THOUGHTS ON MEASLES MODELING

Travis C. Porco

7 June 2016

MEASLES

MEASLES

- The WHO estimated that 158,000 deaths were caused worldwide by measles (2011).
- Safe effective vaccine has reduced this from 630,000 cases in 1990.
- Most WHO regions have set elimination targets

- USA declared elimination
- Handful of cases every year in California

CALIFORNIA OUTBREAK

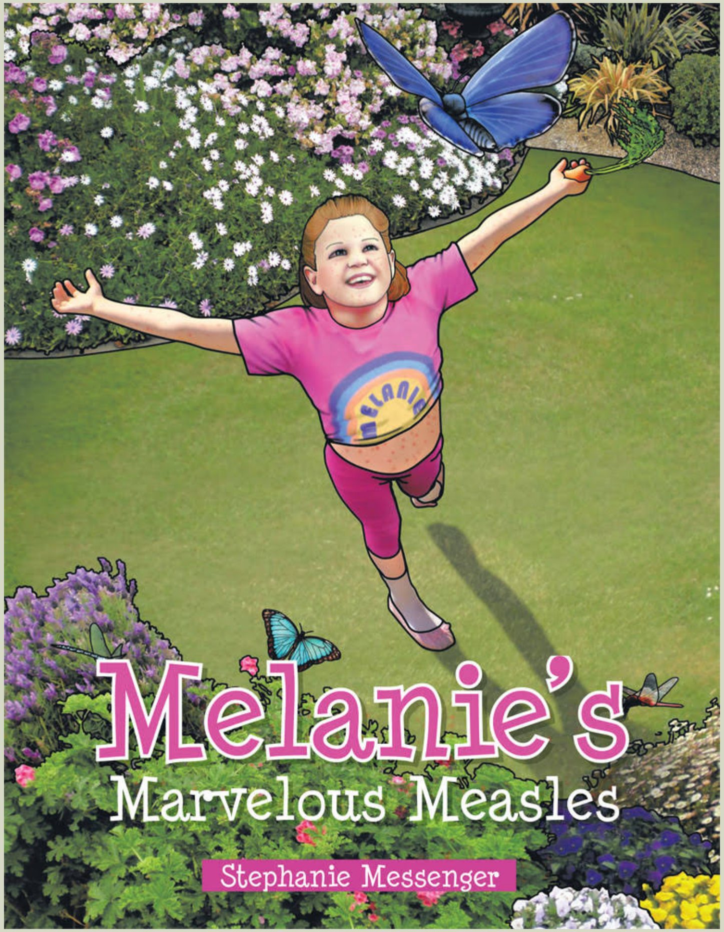
- 6 Jan 2015 – public notified of outbreak
- By early March, 134 cases in California
- Forty people who visited the Disneyland theme resort in Anaheim between 17 Dec 2014 and 20 Dec 2014 were believed exposed there
- Source case or cases never identified
- Nearly all the cases genotyped were of type B3; 3 were not of this type
- Presumption is that 131 of the cases were from the outbreak

Source: CDPH.

REACTION

- Media reaction – international coverage
- Public took considerable interest

WHY



“ANTIVAXXERS”

- Small but vocal group of individuals who are opposed to vaccination
- Many such individuals do not believe scientific studies showing the safety and efficacy of vaccines in general and MMR in particular
- Through courts or public opinion, could a small group ultimately imperil measles elimination?

- Some in the press blamed antivaxxers for the outbreak, and the state of California changed the laws. Children are now required to have measles vaccinations to attend school.

MODELING

- What does it really mean to go from 40 exposures to 131 cases?

- Break the epidemic up into two parts:
  - Disneyland
  - After disneyland

IN DISNEYLAND

- All we know is that 40 people were exposed.
- We do not know in what location.
- We do not know whether it was all at one time or spread over several days.
- We do not know whether the source case was a tourist, a family of tourists, or even an employee.
- We do not know if it was a single source case or more than one.

- We DO know that measles is highly contagious, and that Disneyland is crowded.

IN DISNEYLAND

- Tthere are a few percent of people who are not vaccinated or who never were infected.
- Some people may have received only one shot.
- Protection is very high, but not absolutely perfect.

- Over the course of a day or two, a sick person could come into “contact” with many hundreds, if not thousands of people, at a crowded resort.

- In that setting, 40 cases does not seem surprising.

AFTER DISNEYLAND

- Afterwards…
- 40 cases become 131 total cases before the outbreak is over.
- What does this mean?

AFTER

- With 40 cases to start with, and 131 total, we have 131/40 = 3.275 ultimate cases per initial case.

- Every case has caused 3.275-1 = 2.275 new cases ULTIMATELY, before the outbreak was over.

R

- We had 40 cases causing 131.
- That means 91 transmissions out of 131 people.
- The number of cases per case is 91/131=0.69, approximately.

LET’S LINGER ON THIS

- But the transmission did not happen all at once.
- There were several rounds or generations of transmission.
- The epidemic went on for two months.

TRANSMISSION OVER GENERATIONS

- We want to relate the transmission by generation to the total outbreak size.

- One case -> More cases -> More cases -> …
- How many total?

GENERATIONS

- Imagine I start with 100 cases. Thought experiment.
- Say each case only causes ½ a case on average.

- Work with AVERAGES:
- $E[X(t+1)] = r\, E[X(t)]$ with $r=1/2$

- One of the simplest dynamic models

GENERATIONS

- So I start with 100 cases, and say each case only causes ½ a case on average.
- Then in generation TWO, our 100 cases have given us 50 more.
- The total outbreak is at size 100 + 50 = 150.

- In generation THREE, our 50 cases give us 25 new cases.
- The total outbreak size is 100 + 50 + 25 = 175.

- In generation FOUR, say we get 12.5 new cases.
- The total outbreak size is 100 + 50 + 25 + 12.5 = 187.5

…CONTINUING

- We are now departing from reality by thinking about fractions of people.
- But let us continue thinking about this outbreak.
- Generation five brings us 12.5/2 = 6.25 cases, and a total outbreak size of 193.75.
- We can just use a statistics package to keep this going for us:
-
  ```
  > cumsum(100 * (0.5)^(0:20))
  ```
-
  ```
  100.0000 150.0000 175.0000 187.5000 193.7500 196.8750
  198.4375 199.2188 199.6094 199.8047 199.9023 199.9512
  199.9756 199.9878 199.9939 199.9969 199.9985 199.9992
  199.9996 199.9998 199.9999
  ```
- “approaching” 200

…CONTINUING

- Let's now imagine that instead of each case causing ½ a case, each case causes 2/3 of a case:
- `cumsum(100 * (2/3)^(0:20))`
- 100.0000 166.6667 211.1111 240.7407 260.4938 273.6626 282.4417 288.2945 292.1963 294.7975 296.5317 297.6878 298.4585 298.9724 299.3149 299.5433 299.6955 299.7970 299.8647 299.9098 299.9399
- …approaching 300, it looks like
- So if each case causes 2/3 of a case EACH ROUND, then if we start with 100 cases, by the END of the epidemic, we should have 300 cases. One case gives rise to 2/3 each round, and 2 new cases ULTIMATELY.

GEOMETRIC SERIES

- Some of you will recognize this as the geometric series:
- $S=1+r+r^2 + r^3 + \ldots = 1/(1-r)$
- …if $r < 1$!

- If $r > 1$ – each case produces more than 1 new case, the terms get larger and larger. You can’t add them up. In this simple model, the outbreak NEVER ENDS.

CLUSTER SIZE

- But S is the cluster size, and r is the number of new cases a case can cause!

- If we observe the mean cluster size, we can compute r.
- $r = 1 - 1/s$

- In the Disneyland outbreak, s, the cluster size, was 131/40 = 3.325.

- This gives r = 0.694 or so.
- Each case causes about 0.7 of a new case.

INTERPRETATION

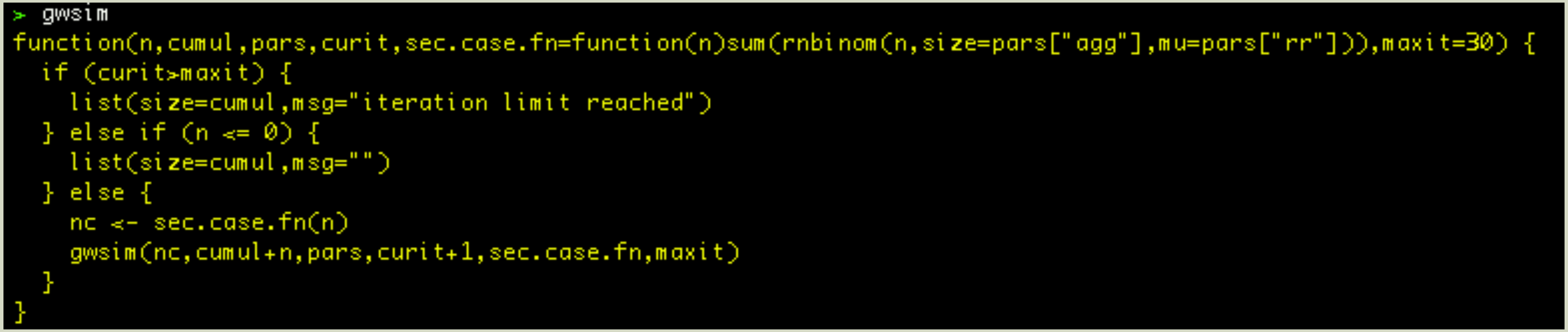
- Some people believe that if no one were vaccinated, a case of measles would cause 7 – 30 new cases each generation.
- If about 90% of people were vaccinated (which is about what the CDC says is true for the US), and the vaccine were perfect, what would this be?
- If only 10% of people are susceptible, then instead of 7 cases you should get 10% of 7 cases, or 0.7.

RANDOMNESS

- Treat the process as random to get an idea of the precision
- We won’t go thru this in detail

GALTON-WATSON PROCESS

- Say you have X cases at time t.
- Each case has a random number of secondary cases, assumed independent and identical
- The new X at time t+1 is the sum of these X(t) independent secondary case distributions



```
> gwsim
function(n,cumul,pars,curit,sec.case.fn=function(n)sum(rnbinom(n,size=pars["agg"],mu=pars["rr"])),maxit=30) {
  if (curit>maxit) {
    list(size=cumul,msg="iteration limit reached")
  } else if (n <= 0) {
    list(size=cumul,msg="")
  } else {
    nc <- sec.case.fn(n)
    gwsim(nc,cumul+n,pars,curit+1,sec.case.fn,maxit)
  }
}
```

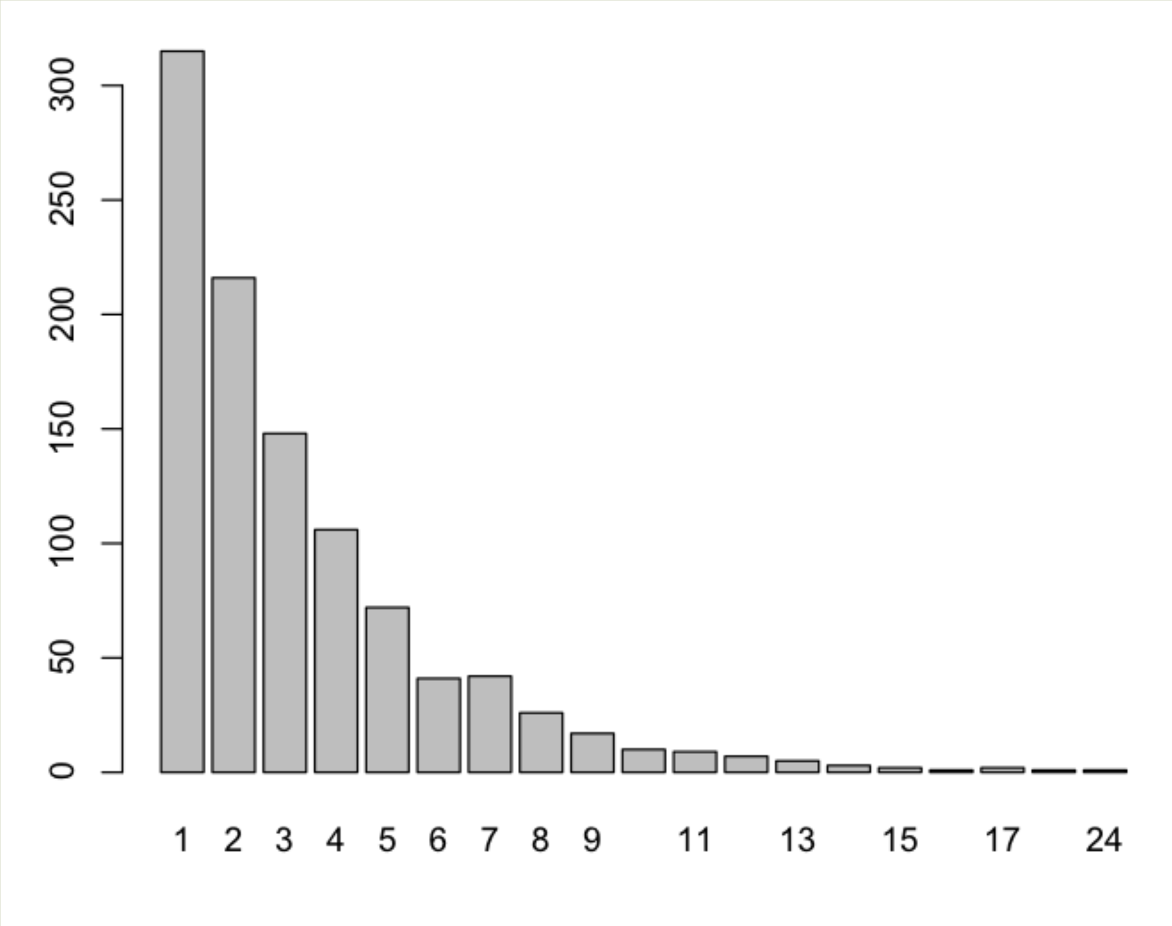
R

```
gwsim <- function(n, cumul, pars, curit, sec.case.fn=function(n)sum(rnbinom(n,size=pars["agg"], mu=pars["rr"])),maxit=30) {
  if (curit>maxit) {
        list(size=cumul,msg="iteration limit reached")
  } else if (n <= 0) {
        list(size=cumul,msg="")
  } else {
        nc <- sec.case.fn(n)
        gwsim(nc,cumul+n,pars,curit+1,sec.case.fn,maxit)
  }
}
```

SIMPLE

- We could assume every case produces a single new case 70% of the time, and none 30% of the time. That would give us r=0.7.

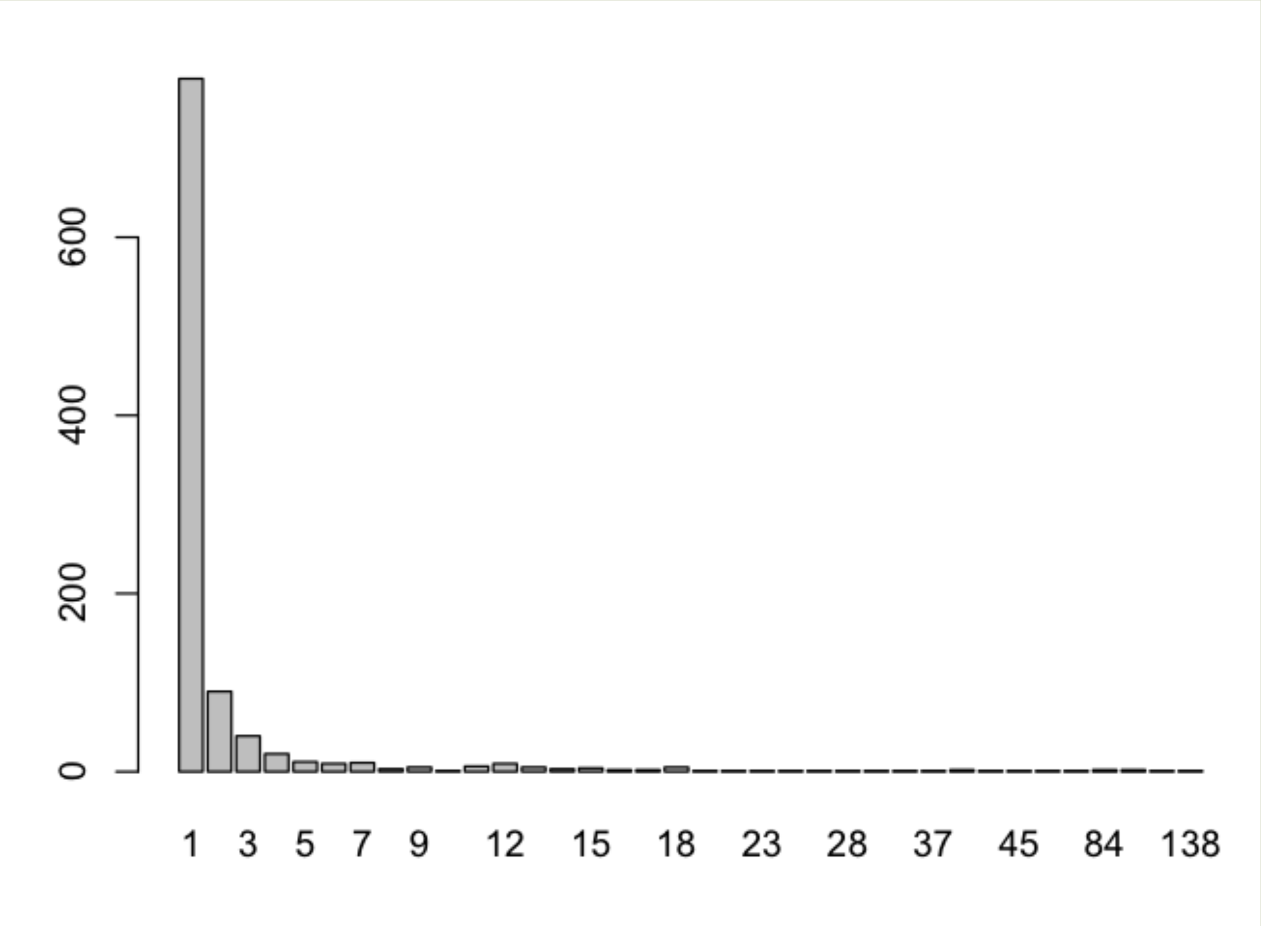
BARPLOT OF 1024 OUTBREAK SIZES



A DIFFERENT SECONDARY CASE DISTRIBUTION

- Assume that the distribution of secondary cases allows for superspreaders
- Without details, the negative binomial will do the trick, with an aggregation parameter much smaller than 1, say 0.15.

BARPLOT OF 1024 OUTBREAK SIZES



WHAT IS THE DIFFERENCE?

- *For the same mean* (R_eff), superspreading means more large outbreaks but also more clusters of size 1.

MOVING FURTHER

- If we compute Reff from data, we must have a precision estimate.
- Some sort of error bar, confidence interval, credible interval, …
- We obtained that using a random process version of the simple model (Galton-Watson process)
- Blumberg, Worden, Enanoria et al, PLoS Currents: Outbreaks, May 7, 2015.

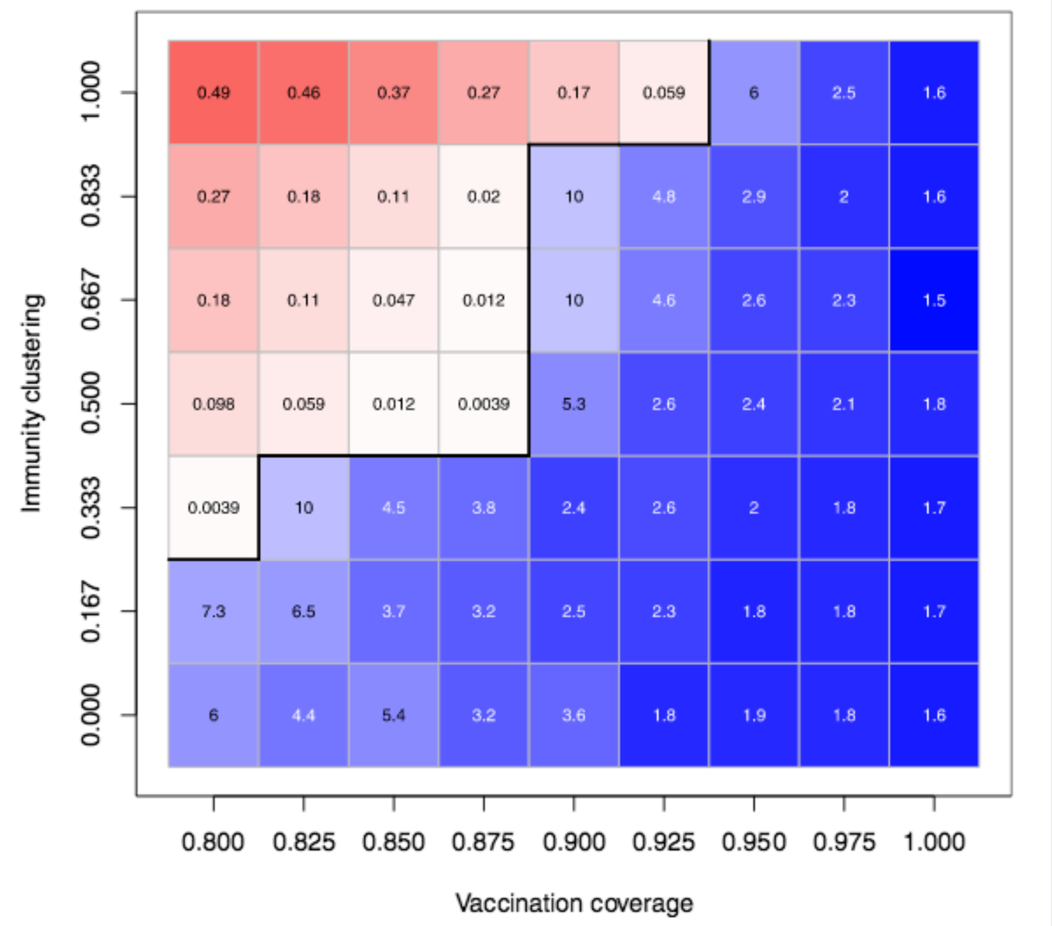
WHAT DOES THE MODEL TELL YOU?

- Couldn’t we just say that the outbreak is over, and that therefore there was nothing to worry about?
- The model gives you an idea of how close you are to r=1, and a confidence interval.
- The model can give you predictions of what future cluster sizes would be. But very difficult to predict “superspreader” events.
- How can we test the model?

OBJECTIONS

- Isn’t this a very simple model?
- Aren’t we leaving out:
  - Households
  - Nonrandom vaccination status
  - Changes in behavior over time during the epidemic

GOING FURTHER



Immunity clustering	0.800	0.825	0.850	0.875	0.900	0.925	0.950	0.975	1.000
1.000	0.49	0.46	0.37	0.27	0.17	0.059	6	2.5	1.6
0.833	0.27	0.18	0.11	0.02	10	4.8	2.9	2	1.6
0.667	0.18	0.11	0.047	0.012	10	4.6	2.6	2.3	1.5
0.500	0.098	0.059	0.012	0.0039	5.3	2.6	2.4	2.1	1.8
0.333	0.0039	10	4.5	3.8	2.4	2.6	2	1.8	1.7
0.167	7.3	6.5	3.7	3.2	2.5	2.3	1.8	1.8	1.7
0.000	6	4.4	5.4	3.2	3.6	1.8	1.9	1.8	1.6

Vaccination coverage

DEVELOPING MODELS

- Insight through simplicity
- Realism through detail
- Validation through prediction

ACKNOWLEDGMENTS

- CDPH (Jennifer Zipprich, Kathleen Harriman, Sarah Royce)
- UCSF team (Wayne Enanoria, Seth Blumberg, Sarah Ackley, Lee Worden, Fengchen Liu, Tom Lietman)
- NIH NIGMS MIDAS Program