Sub-national estimation of surveillance sensitivity to inform declaration of disease elimination

A retrospective validation against the elimination of wild poliovirus in Nigeria

Emily S Nightingale[1], Ly Pham-Minh[2], Isah Mohammed Bello[2], Samuel Okrior[3], Tesfaye Bedada Erbeto[4], Marycelin Baba[5], Adekunle Adeneji[5], Megan Auzenbergs[1], W John Edmunds[1], Kathleen M O'Reilly[1]

Affiliations:

1. Centre for Mathematical Modelling of Infectious Diseases, Faculty of Epidemiology and Population Health, London School of Hygiene and Tropical Medicine, London, UK.
2. Polio Eradication Department, World Health Organization, Geneva, Switzerland.
3. The Bill and Melinda Gates Foundation, Seattle, Washington, USA.
4. World Health Organization, Country Office, Abuja, Nigeria.
5. National Polio Laboratory, Ibadan, Nigeria.

Abstract

A fundamental question in the global commitment to polio eradication is how long a period of absence would be consistent with regional elimination, and the safe withdrawal of the oral polio vaccine is contingent on the answer. We present a statistical framework to estimate time-varying sensitivity of two key components of polio surveillance - environmental sampling and clinical cases of acute flaccid paralysis - for detecting infection at the local government authority level. We estimate the probability of freedom from infection (FFI) at a critical prevalence level that is consistent with interruption of transmission, given the absence of virus in collected samples. We validated this framework against two periods of poliovirus absence in Nigeria (2014-2016 and 2016-2020).

We observed substantial heterogeneity in surveillance sensitivity over time and space and, given this, concluded an 85% probability (95% uncertainty interval: 77.1-90.2%) of the country being free from WPV1 infection after 23 months without detection from July 2014. Detection of WPV1 in July 2016 demonstrated that circulation had indeed persisted during this time. In contrast, we conclude a probability of 98% (97.9-98.7%) by the time elimination of the serotype was officially declared in 2020.

The inferred probability of FFI during both time periods was found to be consistent with the retrospectively known status of regional elimination. This supports the validity of applying this framework prospectively to inform the certification of wild poliovirus elimination from remaining endemic regions, and to determine the resolution of cVDPV2 outbreaks.

Introduction

Since the World Health Assembly's resolution for a global commitment to polio eradication, an important question has been how long of a period of absence would be consistent with local elimination. Confidence in the attainment of elimination must be inferred from surveillance data and our understanding of the system's sensitivity to detect circulating viruses. The latter is challenging to quantify for complex systems consisting of multiple components and data streams, such as that for poliovirus surveillance.

In contrast to smallpox (which was eradicated in 1980) poliovirus is known to cause asymptomatic infection, meaning that infection can effectively spread 'silently' within a population for some time until a paralytic case is reported. In a paper by Eichner and Dietz from 1996 (1), a transmission model of both symptomatic (causing acute flaccid paralysis, or AFP) and asymptomatic infections was used to estimate the probability of unobserved, asymptomatic infections with increasing time since the last observed paralytic case. The authors concluded that a three-year time period would be sufficient to reach a 95% probability that the number of infections had been brought to zero. This inference was utilised by the Global Polio Eradication Initiative's Global Certification Commission (GPEI, GCC) to define initial criteria for declaration of national polio elimination. The modelling, however, assumed that all paralytic cases would be reported and did not account for environmental surveillance (ES) for circulating virus in wastewater, which is now an important tool in at least 45 countries (2).

More recent analyses have accounted for a more nuanced combination of surveillance strategies. Assuming varying degrees of "quality" of both AFP and ES (3–7) is found to strongly influence confidence in interruption of transmission given an absence of observed positives. The GCC's interpretation of elimination therefore takes into account sub-national indicators of surveillance quality, which are monitored with respect to defined thresholds for "adequate" performance. Key indicators of interest are the rate of reported non-polio AFP among children under 15 (aiming for at least 2 per 100,000 in endemic countries), the proportion of reported AFP cases for whom adequate stool sample is collected (at least 85%), and the detection rate of non-polio enteroviruses in collected wastewater samples (at least 50%) (8,9). The fact that these indicators do not directly translate to a formal definition of sensitivity, and vary substantially over time and space, complicates interpretation of what performance is adequate to conclude elimination.

Several previous studies describe the "sensitivity" of polio surveillance, but few refer to the formal definition of the term. Sensitivity is referred to as a status achieved by meeting the WHO-defined thresholds for the above indicators (10,11), or by the presence of correlation between vaccine doses distributed and detection rate in collected samples (12). The detection rate of non-polio enteroviruses in ES has been used as a relative measure of sensitivity with which to rank sampling sites (5), and the relative sensitivity of ES and AFP explored according to which method detected an outbreak or cluster first (or at all) (13).

Ranta (14) and O'Reilly (4) defined epidemiological models of polio transmission to infer a conditional probability of detection given infection status - the former via a simulated population and the latter fit to observed data from Pakistan - and Kroiss (15) defines a similar probability via the detection rate of vaccine virus at ES sites in proximity to known vaccination activities. Watkins (16) employed a scenario tree approach to model the AFP surveillance system and estimate its sensitivity relative to an assumed infection prevalence. Although the latter four studies all aim for the formal definition of sensitivity as opposed to a proxy indicator of surveillance performance, differences in definitions, assumptions, context and data sources result in highly variable estimates, from 1-30% for AFP surveillance and 30-55% for ES.

If surveillance can be quantified with respect to the probabilistic definition of sensitivity for detecting infection, this can be inverted via standard probability rules to obtain the probability of a negative outcome given absence of infection. The trajectory with which this conditional probability increases with each month of negative observations can provide a timeline for declaring elimination. This approach, termed 'freedom from infection' (FFI), was proposed by Martin et al. in 2007 (17) has been applied widely in veterinary medicine (18–22) and for a number of human infectious diseases, including polio (4,16,23–26). Here, we apply and validate the estimation of regional freedom from wild poliovirus infection using the retrospective example of wild poliovirus serotype 1 (WPV1) elimination in Nigeria.

Prior to 2018, Nigeria was endemic for wild poliovirus. Elimination was almost declared in 2016, when WPV1 had not been detected in any local government authority (LGA, equivalent to district) for a period of nearly two years (27,28). A modelling study based on data up to 31st March 2015 concluded an 84% probability that WPV1 had been eliminated in Nigeria, and that if elimination *hadn't* been achieved a case would be detected with >99% probability by the end of 2015 (29). This study crudely accounted for imperfect surveillance sensitivity by reducing the assumed case-to-infection (of 1:200) ratios by 50%, and did not account for evidence gathered from ES. In fact, WPV1 re-emerged in July 2016 in four clinical cases and four healthy contacts, all within the northeastern state of Borno. No evidence of WPV1 has been detected through either AFP or environmental surveillance since 2016, and interruption of transmission was officially certified in August 2020 (28).

We describe and apply a scenario tree approach to model the sensitivity of surveillance and estimate the probability of FFI for Nigeria during these two significant periods without detection of WPV1 (2014-16 and 2016-20). We consider the evident persistence of circulation as of July 2016 and the official elimination as of August 2020 as known outcomes, against which we validate our inferred confidence in FFI. We illustrate the monthly accumulation of evidence from negative surveillance observations, conditional on time-varying surveillance sensitivity during these two periods, and discuss what value this approach offers for the GCC's declaration of global poliovirus eradication.

Materials and Methods

Polio surveillance and data collection

Children under the age of 15 years who present to primary care with sudden weakness or paralysis (ie. AFP) in their limbs should be referred for stool testing to determine whether poliovirus infection is present. Every notification of AFP is recorded by the national polio surveillance programme and entered into the GPEI's Polio Information System (POLIS) (30). It is flagged as to whether the patient had adequate stool collected for analysis - defined as two separate samples collected >24h apart, within 14 days of onset of paralysis and which arrive at the lab in good condition (31) - and this should be achieved for at least 85% of cases. The outcome of the stool analysis defines the case as either non-polio AFP - the vast majority, or a clinical case of paralytic polio.

Samples of wastewater are collected regularly from defined sites ranging from formal sewage treatment plants to open drains and natural waterways. Sites have been selected to capture a population of approximately 100,000 people; however, in practice, the catchment is likely to vary substantially depending on the nature of the site and may not reach this intended catchment size. Collected samples are concentrated and analysed for the presence of polioviruses and non-polio enteroviruses (NPEV). Samples are refrigerated for up to 3 days or frozen if more time is needed, and arrive at the lab within 3 days of collection, with adequate quality interpreted as an NPEV detection rate of at least 50% (2). The location, timing and outcomes from each sample are also recorded within the POLIS system.

For the present analysis, reported cases of AFP and collected environmental samples between January 2013 and August 2020 in Nigeria were obtained from POLIS. We consider a monthly time scale and a spatial scale of LGAs, ie. administrative level two. Gridded ($100m^2$) population count estimates were obtained from WorldPop (32), for inference of proportional population catchment around environmental surveillance sites in each LGA. The proportion of under-15s per LGA was taken from sex and age-specific estimates for the time period of interest, projected from the 2006 census (33).

Immunity modelling

Monthly estimates of mucosal immunity against paralytic poliomyelitis per LGA were estimated using the same methods as used in a previous analysis of cVDPV2 outbreak risk on a sub-national level (34). These are informed by a regression analysis of caregiver-reported vaccine doses relative to recent Supplementary Immunisation Activities (SIAs) among reported non-polio AFP cases across 47 countries, and estimates of routine immunisation coverage from the Institute for Health Metrics and Evaluation (IHME) (35).

Statistical framework

Our goal is to understand whether poliovirus is truly absent from a region when no evidence of infection has been observed through the surveillance system. If T^- denotes the *observed* absence of infection and I^- denotes the underlying *true* absence of infection, then this is the conditional probability $P[T^-|I^-]$. An analogy in the medical literature is the negative predictive value (NPV) of an individual diagnostic test, i.e.

$$NPV = P[T^-|I^-] = \frac{True\,negatives}{True\,negatives + False\,negatives}$$

Via standard probability rules, the NPV is linked to the sensitivity Se and specificity Sp of the test, and to the underlying prevalence (p) of the condition being tested for:

$$NPV = \frac{Sp * (1-p)}{Sp * (1-p) + (1-Se) * p}$$

Here, we follow the approach from (17), extending this concept to a population level and construct a multi-stage, multi-component model of a “test” that represents the complexity of the polio surveillance system. In this scenario, a negative result is defined when *all* AFP cases and environmental samples tested within a particular time step are found to be negative for poliovirus. Conversely, a positive result is defined if poliovirus is detected in *any* tested sample.

Freedom From Infection

By assuming a critical prevalence level which would be consistent with interruption of transmission (the *design prevalence*, Dp), quantifying the probability of a positive result from this test given such a prevalence (the test sensitivity Se) and assuming that false positive results are not possible (all positive results are followed up to complete confirmation, resulting in test specificity Sp of 100%), we are able to calculate the NPV defined above. This is interpreted as the probability that a region is genuinely free from infection above the design prevalence (i.e. elimination has been achieved) during a particular time step, given that the surveillance system has returned a negative result. This probability is defined as the *probability of freedom from infection*, or $P[FFI]$.

On a regional level, we specify a design prevalence such that if the country overall had ongoing circulation, a positive result would be observed in *at least one* LGA, $Dp^{region} = 1/774$. On an LGA level, we specify a design prevalence of one infection per 100,000 population, $Dp = 1/100,000$, and compare this to alternative, less stringent, values in a supplementary analysis.

Modelling surveillance sensitivity

The general definition of sensitivity is the conditional probability that a test (T) yields a positive result, given that the condition being tested for (I) is truly present. Formally,

$$Se = P[T^{+}|I^{+}] = \frac{True\,positives}{True\,positives + False\,negatives}$$

To estimate this probability for the entire surveillance system as a "test" for underlying poliovirus infection, the two components (AFP and ENV) are broken down into the sequence of events that lead to a positive result, i.e. a reported detection of poliovirus (*Figure 1*). We assume that these surveillance components operate independently, and are therefore treated as independent probabilities.

The assumed stages of detection of poliovirus via AFP surveillance are as follows: an infection must first develop into clinical disease (i.e. with symptoms of paralysis; $P^{Clinical}$), then such symptoms must be recognised and presented to a healthcare provider ($P^{Notified}$). The provider must collect adequate stool samples from the patient ($P^{StoolAdeq}$) to be sent for testing. Finally, poliovirus must be detected in the sample via lab analysis ($P^{StoolTest}$).

For environmental surveillance, an infected individual will only be captured if they are defecating and shedding virus into the wastewater within the catchment area of an active site ($P^{Catchment}$). The duration of shedding must coincide with collection of a sample at that site ($P^{Sampled}$), and the sample collected must be of adequate quality to detect enteroviruses ($P^{SampleAdeq}$). Finally, as before, presence of the virus must be detected via lab analysis (assumed to be virus culture and intratypic differentiation as per the protocol for the Global Polio Laboratory Networks (36); $P^{SampleTest}$). We define the catchment of ES as the area within 5km of an actively sampling site, within the boundary of the LGA in which the site falls. This definition excludes areas within 5km which extend into neighbouring LGAs, to which no samples are attributed. This will result in some under-estimation of the overall coverage of ES, but allows estimation of sensitivity within the specific LGA in which ES is active. A supplementary analysis explores the influence of the choice of catchment radius.

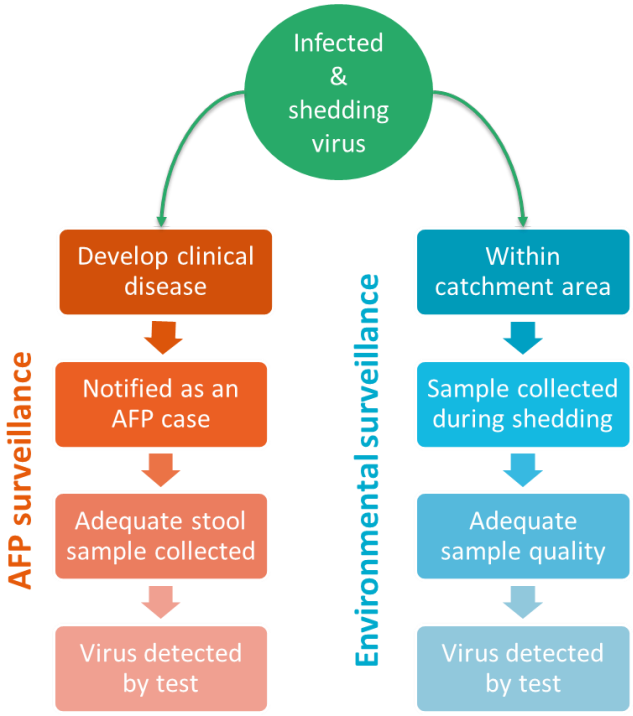

Figure 1: *Illustration of the two components of the polio surveillance system - AFP surveillance and environmental surveillance - and the sequence of events which must occur for a single infection to result in a reported detection of poliovirus.*

Given this set of events, we define the individual-level sensitivity as the probability that a *single* infected individual would be detected through AFP surveillance at time t in LGA i as

$$P_{it}^{unit,AFP} = P_i^{Clinical} * P_i^{Notified} * P_{it}^{StoolAdeq} * P_i^{Stooltest}$$

and through ENV surveillance,

$$P_{it}^{unit,ENV} = P_{it}^{Catchment} * P_{it}^{Sampled} * P_i^{SampleAdeq} * P_i^{SampleTest}$$

using the same notation. A distribution for the probability of each of these events is specified, either based on assumptions from the literature or estimated from recent data in the LGA. A detailed explanation of how the probabilities for each event were specified is given in the *Supplementary Materials*, and a summary is provided in *Table 1*.

Table 1: *Specification of event probabilities within each surveillance component. Each event must occur for a given infected individual to result in a reported detection of poliovirus. For all parameters except catchment of ES, a Beta distribution is fit to the given value and uncertainty interval from which random draws are used in the subsequent analysis. Uncertainty in the catchment population is explored through a sensitivity analysis of the assumed radius.*

Surveillance component	Event	Fixed / Time-updated	National / LGA-specific	Estimate (95% certainty interval, where applicable)
AFP	An infection develops into clinical case of AFP: $P^{Clinical}$	Fixed	National	0.005 (0.004 - 0.006)
	A clinical case of AFP is notified to the health system: $P_{it}^{Notified}$	Time-updated	LGA-specific	0.9 (0.6 - 0.99) *Scaled by LGA rate of reported AFP relative to 2/100,000 target, prior 12m.*
	Adequate stool sample is collected from the case for laboratory testing: $P_{it}^{StoolAdeq}$	Time-updated	LGA-specific	*Proportion of reported AFP cases with adequate stool sampling, prior 12m (with 95% binomial confidence interval).*
	Laboratory testing detects poliovirus in the sample: $P^{StoolTest}$	Fixed	National	0.97 (0.95 - 0.999)
Environmental	Infected individual is within the catchment area of active ES site(s): $P_{it}^{Catchment}$	Time-updated	LGA-specific	*Proportion of LGA population within 5km of currently active ES site(s), with uncertainty interval of +/- 15%*
	A sample is collected while the individual is shedding virus: $P_{it}^{Sampled}$	Time-updated	LGA-specific	*Median and 95% quantile interval across simulations, given mean sampling frequency, prior 12m.*
	Collected sample is of adequate quality to detect virus: $P_{it}^{SampleAdeq}$	Time-updated	LGA-specific	*Proportion of samples in which NPEV was detected per LGA, in the last 12m (with 95% binomial confidence interval).*
	Laboratory testing	Fixed	National	0.9 (0.7 - 0.99)

	detects poliovirus in the sample: $P^{SampleTest}$			

Sensitivity per LGA

The LGA-specific sensitivity of AFP and ENV surveillance is the probability that any infected individual in the assessed population will yield a positive outcome, given that there are infected individuals in the population at the assumed design prevalence, Dp. Here, we consider the number of individuals "assessed" by the surveillance system, N_i, to be the under-15 population size. Although the ES system in theory captures all ages, those older than 15 years have minimal risk of poliovirus infection due to built-up immunity (37). They would have minimal influence on sensitivity since their probability of being infected is negligible, therefore we chose to exclude the older population from our calculations.

Considering all under-15s in the LGA as independent, the probability of *any* assessed individual yielding a positive outcome, given that each is infected with probability Dp is

$$Se_{it}^{AFP/ENV} = 1 - (1 - Dp * P_{it}^{unit,AFP/ENV})^{N_i}$$

since we assume a specificity of 100%.

Relative risk of circulation per LGA

To accommodate differential risk of poliovirus circulation between LGAs, we adjust our calculation of sensitivity according to a relative risk based on data from the past 12 months. Our estimation of the underlying relative risk of circulation between LGAs followers the approach described by Voorman et al. (34,38), which incorporated estimated mucosal immunity, positive observations in the past 12 months (both AFP cases and environmental samples) and a radiation model of risk between neighbouring LGAs (suggested to reflect the dynamics of WPV spread in Nigeria better than distance or gravity models (39)). This relative risk is then transformed into an *adjusted* risk (AR) for each LGA, which is used to scale the estimated sensitivity in each LGA. In principle, inclusion of the adjusted risk means that increased surveillance effort in higher risk LGAs corresponds to a higher estimate of sensitivity than focussing this effort on lower risk LGAs.

Overall system sensitivity

The sensitivity of each component across *all* LGAs (ie. nationally) is one minus the probability of not detecting infection in *any* LGA (adjusted for variable underlying risk), given that infection is present at the design prevalence in at least one LGA. This is one minus the product of the probabilities that each LGA would not detect infection, $Pneg_{it}^{AFP/ENV}$:

$$Se_t^{AFP} = 1 - \prod_{i=1}^{I}(Pneg_{it}^{AFP})$$

$$Se_t^{ENV} = 1 - \prod_{i=1}^{I}(Pneg_{it}^{ENV})$$

where

$$Pneg_{it}^{AFP} = 1 - (Dp^{region} * Se_{it}^{AFP} * AR_i)$$

$$Pneg_{it}^{ENV} = 1 - (Dp^{region} * Se_{it}^{ENV} * AR_i)$$

If an LGA has no operating ES the Se^{ENV} is equal to zero. The total, national-level sensitivity of the entire system is then

$$Se_t = 1 - (1 - Se_t^{AFP}) * (1 - Se_t^{ENV})$$

This can be interpreted as the probability that the surveillance system would yield a positive observation if infection was present at the design prevalence in at least one LGA.

To propagate uncertainty in the underlying step-wise probabilities, sensitivity is estimated across 1,000 draws from beta distributions fit to the central values and certainty intervals given in *Table 1*.

FFI probability over time

The probability of Freedom From Infection (FFI) is defined as the probability that infection is below the design prevalence, given that no positive AFP cases or environmental samples have been observed in *any* LGA. This is conditional on the prior months with no positives, and the possibility of re-introduction, hence we calculate it by applying Bayes theorem.

Specifically, the probability of freedom from infection in month *t+1*, given infection status and surveillance sensitivity in month *t* and incorporating risk of introduction, is defined as

$$P_{t+1}^{Free} = (1 - P_t^{Infect})/(1 - P_t^{Infect} * Se_t)$$

P_t^{Infect} is the probability that the country was actually infected at time *t*, despite no positives being detected. This could be either because infections were missed and a false negative conclusion was drawn ($1 - P_t^{Free}$), or because infection was re-introduced (p^{Intro}). As WPV1 at the time had been declared eliminated from all other African countries, and remained endemic only in Afghanistan and Pakistan, this risk of introduction was assumed to be low but

non-zero. For this parameter we therefore defined a beta distribution centred on 1/1,000 with 95% quantile interval between 1/500 and 1/5000.

The union probability of these two scenarios is

$$P_t^{Infect} = (1 - P_t^{Free}) + p^{Intro} - (1 - P_t^{Free}) * p^{Intro}$$

P_1^{Free} is the assumed prior probability that infection is below the design prevalence, given that the last positive detection was observed in month 0. For the primary analysis we assume a beta distribution centred on 0.5 for this value, but explore alternatives in sensitivity analyses. For subsequent time steps, the prior probability is defined sequentially as the previously calculated probability P_{t-1}^{Free}.

Availability of data and materials

Detailed disease surveillance data on which this research is based are available from the WHO Institutional Data Access/Ethics Committee for Global Polio Eradication Initiative research partners who meet the criteria for access to confidential data. The code used to conduct this analysis is available in the following repository, alongside a test dataset of randomised values: https://github.com/kath-o-reilly/polio-spec-models/

Ethical statement

This study was approved by the ethics committee of the London School of Hygiene and Tropical Medicine (approval number: 29381, granted 18th April 2023).

Results

Observed surveillance activity

Detection of WPV1

Wild poliovirus serotype 1 had been consistently detected among reported cases of AFP during the years prior to 2014 and occasionally detected in environmental samples. Between August 2014 and June 2016 (a period of 23 months), all observations from both AFP and ES were negative for WPV1 (*Figure 2*). In July 2016, two cases of AFP reported in the north-eastern state of Borno were found to be WPV1-positive. Genetic analysis showed little similarity between the isolates but identified a weak link to a virus last detected in the region in 2013, suggesting a prolonged period of undetected transmission rather than importation of infection (40). Known security concerns, inaccessibility and population displacement in the region also supports this conclusion.

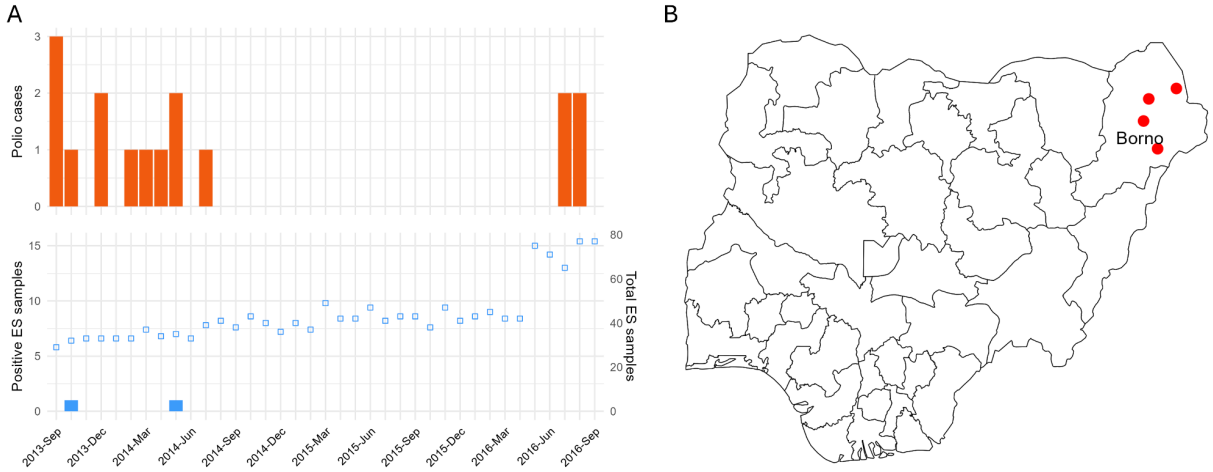

Figure 2: (A) *Detections of wild poliovirus serotype 1 (WPV1) among cases of AFP and environmental samples. The total number of environmental samples collected is indicated by unfilled blue squares.* ***(B)*** *The recorded locations of four WPV1-positive cases of AFP detected in Borno state in July 2016, after 23 months of surveillance without detection (AFP cases are plotted randomly within LGA boundaries).*

The 12-month rolling reported AFP rates among under-15s reflect extensive surveillance efforts during this time, with only a handful of LGAs not meeting the target rate of 2 per 100,000 for at least half of months (*Figure 3 (A,B)*). In fact, the majority of LGAs were consistently exceeding the target rate by ten-fold or even higher. This is also the case for adequacy of stool sampling among reported AFP cases (*Figure 3 (C,D)*).

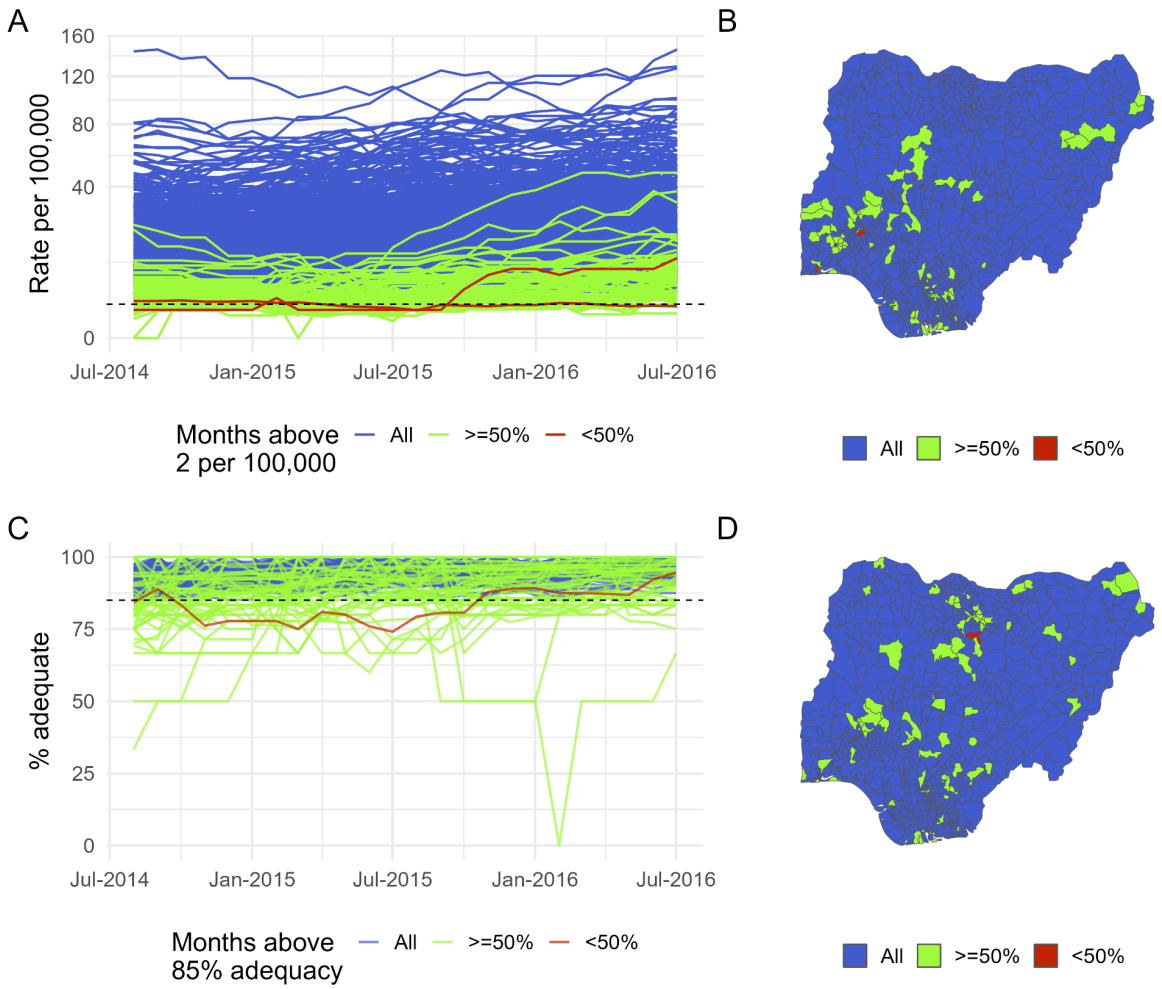

Figure 3: *Summary of observed AFP surveillance performance per LGA.* ***(A,B)*** *The 12-month rolling rate of reported AFP per LGA for the period of Aug 2013-July 2016. LGAs are classified by the proportion of time in which the reported rate exceeds the target threshold of 2 cases per 100,000 under-15s.* ***(C,D)*** *The 12-month rolling percentage of adequate stool sample collection among AFP cases reported by each LGA, in this case classified with respect to the target of 85%. LGAs are mapped by classification in the two right-hand figures.*

A number of new environmental surveillance sites were established during this period, increasing the total number of LGAs with *any* active ES but decreasing the average population coverage per LGA (*Figure 4*). Although the median population size within 5km of a collected sample is 169,000 (IQR [42,500 - 402,700]), within each LGA only 3% of the population is estimated to fall within the catchment of an active ES site in a given month on average. Exploring alternative catchment radii in sensitivity analyses, this percentage varies from 1.2% with a radius of 2km to 4.2% with a radius of 10km (*Supplementary Materials; Figure S5*).

Adequacy of sampling at ES sites in each LGA, with respect to detection of NPEV in collected samples, is more variable between LGAs than stool sampling adequacy among reported AFP cases (*Figure 4C*).

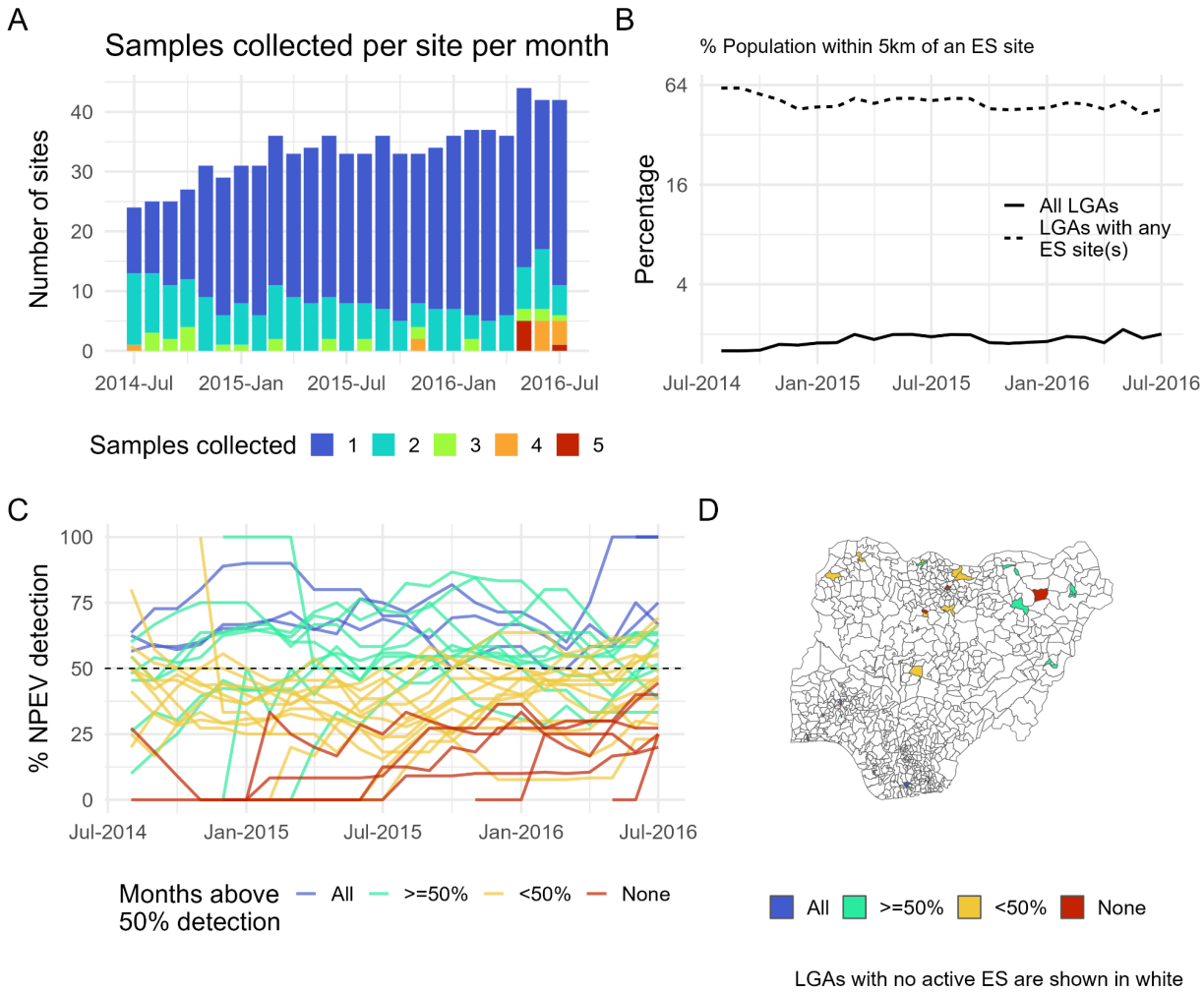

Figure 4: *Summary of observed ENV surveillance performance by LGA.* ***(A, B)*** *The coverage of ES by LGA and 12-month-rolling coverage over time.* ***(C, D)*** *The 12-month rolling detection rate of non-polio enterovirus (NPEV) among samples collected per LGA, classified with respect to the target of 50%. Again, classified LGAs are mapped on the right-hand side.*

Estimating sensitivity and freedom from infection

The effect of increasing ES capacity over time and improved quality of AFP surveillance is observed in the resulting sensitivity estimates for the entire country (*Figure 5A*), which increase on average for both components between 2014 and 2016. However, these gains are minimal and the absolute probability of either component detecting infection (if prevalent at the design prevalence) remains small. This is to be expected given that AFP surveillance is limited by high asymptomatic infection rate, and ES by low overall coverage of the population. Considering the underlying LGA-specific estimates, the mean monthly sensitivity of ES is 2.6% per LGA (95% quantile interval: 0-41%) across all LGAs/months and 96% (65-100%) across only LGA-months with greater than 10% population coverage.

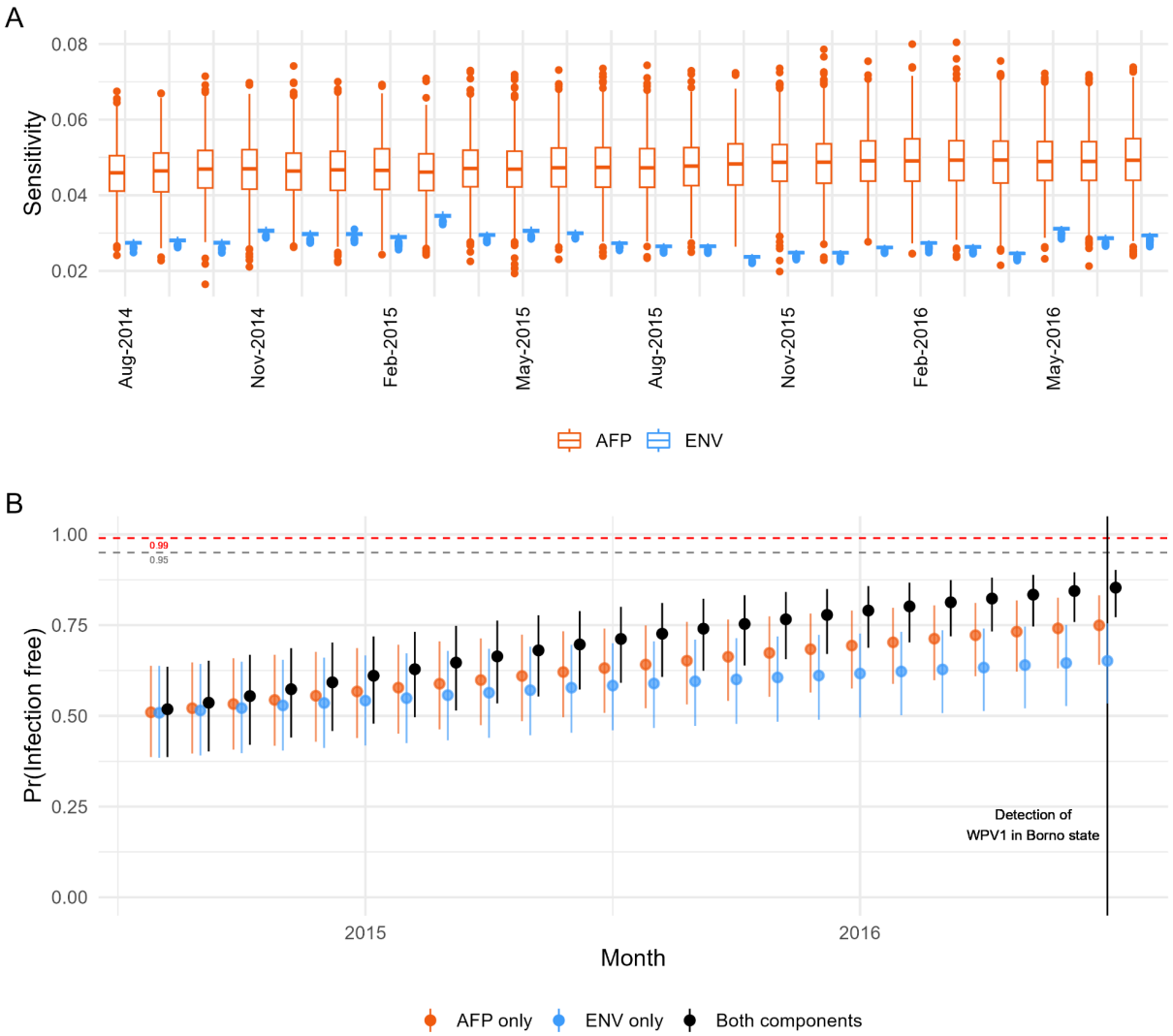

Figure 5: (A) *Estimated sensitivity of AFP and ENV surveillance per month, for detecting infection at the specified design prevalence of 1 per 100,000 in any LGA. The box plots illustrate uncertainty across 1,000 draws for each probability in the scenario tree.* ***(B)*** *Inferred probability of freedom from infection for each accumulating month without detection of WPV1 through either AFP or ENV surveillance. Uncertainty is represented by 95% quantile intervals across 1,000 draws. Thresholds of 95% and 99% probability are illustrated with dashed lines. Results from the combined surveillance system are compared to those obtained by considering each component alone.*

By July 2016, when WPV1 was again detected in Borno state, we estimated a national probability of freedom from infection of 85.8% (95% uncertainty interval: 78.2-90.5%). Considering evidence from AFP surveillance alone this falls to 75% (64.5-83.1%) and ES alone to 66% (54.3-76.2%).

Declaration of elimination: 2016-2020

If we apply the same analysis to the subsequent period from 2016 to 2020, we find that the FFI probability that is attained by the official declaration of WPV elimination in August 2020 (41,42) is 98.2% (97.6-98.6%). Our estimated trajectory for reaching the threshold of 95% is approximately two months shorter than Eichner and Dietz's estimated timeline of 2.9 years after the last observed case (1), which did not account for evidence from environmental surveillance (*Figure 6(B)*).

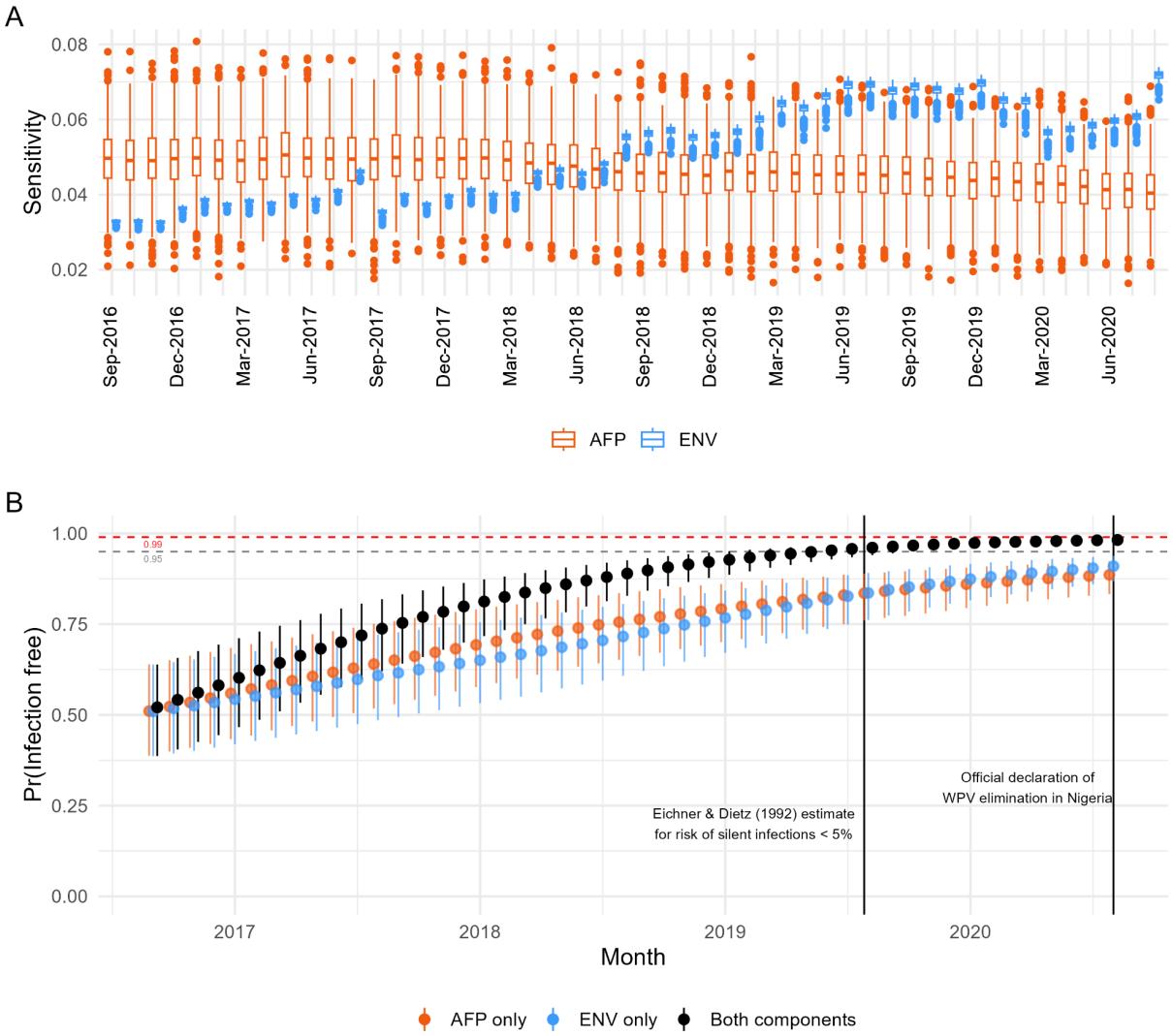

Figure 6: *Equivalent illustrations of estimated sensitivity and probability of freedom from infection as in Figure 5, for the subsequent period from 2016-2020. The estimated timeline from Eichner and Dietz (1) for the risk of continued "silent" infections to fall below 5% after detection of the last clinical case is marked in (B), along with the official declaration of WPV elimination from Nigeria in August 2020.*

During this time we see a steady increase in the sensitivity of ENV surveillance (*Figure 6(A)*), as the network of sites was substantially expanded (*Supplementary Materials; Figure S3*). This is

offset by a slight decrease in the sensitivity of AFP surveillance as the average adequacy of stool sample collection declined from 99% to 95% and the notification rate of AFP declined from 18 to 6.7 per 100,000 (*Supplementary Materials; Figure S4*). The sensitivity of ES overtakes AFP surveillance in mid-2018, and considering ENV surveillance alone results in a level of confidence that is approaching that of the combined system (91% (86.8-93.9%); *Figure 6(B)*) by August 2020. Not accounting for this variation in surveillance sensitivity over time results in an under-estimation of the evidence for freedom from infection in this case (*Supplementary Materials; Figure S8*), and a further five months until the estimated probability exceeds 95%.

Discussion

Undetected transmission and asymptomatic infection are a fundamental concern for the global eradication of wild poliovirus. We may only define elimination with a lag from the last observed paralytic case - beyond the feasible limits of silent transmission - and the question of how long to wait is fundamental to defining the success of the programme and the safe withdrawal of the oral polio vaccine. Premature declaration of eradication based on insufficient evidence that transmission has ceased would have disastrous consequences, potentially setting back the Global Polio Eradication Initiative's strategic plan by decades.

In this analysis, we demonstrated a statistical framework within which routine indicators of surveillance performance feed into our interpretation of the absence of detected virus. We expand on previous applications of the approach with more spatially- and temporally-explicit scenario tree probabilities, informed by these indicators. We also explore more realistic assumptions than previous applications as to the population catchment of ES, based on the point location of active sites as opposed to a fixed district-level proportion. We estimate a monthly combined sensitivity of ES and AFP surveillance to detect infection across Nigeria that varies between 6-9% from 2014-16, and 7-12% from 2016-20. From this we are able to determine timelines for the declaration of regional elimination after the last reported case or positive environmental sample.

We evaluated this approach against known positive and negative scenarios of WPV1 elimination in Nigeria, and found that our estimated timelines for high confidence in freedom from infection were consistent with both the re-emergence of WPV1 in July 2016 and the official declaration of elimination four years later. This supports prospective use of the approach to inform timelines for certification of the remaining WPV-endemic countries of Afghanistan and Pakistan, accounting for the specific characteristics of surveillance in those settings.

We further observe that, likely due to the expansion of ES after 2018, considering evidence from ES alone yields similar levels of confidence as pre-2018 data from the combined surveillance system. This could suggest that - given sufficient investment - ES would have the potential to supersede AFP surveillance in near-elimination settings. Especially given that the WHO now recommends the use of inactivated polio vaccine in the majority of settings (43) - which prevents

paralytic poliomyelitis but not asymptomatic infection - extensive ES may be crucial to detect silent transmission.

Average monthly sensitivity of ES surveillance to detect infection is estimated at 3% between 2014-16, increasing to around 5% between 2016-20. These national level estimates (incorporating districts with no ES at all) fall considerably lower than some previous estimates. Kroiss' (6) analysis looked at the rate of sabin virus detection in ES samples following SIAs on a district level, not accounting for the limited geographical catchment of sites within district boundaries. O'Reilly's approach in (4) estimated sensitivity on a population level - as we have done - but only within districts with active ES and those neighbouring. Our estimates are similar in magnitude to those obtained from applying the same scenario tree framework in Australia and in England and Wales (16,44), but there are still important differences in assumptions that limit comparability. Watkins et al. (16) only accounted for variation in scenario tree probabilities between jurisdictions - rather than additionally over time - in their estimates of AFP sensitivity, and O'Reilly et al. (44) assumed a high ES population catchment of 80% given the large convergent sewerage systems in England and Wales.

Our estimation of AFP sensitivity is limited by uncertainty in the probability of a symptomatic case being reported. Nuance between LGAs is lost among consistently high AFP rates relative to the target of 2 per 100,000; it is not clear what factors drive the variability at higher levels, at what point higher rates do not translate to higher probability of reporting, or how high regional rates may mask care-seeking barriers in certain communities. In particular, the persistence of undetected WPV1 circulation in northeastern Nigeria was attributed to conflict and inaccessibility in the region at the time, which was expected to have disrupted surveillance, vaccination and health service utilisation (40,45), yet AFP rates remained substantially above target levels in the vast majority of LGAs. This suggests that important dynamics may be missed when interpreting the indicator against this minimal threshold. The extent to which the burden of conflict is evident across routine measures of surveillance performance should be investigated more formally, for example utilising publicly-available data on conflict events and fatalities from the ACLED project (46).

Our conclusions are sensitive to some key assumptions. The time to accumulate high probability of freedom from infection is strongly influenced by both the assumed prior probability and by the critical prevalence level we require the surveillance system to detect (the design prevalence). A prior FFI probability around 0.5 could be considered optimistic if a positive was observed in the previous month. We would argue that this depends on whether the last detection was - for example - the tail of a recent epidemic or the final blip in a stuttering chain, and suggest that the prior is specified in context of that recent pattern. Further work to elicit and incorporate expert opinion on these points (for example, via a Delphi study (47)) would add validity to the choice of assumption.

Previous applications of this framework to poliovirus surveillance systems have used a design prevalence motivated by the global elimination target of 1 case of AFP per 100,000 under-15s (16,44), while an application to dengue fever chose a value corresponding to a single case

among the entire population (25). We found that specifying a less stringent design prevalence of 1 per 10,000 resulted in a high probability of freedom from infection as early as mid-2015, which does not align with the known persistence of transmission during this time. Overall, the consistency of our conclusions with what was observed in Nigeria between both 2014-16 and 2016-20 suggests that our assumptions for these two elements are reasonable.

For environmental surveillance, the greatest limiting factor for sensitivity is population catchment. A previous application of this framework (4) assumed a static catchment of around 60% of the population for each sub-region with active ES, and guidelines suggest that sites should be placed to have a catchment of around 100,000 people. Although this target is met by most collection sites when assuming a catchment radius of 5km, only around 3% of an LGA's population is captured by actively-sampling sites each month on average. This equates to approximately 12,500 people, increasing to 19,000 people between 2016-20. The assumed 5km radius is arguably generous in comparison to another analysis of ES performance in Nigeria (5). Sensitivity analyses considering alternative catchments indicate that the area would have to be expanded to an unrealistic extent to yield any consequential increase in sensitivity. More realistic catchment areas may be inferred from geographic contours and water flow directions (5,48), but this work is yet to be adequately linked to the sample locations recorded in POLIS.

Having ES within sewage treatment works would likely increase the size of catchment areas. However, households connected to formal sewage networks are thought to be at lower risk of being infected with poliovirus (due to the prominence of faecal-oral transmission in the virus' persistence), and likely to have a higher income along with other socioeconomic advantages (49,50). This bias would be expected in most polio-affected settings in which this framework could be applied (51). Water flow, dilution, temperature, chemical contamination and recent weather will also influence what is detected in a collected sample (52,53). The detection rate of NPEVs likely reflects some of this natural variability and serves as a proxy in our analysis, but routine measurement of key indicators during collection would allow future analyses to account for these sources of variability directly.

Conclusion

The GGC have agreed that the definition of timelines for declaring elimination of poliovirus should be data-driven and informed by the performance of surveillance in the specific setting of interest. This performance should ideally be evaluated quantitatively in terms of statistical sensitivity to detect infection on a population level since - as we observe in this example - the meeting of predefined threshold levels may be insufficient to highlight blind spots in surveillance. Using the approach demonstrated here, temporal changes and localised disruption to surveillance may directly and quantitatively inform our interpretation of the absence of detection, and the accumulating evidence that transmission has been interrupted. This is applicable to both WPV eradication goals and determining the resolution of cVDPV2 outbreaks. However, there is a need to more thoroughly define the impact of variable conditions of environmental sample collection and the disruption of AFP surveillance caused by conflict, in terms of the probability of persisting virus circulation being detected. With the conclusion of 40 years of

global commitment in sight but balancing on the edge of significant resurgence, this level of detail in inference will be crucial to confidently certify the last endemic areas as free from wild poliovirus.

Acknowledgements

This research is supported by grants from the Polio Research Committee (Surveillance Modelling to Support Polio Elimination and Certification (SPEC), ref: 1432457-1) and the Bill and Melinda Gates Foundation (INV-049298). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

References

1. Eichner M, Dietz K. Eradication of poliomyelitis: when can one be sure that polio virus transmission has been terminated? Am J Epidemiol. 1996 Apr 15;143(8):816–22.
2. Global Polio Eradication Initiative. Field guidance for the implementation of environmental surveillance for poliovirus. World Health Organization; 2023.
3. Kalkowska DA, Thompson KM. Modeling Undetected Live Poliovirus Circulation After Apparent Interruption of Transmission: Borno and Yobe in Northeast Nigeria. Risk Anal Off Publ Soc Risk Anal. 2021 Feb;41(2):303–11.
4. O'Reilly KM, Verity R, Durry E, Asghar H, Sharif S, Zaidi SZ, et al. Population sensitivity of acute flaccid paralysis and environmental surveillance for serotype 1 poliovirus in Pakistan: an observational study. BMC Infect Dis. 2018 Apr 13;18(1):176.
5. Hamisu AW, Blake IM, Sume G, Braka F, Jimoh A, Dahiru H, et al. Characterizing Environmental Surveillance Sites in Nigeria and Their Sensitivity to Detect Poliovirus and Other Enteroviruses. J Infect Dis. 2022 Apr 15;225(8):1377–86.
6. Kroiss SJ, Ahmadzai M, Ahmed J, Alam MM, Chabot-Couture G, Famulare M, et al. Assessing the sensitivity of the polio environmental surveillance system. PloS One. 2018;13(12):e0208336.
7. Manyanga D, Byabamazima C, Masvikeni B, Daniel F. Assessment of acute flaccid paralysis surveillance performance in East and Southern African countries 2012 - 2019. Pan Afr Med J. 2020;36:71.
8. World Health Organization. World Health Organization - Regional Office for the Eastern Mediterranean. [cited 2024 Mar 26]. Polio containment and certification. Available from: http://www.emro.who.int/polio-eradication/about-eradication/polio-containment-and-certification.html
9. GPEI-Surveillance Indicators [Internet]. [cited 2024 May 16]. Available from: https://polioeradication.org/polio-today/polio-now/surveillance-indicators/
10. Howard W, Moonsamy S, Seakamela L, Jallow S, Modiko F, du Plessis H, et al. Sensitivity of the acute flaccid paralysis surveillance system for poliovirus in South Africa, 2016-2019. J Med Microbiol. 2021 Oct;70(10):001441.
11. Tegegne AA, Maleghemi S, Anyuon AN, Zeleke FA, Legge GA, Ferede MA, et al. The sensitivity of acute flaccid paralysis surveillance - the case of South Sudan: retrospective secondary analysis of AFP surveillance data 2014-2019. Pan Afr Med J. 2022;42(Suppl 1):12.
12. Chen P, Liu Y, Wang H, Liu G, Lin X, Zhang W, et al. Environmental Surveillance Complements Case-Based Surveillance of Acute Flaccid Paralysis in Polio Endgame

Strategy 2019-2023. Appl Environ Microbiol. 2020 Jul 20;86(15).
13. Tushabe P, Bwogi J, Eliku JP, Aine F, Birungi M, Gaizi J, et al. Environmental surveillance detects circulating vaccine-derived poliovirus type 2 that was undetected by acute flaccid paralysis surveillance in 2021 in Uganda. Arch Virol. 2023 Apr 15;168(5):140.
14. Ranta J, Hovi T, Arjas E. Poliovirus Surveillance by Examining Sewage Water Specimens: Studies on Detection Probability Using Simulation Models. Risk Anal. 2001;21(6):1087–96.
15. Kroiss SJ, Ahmadzai M, Ahmed J, Alam MM, Chabot-Couture G, Famulare M, et al. Assessing the sensitivity of the polio environmental surveillance system. PLOS ONE. 2018 Dec 28;13(12):e0208336.
16. Watkins RE, Martin PAJ, Kelly H, Madin B, Watson C. An evaluation of the sensitivity of acute flaccid paralysis surveillance for poliovirus infection in Australia. BMC Infect Dis. 2009 Sep 30;9:162.
17. Martin P a. J, Cameron AR, Greiner M. Demonstrating freedom from disease using multiple complex data sources 1: a new methodology based on scenario trees. Prev Vet Med. 2007 May 16;79(2–4):71–97.
18. Martin P a. J, Cameron AR, Barfod K, Sergeant ESG, Greiner M. Demonstrating freedom from disease using multiple complex data sources 2: case study--classical swine fever in Denmark. Prev Vet Med. 2007 May 16;79(2–4):98–115.
19. Frössling J, Agren ECC, Eliasson-Selling L, Lewerin SS. Probability of freedom from disease after the first detection and eradication of PRRS in Sweden: scenario-tree modelling of the surveillance system. Prev Vet Med. 2009 Oct 1;91(2–4):137–45.
20. Nathues C, Hillebrand A, Rossteuscher S, Zimmermann W, Nathues H, Schüpbach G. Evaluating the surveillance for swine dysentery and progressive atrophic rhinitis in closed multiplier herds using scenario tree modelling. Porc Health Manag. 2015 Jun 10;1(1):7.
21. Foddai A, Stockmarr A, Boklund A. Evaluation of temporal surveillance system sensitivity and freedom from bovine viral diarrhea in Danish dairy herds using scenario tree modelling. BMC Vet Res. 2016 Jun 21;12(1):118.
22. Hou K, Firestone S, Stevenson M. Scenario tree modelling to inform surveillance design for maintaining freedom from *Coxiella burnetii* infection in Australian commercial dairy goat herds. Prev Vet Med. 2023 Oct 1;219:106024.
23. Thompson RN, Morgan OW, Jalava K. Rigorous surveillance is necessary for high confidence in end-of-outbreak declarations for Ebola and other infectious diseases. Philos Trans R Soc B Biol Sci. 2019;374(1776).
24. Michael E, Smith ME, Katabarwa MN, Byamukama E, Griswold E, Habomugisha P, et al. Substantiating freedom from parasitic infection by combining transmission model predictions with disease surveys. Nat Commun. 2018 Oct 18;9(1):4324.
25. Bannister-Tyrrell M, Hillman A, Indriani C, Ahmad RA, Utarini A, Simmons CP, et al. Utility of surveillance data for planning for dengue elimination in Yogyakarta, Indonesia: A scenario-tree modelling approach. BMJ Glob Health. 2023;8(11).
26. Stresman G, Cameron A, Drakeley C. Freedom from Infection: Confirming Interruption of Malaria Transmission. Trends Parasitol. 2017 May 1;33(5):345–52.
27. WHO Removes Nigeria from Polio-Endemic List [Internet]. [cited 2024 May 31]. Available from: https://www.who.int/news/item/25-09-2015-who-removes-nigeria-from-polio-endemic-list
28. Fomban Leke RG, King A, Pallansch MA, Tangermann RH, Mkanda P, Chunsuttiwat S, et al. Certifying the interruption of wild poliovirus transmission in the WHO African region on the turbulent journey to a polio-free world. Lancet Glob Health. 2020 Oct 1;8(10):e1345–51.
29. Famulare M. Has Wild Poliovirus Been Eliminated from Nigeria? PLOS ONE. 2015 Aug 28;10(8):e0135765.
30. World Health Organisation. Polio Information System [Internet]. [cited 2023 Oct 26]. Available from: https://extranet.who.int/polis/public/CaseCount.aspx

31. Global Polio Eradication Initiative. Global guidance for conducting acute flaccid paralysis (AFP) surveillance in the context of poliovirus eradication [Internet]. Geneva: World Health Organization; 2024 [cited 2024 Jun 3]. Available from: https://polioeradication.org/tools-and-library/resources-for-polio-eradicators/gpei-tools-protocols-and-guidelines/
32. WorldPop. Global 100m Population total adjusted to match the corresponding UNPD estimate [Internet]. University of Southampton; 2020 [cited 2022 Apr 14]. Available from: https://www.worldpop.org/doi/10.5258/SOTON/WP00660
33. Nigeria - Subnational Population Statistics - Humanitarian Data Exchange [Internet]. [cited 2024 Feb 12]. Available from: https://data.humdata.org/dataset/cod-ps-nga?
34. Voorman A, O'Reilly K, Lyons H, Goel AK, Touray K, Okiror S. Real-time prediction model of cVDPV2 outbreaks to aid outbreak response vaccination strategies. Vaccine. 2023 Apr 6;41 Suppl 1(Suppl 1):A105–12.
35. Voorman A, Lyons HM. Measuring polio immunity to plan immunization activities. Vaccine. 2016 Nov 21;34(48):5946–52.
36. Global Polio Laboratory Network. Polio laboratory manual [Internet]. World Health Organization; 2004. Available from: https://polioeradication.org/wp-content/uploads/2017/05/Polio_Lab_Manual04.pdf
37. Wagner BG, Behrend MR, Klein DJ, Upfill-Brown AM, Eckhoff PA, Hu H. Quantifying the Impact of Expanded Age Group Campaigns for Polio Eradication. PLOS ONE. 2014 Dec 1;9(12):e113538.
38. Voorman A, Lyons H, Bennette C, Kovacs S, Makam JK, F Vertefeuille J, et al. Analysis of population immunity to poliovirus following cessation of trivalent oral polio vaccine. Vaccine. 2023 Apr 6;41 Suppl 1:A85–92.
39. Mangal TD, Aylward RB, Shuaib F, Mwanza M, Pate MA, Abanida E, et al. Spatial Dynamics and High Risk Transmission Pathways of Poliovirus in Nigeria 2001-2013. PLOS ONE. 2016 Sep 26;11(9):e0163065.
40. Nnadi C. Continued Endemic Wild Poliovirus Transmission in Security-Compromised Areas — Nigeria, 2016. MMWR Morb Mortal Wkly Rep [Internet]. 2017 [cited 2024 May 31];66. Available from: https://www.facebook.com/CDCMMWR
41. Khan F, Datta SD, Quddus A, Vertefeuille JF, Burns CC, Jorba J, et al. Progress Toward Polio Eradication - Worldwide, January 2016-March 2018. MMWR Morb Mortal Wkly Rep. 2018 May 11;67(18):524–8.
42. Global Polio Eradication Initiative. Annual Report [Internet]. 2020 [cited 2024 May 31]. Available from: https://polioeradication.org/tools-and-library/policy-reports/annual-reports/
43. World Health Organisation. Weekly Epidemilogical Report. 2020;95(22):241–56.
44. O'Reilly KM, Grassly NC, Allen DJ, Bannister-Tyrrell M, Cameron A, Carrion Martin AI, et al. Surveillance optimisation to detect poliovirus in the pre-eradication era: a modelling study of England and Wales. Epidemiol Infect. 2020 May 13;148:e157.
45. Wiesen E, Dankoli R, Musa M, Higgins J, Forbi J, Idris J, et al. Conducting public health surveillance in areas of armed conflict and restricted population access: a qualitative case study of polio surveillance in conflict-affected areas of Borno State, Nigeria. Confl Health. 2022 May 7;16(1):20.
46. ACLED [Internet]. 2024 [cited 2024 May 16]. ACLED | Bringing Clarity to Crisis. Available from: https://acleddata.com/
47. Taylor E. We Agree, Don't We? The Delphi Method for Health Environments Research. HERD Health Environ Res Des J. 2020 Jan 1;13(1):11–23.
48. Environmental Surveillance [Internet]. [cited 2024 May 31]. Available from: https://es.world/catalog
49. Abubakar IR. Access to Sanitation Facilities among Nigerian Households: Determinants and Sustainability Implications. Sustainability. 2017 Apr;9(4):547.

50. Abubakar IR. Understanding the socioeconomic and environmental indicators of household water treatment in Nigeria. Util Policy. 2021 Jun 1;70:101209.
51. Cohen A, Vikesland P, Pruden A, Krometis LA, Lee LM, Darling A, et al. Making waves: The benefits and challenges of responsibly implementing wastewater-based surveillance for rural communities. Water Res. 2024 Feb 15;250:121095.
52. Guo Y, Sivakumar M, Jiang G. Decay of four enteric pathogens and implications to wastewater-based epidemiology: Effects of temperature and wastewater dilutions. Sci Total Environ. 2022 May 1;819:152000.
53. Nguyen Quoc B, Saingam P, RedCorn R, Carter JA, Jain T, Candry P, et al. Case Study: Impact of Diurnal Variations and Stormwater Dilution on SARS-CoV-2 RNA Signal Intensity at Neighborhood Scale Wastewater Pumping Stations. ACS EST Water. 2022 Nov 11;2(11):1964–75.