The RAPIDD Ebola forecasting challenge: Model description and synthetic data generation

Supplementary Information

Marco Ajelli[a,b], Qian Zhang[a], Kaiyuan Sun[a], Stefano Merler[b], Laura Fumanelli[b], Gerardo Chowell[c,d], Lone Simonsen[d,e], Cecile Viboud[d], and Alessandro Vespignani[a,f]

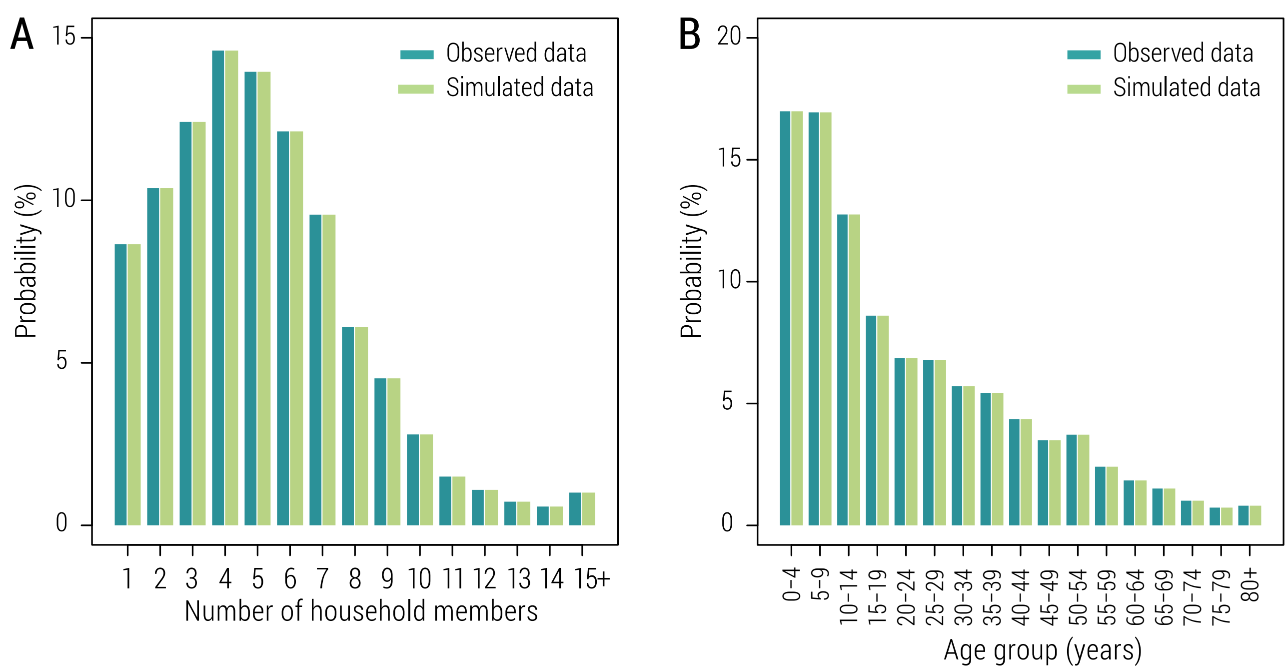
[a]Laboratory for the Modeling of Biological and Socio-technical Systems, Northeastern University, Boston, USA
[b]Bruno Kessler Foundation (FBK), Trento, Italy
[c]School of Public Health, Georgia State University, Atlanta, USA
[d]Fogarty International Center, National Institutes of Health, Bethesda, USA
[e]Department of Public Health, University of Copenhagen, Denmark
[f]Institute for Scientific Interchange Foundation, Turin, Italy

September 5, 2017

1 Synthetic population

In order to construct the synthetic population of each county, we generated villages of about 400 inhabitants on average, randomly located over the territory, until the total population of the county was obtained. For each county we also placed the corresponding capital in the exact location given by GPS coordinates and with the exact number of inhabitants as obtained from census data. We reconstructed population density from county-level census data. Alternative choices, such as using the Gridded Population of the World [1] or LandScan [2], would have favored challenge participants using the same dataset while penalizing groups using a different source. HCW were randomly chosen among all individuals aged at least 18 yrs and living within a radius of 10 km from that specific hospital. During the 2014-15 epidemic in West Africa, health care facilities specifically dedicated to Ebola patients were gradually opened in some counties; we consider these Ebola treatment units (ETUs) in our model as detailed in Sections **??**. Simulated individuals were grouped into households in order to match demographic information derived from the 2007 Demographic Health Surveys [3] on household size and demographics for Liberia (see Fig. S1A,B). Households were then grouped into clusters of six households each – i.e., every household was linked to five other households to mimic the structure of extended family that is typical of West Africa. We assumed that the maximum distance between households in the same cluster of extended households could not be in villages more than 50 km apart, and that on average 50% of these households belonged to the same village. Note that a household could belong to several extended families.

Figure S1: **A** Predicted and observed household size distributions. **B** Predicted and observed age distributions of the synthetic population.



2 Transmission by setting

At every time t of the simulation (time step $\Delta t = 1$ day), every susceptible individual i has a probability $p_i(t;z) = \delta(a_i)\left[1 - \exp\left(-\Delta t \sum_j \lambda_j(t;z)\right)\right]$ of getting infected from each infectious individual j in the population, where a_i is the age of individual i, $\delta(a_i)$ is the age-dependent risk of infection, and $\lambda_j(t;z)$ is the force of infection in setting z, where $z \in \{\text{household, extended family, hospital, ETU, funeral}\}$.

2.1 Transmission in household

At time t, a non-hospitalized infectious individual j is able to transmit EVD to all other members of his/her household with the following force of infection:

$$\lambda_j(t;\text{household}) = r(t)\frac{\nu_j \beta_f}{N_{f_j}(t)}$$

where β_f is the transmission rate in households (the same for all households), $N_{f_j}(t)$ is the household size at time t (thus excluding deceased and hospitalized members), ν_j is a scalar factor accounting for the heterogeneity in infectiousness among individuals and $r(t)$ accounts for behavioral changes.

2.2 Transmission in the extended family

Let α_j be the set of additional households (i.e., belonging to the extended family) for an infectious (non-hospitalized) individual j. Individual j transmits infection to the members of α_j with force of infection

$$\lambda_j(t;\text{extended family}) = r(t)\frac{\nu_j \sigma \beta_f}{N_{\alpha_j}(t)}$$

where $N_{\alpha_j}(t)$ is number of individuals in α_j at time t, $0 \leq \sigma \leq 1$ is the reduction of the transmission rate in the extended family.

2.3 Transmission during burial ceremonies

In case of unsafe burial, deceased individual j transmits EVD to his/her household members and members of the extended family α_j similarly to transmission in households, namely:

$$\lambda_j(t;\text{funeral}) = r(t)\frac{\nu_j \beta_b}{N_{f_j}(t)}$$

to household members and

$$\lambda_j(t;\text{funeral}) = r(t)\frac{\nu_j \sigma \beta_b}{N_{\alpha_j}(t)}$$

to members of the additional households, where β_b is the transmission rate in unsafe burials. Note that the same set of households involved in the extended family transmission is assumed also for burial ceremonies.

2.4 Transmission in hospital

An infectious individual j, admitted to a general hospital, transmits the infection to both susceptible hospitalized individuals and to health care workers with force of infection

$$\lambda_j(t;\text{hospital}) = r(t)\frac{\nu_j \beta_h}{N_{h_j}(t)}$$

where β_h is the transmission rate in hospital, and $N_{h_j}(t)$ is the overall number of hospitalized individuals and HCW in hospital h_j.

2.5 Transmission in ETU

An infectious individual j, admitted to an ETU, transmits the infection only to HCW, with force of infection equal to

$$\lambda_j(t; \mathrm{ETU}) = r(t)\frac{\nu_j k \beta_h}{N_{hcw_j}(t)}$$

where $N_{hcw_j}(t)$ is the overall number of HCWs in ETU h_j, and k is a scaling factor accounting for reduced transmission in ETUs with respect to hospitals.

3 Parameters values used in the 4 synthetic scenarios

All transmission parameters and intervention strategies specifications used in the four scenarios are listed in Table S1 and Table S2 . Figure S2 describes the evolution of the implemented interventions over time. In each scenario, contact tracing, safe burials in the community, and behavioral changes are implemented after the opening of the first ETU since this event usually triggers other interventions. Scenarios 1 and 2 are characterized by greater variability while the pattern of interventions in scenarios 3 and 4 is more regular but characterized by abrupt changes. In particular, while the opening of ETUs follows the spatiotemporal evolution of the selected epidemics (Figure S2 A), the number of contacts traced daily and the probability of safe burial in the community are rather variable in the first two scenarios and constant (after the initial period when they are both zero) in the remaining two (Figure S2 B,C). Figure S2 D depicts the behavioral change effect over time, which is marked in scenario 1, mild in scenario 2, and null in scenarios 3 and 4.

4 ETUs locations and size

In each scenario Ebola Treatment Units, each with a specific number of beds and HCWs, have been opened at different location and time. We report in Table S2 the information concerning Ebola Treatment Units in each scenario.

5 Structure of the patient record database

The patient record database included the following fields for each patient:

- Patient ID: Patient id number (patient A).
- Infector ID: Patient who infected patient A.
- Infection Time: Date of infection.
- Symptom Time: Date of symptoms onset.
- Death Time: Date of death (NA if the individual survived).
- Burial Time: Date of burial (NA if the individual survived).
- Age Group: Age of patient, three classes available; 0-14 yrs; 15-44yrs ; $\geq$45 yrs.
- HCW: Health Care Worker status (1 true; 0 false).
- County: County of residence of the patient
- Hospital Admission Time: Date of admission to the hospital/holding center (otherwise NA if patient never went to a hospital/holding center; N.B. an individual could have NA also because the patient went straight to an ETU).

- ETU Admission Time: Date of admission to ETU (otherwise NA if never admitted to an ETU. Note that an individual could be first admitted to a hospital/holding center and then to an ETU).
- Discharge Time: Date of discharge from the most recent health facility visited. This could be an ETU or a hospital/holding center (otherwise NA if never discharged because of death or never admitted to hospital/ETU)
- Contact Tracing:=1 if the individual was a contact traced from another case; 0 if not identified through contact tracing.
- Safe Burial: If individual had received a safe burial, NA otherwise.
- New traced contacts: Number of contacts traced from the patient

6 "Fog of war" noise in the patient record database

The records of all contact-traced patients were fully reported, but we randomly removed infection time and infector information for 80% of records without history of contact tracing. For the other records we randomly added a shift of between -1, 0 and +1 day to their infection time. In records from non-contact traced patients, symptom onset time, admission time and discharge time each had a 40% chance to be omitted (independent draws). If symptom time, admission time and discharge time were reported, then the record was adjusted with a uniform probability in the interval from -1 to +1 day with respect to the exact times. Cases not admitted to medical units and not traced were not reported in the patient database. Computationally the noise has been generated by individual level stochastic processes; i.e. each record or missing data was altered, or kept as in the original time output, according to the probability listed above.

7 Summary Reports of outbreak situations

In order to mimic a real world situation, for each scenario, we provide summary reports containing contextual information on case count, geographic location, planned and ongoing interventions, etc. Figure S3 shows a summary report provided to the participants.

References

[1] Columbia Univerity. Socioeconomic data and Applications center, 2015.

[2] Oak Ridge National Laboratory. LandScan, 2010.

[3] The DHS Program. Demographic and Health Survey, 2007.

Table S1: Model parameters used in the 4 synthetic scenarios.

Parameter	**Scenario 1**	**Scenario 2**	**Scenario 3**	**Scenario 4**
Seeding county	Grand Cape Mount	Grand Gedeh	Nimba	Montserrado
Transmission rate in household, β_f	0.63	0.48	0.42	0.45
Transmission rate in hospital/ETU, β_h	0.05	0.089	0.11	0.081
Transmission rate at unsafe burials, β_b	1.56	1.084	0.38	0.861
Scaling factor for transmission rate in the extended family, σ	0.42	0.51	0.45	0.47
Scaling factor for transmission in ETU with respect to transmission in hospital setting, k	0.05	0.05	0.05	0.05
Susceptibility to infection of children (0-14 years old) with respect to adults	0.24	0.18	0.3	0.2
Shape of gamma distribution for infectivity	0.277	0.391	0.4	1.235
Case fatality ratio	0.78	0.76	0.82	0.68
Hospitalization rate	0.84	0.75	0.75	0.72
Incubation period, days [gamma distributed; mean (sd)]	9.6 (2.4)	11 (2.3)	8.5 (2)	12.5 (4.2)
Time interval from symptom onset to admission, days [gamma distributed; mean (sd)]	4.3 (1.6)	4.4 (2.1)	4.8 (1.25)	5.7 (1.2)
Time interval from symptom onset to death for non-admitted individuals, days [gamma distributed; mean (sd)]	7.4 (2.7)	8.6 (1.6)	5.7 (3)	6.8 (1.7)
Time interval from symptom onset to recovery for non-admitted individuals, days [gamma distributed; mean (sd)]	7.7 (1.4)	9.4 (1.9)	7.8 (2)	7.4 (1.2)
Time interval from admission to death, days [gamma distributed; mean (sd)]	5.5 (2.1)	5.1 (2.7)	4.5 (2)	3.7 (2.1)
Time interval from admission to recovery, days [gamma distributed; mean (sd)]	5.6 (1.2)	5.6 (0.9)	5.5 (3)	2.5 (1.4)
Time interval from recovery to discharge, days [gamma distributed; mean (sd)]	9.3 (2.9)	8.9 (2.1)	7.5 (3.2)	10 (2.7)

Table S2: Opening dates, geographic locations, number of beds and number of HCW of ETUs in the four simulated scenarios.

Scenario	Opening day*	Beds	HCW	County
1	160	10	12	Montserrado
	191	25	42	Grand Gedeh
	205	25	35	Nimba
	214	40	70	Sinoe
	269	100	150	Montserrado
	282	75	85	Nimba
	293	30	25	Bomi
	297	80	90	Grand Gedeh
2	165	60	45	Grand Cape Mount
	180	55	60	Montserrado
	187	35	55	Bong
	204	20	30	Nimba
	218	30	35	Grand Bassa
	268	95	125	Montserrado
	275	50	65	Lofa
3	150	40	30	Nimba
	180	80	100	Montserrado
	189	20	25	Bomi
	194	30	20	Lofa
	200	40	50	Grand Gedeh
4	170	35	20	Montserrado
	200	25	30	Margibi
	215	40	60	Bomi
	235	110	130	Montserrado
	240	40	30	Nimba
	250	50	70	Lofa
	265	20	24	Gbarpolu
	278	90	105	Grand Gedeh
	291	50	71	Grand Bassa

* days since symptom onset in the index case of the epidemic

Figure S2: Intervention strategies over time for the four synthetic scenarios. **A** Cumulative number of beds in ETUs. **B** Average daily number of traced contacts for each index case. **C** Probability of safe burial in the community. **D** Reduction of transmission as a consequence of behavioral changes.

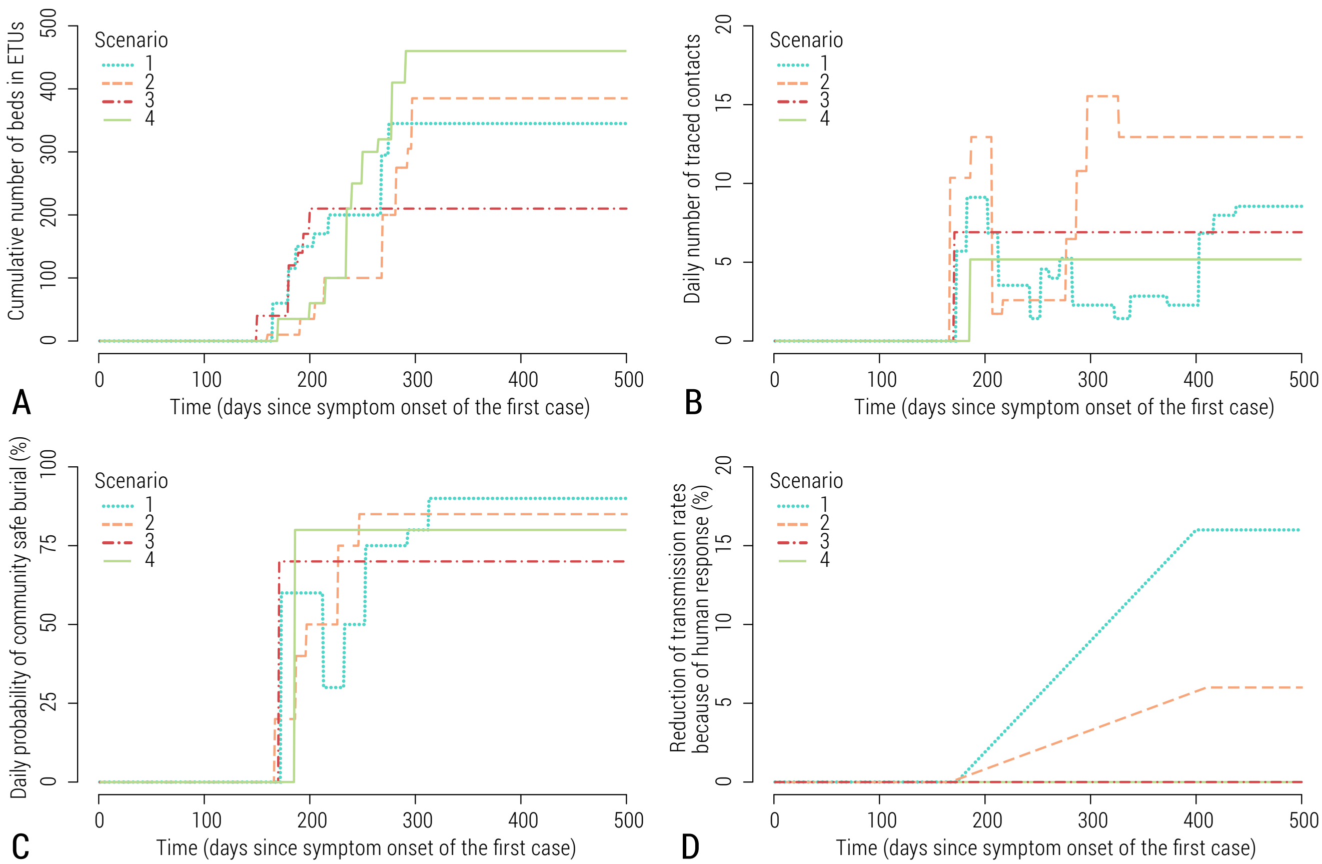


Figure S3: **Summary report of outbreak situation:** an example (scenario 1, week 13) of summary report provided to the participants.

EBOLA CHALLENGE

SUMMARY REPORT #3 (SCENARIO 1)

Date: week 26
Scenario 01

- A total number of 1842 EVD confirmed and probable cases was reported from week 20 till week 26 across Liberia, 553 cases were reported in Grand Cape Mount, 53 cases in Gbarpolu, 27 cases in Grand Bassa, 97 cases in Margibi, 81 cases in Nimba, 240 cases in Bomi, 63 case in Lofa, 1 case in Rivercess, 442 cases in Montserrado, 285 cases in Bong.
- A total number of 1341 deaths were reported from week 20 till week 26 across Liberia, among which 430 were in Grand Cape Mount, 39 in Gbarpolu, 30 in Grand Bassa, 79 in Margibi, 39 in Nimba, 178 in Bomi, 35 in Lofa, 1 in Rivercess, 295 in Montserrado, 215 in Bong.
- A total number of 81 health care workers(HCW) confirmed and probable cases were reported, 52 in Grand Cape Mount, 2 in Margibi, 17 in Montserrado, 8 in Bong. 53 of new deaths of HCWs were reported among the fatalities count from week 13 up to week 20.
- An effective safe burial protocol has started being enforced as of week 21.
- Community engagement and social mobilization programs as well as the distribution of home disinfection kits are starting in several parts of the country. It is not yet possible to estimate the effect of these interventions.
- A total number of 10292 individuals has been contact traced from week 20 till week 26. Among which 2830 were in Grand Cape Mount, 435 in Gbarpolu, 146 in Grand Bassa, 480 in Margibi, 604 in Nimba, 1125 in Bomi, 358 in Lofa, 2727 in Montserrado, 1587 in Bong.
- Here is a list of ETUs in operation:
    - An ETU at Grand Cape Mount with 60 beds and 45 HCWs began operating since week 20.
    - An ETU at Montserrado with 55 beds and 60 HCWs began operating since week 22.
    - An ETU at Bong with 35 beds and 55 HCWs began operating since week 24.
- Here is a list of ETUs in plan for opening:
    - An ETU at Nimba with 20 beds and 30 HCWs will begin operating sometime between week 25 and week 27
    - An ETU at Grand Bassa with 30 beds and 35 HCWs will begin operating sometime between week 27 and week 29