Supporting Information for

(Re)Emerging Disease and Conflict Risk in Africa, 2000 – 2018

Contents

Overview of G-ZOD

This section, discusses in detail the construction of the Geolocated Zoonotic Disease Database (G-ZOD). For assisting with the coding procedures, two highly qualified research assistants were selectively recruited and thoroughly trained. The research assistants followed a detailed codebook (provided as a separate online attachment) in deciding what incidents to code. In choosing the pathogens for coding, several parameters were emphasized. First, examining reports from the World Health Organization, we wanted to ensure that the pathogens included were those that have been especially relevant within the African context, such as Ebola, Lassa, and Yellow Fever. We also wanted to choose some pathogens that are likely to be carried due to animal migration from other world regions and can hence induce a disruption, such as dangerous flu pathogens carried from East and Central Asia. Accordingly, and as listed in the main report, the final list of 23 pathogens included the following diseases: Ebola-Zaire, Ebola-Sudan, Ebola-Täi forest, Ebola-Bundibugyo, Marburg, yellow fever, Rift Valley fever, West Nile fever, H1N1 flu, H5N1 flu, SARS, MERS, chikungunya, Lassa, dengue, monkeypox, septicemic plague, bubonic plague, Crimean-Congo hemorrhagic fever, shigellosis, rabies, zika, and anthrax.

The result is a database that, in addition to its specific advantages for conflict analysis (discussed in the main paper) offers several additional new features compared with existing datasets, such as WHO, EM-DAT, Torres et al. (2022) [1], and ISID. First, G-ZOD is *geolocated by design and construction*. All events included in analysis were recorded at the first administrative unit (province) levels or below, although for those interested in country level analysis, G-ZOD also contains country level data that were not used for analysis in the paper. In contrast, all other aforementioned datasets are country level, and rarely allow for local level assessment, despite the crucial need for such assessment in the role of environmental indicators' impact on armed conflict [2].

Second, similarly to widely used conflict datasets, G-ZOD is an *event-based datasets*, which includes – in addition to location – information on the month and year (or exact date, if available) of the event, and provide for a more data driven operationalization of disease outbreak risk and frequency. This in contrast to all other aforementioned datasets, which other used a high (100 or more) casualty thresholds, or really on qualitative definitions (e.g., the WHO defines an outbreak as being an epidemic) in coding data. G-ZOD theretofore codes not only extreme epidemic and pandemic events, but also a higher resolution of such incidents as well as outbreak events that ultimately did not – luckily – deteriorate into an epidemic, even though *they posed the risk for becoming one*. Third, in addition to the specific pathogen, we coded – where available – additional relevant information for each outbreak events, including the number of people affected and dead and the direct cause of the infection.

To illustrate the process used to code all relevant zoonotic outbreak events, the following procedure is provided. As mentioned in the main article, the coding effort focused only on (i) reputable English language sources that (ii) were available in LexisNexis over the entire period, including NYT, Associated Press/AP, Reuters, CNN, and CBS as well as situation reports from leading health agencies (CDC, WHO).[1] We used local English language dailies as well, but made sure to focus on reputable ones and made sure – to the best of our ability – to verify that the same incident was confirmed via another reputable source (e.g., AP).

To illustrate how the process of coding worked, I selected two events at random from my database. To this end, Figure S1 shows the relevant information (highlighted) collected from a story on an Ebola outbreak. Figure S2 then reports a similar report for coding an H5N1 flu outbreak case. Each outbreak case is assigned a unique ID to reduce the risk of the same outbreak being recorded twice.

[1]For this reason, no information from the *Washington Post* was included – it was removed from Lexis-Nexis, and so reporting in this case was imbalanced.

41 deaths from Ebola in Congo; health officials 'worried'

CNN Wire

August 14, 2018 Tuesday 4:04 PM GMT

Length: 834 words

Byline: By Jacqueline Howard and Debra Goldschmidt, CNN

Dateline: (CNN)

Body

(CNN) -- An ongoing ***Ebola*** outbreak is spreading across the ***Democratic Republic of Congo***'s North Kivu province, and it has health officials gravely concerned due to conflict in the area along with other factors.

The country's Ministry of Health announced Monday that 57 cases of hemorrhagic fever have been reported in the region: 30 confirmed and 27 probable.

The outbreak has left 41 people dead, according to the World Health Organization.

WHO Director-General Dr. Tedros Adhanom Ghebreyesus visited the city of Beni and the Mangina area in North Kivu. Mangina, which is about 30 kilometers or 19 miles from Beni, sits at the epicenter of the outbreak.

"After the visit, I am actually more worried," Tedros said at a news conference Tuesday.

"What makes actually the outbreak in eastern DRC or northern Kivu more dangerous is, there is a security challenge. There is active conflict in that area. Since January, there were 120 violent incidences," he said.

Thousands in the region have been displaced by fighting between rebel groups and government forces.

"That environment is really conducive for ***Ebola*** actually to transmit freely, because in that area, there are places called red zones, inaccessible areas, because there are many armed groups that operate in that region or in that province," he said.

In other words, those areas remain difficult to access, and people may have difficulty getting health care and support. Also, he added, many cases are among women and children.

"There are other factors that complicate the issue. One is high population density," Tedros said.

North Kivu is among Congo's most populated provinces, with 8 million inhabitants, according to WHO and the United Nations. Its capital is the city of Goma. The region has been experiencing intense insecurity and a worsening humanitarian crisis, with over 1 million internally displaced people and a continuous outflow of refugees to neighboring countries, including Uganda, Burundi and Tanzania.

"The other worrisome problem, especially in Mangina, is the number of health workers that are already effected with ***Ebola***, and we have, so far, seven cases," Tedros said.

This new outbreak is the second of the summer for Congo.

Figure S1: Example of a story underlying the coded case (Ebola)

Bird flu strain found in Ghana may have come from Ivory Coast

BBC Monitoring Africa - Political

Supplied by BBC Worldwide Monitoring

May 14, 2007 Monday

Length: 602 words

Body

Text of report by Ghanaian independent Joy FM radio on 14 May

[Presenter] Joy news has learnt that a strain of bird flu discovered in the country last week is believed to have come from Ivory Coast.

The reports are said to have been conclusions from analysis performed on the virus found in a farm in Tema by the World Organization for Animal Health. According to the source, analysis found the strains similar to when discovered in Ivory Coast. Meanwhile the results of the analysis are said to have been confirmed by other reports linking a poultry farm here in ***Ghana*** with another farm in the Ivory Coast.

Dr Kiki Moto [phonetic] is a veterinarian with the veterinary services department. He joins us on the line with some more details.

Thank you very much for your time sir. Can you confirm that indeed the strain of bird flu found in the farm in Tema last week is from the Ivory Coast?

[Dr Moto] Excuse me I think that we might be going ahead to make a hasty conclusion. All that we know is that, there is a (?phalogenetic tendogram) which normally has the HH sequence of the ***H5N1*** isolate that we have got in ***Ghana***, and this one has been confirmed by Italy and they have stated that this isolate has 99.6 per cent of amino acid sequence which is related to the isolate from Ivory Coast they are not saying that the thing has come from Ivory Coast, but this is a sub-Saharan virus which has similarity in the amino acid sequence of 99.6 per cent.

[Presenter] In other words the virus found in ***Ghana*** is about 99.6 per cent similar to what was found in Ivory Coast

[Dr Moto] You are very correct.

[Presenter]: So which other country in the subregion has that similarity?

[Dr Moto] : If you look at the report from Cordevole Italy, the sub-Saharan African isolate all fall within this and they range from 98.8 to 99.6. So it is just a range, so all the isolates from the sub-Saharan African fall within that range.

[Presenter] So in other words it could come from the Ivory Coast but there are other countries in the sub region that it could come from.

[Dr Moto] It is possible. While we are thinking the likelihood of it being nearer to that of Ivory Coast because that one is the highest that 99.6 per cent as compared to the one which was isolated in Burkina, Nigeria, Niger they are a little bit far they are about 98.8 per cent.

[Presenter] Sir would you be conducting further tests to attempt to find out were exactly it would be coming from apart from the similarities in the strain?

Figure S2: Example of a story underlying the coded case (H5N1)

Summary Statistics

Table S1: Summary Statistics

	Min.	**Median**	**Mean**	**Max.**	**Std. Dev.**
Dependent variables					
State conflict$_{it}$	0	0	0.008	75	0.258
Rebel conflict$_{it}$	0	0	0.003	37	0.130
Pol. militia conflict $_{it}$	0	0	0.004	21	0.111
Id. militia conflict $_{it}$	0	0	0.002	19	0.076
Independent variables					
ZDO events$_{it}$	0	0	0.001	14	0.034
NTL$_{it}$ [1]	0	0	1.404	5.344	1.673
Population$_{it}$ [1]	0	4.211	3.913	7.267	1.295
Prec. anom$_{it}$	-4.032	-0.025	0.054	5.295	0.837
Drought$_{it}$	-9.004	-0.195	-0.273	7.887	0.998
Temp. anom$_{it}$	-4.397	0.476	0.468	5.295	0.904
Life exp.$_{it}$ [1]	4.727	7.093	7.206	10.041	1.072
Gov. eff.$_{it}$	0	0.857	1.977	50.102	3.045
GDPpc$_{it}$ [1]	3.747	4.110	4.113	4.355	0.130
State conflict$_{jt}$	0	0	0.007	1	0.085
Rebel conflict$_{jt}$	0	0	0.003	1	0.057
Pol. mil. conflict$_{jt}$	0	0	0.004	1	0.064
Id. mil. conflict$_{jt}$	0	0	0.003	1	0.054

i refers to a given grid cell, j refers to any adjacent grid cells, and t refers to month in a given year

[1] Natural log

Sensitivity Analyses

Table S2: Determinants of Armed Conflict in African Locations – Minimalist Specifications

	Civil War		*Social Conflict*	
	State	**Rebel**	**Pol. Mil.**	**Id. Mil.**
	(17)	(18)	(19)	(20)
ZDO events$_{it}$	-0.015^{**}	-0.002	0.018	0.006^{*}
	(0.006)	(0.006)	(0.013)	(0.003)
Observations	2,561,760			
R^2	0.197	0.152	0.157	0.035
Adjusted R^2	0.194	0.149	0.154	0.031

Standard errors clustered on grid cell in parentheses; fixed effects by grid cell were included in each regression, but not reported here.

*** p<0.01, ** p<0.05, * p<0.1

[1] Natural log.

The removal of controls allowed ZDO event data to be included for the entire Jan. 2000 – Dec. 2019 period.

Table S3: Determinants of Armed Conflict in African Locations – Pestilence

	Civil War		*Social Conflict*	
	State	**Rebel**	**Pol. Mil.**	**Id. Mil.**
	(21)	(22)	(23)	(24)
ZDO events$_{it}$	-0.013^{**}	0.002	0.010	0.012^{**}
	(0.006)	(0.006)	(0.014)	(0.005)
NTL$_{it}$[1]	0.0001	0.0001	0.0004^{**}	0.0002^{**}
	(0.0003)	(0.0001)	(0.0002)	(0.0001)
Population$_{it}$[1]	-0.001	-0.0003	0.003^{*}	-0.0003
	(0.003)	(0.001)	(0.001)	(0.001)
DV_{it-1}	0.453^{***}	0.323^{***}	0.305^{***}	0.092^{***}
	(0.090)	(0.104)	(0.053)	(0.016)
Prec. anom$_{it}$	0.0001	-0.0002^{**}	-0.0002^{*}	-0.0002^{**}
	(0.0002)	(0.0001)	(0.0001)	(0.0001)
Prec. anom$_{it}$ × *ZDO events*$_{it}$	-0.001	-0.003	0.001	-0.015^{**}
	(0.003)	(0.003)	(0.009)	(0.007)
Drought$_{it}$	-0.0001	0.0003^{***}	0.0001	0.0002^{**}
	(0.0002)	(0.0001)	(0.0001)	(0.0001)
Temp. anom$_{it}$	-0.0003	0.00002	-0.0001	-0.00001
	(0.0003)	(0.0001)	(0.0001)	(0.0001)
Life exp.$_{it}$[1]	-0.004^{**}	-0.001^{***}	-0.001	-0.0004
	(0.002)	(0.0003)	(0.001)	(0.0005)
Gov. eff.$_{it}$	-0.001	0.001^{***}	0.0002	-0.0002
	(0.0004)	(0.0004)	(0.0002)	(0.0001)
GDPpc$_{it}$[1]	-0.061^{***}	-0.011^{***}	-0.016^{***}	-0.026^{***}
	(0.009)	(0.004)	(0.005)	(0.004)
τ_t	0.0001^{***}	0.00002^{***}	0.00003^{***}	0.00004^{***}
	(0.00002)	(0.00000)	(0.00001)	(0.00001)
Observations	1,473,765			
R^2	0.641	0.152	0.166	0.067
Adjusted R^2	0.639	0.146	0.160	0.061

Standard errors clustered on grid cell in parentheses; fixed effects by grid cell and month were included in each regression, but not reported here.

*** p<0.01, ** p<0.05, * p<0.1

[1] Natural log

Table S4: Determinants of Armed Conflict in African Locations – Conflict by other Actors

	Civil War		*Social Conflict*	
	State	**Rebel**	**Pol. Mil.**	**Id. Mil.**
	(25)	(26)	(27)	(28)
ZDO events$_{it}$	-0.014^{**}	0.001	0.009	0.011^{**}
	(0.006)	(0.005)	(0.013)	(0.005)
NTL$_{it}$[1]	-0.00000	0.0001	0.0004^{**}	0.0002^{**}
	(0.0003)	(0.0001)	(0.0002)	(0.0001)
Population$_{it}$[1]	-0.002	-0.001	0.003^{**}	-0.0004
	(0.003)	(0.001)	(0.001)	(0.001)
DV_{it-1}	0.448^{***}	0.317^{***}	0.297^{***}	0.090^{***}
	(0.087)	(0.094)	(0.052)	(0.015)
Prec. anom$_{it}$	0.0001	-0.0002^{**}	-0.0002	-0.0002^{**}
	(0.0002)	(0.0001)	(0.0001)	(0.0001)
Drought$_{it}$	-0.0001	0.0003^{***}	0.0001	0.0002^{**}
	(0.0002)	(0.0001)	(0.0001)	(0.0001)
Temp. anom$_{it}$	-0.0003	0.00002	-0.0001	-0.00000
	(0.0003)	(0.0001)	(0.0001)	(0.0001)
Life exp.$_{it}$[1]	-0.004^{**}	-0.001^{**}	-0.001	-0.0003
	(0.002)	(0.0003)	(0.001)	(0.0005)
Gov. eff.$_{it}$	-0.0004^{**}	0.001^{***}	0.0002^{*}	-0.0002
	(0.0002)	(0.0004)	(0.0001)	(0.0001)
GDPpc$_{it}$[1]	-0.060^{***}	-0.013^{***}	-0.011^{**}	-0.025^{***}
	(0.008)	(0.004)	(0.004)	(0.004)
τ_t	0.0001^{***}	0.00002^{***}	0.00003^{***}	0.00004^{***}
	(0.00001)	(0.00000)	(0.00001)	(0.00001)
State conflict$_{it}$		-0.019	0.023^{*}	0.008^{***}
		(0.021)	(0.012)	(0.003)
Rebel conflict$_{it}$	-0.126		0.073^{**}	-0.007
	(0.137)		(0.035)	(0.004)
Pol. mil. conflict$_{it}$	0.100^{*}	0.048^{*}		0.027^{***}
	(0.058)	(0.027)		(0.009)
Id. mil. conflict$_{it}$	0.054^{***}	-0.006	0.039^{***}	
	(0.018)	(0.005)	(0.015)	
Observations	1,473,765			
R^2	0.644	0.157	0.172	0.069
Adjusted R^2	0.641	0.151	0.166	0.063

Standard errors clustered on grid cell in parentheses; fixed effects by grid cell and month were included in each regression, but not reported here.

*** p<0.01, ** p<0.05, * p<0.1

[1] Natural log

Table S5: Determinants of Armed Conflict in African Locations – Conflict by other Actors with Conflict Lags

	Civil War		*Social Conflict*	
	State	**Rebel**	**Pol. Mil.**	**Id. Mil.**
	(29)	(30)	(31)	(32)
ZDO events$_{it}$	-0.014^{**}	0.001	0.009	0.011^{**}
	(0.006)	(0.005)	(0.012)	(0.005)
NTL$_{it}$[1]	-0.00004	0.0001	0.0004^{**}	0.0002^{*}
	(0.0003)	(0.0001)	(0.0002)	(0.0001)
Population$_{it}$[1]	-0.002	-0.001	0.003^{**}	-0.0004
	(0.003)	(0.001)	(0.001)	(0.001)
Prec. anom$_{it}$	0.0001	-0.0002^{**}	-0.0002	-0.0002^{**}
	(0.0002)	(0.0001)	(0.0001)	(0.0001)
Drought$_{it}$	-0.0001	0.0003^{***}	0.0001	0.0002^{**}
	(0.0002)	(0.0001)	(0.0001)	(0.0001)
Temp. anom$_{it}$	-0.0002	0.00002	-0.0001	0.00000
	(0.0003)	(0.0001)	(0.0001)	(0.0001)
Life exp.$_{it}$[1]	-0.004^{**}	-0.001^{**}	-0.001	-0.0003
	(0.001)	(0.0003)	(0.001)	(0.0005)
Gov. eff.$_{it}$	-0.0003^{*}	0.001^{***}	0.0001^{*}	-0.0002
	(0.0002)	(0.0004)	(0.0001)	(0.0001)
GDPpc$_{it}$[1]	-0.060^{***}	-0.013^{***}	-0.009^{**}	-0.025^{***}
	(0.008)	(0.004)	(0.004)	(0.004)
τ_t	0.0001^{***}	0.00002^{***}	0.00002^{***}	0.00004^{***}
	(0.00001)	(0.00000)	(0.00001)	(0.00001)
State conflict$_{it}$		-0.019	0.018^{*}	0.008^{**}
		(0.019)	(0.010)	(0.003)
Rebel conflict$_{it}$	-0.108		0.063^{**}	-0.006
	(0.113)		(0.028)	(0.004)
Pol. mil. conflict$_{it}$	0.069^{**}	0.043^{*}		0.024^{***}
	(0.033)	(0.025)		(0.008)
Id. mil. conflict$_{it}$	0.050^{***}	-0.006	0.037^{**}	
	(0.018)	(0.005)	(0.015)	
State conflict$_{it-1}$	0.444^{***}	-0.0003	0.012	-0.002
	(0.083)	(0.006)	(0.008)	(0.002)
Rebel conflict$_{it-1}$	-0.074	0.315^{***}	0.036^{**}	-0.005^{*}
	(0.067)	(0.093)	(0.017)	(0.003)
Pol. mil. conflict$_{it-1}$	0.112	0.019^{*}	0.293^{***}	0.013^{*}
	(0.097)	(0.010)	(0.052)	(0.007)
Id. mil. conflict$_{it-1}$	0.010	-0.008	0.022^{*}	0.090^{***}
	(0.017)	(0.006)	(0.012)	(0.015)
Observations	1,473,765			
R^2	0.645	0.157	0.173	0.069
Adjusted R^2	0.643	0.151	0.167	0.063

Standard errors clustered on grid cell in parentheses; fixed effects by grid cell and month were included in each regression, but not reported here.

*** p<0.01, ** p<0.05, * p<0.1

[1] Natural log

Table S6: Determinants of Armed Conflict in African Locations – Spatial Conflict Indicator Lags

	Civil War		Social Conflict	
	State	**Rebel**	**Pol. Mil.**	**Id. Mil.**
	(33)	(34)	(35)	(36)
ZDO events$_{it}$	-0.013^{**}	0.002	0.010	0.011^{**}
	(0.006)	(0.006)	(0.013)	(0.005)
NTL$_{it}$[1]	0.0001	0.0001	0.0004^{***}	0.0002^{**}
	(0.0003)	(0.0001)	(0.0002)	(0.0001)
Population$_{it}$[1]	-0.001	-0.0003	0.003^{*}	-0.0004
	(0.003)	(0.001)	(0.002)	(0.001)
DV_{it-1}	0.464^{***}	0.314^{**}	0.307^{***}	0.058^{***}
	(0.098)	(0.136)	(0.064)	(0.018)
Prec. anom$_{it}$	0.0001	-0.0002^{**}	-0.0002^{*}	-0.0002^{**}
	(0.0002)	(0.0001)	(0.0001)	(0.0001)
Drought$_{it}$	-0.0001	0.0003^{***}	0.0001	0.0002^{**}
	(0.0002)	(0.0001)	(0.0001)	(0.0001)
Temp. anom$_{it}$	-0.0003	0.00001	-0.0001	-0.00002
	(0.0003)	(0.0001)	(0.0001)	(0.0001)
Life exp.$_{it}$[1]	-0.004^{**}	-0.001^{***}	-0.001	-0.0004
	(0.002)	(0.0003)	(0.001)	(0.0005)
Gov. eff.$_{it}$	-0.001	0.001^{***}	0.0002	-0.0002
	(0.0004)	(0.0004)	(0.0002)	(0.0001)
GDPpc$_{it}$[1]	-0.066^{***}	-0.011^{***}	-0.016^{***}	-0.024^{***}
	(0.008)	(0.004)	(0.005)	(0.004)
τ_t	0.0001^{***}	0.00002^{***}	0.00003^{***}	0.00004^{***}
	(0.00002)	(0.00000)	(0.00001)	(0.00001)
State conflict$_{jt}$	-0.068			
	(0.072)			
Rebel conflict$_{jt}$		0.026		
		(0.094)		
Pol. mil. conflict$_{jt}$			-0.007	
			(0.040)	
Id. mil. conflict$_{jt}$				0.086^{***}
				(0.013)
Observations	1,473,765			
R^2	0.642	0.152	0.166	0.070
Adjusted R^2	0.639	0.146	0.160	0.064

Standard errors clustered on grid cell in parentheses; fixed effects by grid cell and month were included in each regression, but not reported here.

*** p<0.01, ** p<0.05, * p<0.1

[1] Natural log

References

[1] Juan Armando Torres Munguía, Florina Cristina Badarau, Luis Rodrigo Díaz Pavez, Inmaculada Martínez-Zarzoso, and Konstantin M Wacker. A global dataset of pandemic-and epidemic-prone disease outbreaks. *Scientific data*, 9(1):683, 2022.

[2] Ole Magnus Theisen, Nils Petter Gleditsch, and Halvard Buhaug. Is climate change a driver of armed conflict? *Climatic change*, 117(3):613–625, 2013.