proposed figs 1-3 for imperfect data

Eric Marty

August 11, 2017

Reading over the introduction, it seems to me that in order to illustrate the aggregation progess and the two sources of reporting error (aggregation-based mismatch and imperfect reporting), 3 figures are required instead of just 2.

I'd suggest these three figs:

1. Wide view showing time series of I and time series of imperfect case reports, showing the 10 years approaching epidemic threshold.
2. The aggregation process figure, showing time series of I and 7-day and 30-day case reports (under perfect reporting conditions), with the individual cases graphed below as segments.
3. Example of the difference between perfect and imperfect reporting ($\rho < 0.125, \text{dispersion} = .1$) without aggregation-based mismatch

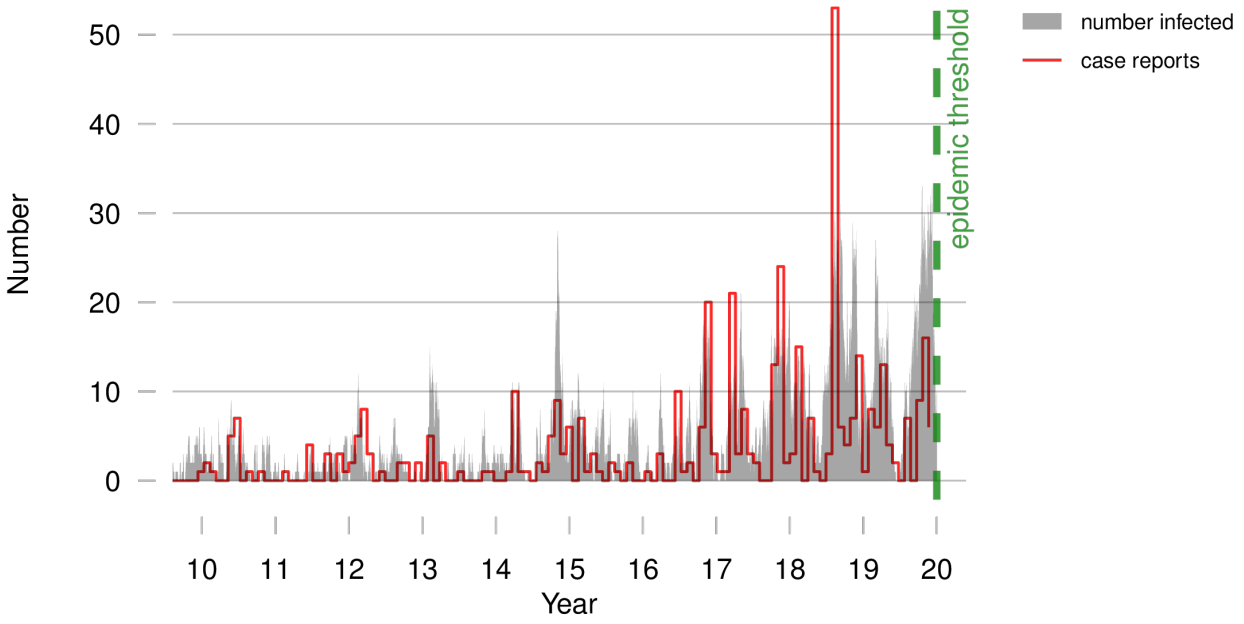

Figure 1: Figure 1. Simulated epidemic approaching epidemic threshold (grey) and simulated monthly case reporting (red). Average infectious period = 7 days. Aggregation period = 30 days. Error arises from a) mismatch between infectious period and aggregation period, and b) and reporting error (such as unreported cases and false positives). In this simulation, reporting probability = 0.125 and the dispersion parameter of the negative binomial = 1.

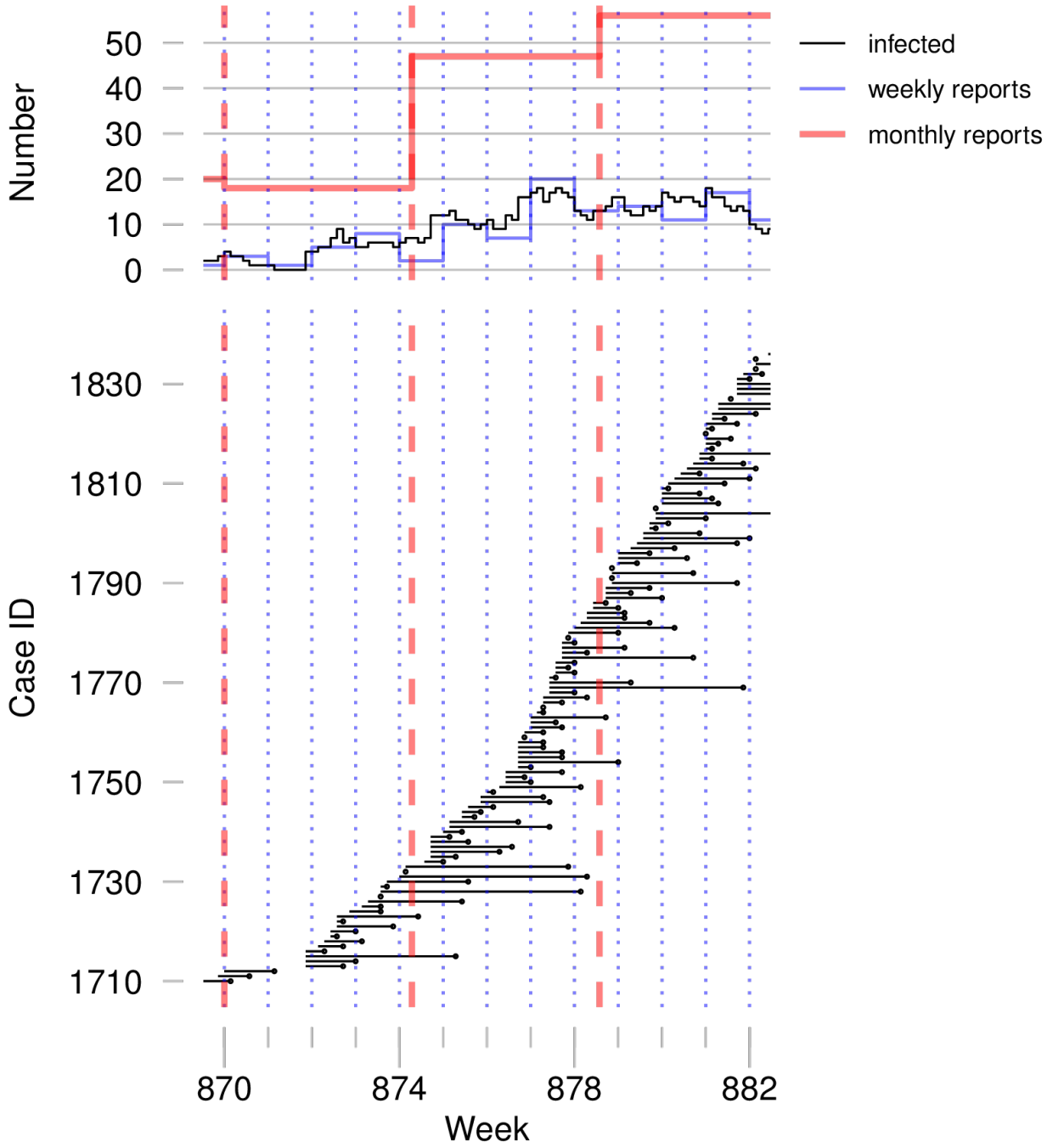

Figure 2: Figure 2. Example of aggregation process and agregation-based mismatch. Illustration of the aggregation of case reports for a disease with average infectious period of 7 days. In the lower panel, each black line segment represents the infectious period of a simulated individual, from onset of infection to recovery. Each black dot represents a case report, which in this simulation corresponds to recovery of the individual ($I \to R$ transitions). Dotted blue verticals delineate one-week reporting windows; dashed red verticals delineate 30-day reporting windows. Counting the number of segments present at any given time gives the true number of infected, plotted as a black trace in the upper plot. Counting the number of dots in a given window gives the number of case reports for that window, plotted as blue (7-day) and red (30-day) traces in the upper plot.

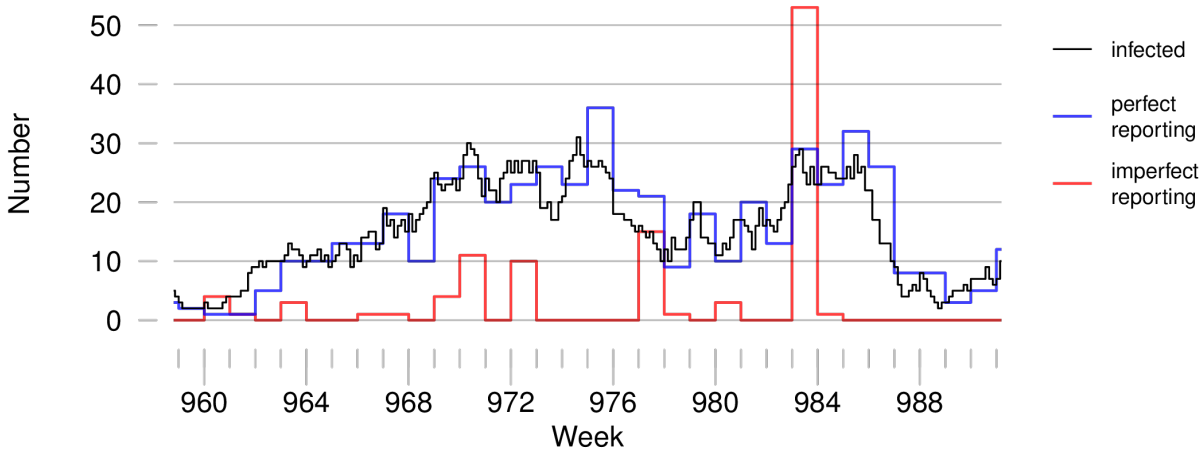

Figure 3: Figure 3. Example of reporting error. Simulation of an SIR model with average infectious period of 7 days, showing abundance of the infected class I (black trace), case reports aggregated weekly under "perfect reporting" conditions (blue trace), and imperfect conditions (red trace). Under "perfect" conditions, the reporting probability = 1. Under imperfect conditions, reporting probability = 0.125 and the dipsersion parameter of the negative binomial is 0.1.