 (2.1, 4.3)

Table 1: Summary of the data. Columns give the summary for district-months with and without cVDPV2 detections

Variable	*Relative risk (95% CI)*	*Wald (Chi-square) statistic*	*p-value*
Susceptibility (log10)	1.77 (1.65, 1.89)	271.5	< 2e-16
DPT3 (%)	0.61 (0.45, 0.81)	11.3	0.000755
Exposure (log10)	1.58 (1.55, 1.60)	3257.4	< 2e-16
ES Sample (yes vs no)	7.50 (6.40, 8.79)	621.5	< 2e-16
Population (100k)	1.06 (1.00, 1.13)	3.5	0.063048
WPV Risk (relative risk)	1.87 (1.66, 2.10)	109.4	< 2e-16

Table 2: Parameter estimates, confidence intervals, and significance, for model fit.

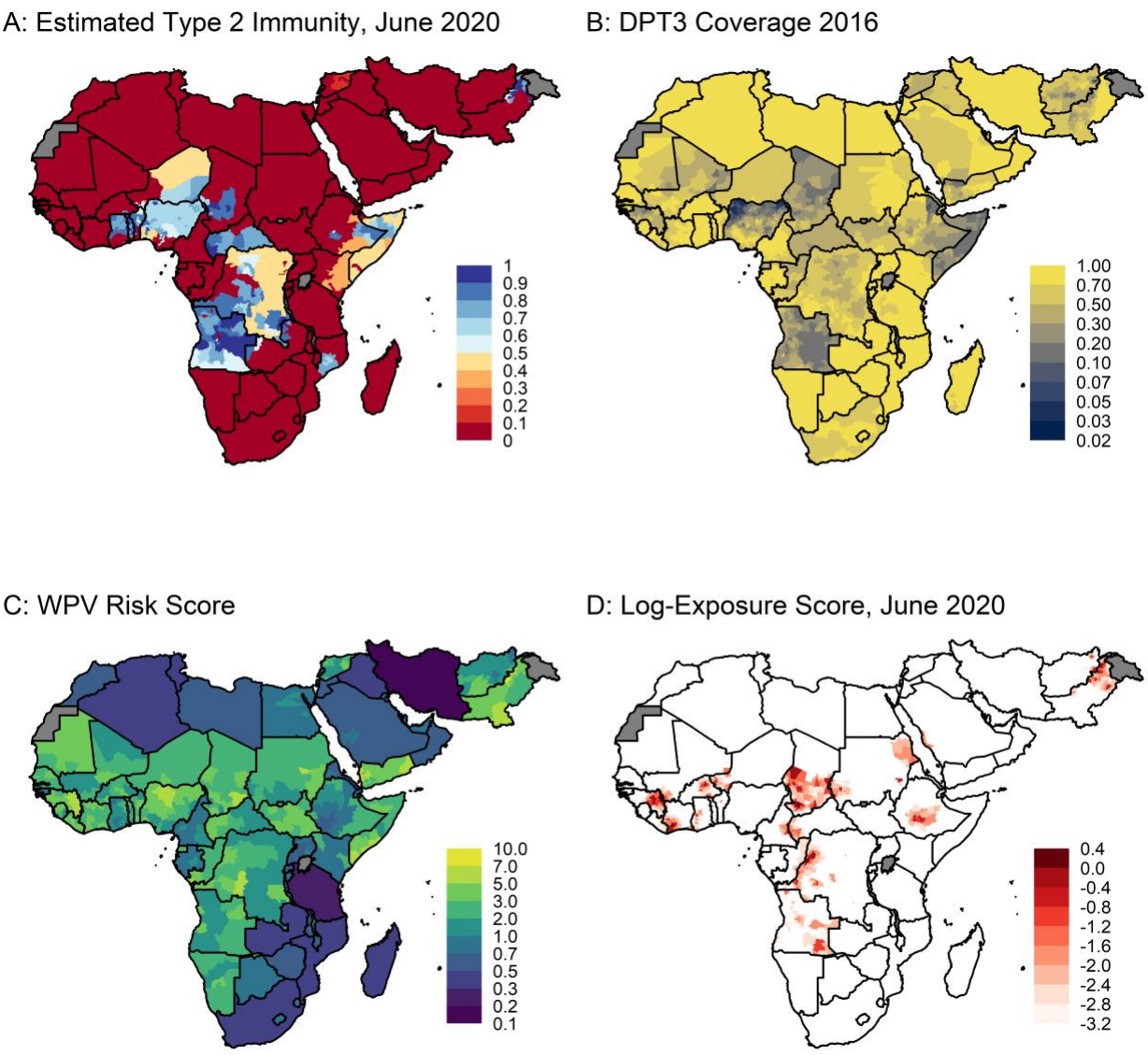


Figure 1: Spatial distribution of the variables used to estimate cVDPV2 risk. Counter-clockwise from top left A: WPV Risk score, B: DPT3 Coverage, C: Immunity, and D: Exposure, for June 2020. Disputed areas shown in grey.

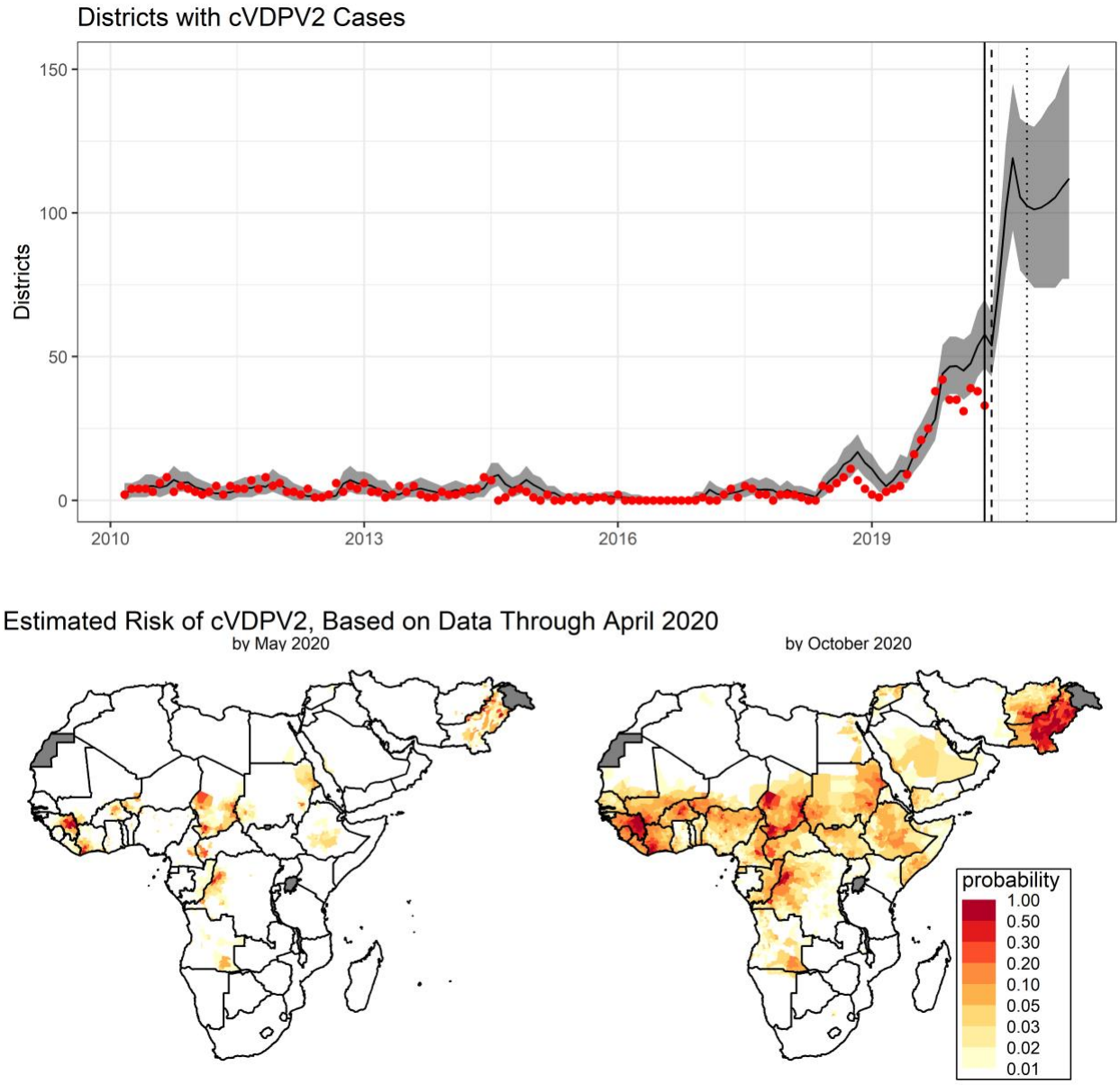
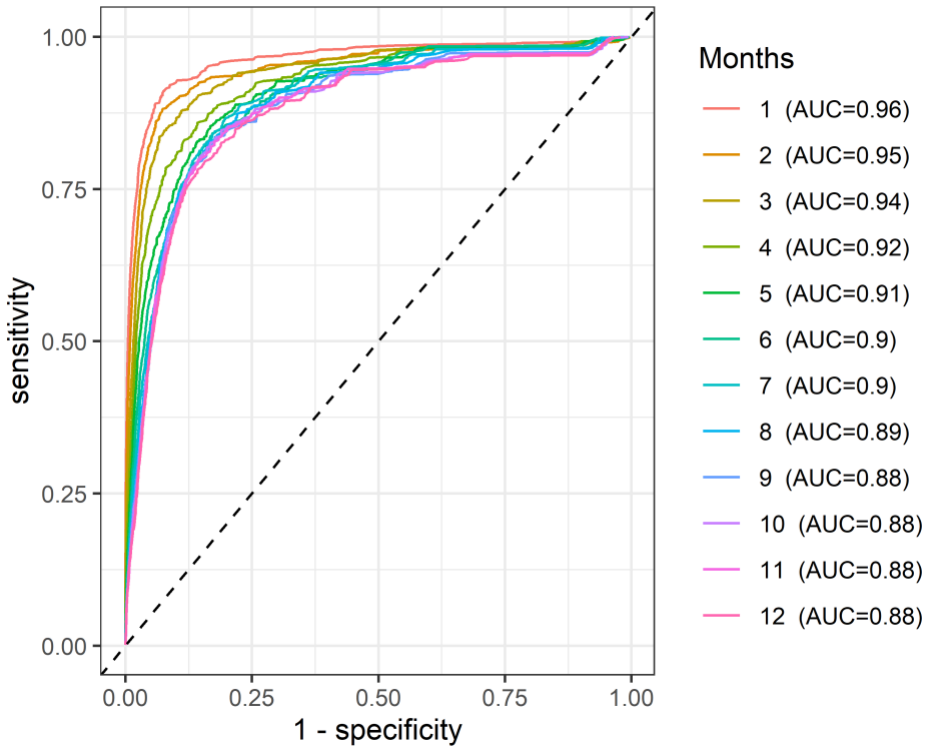


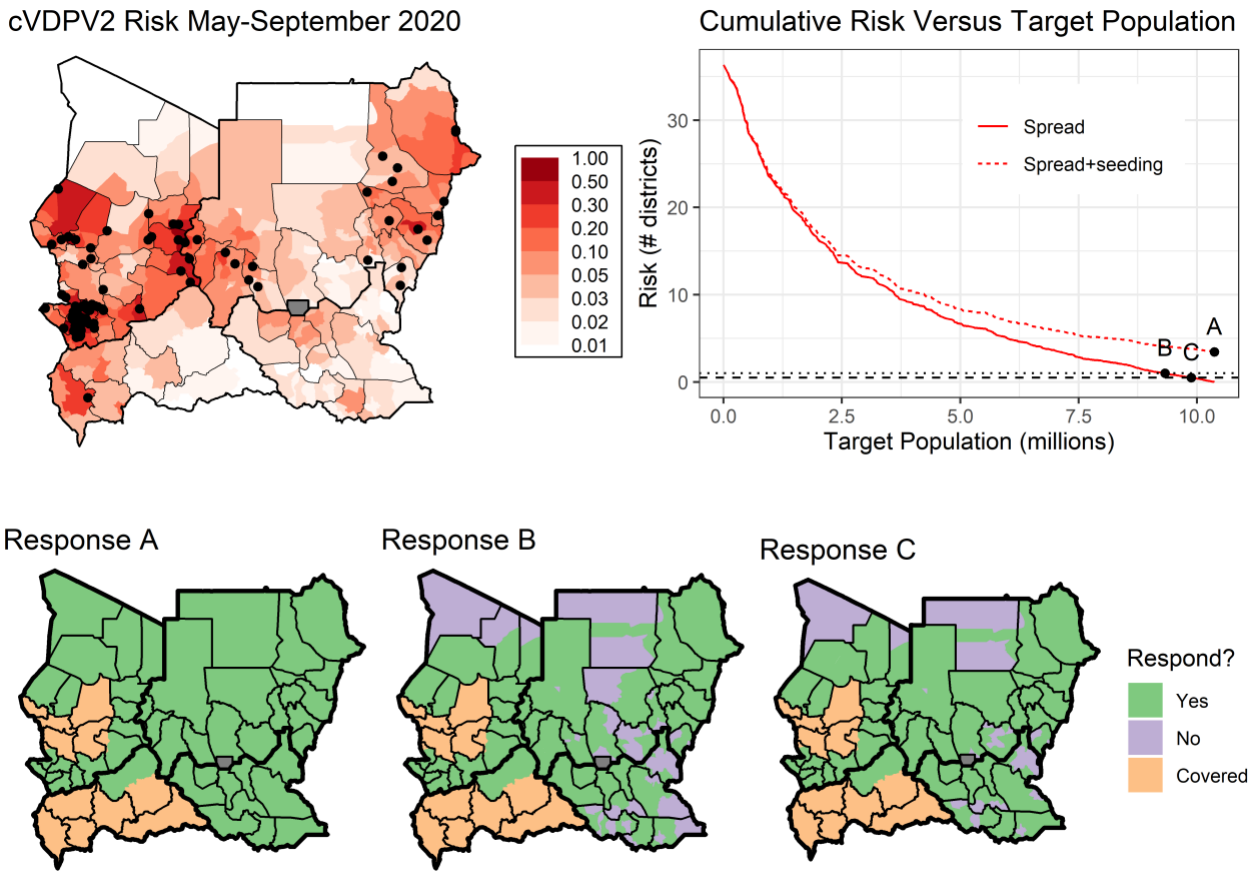
Figure 2: Estimated risk of cVDPV2. Top panel: predicted cVDPV2-infected districts, 2010-2020, with Poisson prediction intervals in grey. Red dots give observed districts with cVDPV2 cases in a given month. Dotted vertical lines indicate June and November 2020. Bottom panel: estimated Risk of cVDPV2 spread, looking one month ahead of the available data (June 2020) and six months ahead (November 2020).



357

358 **Figure 3: Sensitivity and specificity of district-level forecasts of cVDPV2 case locations, from January**
359 **2016 – June 2020. Curves indicate forecasts for different lengths of time in the future (1-12months).**

360



361

362 **Figure 4: Example response options. The top left panel give the reported cVDPV2 cases (black circles)**
363 **and a map of district probabilities of one or more cVDPV2 case in May-September. The top right panel**
364 **gives the cumulative risk of one or more cVDPV2 detections outside of a response area as a function**
365 **of response size (solid line), and also considers the risk of seeding a new outbreak with the response**
366 **(dashed line). The lower panels give the recommended response options that minimize total risk of**
367 **spread and seeding (Response A), or that reduce the expected risk of cVDPV2 spread to 1 district**
368 **(Response B), or 0.5 districts (Response C).**

369

370

371