Faster detection of poliomyelitis outbreaks to support the polio end-game

Isobel M Blake, Paul Chenoweth, Hiro Okayasu, Christl A Donnelly, R Bruce Aylward and Nicholas C Grassly

http://dx.doi.org/10.3201/eid2203.151394

Technical Appendix

Technical Appendix: Supplementary Methods

Space-time analysis of nonpolio AFP data: Details of the spatiotemporal mixed effects regression model

The number of nonpolio AFP cases reported in administrative area *i* at year *t* is denoted $O_{i,t}$. We firstly assumed that $O_{i,t}$ is Poisson distributed with the mean and variance given as $E_i \exp(\eta_{i,t})$ (the population size of location *i* , E_i, multiplied by the exponent of the linear predictor $\eta_{i,t}$). We secondly assumed that $O_{i,t}$ follows a Negative Binomial distribution to allow for overdispersion, with mean $E_i \exp(\eta_{i,t})$ and variance $E_i \exp(\eta_{i,t}) (1 + E_i \exp(\eta_{i,t}) /b)$, where *b* is the size parameter determining the level of overdispersion.

The linear predictor $\eta_{i,t}$ is defined by the following equation, based on the Besag, York and Mollié (BYM) model (1) with an additional temporal trend:

$$\eta_{i,t} = \alpha + \mu_i + \nu_i + \omega_t \quad (1)$$

.

The spatially unstructured random effect μ_i, was assumed to be Gaussian distributed:

$$\mu_i \sim Norm\left(0, \sigma_\mu^2\right) \quad (2)$$

.

The spatially structured (conditional autoregressive) effect ν_i was defined as:

$$\nu_i \sim Norm\left(\frac{\sum_{j\sim i} \nu_j}{m_i}, \frac{\sigma_\nu^2}{m_i}\right) \quad (3)$$

, where the number of adjacent districts of location *i* is given as m_i. The notation $j \sim i$ denotes districts *i* and *j* are neighbours. The parameter ν_j is the spatially structured effect in a neighbouring district j. A few districts in certain countries did not share physical boundaries with other districts (e.g. islands). For these districts a neighbour was defined as the closest district (by straight line distance). The spatially structured effect accounts for the fact that geographically adjacent districts

may have similar reporting rates of nonpolio AFP cases. This means, the greater the number of neighbours a district has, the greater the information in the data of its spatial random effect.

The national temporal effect was modelled as a second-order random walk (2):

$$\left(\omega_t - 2\omega_{t+1} + \omega_{t+2}\right) \sim Norm\left(0, \sigma_\omega^2\right) \quad (4)$$

The parameter α determines the intercept of the model.

The precision for each of the three random effects were estimated ($1/\sigma_\mu^2$, $1/\sigma_v^2$ and $1/\sigma_w^2$) as well as α and b where applicable and each model component was scaled to have an average variance equal to one, to allow comparison between the precision parameters of the different models (3).

The three hyperparameters (log of each precision parameter) were each assigned a non-informative LogGamma(a,b) prior distribution where a = 1 and b = 0.0001. The most parsimonious yet best fitting model was selected according to the Deviance Information Criterion (DIC) (4), when comparing if the count data was Poisson or negative binomially distributed. The mean and 95% credible interval from the posterior marginal distributions are presented.

The proportion of variance in the data explained by the structured spatial component compared to the unstructured spatial component, p_v, was evaluated (for full details see (5)). Briefly, an empirical estimate of the posterior marginal variance for the structured effect was obtained (S_v^2) and

$$p_v = S_v^2 \Big/ \left(S_v^2 + \sigma_v^2\right) \quad (5)$$

Where

$$S_v^2 = \frac{\sum_{i=1}^{d} \left(v_i - \bar{v}\right)}{d - 1} \quad (6)$$

, and d corresponds to the total number of districts.

The expected annual number of nonpolio AFP cases reported per district independent of time, A_i , was calculated as

$$A_i = E_i * \exp(\alpha + \mu_i + \nu_i) \quad (7)$$

where the values of α, μ_i and ν_i were taken as the mean values from the posterior marginal distributions.

Testing of real-time AFP databases to detect polio outbreaks

This section follows from '*Testing of real-time AFP databases to detect polio outbreaks*' methods section in the main text.

The likelihood ratio function defining how likely there is an elevated risk within the cylinder compared to outside the cylinder was defined as:

$$\left(\frac{n_z}{\gamma(z)}\right)^{n_z}\left(\frac{N-n_z}{N-\gamma(z)}\right)^{N-n_z} \quad (8)$$

(as described in detail in (6)). *N* is the total number of AFP cases in the last two years (from the current surveillance date), n_z is the number of AFP cases within the cylinder, and $\gamma(z)$ is the expected number of cases under the null hypothesis in cylinder *z*. The likelihood ratio function was maximised across cylinders of all locations and sizes whereby the cylinder with the maximum likelihood ratio corresponds to the most likely cluster.

The expected number of cases in a given cylinder, $\gamma(z)$, was calculated as the proportion of the country's expected nonpolio AFP rate that is contained within the circle defining the cylinder, multiplied by both the total number of cases in the last two years and the fraction of 2-years the height of the cylinder occupied (*f*). *i.e.*

$$\gamma(z) = \frac{\sum_{\frac{i \in Z}{d}} A_i}{\sum_{i=1} A_i} fN \quad (9)$$

, where *d* is the total number of districts within the given country and $i \in Z$ is the set of district indexes whose centroids are contained within the radius of the cylinder *z*.

The definition of *N* spanned two years since the current surveillance date to enable a relatively stable estimate of $\gamma(z)$ without being confounded by shifts in populations size (which would be likely if a longer time period had been chosen).

The occurrence of cases was assumed to be Poisson distributed and with a constant rate over time under the null hypothesis. The space-time scan statistic was evaluated using SaTScan™ (v9.3) (7), which was called using the R programming language.

The p-value level of significance of the cluster was determined by Monte Carlo hypothesis testing. Random data sets were simulated under the null hypothesis (from a Poisson distribution with mean $\gamma(z)$), and the maximum likelihood ratios were calculated for each simulated data-set. The rank of the maximum likelihood estimate for the real dataset was compared to those from the simulated data set to obtain the p-value. To adjust for multiple testing associated with prospective surveillance, the likelihood for random datasets was maximised over all cylinders used in previous analyses within the last year of the surveillance date in addition to the current cylinders. This was performed also using SaTScan™ (v9.3) (7), which was called using the R programming language and the computation was parallelised over a 16-core, high performance cluster.

An R script is included at the end of this document providing an example on how this analysis is run.

Space-time permutation model

We have so far described the Poisson space-time scan statistic to identify significant clusters of AFP cases. We also tested the space-time permutation scan statistic (8) as alternative method to detect clusters of AFP cases. This method only relies on case data and compares cases in a cylinder to the expected number if the spatial and temporal locations of all cases were independent of each other.

In this scenario, the likelihood ratio function defining how likely there is an elevated risk within the cylinder compared to outside the cylinder remains the same as equation 8 and the likelihood ratio function was maximised across cylinders of all locations and sizes as before. The difference is of this method is the calculation of $\gamma(z)$. The expected number of cases in district i on day d, γ_{id}, is defined as the proportion of all cases (N) in district i, multiplied by the total number of cases that occurred on day d. The expected number of cases in cylinder z in the absence of an outbreak is then defined as:

$$\gamma(z) = \sum_{(i,d)\in Z} \gamma_{id} \tag{10}$$

Sensitivity analysis to the level of significance determining an early warning alarm

We assumed an early warning alarm of a polio outbreak would be raised when the significance of the detected cluster was $p<0.05$. Sensitivity analysis to this threshold (Technical Appendix Tables 2 and 3) found the dates of alarms raised for the major outbreaks to be insensitive to the threshold as the alarms were highly significant. For the smaller outbreaks increasing the threshold to $p<0.1$ would

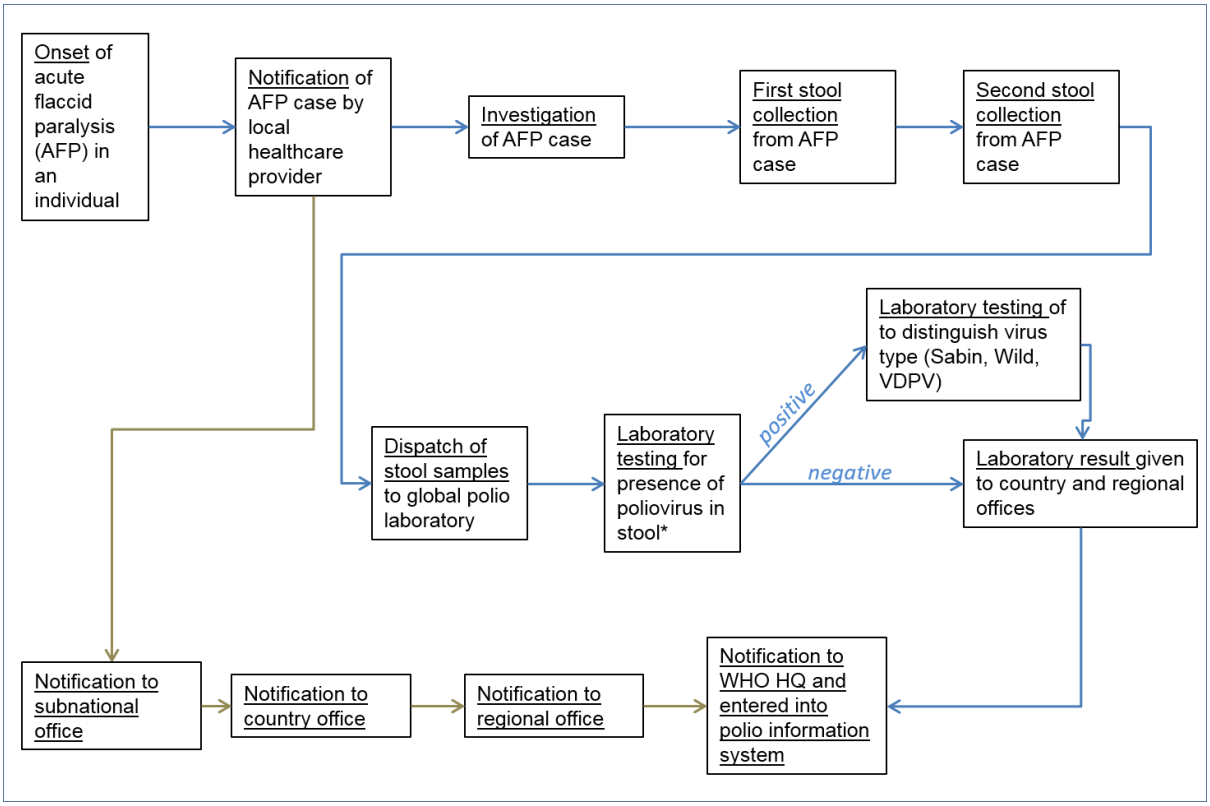
not have resulted in earlier detection of outbreaks, whilst reducing the threshold to $p<0.001$ would have lowered the number of false alarms but at the cost of detecting outbreaks later or not at all.

Spatial and temporal characteristics of poliomyelitis cases

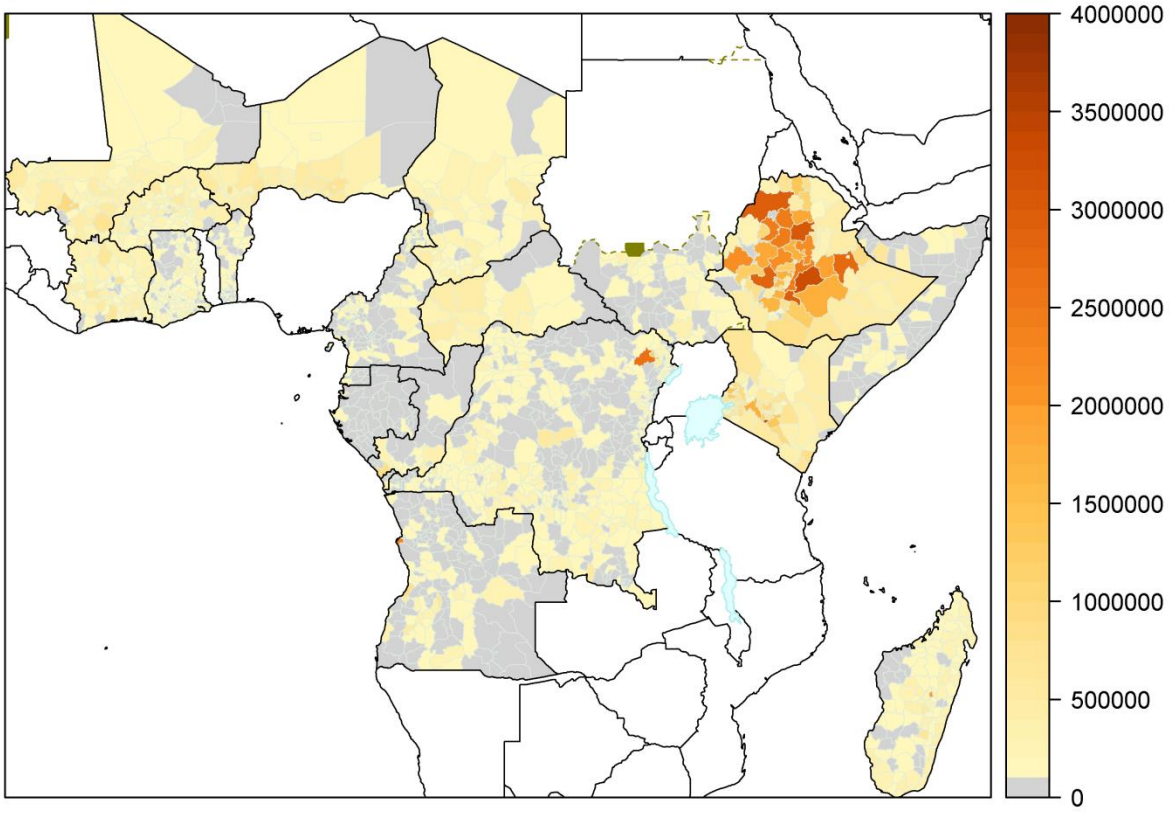
The spatial and temporal distributions of serotype 1 confirmed poliomyelitis cases in the four most recent large-scale polio outbreaks (Tajikistan 2010, Republic of Congo 2010, Somalia 2005 and Somalia 2013) were compared to the respective distributions of nonpolio AFP cases.

To assess the spatial distribution, thirty initial confirmed cases of poliomyelitis (defined by date of onset) were assigned to their district centroid and the median 'great-circle-distance' distance (9) between a case and the subsequent four cases in time was calculated and the distribution of these median distances was computed. This was repeated for nonpolio AFP cases which occurred during the two-year period prior to each respective outbreak. To assess the temporal distributions the median number of days between a case and the subsequent four cases in time was calculated and the median distributions compared.

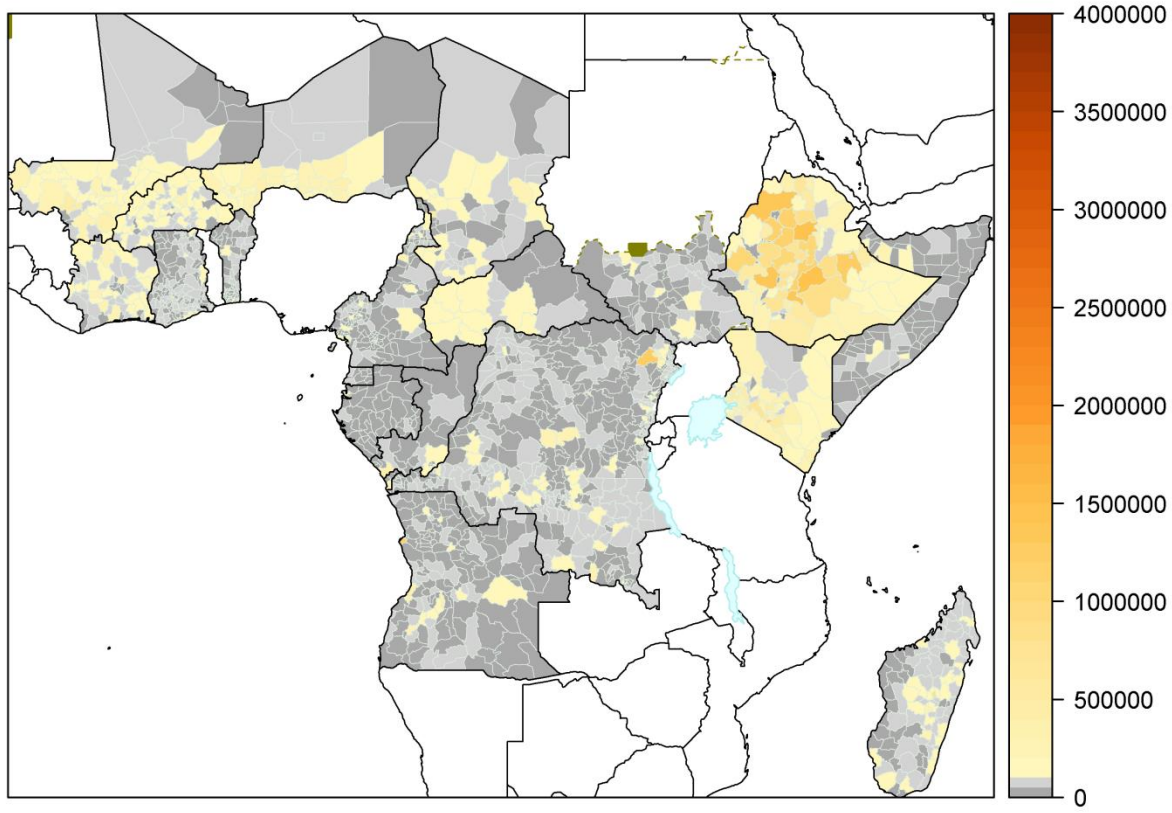
Supplementary Figures and Tables



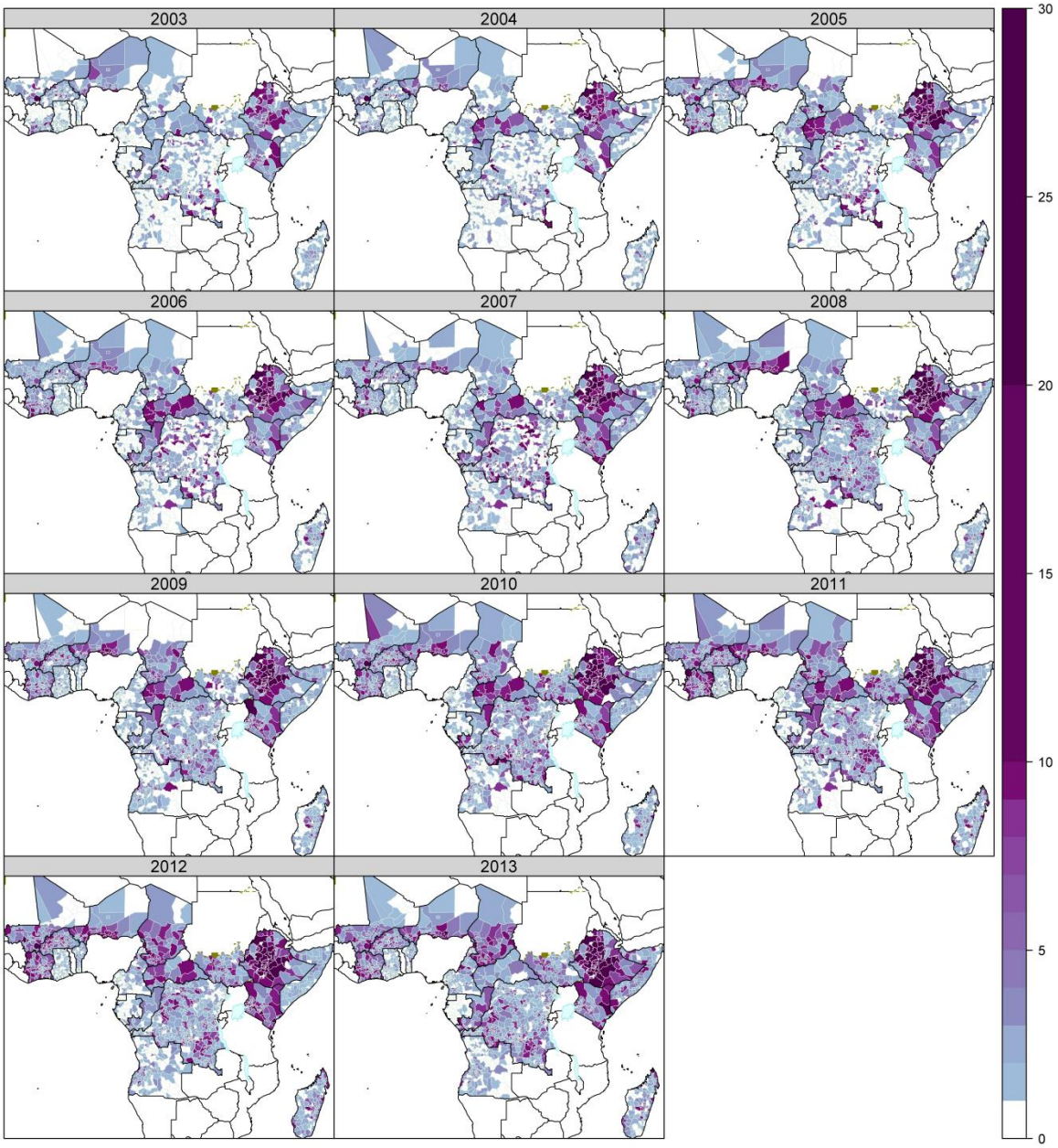
Technical Appendix Figure 1 Schematic of global acute flaccid paralysis (AFP) surveillance for detecting poliomyelitis cases. Blue arrows indicate the pathway of laboratory testing of stool samples from reported AFP cases for poliovirus. Brown arrows indicate the pathway of reporting of AFP cases irrespective of laboratory testing of stool samples. *Indicates testing occurs on stool samples that have been collected within 14 days after the onset of paralysis, 24 hours apart and that arrived in the laboratory by reverse cold chain. If 80% of AFP cases have adequate stool samples collected then surveillance is determined sufficient.



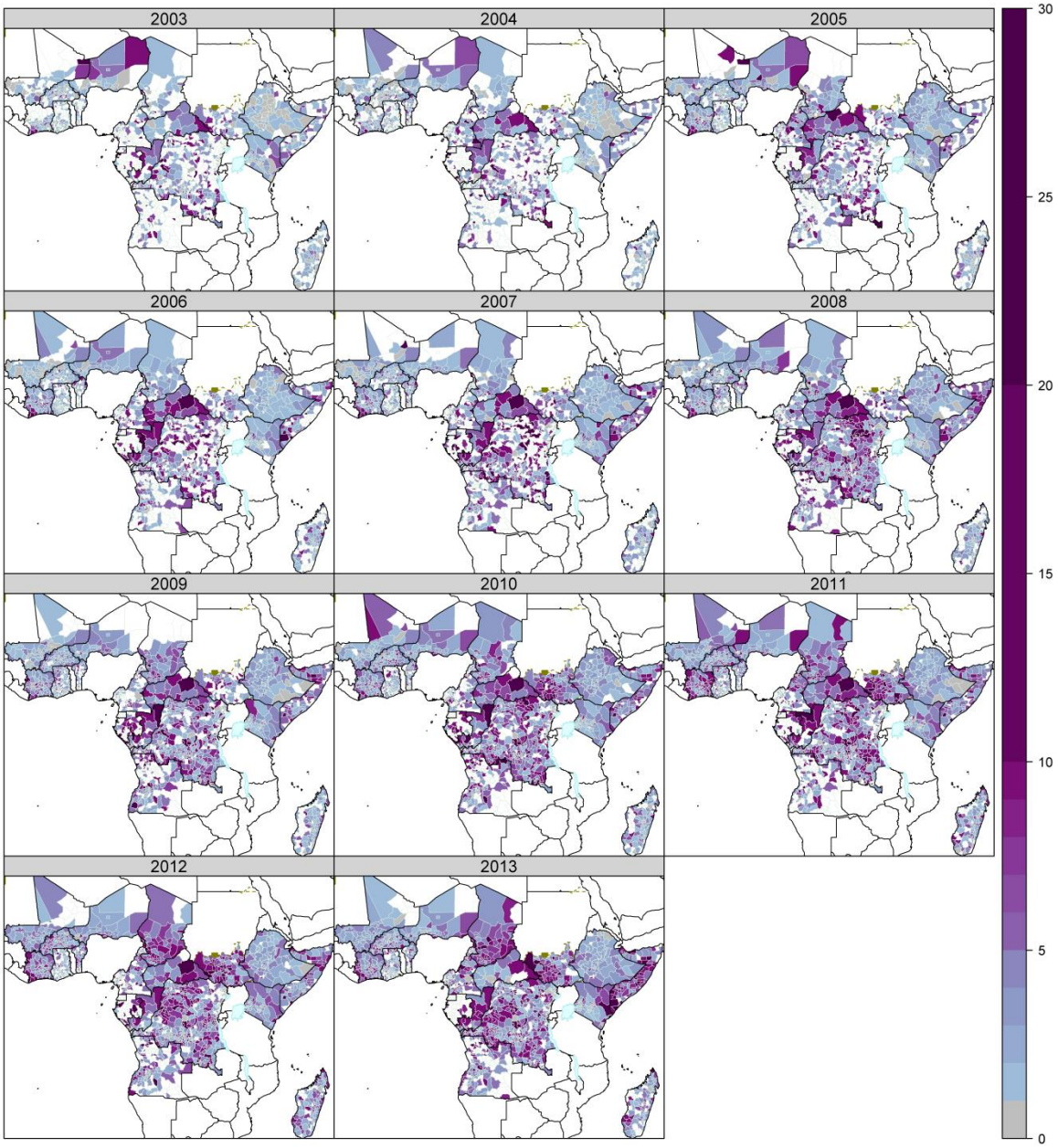
Technical Appendix Figure 2 Population size per second administrative division (ADM2) area for sub-Saharan African countries that have recently experienced a polio importation or outbreak or are considered to be at high risk for these events. Raster population estimates for the year 2010 were obtained from www.worldpop.org.uk and the population within each ADM2 division was summed. The publication of this map does not imply the expression of any opinion whatsoever on the part of WHO concerning the legal status of any territory, city or area or of its authorities, or concerning the delimitation of its frontiers or boundaries. WHO does not endorse or approve the use of sub-national boundaries in this map. Disputed borders and areas are shown in green and lakes at borders are shown in pale blue.



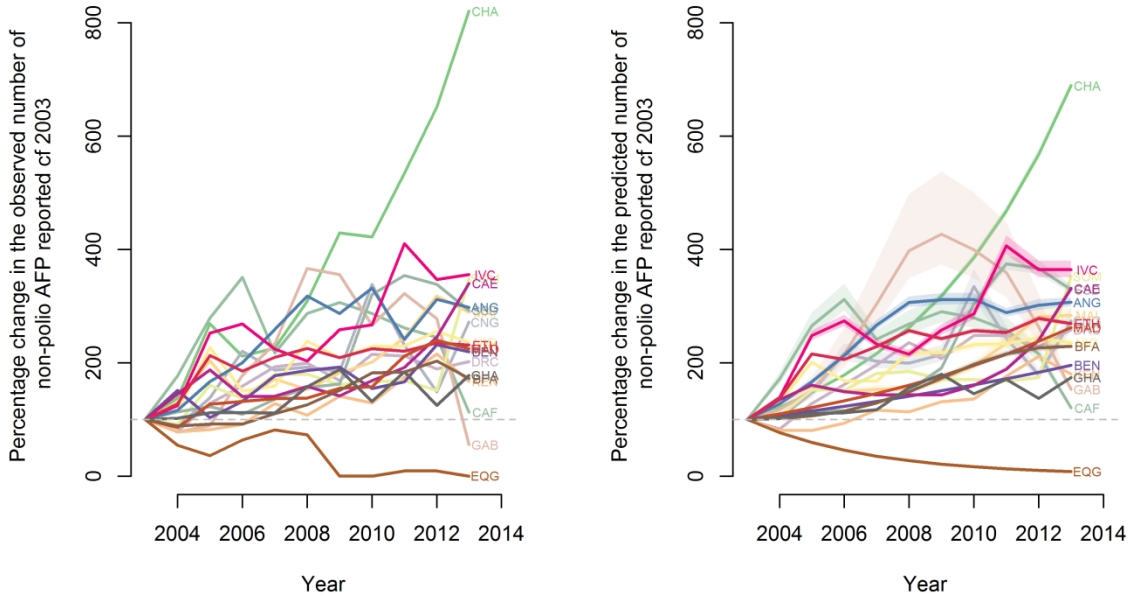
Technical Appendix Figure 3 Population size of <15 year-olds per second administrative division (ADM2) area for sub-Saharan African countries that have recently experienced a polio importation or outbreak or are considered to be at high risk for these events. Raster population estimates for the year 2010 were obtained from www.worldpop.org.uk, the population within each ADM2 division was summed and multiplied by the estimated national proportion of the population in that age group http://www.census.gov/population/international/data/idb/informationGateway.php. Dark grey indicates < 50,000 people. The publication of this map does not imply the expression of any opinion whatsoever on the part of WHO concerning the legal status of any territory, city or area or of its authorities, or concerning the delimitation of its frontiers or boundaries. WHO does not endorse or approve the use of sub-national boundaries in this map. Disputed borders and areas are shown in green and lakes at borders are shown in pale blue.



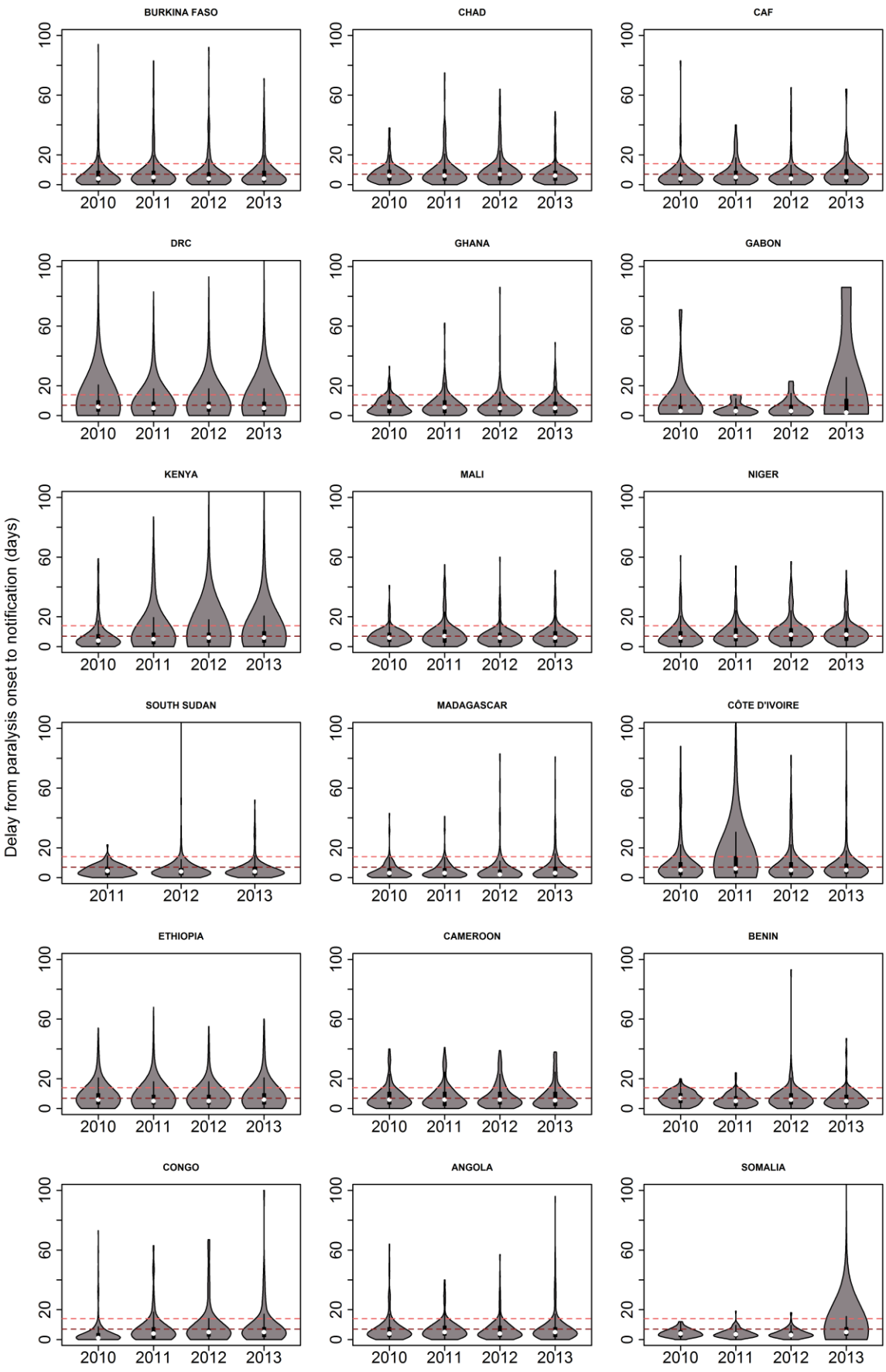
Technical Appendix Figure 4 Annual number of nonpolio acute flaccid paralysis cases reported at the administrative 2 unit in countries in sub-Saharan Africa that have recently experienced a polio importation or outbreak or are considered to be at high risk for these events. Areas that report ≥ 25 annual cases are grouped into the 25-30 category. The Republic of South Sudan gained independence in 2011 but reporting in this area prior to independence is shown for comparison. The publication of this map does not imply the expression of any opinion whatsoever on the part of WHO concerning the legal status of any territory, city or area or of its authorities, or concerning the delimitation of its frontiers or boundaries. WHO does not endorse or approve the use of sub-national boundaries in this map. Disputed borders and areas are shown in green and lakes at borders are shown in pale blue. Maps were plotted using the 'spacetime' R package (10).



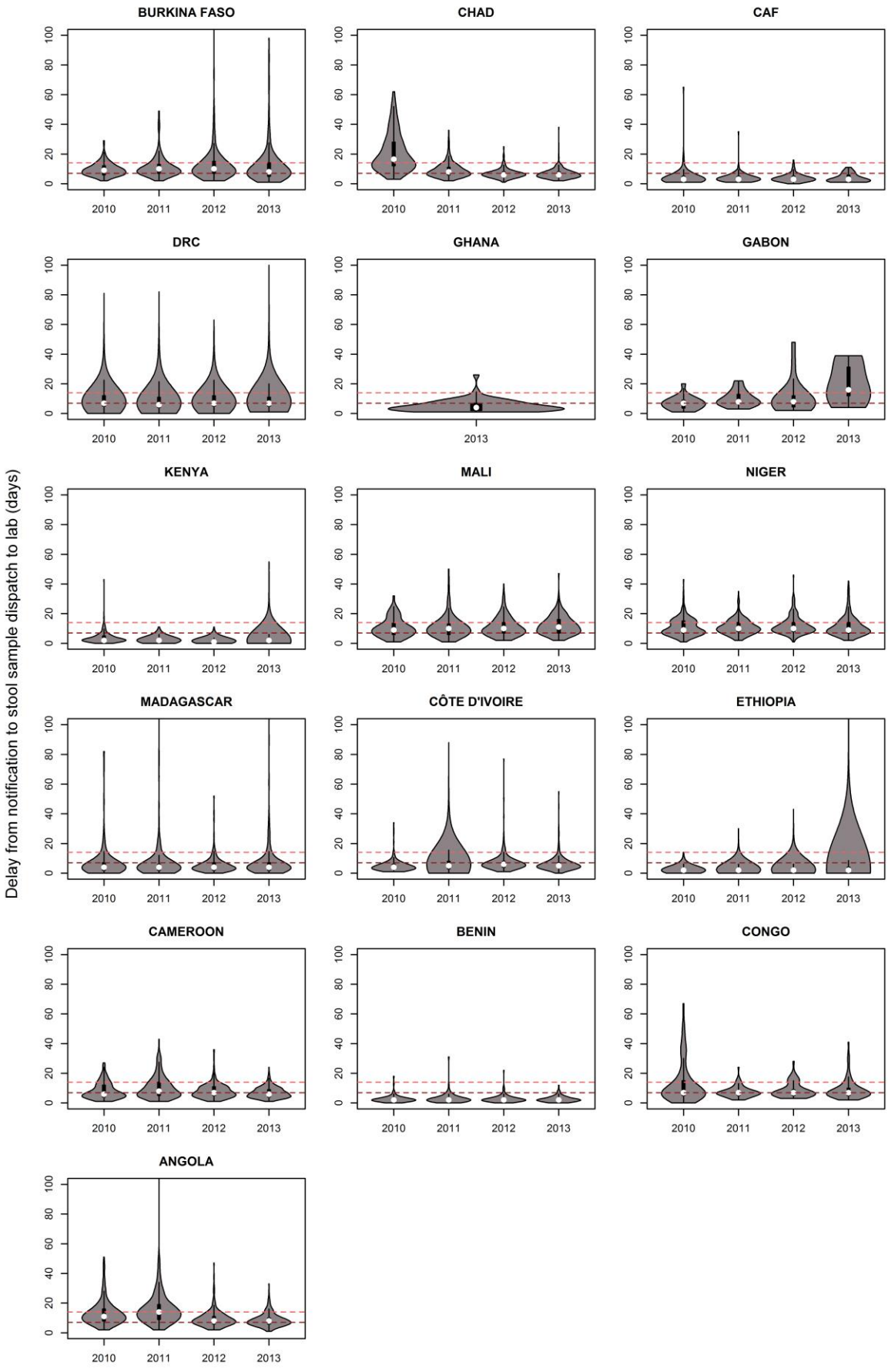
Technical Appendix Figure 5 Annual number of nonpolio acute flaccid paralysis cases reported per 100,000 people <15 years old (population size estimated for 2010) at the ADM 2 unit in countries in sub-Saharan Africa that have recently experienced a polio importation or outbreak or are considered to be at high risk for these events. Areas that report ≥ 25 annual cases per 100,000 people <15 years old are grouped into the 25-30 category. The Republic of South Sudan gained independence in 2011 but reporting in this area prior to independence is shown for comparison. The publication of this map does not imply the expression of any opinion whatsoever on the part of WHO concerning the legal status of any territory, city or area or of its authorities, or concerning the delimitation of its frontiers or boundaries. WHO does not endorse or approve the use of sub-national boundaries in this map. Disputed borders and areas are shown in green and lakes at borders are shown in pale blue. Maps were plotted using the 'spacetime' R package (10).



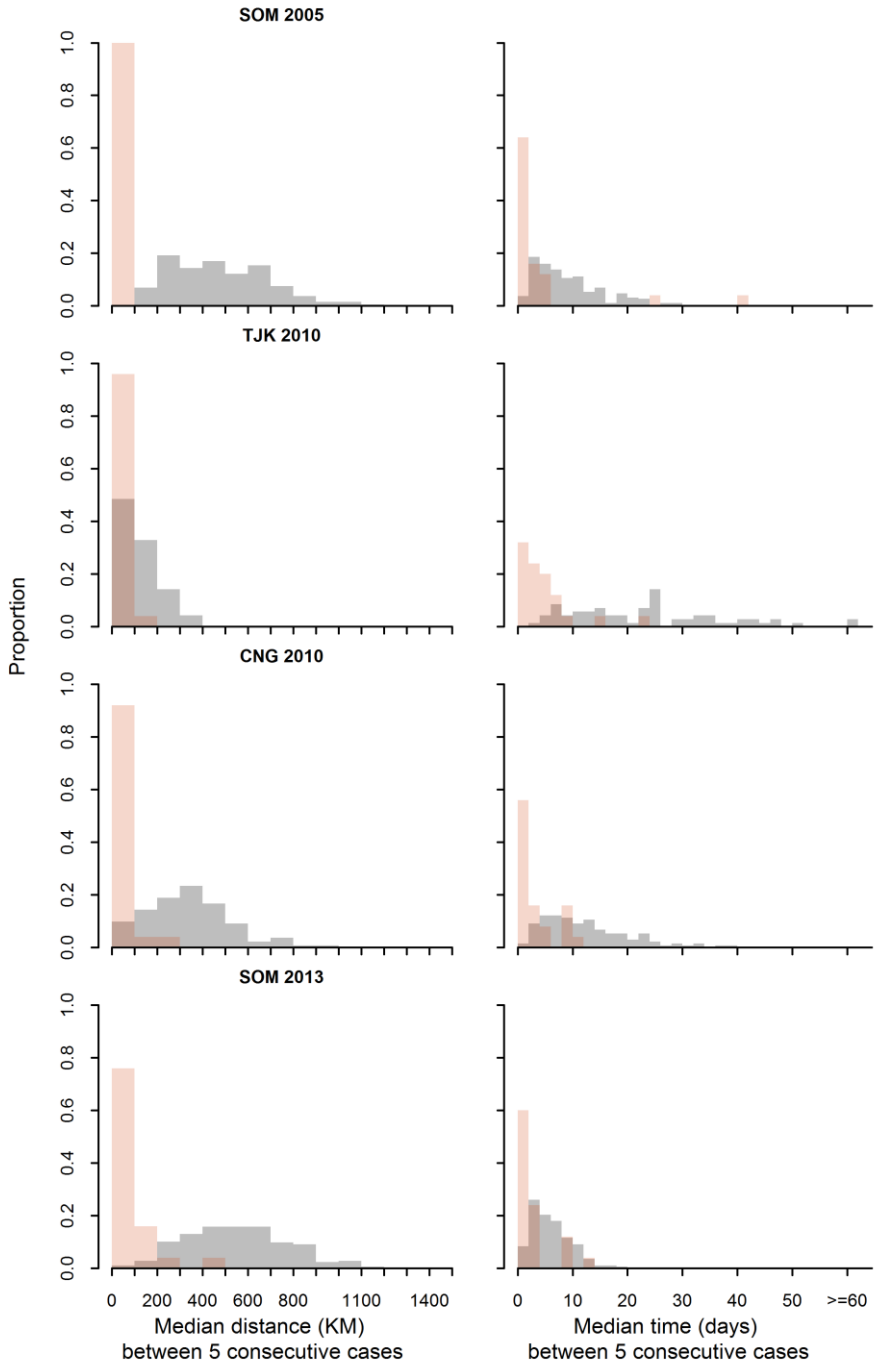
Technical Appendix Figure 6 Left: Percentage change in the observed annual number of nonpolio acute flaccid paralysis cases reported compared to 2003 in sub-Saharan African countries in that have recently experienced a polio importation or outbreak or are considered to be at high risk for these events. Right: Percentage change in the predicted annual number of nonpolio acute flaccid paralysis cases reported compared to 2003, obtained through fitting a spatio-temporal mixed effects regression model to nonpolio acute flaccid paralysis data from 2003-2013. Shading indicates 95% credible intervals. BFA = Burkina Faso, EQG = Equatorial Guinea, GAB = Gabon, GHA = Ghana, CAF = Central African Republic, KEN = Kenya, COD = Democratic Republic of Congo, BEN=Benin, MAD = Madagascar, ETH=Ethiopia, NIG=Niger, CNG = Republic of Congo, SSD = Republic of South Sudan, ANG = Angola, SOM = Somalia, CAE = Cameroon, IVC = Ivory Coast, CHA = Chad.


BURKINA FASO
CHAD
CAF
DRC
GHANA
GABON
KENYA
MALI
NIGER
SOUTH SUDAN
MADAGASCAR
CÔTE D'IVOIRE
ETHIOPIA
CAMEROON
BENIN
CONGO
ANGOLA
SOMALIA
Delay from paralysis onset to notification (days)
0
20
60
100
2010
2011
2012
2013

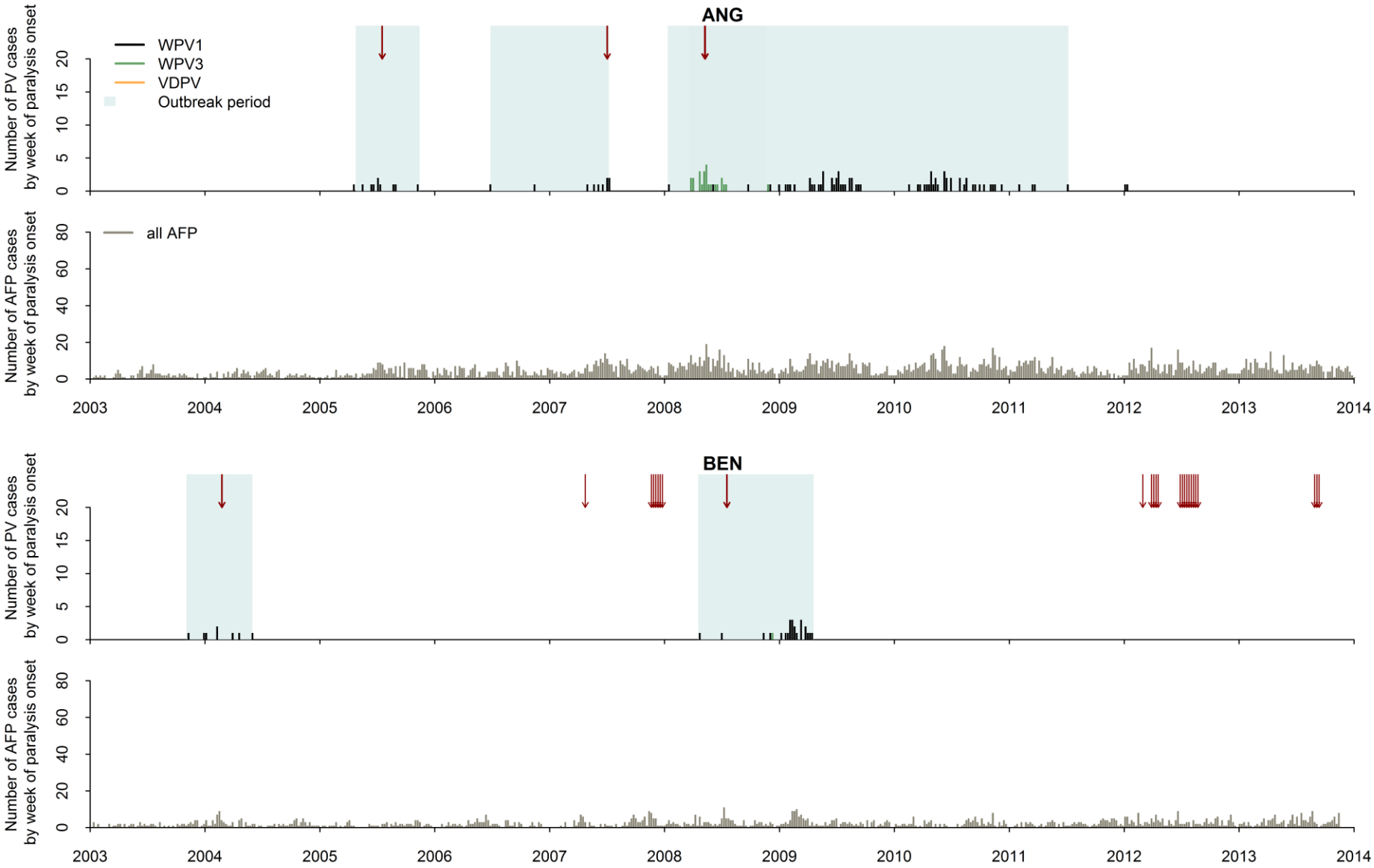
Technical Appendix Figure 7 Distribution (violin plot) of the time delay (days) between onset of acute flaccid paralysis (AFP) and the notification of the case during 2010-2013. White dots correspond to the median value and the black rectangle indicates the interquartile range. The dashed dark red line represents one week and the lighter dashed red line represents two weeks. DRC = Democratic Republic of Congo, CAF = Central African Republic. Of the 30603 AFP cases, 100% had a recorded date of onset and 99.1% had a recorded date of notification. Equatorial Guinea is not included as only two AFP cases were reported during this period.

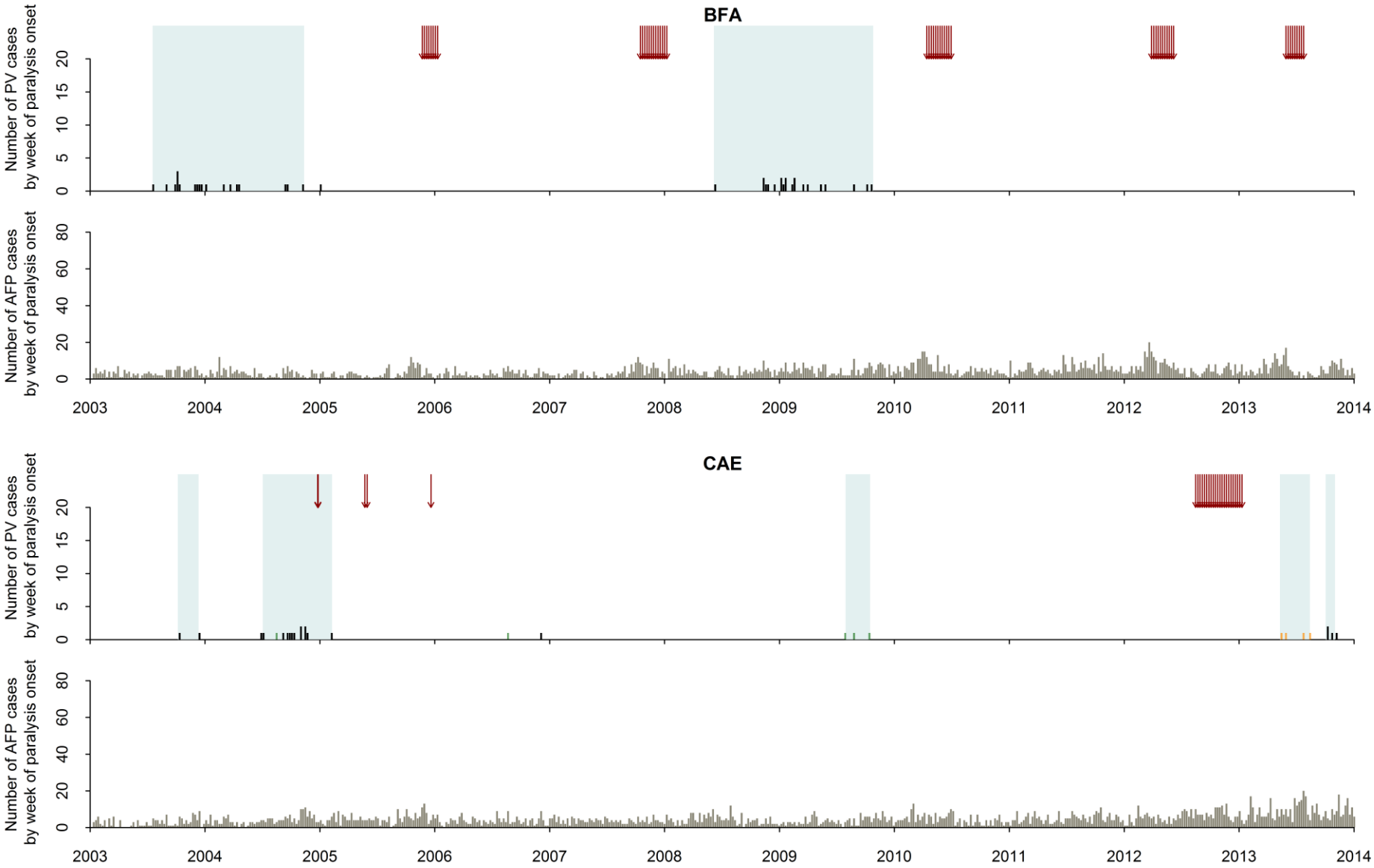

BURKINA FASO
CHAD
CAF
DRC
GHANA
GABON
KENYA
MALI
NIGER
MADAGASCAR
CÔTE D'IVOIRE
ETHIOPIA
CAMEROON
BENIN
CONGO
ANGOLA
Delay from notification to stool sample dispatch to lab (days)
0
20
40
60
80
100
2010
2011
2012
2013

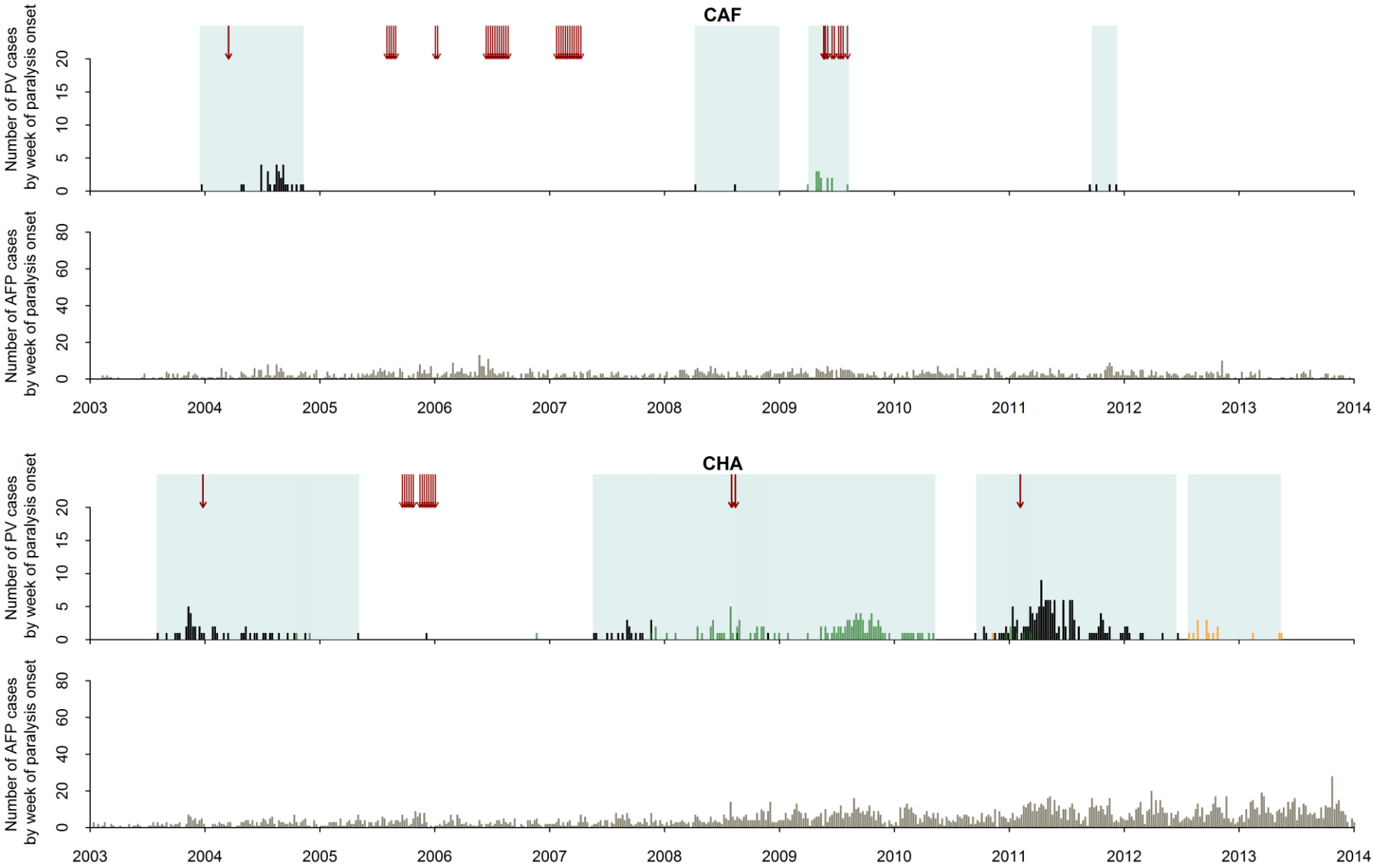
Technical Appendix Figure 8 Distribution (violin plot) of the time delay (days) between case notification of acute flaccid paralysis cases and the date collected stool samples are dispatched to a global polio laboratory during 2010-2013. White dots correspond to the median value and the black rectangle indicates the interquartile range. The dashed dark red line represents one week and the lighter dashed red line represents two weeks. DRC = Democratic Republic of Congo, CAF = Central African Republic. The date that stool samples from Ghana were sent to a laboratory was not available before 2013. Of the 30603 AFP cases, 99.1% had a recorded date of notification and 86.7% had a recorded date that stool samples were dispatched to the laboratory. For the latter, a missing date could have occurred either due to error in data entry or that the stool sample was not collected or dispatched.

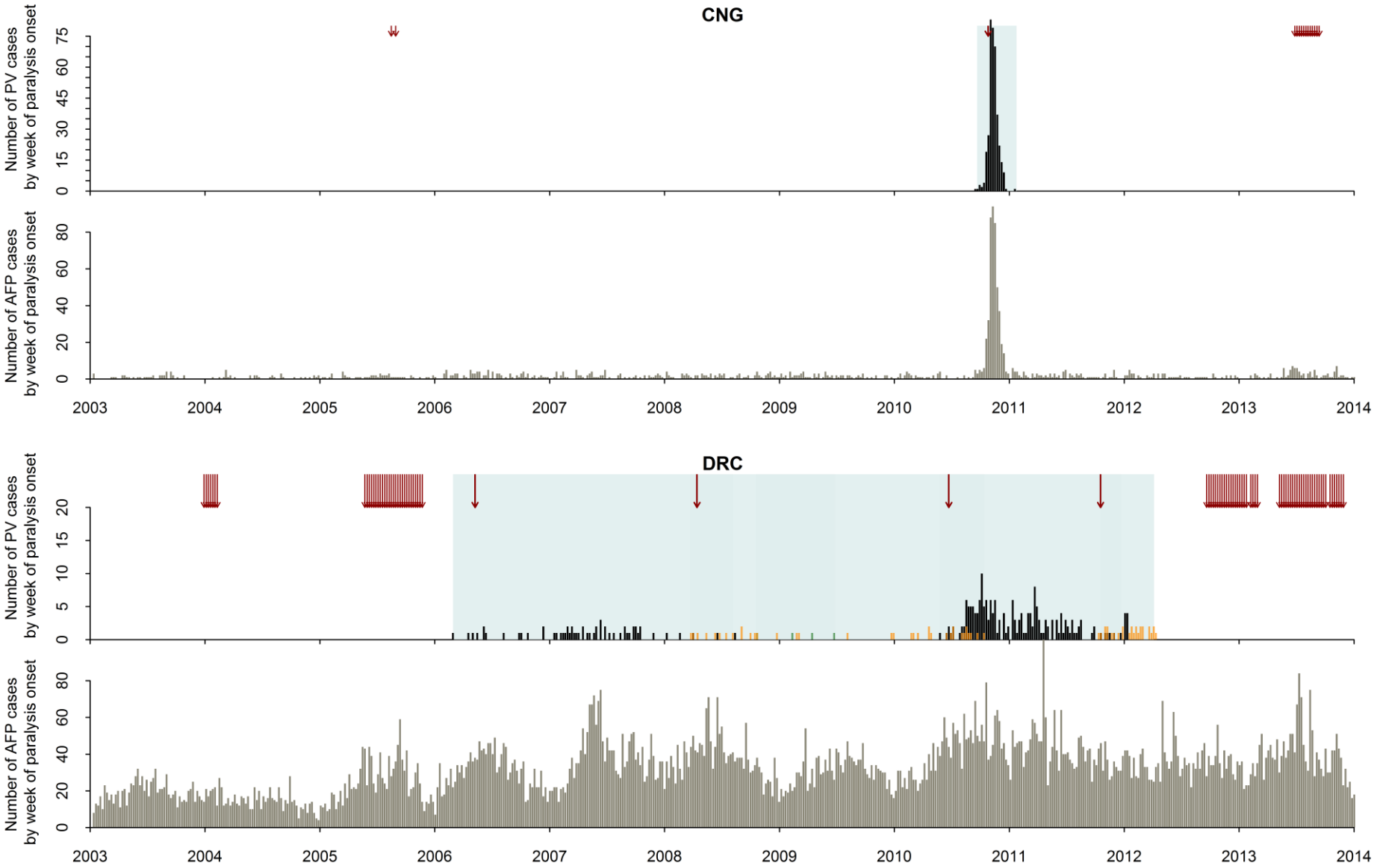


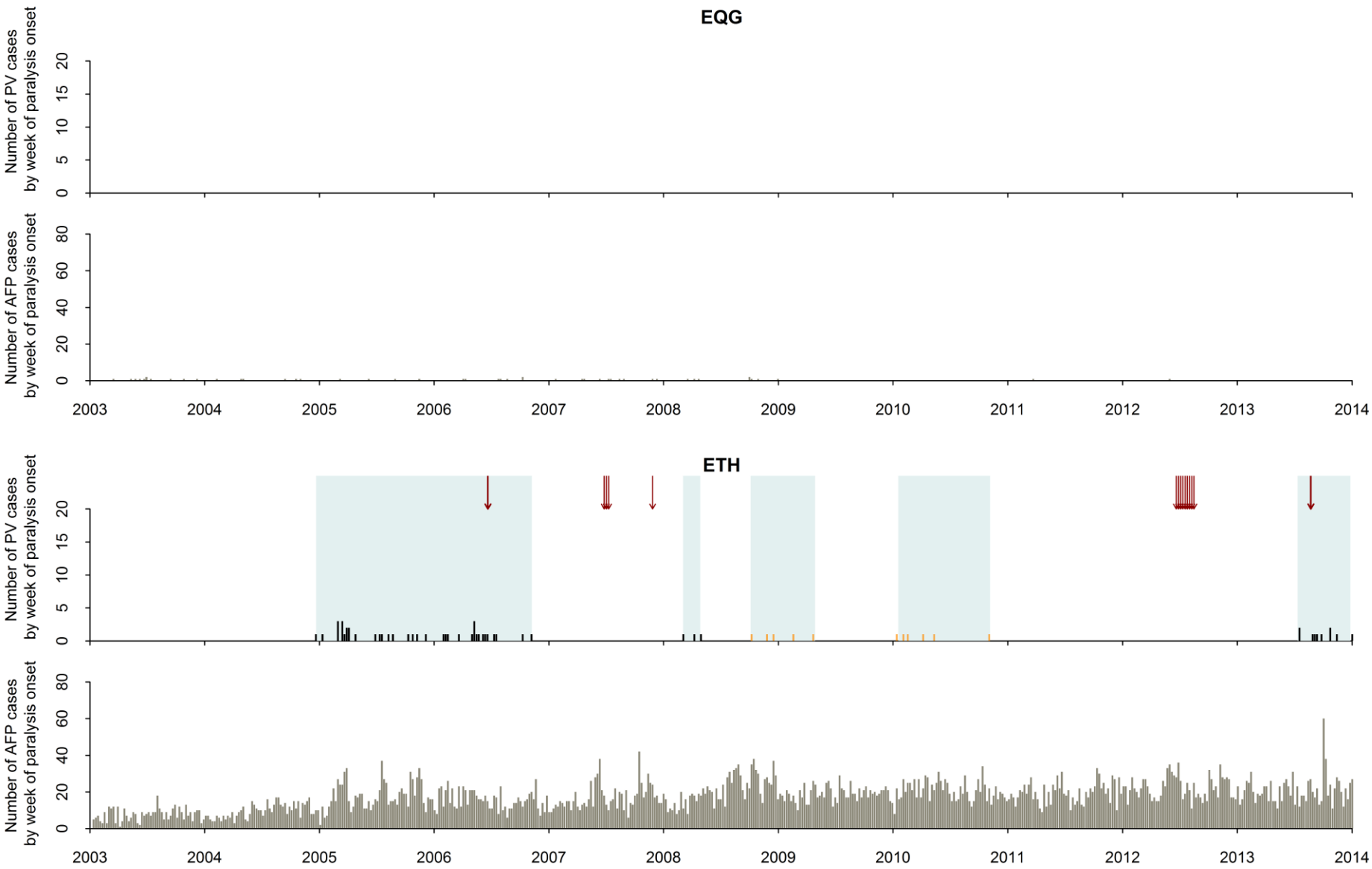
Technical Appendix Figure 9 Distribution of first 30 confirmed poliomyelitis cases (red) during recent large outbreaks in space (left) and time (right) compared to nonpolio acute flaccid paralysis cases (grey) that occurred during the two-year period before the respective outbreaks. Cases were assigned to district centroids and the median great circle distance (km) between a case and the subsequent four cases was computed. Similarly the median time (days) between a case and the subsequent four cases was calculated. SOM 2005 = Somalia 2005 outbreak, TJK 2010 = Tajikistan 2010 outbreak, CNG 2010 = Republic of Congo 2010 outbreak and SOM 2013 = Somalia 2013 outbreak.

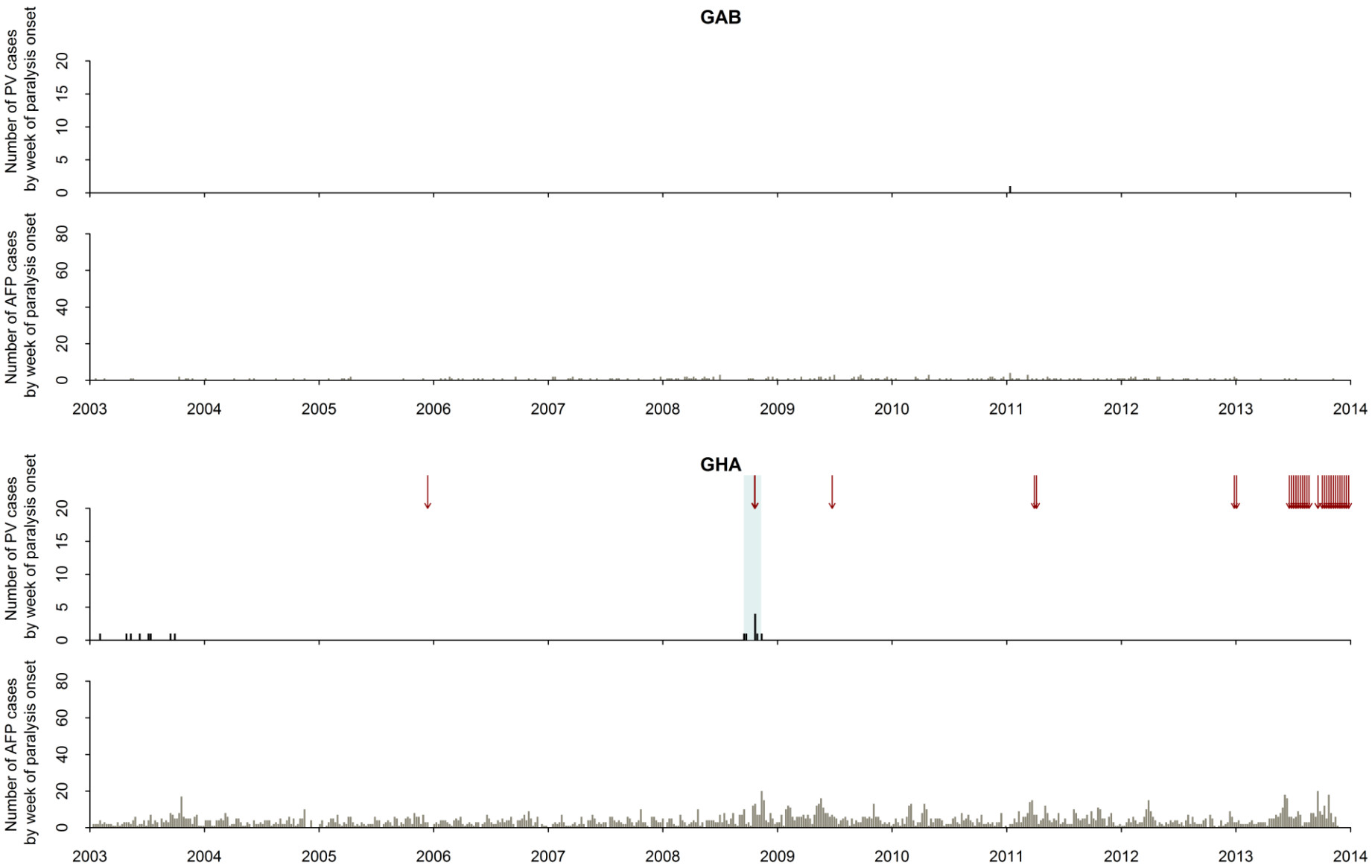

ANG
WPV1
WPV3
VDPV
Outbreak period
Number of PV cases by week of paralysis onset
0
5
10
15
20
all AFP
Number of AFP cases by week of paralysis onset
0
20
40
60
80
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
BEN
Number of PV cases by week of paralysis onset
0
5
10
15
20
Number of AFP cases by week of paralysis onset
0
20
40
60
80
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014

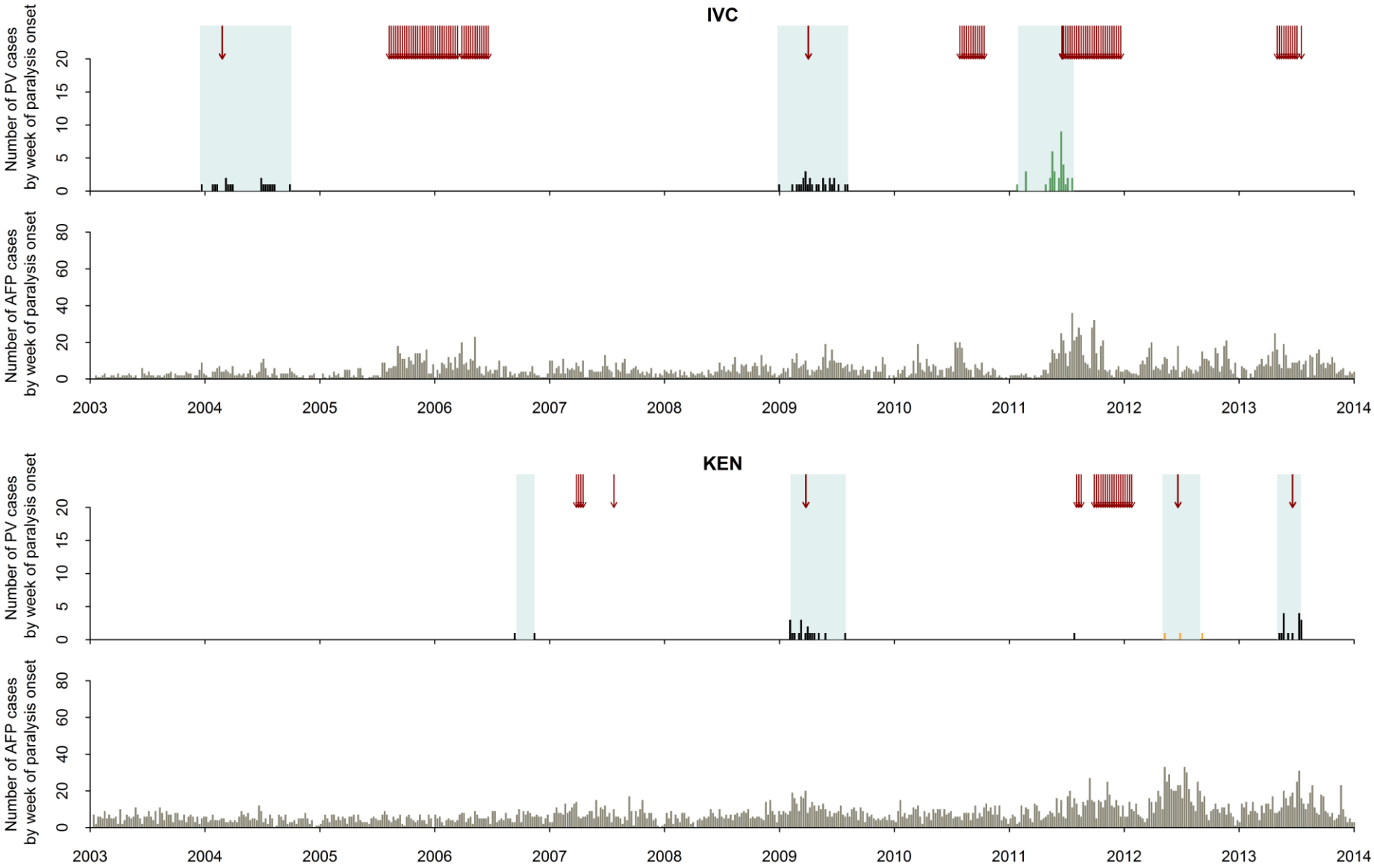

BFA
Number of PV cases by week of paralysis onset
0
5
10
15
20
Number of AFP cases by week of paralysis onset
0
20
40
60
80
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
CAE
Number of PV cases by week of paralysis onset
0
5
10
15
20
Number of AFP cases by week of paralysis onset
0
20
40
60
80
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014

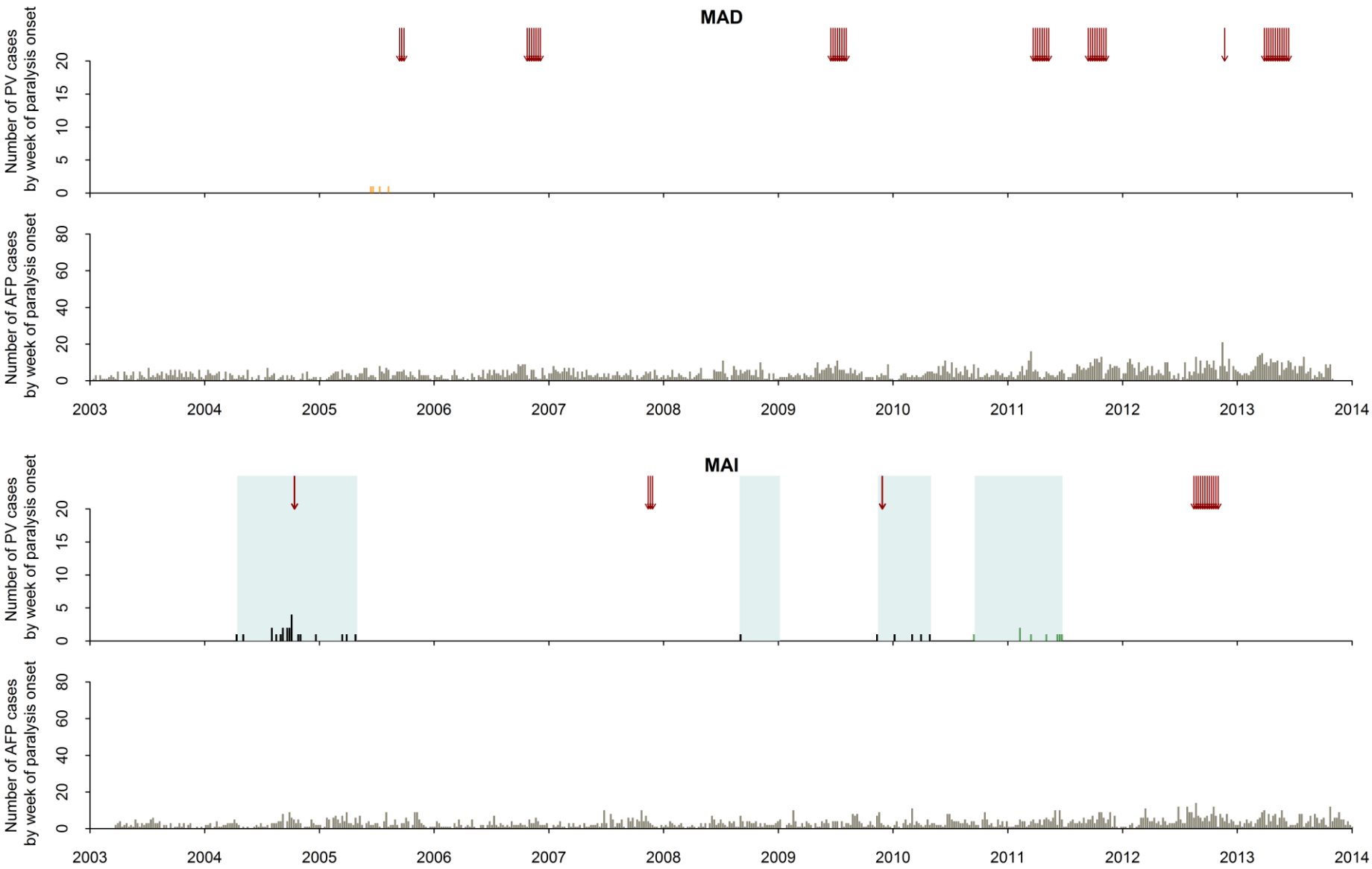

CAF
Number of PV cases by week of paralysis onset
0
5
10
15
20
Number of AFP cases by week of paralysis onset
0
20
40
60
80
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
CHA
Number of PV cases by week of paralysis onset
0
5
10
15
20
Number of AFP cases by week of paralysis onset
0
20
40
60
80
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014

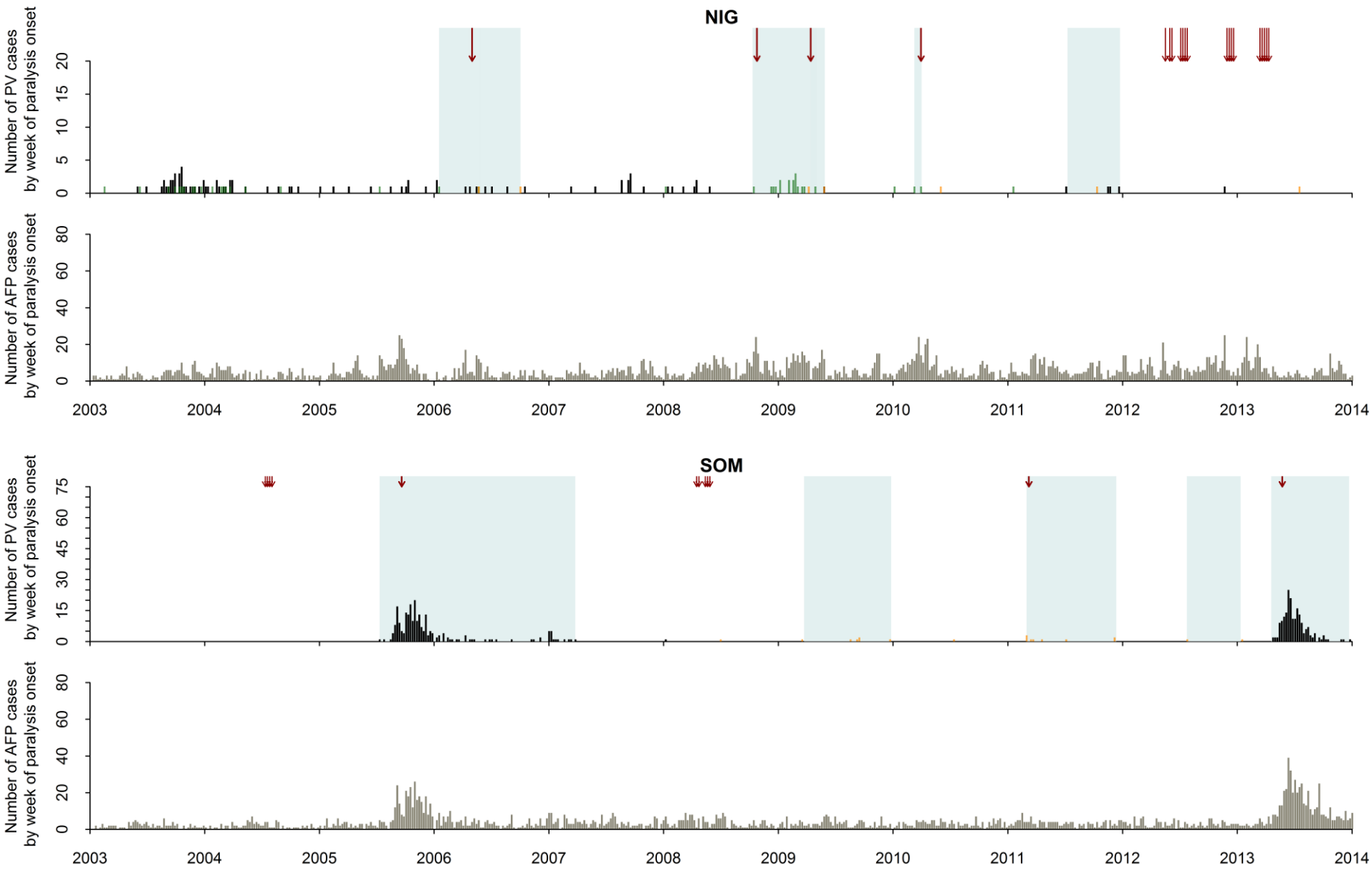

CNG
Number of PV cases by week of paralysis onset
0
15
30
45
60
75
Number of AFP cases by week of paralysis onset
0
20
40
60
80
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
DRC
Number of PV cases by week of paralysis onset
0
5
10
15
20
Number of AFP cases by week of paralysis onset
0
20
40
60
80
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014

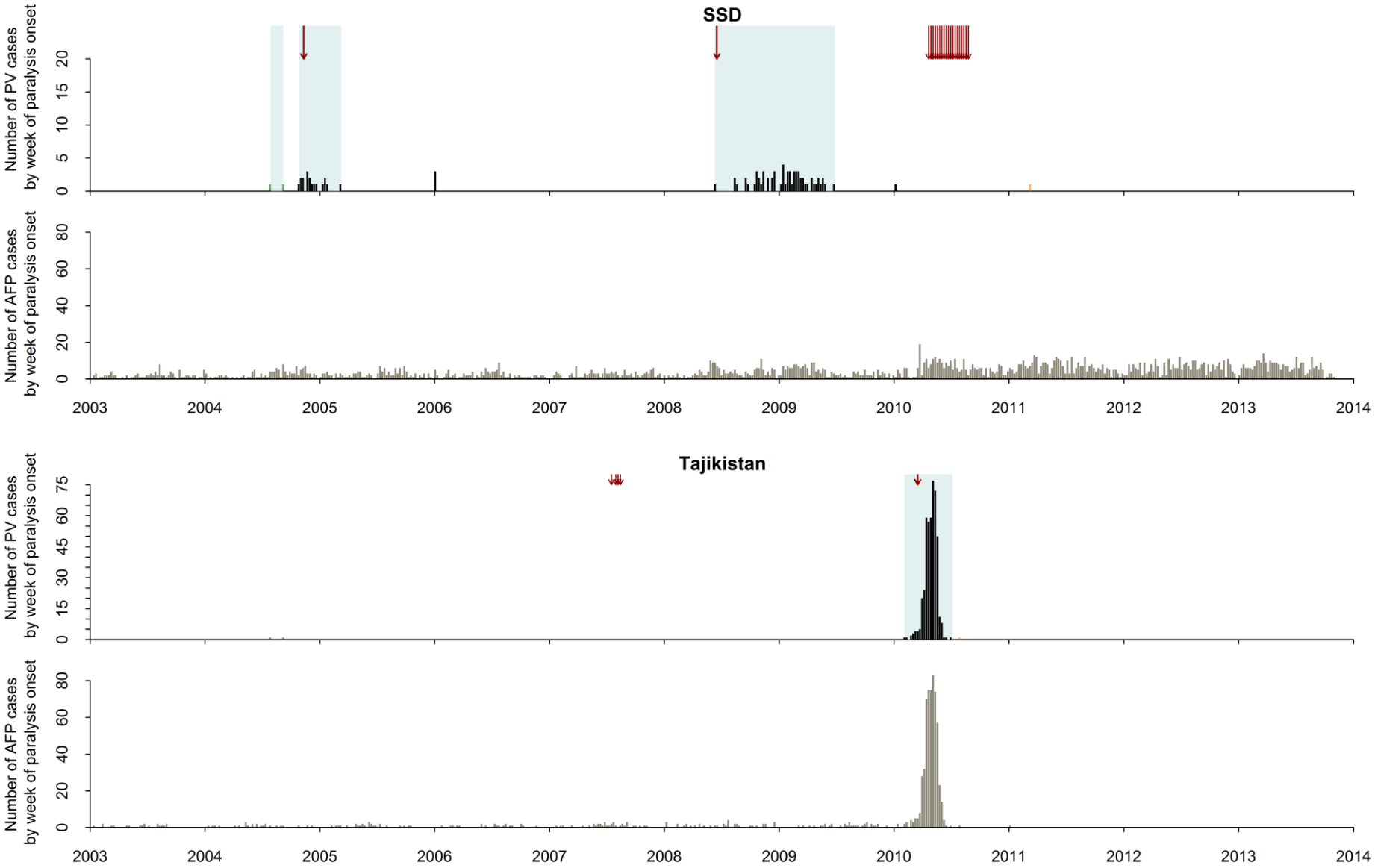

EQG
Number of PV cases by week of paralysis onset
0
5
10
15
20
Number of AFP cases by week of paralysis onset
0
20
40
60
80
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
ETH
Number of PV cases by week of paralysis onset
0
5
10
15
20
Number of AFP cases by week of paralysis onset
0
20
40
60
80
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014


GAB
Number of PV cases by week of paralysis onset
0
5
10
15
20
Number of AFP cases by week of paralysis onset
0
20
40
60
80
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
GHA
Number of PV cases by week of paralysis onset
0
5
10
15
20
Number of AFP cases by week of paralysis onset
0
20
40
60
80
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014

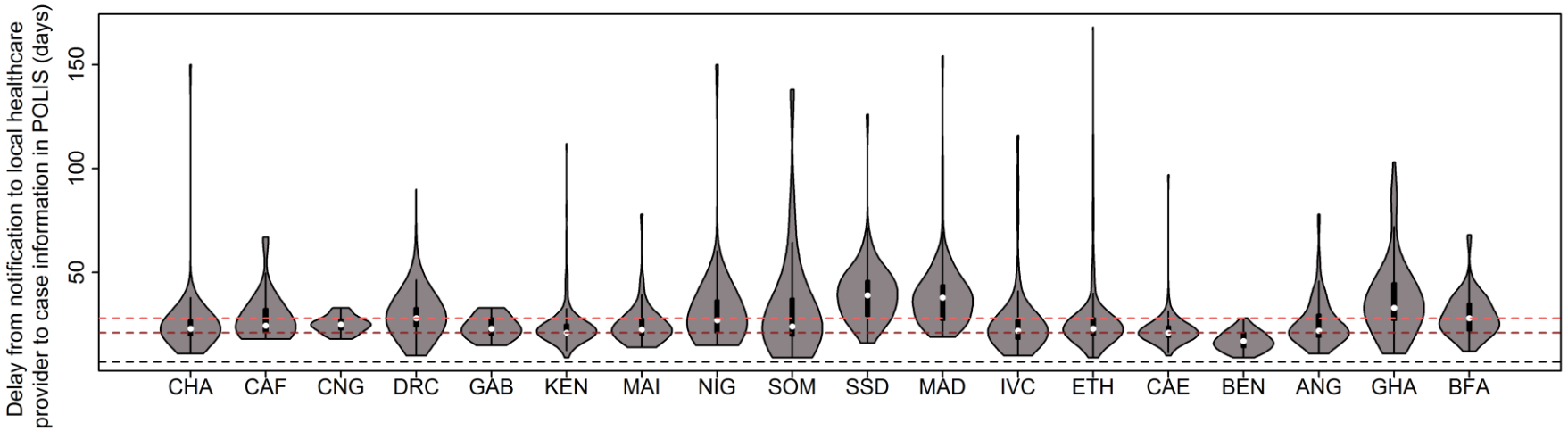

IVC
Number of PV cases by week of paralysis onset
0
5
10
15
20
Number of AFP cases by week of paralysis onset
0
20
40
60
80
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
KEN
Number of PV cases by week of paralysis onset
0
5
10
15
20
Number of AFP cases by week of paralysis onset
0
20
40
60
80
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014


MAD
Number of PV cases by week of paralysis onset
0
5
10
15
20
Number of AFP cases by week of paralysis onset
0
20
40
60
80
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
MAI
Number of PV cases by week of paralysis onset
0
5
10
15
20
Number of AFP cases by week of paralysis onset
0
20
40
60
80
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014


NIG
Number of PV cases by week of paralysis onset
0
5
10
15
20
Number of AFP cases by week of paralysis onset
0
20
40
60
80
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
SOM
Number of PV cases by week of paralysis onset
0
15
30
45
60
75
Number of AFP cases by week of paralysis onset
0
20
40
60
80
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014


SSD
Number of PV cases by week of paralysis onset
0
5
10
15
20
Number of AFP cases by week of paralysis onset
0
20
40
60
80
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
Tajikistan
Number of PV cases by week of paralysis onset
0
15
30
45
60
75
Number of AFP cases by week of paralysis onset
0
20
40
60
80
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014

Technical Appendix Figure 10 Time series of WPV1 (black bars), WPV3 (green bars), cVDPV (orange bars), and all (grey bars) acute flaccid paralysis cases by week of onset in sub-Saharan Africa countries that are currently considered to be at high risk for polio outbreaks, or large outbreak countries. Light blue shading indicates a confirmed wild poliovirus outbreak period (defined as consecutive cases < 6 months apart). Red arrows indicate the date an alarm would have been raised indicating cluster of acute flaccid paralysis cases (An arrow outside of the outbreak period is a false alarm and an arrow within an outbreak period is an 'outbreak alarm'). EQG = Equatorial Guinea, GAB = Gabon, GHA = Ghana, CAF = Central African Republic, KEN = Kenya, DRC = Democratic Republic of Congo, BEN=Benin, BFA = Burkina Faso, MAI = Mali, MAD = Madagascar, ETH=Ethiopia, NIG=Niger, CNG = Republic of Congo, SSD = Republic of South Sudan, ANG = Angola, SOM = Somalia, CAE = Cameroon, IVC = Côte D'Ivoire, CHA = Chad



Technical Appendix Figure 11 Distribution (violin plot) of the time delay between case notification of acute flaccid paralysis cases by local health care providers and the date the case is listed in the centralised polio information system (POLIS). The cases are limited to those who entered POLIS between 25 May 2015 and 10 August 2015 and to those from the African countries considered in this analysis. White dots correspond to the median value and the black rectangle indicates the interquartile range. The dashed black line represents one week, the dashed dark red line represents three weeks and the lighter dashed red line represents four weeks. CHA = Chad, CAF = Central African Republic, CNG = Republic of Congo, DRC = Democratic Republic of Congo, GAB = Gabon, KEN = Kenya, MAI = Mali, NIG=Niger, SOM = Somalia, SSD = Republic of South Sudan, MAD = Madagascar, IVC = Côte D'Ivoire, ETH=Ethiopia, CAE = Cameroon, BEN=Benin, ANG = Angola, GHA = Ghana, BFA = Burkina Faso.

Technical Appendix Table 1 Parameter estimates (and 95% credible intervals) for the spatio-temporal mixed effects model of nonpolio acute flaccid paralysis reporting in sub-Saharan African countries that have recently experienced a polio importation or outbreak or are considered to be at high risk for these events, or countries that have recently experience an outbreak. Reported numbers of cases were assumed to be either Poisson or negative binomial distributed, and the parameter estimates are given for the most parsimonious yet adequate model (*). Both models included population size (all ages) as an offset.

Country	DIC (Poiss)	DIC (NegBin)	Intercept (α)	Size of over-dispersion	$1/\sigma_\mu^2$ (spatially unstructured)	$1/\sigma_v^2$ (spatially structured)	Proportion of spatial variance explained by the structured spatial component	$1/\sigma_\delta^2$ (temporal)
Chad	2673	2660*	-10.9 (-11.0, -10.9)	8.0 (5.9, 10.8)	1918 (130, 6836)	1.4 (0.9, 2.1)	0.998	19827 (1544, 70453)
DRC	24919	22705*	-10.6 (-10.7, -10.5)	2.5 (2.3, 2.7)	1.2 (0.9, 1.5)	1.4 (0.8, 2.4)	0.415	7.7 (2.1, 19.9)
Mali	2298*	2300	-11.5 (-11.6, -11.4)	--	8.9 (5.2, 14.2)	1993 (135, 7365)	0.011	63 (4.1, 332)
Niger	2275	2242*	-11.1 (-11.1, -11.0)	8.3 (6.0, 11.2)	1834 (115, 6751)	3.3 (1.8, 5.3)	0.995	15.2 (1.5, 74.6)
Kenya	3718	3636*	-11.5 (-11.5, -11.4)	7.9 (6.1, 10.0)	1693 (103, 6343)	3.8 (2.2, 5.7)	0.993	16.4 (2.7, 55.2)
South Sudan	3022	2995*	-10.9 (-10.9, -10.8)	5.8 (4.2, 7.9)	13.6 (5.3, 32.7)	4.2 (2.1, 7.1)	0.734	19.7 (3.5, 64.2)
Gabon	710*	718	-11.4 (-11.8, -11.0)	--	0.8 (0.5, 1.4)	2132 (171, 7844)	0.001	18.0 (1.6, 70.4)
CAF	1171*	1174	-10.0 (-10.7, -9.5)	--	1834 (123, 6601)	0.3 (0.2, 0.5)	0.999	8.4 (1.8, 25.2)
Rep. Congo	1113*	1128	-10.3 (-10.9, -9.6)	--	0.4 (0.2, 0.6)	1181 (128, 6833)	0.000	3.2 (0.7, 9.5)
Somalia	3295*	3506	-10.5 (-10.6, -10.5)	--	1862 (126, 6696)	0.1 (0.1, 0.2)	1.000	7.5 (1.5, 23.5)
Tajikistan	355*	368	-14.7 (-16.0, -13.8)	--	0.2 (0.1, 0.5)	1854 (129, 6738)	0.000	18840 (1334, 67750)
Angola	4280*	4293	-11.6 (-11.7, -11.4)	--	1.2 (0.8, 1.6)	1880 (115, 6558)	0.002	79 (8.8, 274)
Ethiopia	4520	4511*	-11.4 (-11.5, -11.4)	17.2 (13.2, 21.8)	1717 (100, 6387)	6.4 (4.1, 9.5)	0.990	17 (3.2, 61)
Benin	2561	2560*	-11.3 (-11.4, -11.2)	5.3 (3.5, 7.9)	6.8 (4.0, 10.8)	1650 (74, 6308)	0.015	18775 (1299, 68030)
Côte D'Ivoire	3127*	3131	-11.0 (-11.0, -10.9)	--	1834 (120, 6631)	1.6 (1.1, 2.2)	0.997	7.0 (1.9, 18)
Cameroon	5311	5283*	-11.3 (-11.5, -11.2)	5.3 (4.0, 7.1)	0.9 (0.7, 1.1)	1676 (86, 6327)	0.002	39 (6.4, 131)
Madagascar	3347*	3371	-11.7 (-11.8, -11.6)	--	2430 (231, 9184)	1.05 (0.72, 1.48)	0.999	27350 (2815, 105600)
Equatorial Guinea	176*	198	-13.4 (-14.3, -12.7)	--	1.7 (0.3, 5.6)	2095 (162, 7735)	0.002	18590 (1250, 67460)
Burkina Faso	2930	2910*	-11.2 (-11.3, -11.1)	7.0 (5.1, 9.5)	6.1 (3.8, 9.2)	1989 (144, 7386)	0.010	407 (25, 2166)
Ghana	4501*	4538	-11.6 (-11.7, -11.5)	--	6.8 (4.3, 10.2)	45.8 (9.5, 155.1)	0.276	9.8 (1.5, 37.3)

Reporting of nonpolio AFP at the district level in Chad, CAF, Niger, Kenya, Somalia, Ethiopia, Madagascar and Côte D'Ivoire was dependent upon reporting in adjacent districts (as the proportion of spatial variance explained by the spatially structured random effect was relatively high) whereas in Angola, Mali, Gabon, Republic of Congo, Tajikistan, Benin, Equatorial Guinea, Burkina Faso and Cameroon the heterogeneity in reporting was randomly distributed across the country (as the proportion of spatial variance explained by the spatially unstructured random effect was relatively high). DRC, Ghana and South Sudan contained areas of spatial dependence and other areas of randomly distributed reporting (as the spatial variance was explained by both the spatially unstructured and structured random effects).

Technical Appendix Table 2 Sensitivity analysis to the significance p-value determining when an early warning alarm would have been raised for the large major recent outbreaks [Specificity is calculated across the outbreak free period of a given country therefore there is only one estimate per country and significance level]

	Date of alarm			Specificity (%)		
Significance level of alarm	0.001	0.05	0.1	0.001	0.05	0.1
Somalia 2005	19 Sep 2005	19 Sep 2005	19 Sep 2005	100	97	97
Somalia 2013	20 May 2013	20 May 2013	20 May 2013			
Tajikistan 2010	15 March 2010	15 March 2010	15 March 2010	100	99	99
Republic of Congo 2010	25 Oct 2010	25 Oct 2010	25 Oct 2010	97	96	96

Technical Appendix Table 3 Sensitivity analysis to the significance p-value determining when an early warning alarm would have been raised in sub-Saharan African countries during 2004-2013

	Percentage of outbreaks detected			Specificity (%)		
Significance level of alarm	0.001	0.05	0.1	0.001	0.05	0.1
Chad	50	62	75	94	89	87
Central African Republic	25	50	50	95	91	90
Democratic Republic of Congo	80	80	80	72	63	61
Gabon	NA	NA	NA	100	100	100
Kenya	75	75	75	94	91	90
Mali	25	50	50	98	95	94
Niger	50	67	67	96	87	85
South Sudan	67	67	67	97	96	95
Madagascar	NA	NA	NA	97	92	90
Cote D'Ivoire	33	100	100	85	81	80
Ethiopia	40	40	40	98	93	92
Equatorial Guinea	NA	NA	NA	100	100	100
Cameroon	0	20	20	96	93	92
Benin	100	100	100	98	94	92
Angola	100	100	100	97	96	96
Ghana	100	100	100	94	92	90
Burkina Faso	0	0	0	90	88	86

Technical Appendix Table 4 Sensitivity analysis to the definition of specific of the cluster detection of acute flaccid paralysis cases as an early warning system for detection of polio outbreaks. Specificity was primarily calculated as the percentage of outbreak-free weeks without a false alarm. A second measure of specificity was calculated by omitting 90 days either side of each outbreak period from weeks included in the outbreak-free period (to allow for the potential of misclassification of polio cases at the start of an outbreak or for the continuation of detecting clustering of polio cases at the tail end of the outbreak in a cylinder for a short period after the date of paralysis onset of last case).

Country	Original Specificity (%)	Specificity (%) removing a 90 day window either side of the outbreak period
Somalia	97	96
Tajikistan	99	99
Republic of Congo	96	97
Chad	89	84
Central African Republic	91	90
Democratic Republic of Congo	63	62
Gabon	100	100
Kenya	91	94
Mali	95	96
Niger	87	81
South Sudan	96	95
Madagascar	92	92
Cote D'Ivoire	81	79
Ethiopia	93	95
Equatorial Guinea	100	100
Cameroon	93	93
Benin	94	95
Angola	96	100
Ghana	92	94
Burkina Faso	88	86

Technical Appendix Table 5 Performance of cluster detection of acute flaccid paralysis cases during large outbreaks using the Poisson space-time scan statistic compared to the space-time permutation scan statistic

	Date of alarm		Specificity (%)	
	Poisson	**Permutation**	**Poisson**	**Permutation**
Somalia 2005	19 Sep 2005	19 Sep 2005	97	99
Somalia 2013	20 May 2013	20 May 2013		
Tajikistan 2010	15 March 2010	26 April 2010	99	100
Republic of Congo	25 Oct 2010	1 Nov 2010	96	97

Technical Appendix Table 6 Performance of cluster detection of acute flaccid paralysis cases during 2003 – 2013 as an early warning system for detection of polio outbreaks using the space-time permutation scan statistic (compare to Table 1 for comparison with results from the Poisson space-time scan statistic).

Country	**Number of confirmed (WPV1, WPV3, cVDPV) polio outbreaks 2003-2013**	**Percentage identified**	**Specificity (%)**
Somalia	5	60	99
Tajikistan	1	100	100
Republic of Congo	1	100	97
Chad	8	62	96
Central African Republic	4	50	94
Democratic Republic of Congo	5	100	60
Gabon	0	NA	98
Kenya	4	50	95
Mali	4	50	95
Niger	6	17	95
South Sudan	3	33	100
Madagascar	0	NA	97
Cote D'Ivoire	3	0	93
Ethiopia	5	40	96
Equatorial Guinea	0	NA	100
Cameroon	5	0	97
Benin	2	0	97
Angola	4	50	100
Ghana	1	100	95
Burkina Faso	2	0	98

References

1. Besag J, York J, Mollie A. Bayesian Image-Restoration, with 2 Applications in Spatial Statistics. Ann I Stat Math. 1991 Mar;43(1):1-20.
2. Rue H, Held L. Gaussian Markov random fields: theory and applications: Chapman & Hall/CRC; 2005.
3. Sorbye SH, Rue H. Scaling intrinsic Gaussian Markov random field priors in spatial modelling. Spatial Statistics. 2014;8:39-51.
4. Spiegelhalter DJ, Best NG, Carlin BR, van der Linde A. Bayesian measures of model complexity and fit. J R Stat Soc B. 2002;64:583-616.
5. Blangiardo M, Cameletti M. Spatial modelling. Spatial and spatio-temporal bayesian models with R-INLA. Chichester: John Wiley & Sons, Ltd; 2015. p. 185 - 6.
6. Kulldorff M. A spatial scan statistic. Communications in Statistics-Theory and methods. 1997;26(6):1481-96.
7. Kulldorff M, Information Management Services Inc. SatScan™ v9.3 : Software for the spatial and space-time scan statistics. satscan.org . Accessed 12 June 2014.
8. Kulldorff M, Heffernan R, Hartman J, Assuncao R, Mostashari F. A space-time permutation scan statistic for disease outbreak detection. PLoS medicine. 2005 Mar;2(3):216-24.
9. Hijmans RJ. geosphere: Spherical Trigonometry. R package version 1. 3-11. http://CRAN.R-project.org/package=geosphere. 2014.
10. Pebesma E. spacetime: Spatio-Temporal Data in R. J Stat Softw. 2012 Nov;51(7):1-30.

Technical Appendix: Sample R code

```
#########################################################################
## FASTER DETECTION OF POLIOMYELITIS OUTBREAKS TO SUPPORT POLIO ERADICATION  ##
#
## AUTHORS: Isobel M Blake, Paul Chenoweth, Hiro Okayasu, Christl A Donnelly, R Bruce Aylward
#and Nicholas C Grassly

#THIS R CODE PROVIDES AN EXAMPLE OF HOW TO RUN THE SPATIO-TEMPORAL REGRESSION AND HOW TO TEST
#SURVEILLANCE DATA AVAIABLE FOR A
#COUNTRY ON A GIVEN DATE FOR EVIDENCE OF CLUSTERING OF ACUTE FLACCED PARALYSIS CLUSTERS

###### SPATIO-TEMPORAL REGRESSION ########################################

#Data access assumptions:
# Acute flaccid paralysis (AFP) cases are reported at the same spatial unit as the shapefile
#available
# (E.G. in this work the second administrative (district) level)

# 1: Read shapfile into R and call this 'Countrymap'
library(rgdal) #package to load shapefile in R (see Bivand, Pebesma and Gomez-Rubio: Applied
#Spatial Data Analysis with R book for further details)

setwd("C:/folder_of_shapefile") # Directory where shapefile is saved

Countrymap<-readOGR(dsn = ".", layer="Name_of_shapefile")

# 2: Read in summary table of number of reported AFP cases per district per year

# This code assumes data cleaning has been done to obtain a table of the number AFP reported
#cases per district per year
# this is called 'summary_table' NOTE,THE ORDER OF DISTRICTS MUST BE THE SAME AS THAT OF THE
#SHAPEFILE
# IT IS ALSO ASSUMED THAT DISTRICT NAMES ARE UNIQUE

setwd("C:/AFP_analysis") # location of where the table is stored as a .csv file

summary_table=read.csv('summary_table.csv',stringsAsFactors =F,row.names=1)

#This will look like this (made up example):
#                 2003  2004  2005  2006  2007  2008  2009  2010 2011  2012  2013
#Name district1   0       1     5     3     1     2     5     6     5     7     6
#Name_district2   1       4     2     5     7     6     5     10    11    9     10
#Name_district3   2       2     8     1     3     5     2     7     8     6     9

names(summary_table)=gsub("X","",names(summary_table)) #remove X from column names

# 3: read in district population sizes

setwd("C:/District_population size") # Directory of where the district population size data is
#located

pop=read.csv('district_population.csv',stringsAsFactors =F,row.names=1)

#This will look like this (made up example):
#                 Pop_size
#Name district1   80000
#Name district2   110000
#Name district3   150000

#NOTE DISTRICTS MUST BE IN THE SAME ORDER AS LISTED IN THE SHAPEFILE

# 4: Create space-time full dataframe (STFDF) for spatio-temporal regression#################

#load relavent packages
library(xts)
library(spacetime)

#create a vector of years of class date
year_dates<-as.Date(strptime(paste(1," ",1," ",as.numeric(names(summary_table)),sep=""),"%d %m
%Y"))

#create stfdf
np_stfdf<-STFDF(as(Countrymap,"SpatialPolygons"),
```

```
                xts(1:ncol(summary_table),year_dates),

data.frame(district_AFP=unlist(summary_table),row.names=1:((dim(summary_table)[1])*(dim(summar
y_table)[2])))))



# 5: Spatio-temporal regression   ############################
library(INLA) # package for Integrated Nested Laplace Approximation for Bayesian regression
#modelling

library(spdep) # package to create district adjacency matrix

setwd("C:/AFP_analysis/space_time_model")

neib<-poly2nb(Countrymap) # creates adjacency matrix
nb2INLA("statesth.adj", neib)

yrs=as.numeric(names(summary_table)) # create vector of years there is data for

#this creates the corresponding year and district index for each line in the space-time
#dataframe
np_stfdf$YEAR<-rep(yrs, each=length(Countrymap$ADM2_NAME)) # Countrymap$ADM2_NAME contains a
#vector of the district names in the shapefile
np_stfdf$STATEID<-rep(1:length(Countrymap$ADM2_NAME), length(yrs))

np_stfdf$pop<-rep(log(pop$Pop_size),length(yrs)) # add population size


#run spatio-temporal model in INLA, see r-inla.org for further information
form<-district_AFP~1+offset(pop)+f(STATEID, model="bym", graph="statesth.adj",scale.model =
TRUE)+f(YEAR, model="rw2",scale.model = TRUE)
inlares<-inla(form, family="nbinomial", data=as.data.frame(np_stfdf),
              control.predictor=list(compute=TRUE, link = 1),
              control.compute=list(dic=TRUE),
              control.results=list(return.marginals.predictor=TRUE),
)
summary(inlares)

#obtain annual AFP reporting independnt of time,'expected_afp'

spatial_iid_and_besag_random_effect<-
inlares$summary.random$STATEID[1:length(Countrymap$ADM2_NAME),2]

expected_afp<-pop$Pop_size *
exp(spatial_iid_and_besag_random_effect+summary(inlares)$fixed[1,1])

library(stringr) # package for formatting strings

district_names_simple=(toupper(str_replace_all(Countrymap$ADM2_NAME, "[^[:alnum:]]", ""))) #
#districts must only contain letters without spaces, dashes or accents

expected_afP_df<-data.frame(District=district_names_simple
                            ,Year=2010, expected_afp_model=expected_afp) # this is in the
#format for SaTScan, the year is the yr population data was obtained

write.table(expected_afP_df, "C:/SaTScan_running/example/expected_reporting.txt", sep="
",row.names = FALSE,
            col.names = FALSE,quote=F)


##############################################################################################
###########
##############################Running SaTScan clustering
algorithm#########################################


# this sections assumes that the program SaTScan has been downloaded from
#http://www.satscan.org/. This code relates to v9.3
# this code assumes the executable file is located in C:\Program Files (x86)\SaTScan


setwd("C:/SaTScan_running/example")

#In this folder should be the expected reporting file created above: expected reporting.txt,
#the geo.txt file described below, the satscan.prm file
#and the AFP data of the last two years back from the current data of surveillance 'data.txt'.
```

```
##geo.txt#####
#a .txt file containing 3 columns: column1 = list of district names,
#column2 = latitude of centroid of each corresponding district,
#column3 = longitude  of centroid of each corresponding district
#District names must not contain spaces, accents or hyphens

#eg
########################################
#Name district1   2.33333333  -0.55555555
#Name_district2   3.11111111  -0.11111111
#Name_district3   2.22222222  -0.44444444

###data.txt#####
#a table of district names (column 1), the number of AFP cases on that day (column 2),
# and the date that cases had onset of paralysis in the district (column 3))
#If there are no AFP cases in a district on a given day there are not lines in the table for
#this district-day
#District names must not contain spaces, accents or hyphens

#eg
############################
#Name_district1 1 2005-01-18
#Name district1 1 2005-03-25
#Name district2 1 2005-06-03


###satscan.prm###
#see http://www.satscan.org/ for details on this file in the user guide. An example is given
#at the end of this file

#Build function to call 'SatScanBatch.exe'

BuildCommand<-function(program_name_str, paramater_filename, my_params) {
  command_str = paste(program_name_str , " " , paramater_filename,sep="")
  for (name in names(my_params)) {
    command_str = paste(command_str , " --" , name , " " , my_params[[name]],sep="")
  }
  return(command_str)
}

Current_surveillance_date= as.Date("2013-05-20")



program_name = "\"C:/Program Files (x86)/SaTScan/SatScanBatch.exe\""
parameter_file_name = "\"C:/SaTScan_running/example/satscan.prm\""
case_file_name="\"C:/SaTScan_running/example/data.txt\""


time_aggregation = 1
max_radius = 500
max_time_clus = 90 # 90 days
mcreps = 999 # for monte carlo testing
popfile="C:/SaTScan_running/example/expected_reporting.txt"
coordfile="C:/SaTScan_running/example/geo.txt"

#time trend adjustment type (0=None, 1=Nonparametric, 2=LogLinearPercentage,
#3=CalculatedLogLinearPercentage, 4=TimeStratifiedRandomization,
#5=CalculatedQuadraticPercentage)
time_adjust = 0

start_date=strftime(Current_surveillance_date-2*365,"%Y/%m/%d") # dates need to be in this
#format

end_date=strftime(Current_surveillance_date,"%Y/%m/%d")
prospect_start=strftime(Current_surveillance_date-365,"%Y/%m/%d") # date to correct for
#multple testing

outfile=paste("\"C:/SaTScan_running/example/res_",Current_surveillance_date,"\"",sep="")


my_params = list(CaseFile = case_file_name, EndDate = end_date, StartDate=start_date,
ResultsFile = outfile, TimeAggregationLength = time_aggregation,
                 MaxSpatialSizeInDistanceFromCenter = max_radius,MaxTemporalSize =
max_time_clus, TimeTrendAdjustmentType = time_adjust,
```

```
                ProspectiveStartDate=prospect_start, MonteCarloReps=mcreps,
PopulationFile=popfile,CoordinatesFile=coordfile)

my_command_str = BuildCommand(program_name,parameter_file_name,my_params)

system(my_command_str,show.output.on.console = TRUE)

#the output file 'res 2006-05-15' contains summary information on whether an outbreak was
#detected.
#The output file 'res 2006-05-15.col.txt' contains statistics on the cluster that can be read
#into R
#The output file 'res_2006-05-15.gis.txt' contains information on the districts contained
#within the identified cluster
#see the SaTScan user guide for further information

#this is how you obtain the following through R: significance level of the cluster, the
#district whose centroid is the center of the cluster,
#the radius of the cluster , and the start date of the cluster

outfile=paste("res_",Current_surveillance_date,".col.txt",sep="")
result<-read.table(outfile,header=F)
prob <- result[1,10]
center <- as.character(result[1,2])
radius <- result[1,5]
date1 <- as.Date(result[1,6])

Cluster_map<-readOGR(dsn = ".", layer=paste("res_",Current_surveillance_date,".col",sep=""))
#plot circle of cluster cyclinder on top of the country shapefile
plot(Countrymap)
plot(Cluster_map,add=T,border="red")

library(rgeos) # package to extract centroid of districts in shapefile
plot(centroids,add=T, col="blue") # only centroids within or touching the circumference of the
#circle are included in the cluster

#END
```

Sample .prm file to run SaTScan http://satscan.org/

```
[Input]
;case data filename - R file sends a string to overwrite this
CaseFile=C:\SaTScan\example\data.txt
;control data filename
ControlFile=
;time precision (0=None, 1=Year, 2=Month, 3=Day, 4=Generic)
PrecisionCaseTimes=3
;study period start date (YYYY/MM/DD) - R file sends a string to overwrite this
StartDate=2003/1/1
;study period end date (YYYY/MM/DD) - R file sends a string to overwrite this
EndDate=2010/04/05
;population data filename
PopulationFile=C:\SaTScan\example\expected_reporting.pop
;coordinate data filename
CoordinatesFile=C:\SaTScan\example\geofile.geo
;use grid file? (y/n)
UseGridFile=n
;grid data filename
GridFile=
;coordinate type (0=Cartesian, 1=latitude/longitude)
CoordinatesType=1

[Analysis]
;analysis type (1=Purely Spatial, 2=Purely Temporal, 3=Retrospective Space-Time, 4=Prospective
Space-Time, 5=Spatial Variation in Temporal Trends, 6=Prospective Purely Temporal)
AnalysisType=4
;model type (0=Discrete Poisson, 1=Bernoulli, 2=Space-Time Permutation, 3=Ordinal,
4=Exponential, 5=Normal, 6=Continuous Poisson, 7=Multinomial)
ModelType=0
;scan areas (1=High Rates(Poisson,Bernoulli,STP); High Values(Ordinal,Normal); Short
Survival(Exponential), 2=Low Rates(Poisson,Bernoulli,STP); Low Values(Ordinal,Normal); Long
Survival(Exponential), 3=Both Areas)
ScanAreas=1
;time aggregation units (0=None, 1=Year, 2=Month, 3=Day, 4=Generic)
TimeAggregationUnits=3
;time aggregation length (Positive Integer)
TimeAggregationLength=1

[Output]
;analysis results output filename - R file sends a string to overwrite this
ResultsFile=C:\SaTScan\example\res
;output Google Earth KML file (y/n)
OutputGoogleEarthKML=y
;output shapefiles (y/n)
OutputShapefiles=y
;output cluster information in ASCII format? (y/n)
MostLikelyClusterEachCentroidASCII=y
;output cluster information in dBase format? (y/n)
MostLikelyClusterEachCentroidDBase=n
;output cluster case information in ASCII format? (y/n)
MostLikelyClusterCaseInfoEachCentroidASCII=y
;output cluster case information in dBase format? (y/n)
MostLikelyClusterCaseInfoEachCentroidDBase=n
;output location information in ASCII format? (y/n)
CensusAreasReportedClustersASCII=y
;output location information in dBase format? (y/n)
CensusAreasReportedClustersDBase=n
;output relative risks in ASCII format? (y/n)
IncludeRelativeRisksCensusAreasASCII=y
;output relative risks in dBase format? (y/n)
IncludeRelativeRisksCensusAreasDBase=n
;output simulated log likelihoods ratios in ASCII format? (y/n)
SaveSimLLRsASCII=y
;output simulated log likelihoods ratios in dBase format? (y/n)
SaveSimLLRsDBase=n

[Multiple Data Sets]
; multiple data sets purpose type (0=Multivariate, 1=Adjustment)
MultipleDataSetsPurposeType=0

[Data Checking]
;study period data check (0=Strict Bounds, 1=Relaxed Bounds)
StudyPeriodCheckType=0
```

```
;geographical coordinates data check (0=Strict Coordinates, 1=Relaxed Coordinates)
GeographicalCoordinatesCheckType=0

[Spatial Neighbors]
;use neighbors file (y/n)
UseNeighborsFile=n
;neighbors file
NeighborsFilename=
;use meta locations file (y/n)
UseMetaLocationsFile=n
;meta locations file
MetaLocationsFilename=
;multiple coordinates type (0=OnePerLocation, 1=AtLeastOneLocation, 2=AllLocations)
MultipleCoordinatesType=0

[Spatial Window]
;maximum spatial size in population at risk (<=50%)
MaxSpatialSizeInPopulationAtRisk=50
;restrict maximum spatial size - max circle file? (y/n)
UseMaxCirclePopulationFileOption=n
;maximum spatial size in max circle population file (<=50%)
MaxSpatialSizeInMaxCirclePopulationFile=50
;maximum circle size filename
MaxCirclePopulationFile=
;restrict maximum spatial size - distance? (y/n)
UseDistanceFromCenterOption=y
;maximum spatial size in distance from center (positive integer) - R file sends a string to
overwrite this
MaxSpatialSizeInDistanceFromCenter=100
;include purely temporal clusters? (y/n)
IncludePurelyTemporal=n
;window shape (0=Circular, 1=Elliptic)
SpatialWindowShapeType=0
;elliptic non-compactness penalty (0=NoPenalty, 1=MediumPenalty, 2=StrongPenalty)
NonCompactnessPenalty=0
;isotonic scan (0=Standard, 1=Monotone)
IsotonicScan=0

[Temporal Window]
;how max temporal size should be interpretted (0=Percentage, 1=Time)
MaxTemporalSizeInterpretation=1
;maximum temporal cluster size (<=90%) - R file sends a string to overwrite this
MaxTemporalSize=30
;include purely spatial clusters? (y/n)
IncludePurelySpatial=n
;temporal clusters evaluated (0=All, 1=Alive, 2=Flexible Window)
IncludeClusters=0
;flexible temporal window start range (YYYY/MM/DD,YYYY/MM/DD)
IntervalStartRange=2000/1/1,2000/12/31
;flexible temporal window end range (YYYY/MM/DD,YYYY/MM/DD)
IntervalEndRange=2000/1/1,2000/12/31

[Space and Time Adjustments]
;time trend adjustment type (0=None, 1=Nonparametric, 2=LogLinearPercentage,
3=CalculatedLogLinearPercentage, 4=TimeStratifiedRandomization,
5=CalculatedQuadraticPercentage)
TimeTrendAdjustmentType=4
;time trend adjustment percentage (>-100)
TimeTrendPercentage=14.8
;time trend type - SVTT only (Linear=0, Quadratic=1)
TimeTrendType=0
;adjust for weekly trends, nonparametric
AdjustForWeeklyTrends=n
;spatial adjustments type (0=No Spatial Adjustment, 1=Spatially Stratified Randomization)
SpatialAdjustmentType=0
;use adjustments by known relative risks file? (y/n)
UseAdjustmentsByRRFile=n
;adjustments by known relative risks file name (with HA Randomization=1)
AdjustmentsByKnownRelativeRisksFilename=

[Inference]
;p-value reporting type (Default p-value=0, Standard Monte Carlo=1, Early Termination=2,
Gumbel p-value=3)
PValueReportType=0
;early termination threshold
EarlyTerminationThreshold=50
;report Gumbel p-values (y/n)
```

```
ReportGumbel=n
;Monte Carlo replications (0, 9, 999, n999)
MonteCarloReps=9999
;adjust for earlier analyses(prospective analyses only)? (y/n)
AdjustForEarlierAnalyses=y
;prospective surveillance start date (YYYY/MM/DD)
ProspectiveStartDate=2008/01/01
;perform iterative scans? (y/n)
IterativeScan=n
;maximum iterations for iterative scan (0-32000)
IterativeScanMaxIterations=10
;max p-value for iterative scan before cutoff (0.000-1.000)
IterativeScanMaxPValue=0.05

[Power Evaluation]
;perform power evaluation - Poisson only (y/n)
PerformPowerEvaluation=n
;power evaluation method (0=Analysis And Power Evaluation Together, 1=Only Power Evaluation
With Case File, 2=Only Power Evaluation With Defined Total Cases)
PowerEvaluationsMethod=0
;total cases in power evaluation
PowerEvaluationTotalCases=600
;critical value type (0=Monte Carlo, 1=Gumbel, 2=User Specified Values)
CriticalValueType=0
;power evaluation critical value .05 (> 0)
CriticalValue05=0
;power evaluation critical value .001 (> 0)
CriticalValue01=0
;power evaluation critical value .001 (> 0)
CriticalValue001=0
;power estimation type (0=Monte Carlo, 1=Gumbel)
PowerEstimationType=0
;number of replications in power step
NumberPowerReplications=1000
;power evaluation alternative hypothesis filename
AlternativeHypothesisFilename=
;power evaluation simulation method for power step (0=Null Randomization, 1=N/A, 2=File
Import)
PowerEvaluationsSimulationMethod=0
;power evaluation simulation data source filename
PowerEvaluationsSimulationSourceFilename=
;report power evaluation randomization data from power step (y/n)
ReportPowerEvaluationSimulationData=n
;power evaluation simulation data output filename
PowerEvaluationsSimulationOutputFilename=

[Spatial Output]
;automatically launch Google Earth - gui only (y/n)
LaunchKMLViewer=n
;create compressed KMZ file instead of KML file (y/n)
CompressKMLtoKMZ=n
;whether to include cluster locations kml output (y/n)
IncludeClusterLocationsKML=y
;threshold for generating separate kml files for cluster locations (positive integer)
ThresholdLocationsSeparateKML=1000
;report hierarchical clusters (y/n)
ReportHierarchicalClusters=y
;criteria for reporting secondary clusters(0=NoGeoOverlap, 1=NoCentersInOther,
2=NoCentersInMostLikely,  3=NoCentersInLessLikely, 4=NoPairsCentersEachOther,
5=NoRestrictions)
CriteriaForReportingSecondaryClusters=0
;report gini clusters (y/n)
ReportGiniClusters=n
;gini index cluster reporting type (0=optimal index only, 1=all values)
GiniIndexClusterReportingType=0
;spatial window maxima stops (comma separated decimal values[<=50%] )
SpatialMaxima=1,2,3,4,5,6,8,10,12,15,20,25,30,40,50
;max p-value for clusters used in calculation of index based coefficients (0.000-1.000)
GiniIndexClustersPValueCutOff=0.05
;report gini index coefficents to results file (y/n)
ReportGiniIndexCoefficents=n
;restrict reported clusters to maximum geographical cluster size? (y/n)
UseReportOnlySmallerClusters=n
;maximum reported spatial size in population at risk (<=50%)
MaxSpatialSizeInPopulationAtRisk Reported=50
;restrict maximum reported spatial size - max circle file? (y/n)
UseMaxCirclePopulationFileOption_Reported=n
```

```
;maximum reported spatial size in max circle population file (<=50%)
MaxSizeInMaxCirclePopulationFile Reported=50
;restrict maximum reported spatial size - distance? (y/n)
UseDistanceFromCenterOption Reported=n
;maximum reported spatial size in distance from center (positive integer)
MaxSpatialSizeInDistanceFromCenter_Reported=1

[Temporal Output]
;output temporal graph HTML file (y/n)
OutputTemporalGraphHTML=n

[Other Output]
;report critical values for .01 and .05? (y/n)
CriticalValue=y
;report cluster rank (y/n)
ReportClusterRank=n
;print ascii headers in output files (y/n)
PrintAsciiColumnHeaders=n

[Elliptic Scan]
;elliptic shapes - one value for each ellipse (comma separated decimal values)
EllipseShapes=1.5,2,3,4,5
;elliptic angles - one value for each ellipse (comma separated integer values)
EllipseAngles=4,6,9,12,15

[Power Simulations]
;simulation methods (0=Null Randomization, 1=N/A, 2=File Import)
SimulatedDataMethodType=0
;simulation data input file name (with File Import=2)
SimulatedDataInputFilename=
;print simulation data to file? (y/n)
PrintSimulatedDataToFile=n
;simulation data output filename
SimulatedDataOutputFilename=

[Run Options]
;number of parallel processes to execute (0=All Processors, x=At Most X Processors)
NumberParallelProcesses=0
;suppressing warnings? (y/n)
SuppressWarnings=n
;log analysis run to history file? (y/n)
LogRunToHistoryFile=n
;analysis execution method  (0=Automatic, 1=Successively, 2=Centrically)
ExecutionType=0

[System]
;system setting - do not modify
Version=9.3.0
```