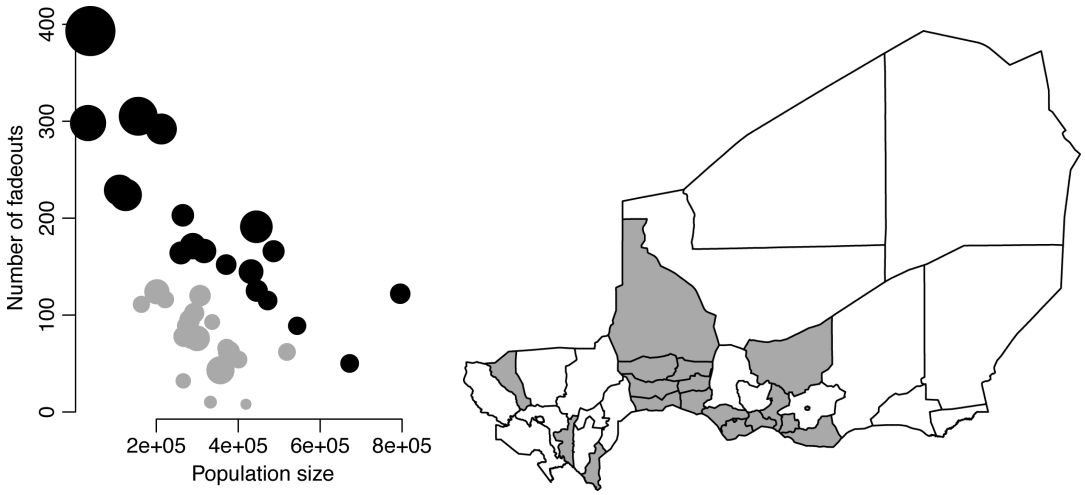
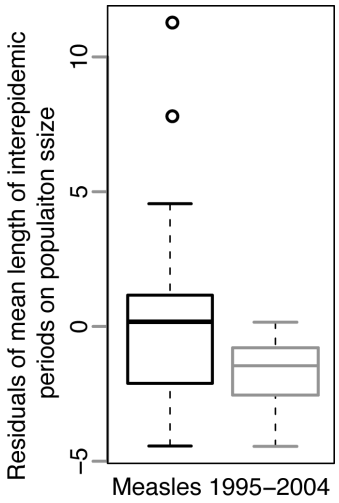
Supplementary Material 2

To account for underreporting, we repeat our analysis on inter-epidemic length using a minimum of four consecutive weeks with zero cases to define a fadeout. Here, inter-epidemic periods are consecutive weeks with zero cases that exceed 4 weeks. The spatial patterns seen from these analyses are nearly identical to ones found for 2-week fadeouts. Thus, we can be fairly certain that these patterns are real spatial signatures as a result of epidemic connectivity and not an artifact of underreporting.



Using 4-week long fadeouts, we plot number of fadeouts against population size (R^2 = 0.2832) and map the residuals (above, size of points is correlated to mean length of inter-epidemic periods). The negative and positive residuals are spatially autocorrelated (Moran *I* statistic = 0.214 and *P* = 0.02) (above, right).



As in the main text, we use a Cox proportional hazard regression model with the length of the inter-epidemic periods as a waiting time to determine the hazard rate of re-introductions (similar to [1]) (Fig. 3*c*, main text). Specifically, population size is the independent variable, inter-epidemic length is the response variable, and hotspots/not

hotspots are factors (same hotspots as in the main text as these were defined by post-SIA incidence, not fadeouts). Accounting for population size, we compare the re-introduction rates of the hotspots identified by the post-SIA analysis against all the other districts.

We again find that population size is strongly negatively correlated on the rate of re-introductions ($P < 0.0001$, independent variable from Cox proportional hazard regression model), indicating that re-introductions are more likely in highly populated districts. We further find that the six hotspots, identified by high numbers of post-SIA cases, have a significantly higher rate of re-introductions relative to population size than do the other 32 districts for 1995-2004 ($P < 0.03$, factor from Cox proportional hazard regression model). This means that the hotspots (above, grey) have significantly shorter waiting time to re-introductions than do non-hotspot districts (above, black) when fadeouts are defined as 4 weeks to account for underreporting.

Reference

1. **Bjornstad ON, Grenfell BT.** Hazards, spatial transmission and timing of outbreaks in epidemic metapopulations. *Environmental and Ecological Statistics* 2008; **15**: 265–277.