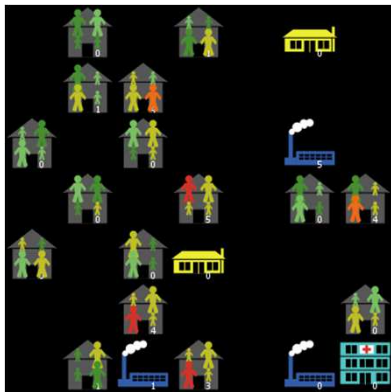


$$\lambda_i = u \cdot \sum_{j \in AG} C_{i,j,t} \cdot (Ip_{j,t} + Ic_{j,t} + f \cdot Is_{j,t}) / N_j$$

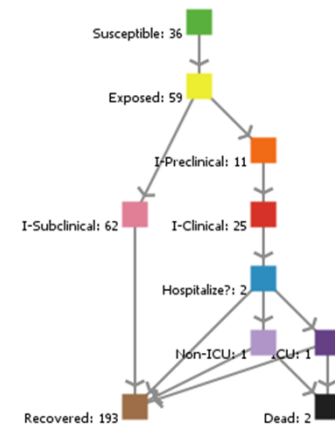
$$n_{S_{i,t} \rightarrow E_{i,t}} \sim \text{Bin}(S_{i,t}, 1 - e^{-\lambda_{i,t} \cdot \delta t})$$

An agent-based modeller looks at an epidemiologist's model



Christopher J. Watts

Social Simulation Week 2020
Wednesday, 10am CET, 16th September



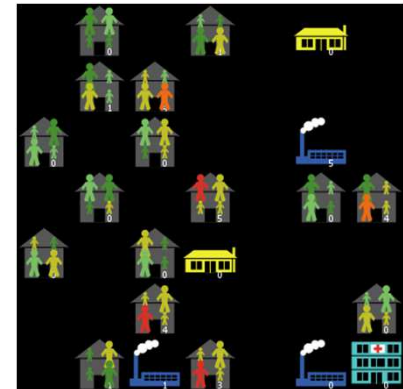
The February Question

- It's February 2020
 - You have a little empirical data about a novel SARS-causing virus
 - How many people will become ill? Need hospital? Die?
 - How soon will the hospitals be full?
 - What can we do and when must we do it?
 - Decisions include to shut schools, shops and workplaces
- In three weeks you must advise the UK Government's Scientific Advisory Committee (SAGE)
- Which modelling tools do you reach for?

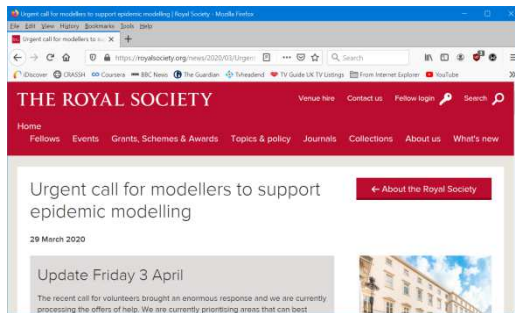
The answer was *not* Agent-Based Modelling.
What was the answer? What did it offer?

How agent-based modellers might see things

- Heterogeneous, Adaptive Agents
 - Perform routine activities
 - In locations (Home, School, Work, Etc.)
 - According to personal schedules (based on survey data?)
 - Interventions (e.g. schools closure) force agents to seek other activities
- Infection occurs between agents in same location at same time
 - A bi-modal, bi-partite transmission network
- Inspiration:
 - Joshua Epstein's keynote at SSC 2019
 - Simulating innovation, organisational learning (routine practices)
 - The team at Hohenheim Uni came up with same ideas



Nigel's question: Why listen to *us*?



March-April, a team of ABMers forms:

- Christopher Watts, independent, Cambridgeshire, UK
- Nigel Gilbert, University of Surrey
- Duncan Robertson, Loughborough University & St. Catherine's College, Oxford
- Daniel Ladley, University of Leicester
- Edmund Chattoe-Brown, University of Leicester
- Anonymous Student, University of Leicester

Nigel's question:

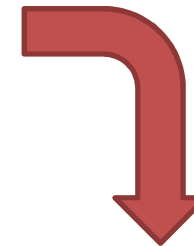
"We're not Epidemiologists. Why should anyone listen to us?"

Nigel's solution (1):

Find some experts with a serious model...

Modellers	Language	Thoughts
Imperial College: Neil Ferguson et al.	C	Too big
London School of <i>Hygiene</i> & Tropical Medicine (LSHTM), CMMID: Nick Davies, W. John Edmunds, Adam Kucharski, et al.	R, C++	Feasible?: <ul style="list-style-type: none">• Download from github• Well-documented(ish)• Runs in “hours” on a laptop
UCL: Dynamic Causal Modelling Karl Friston	Matlab / Octave	Too novel

The “Hygiene Model”




Davies et al. (2020) “Effects of non-pharmaceutical interventions on COVID-19 cases, deaths, and demand for hospital services in the UK: a modelling study”. The Lancet Public Health. Published online June 2, 2020 [https://doi.org/10.1016/S2468-2667\(20\)30133-X](https://doi.org/10.1016/S2468-2667(20)30133-X)

Steps 2,3,4: Dock with the experts' model... then gradually add ABM features

2) NetLogo replication of Hygiene model, using POLYMOD contact matrices

3) Replication model, but using our *own* contact matrices generated from survey data and simulated locations

4) Replication model, but no contact matrices. Contacts based on activity schedules from survey data, locations, adaptive behaviour



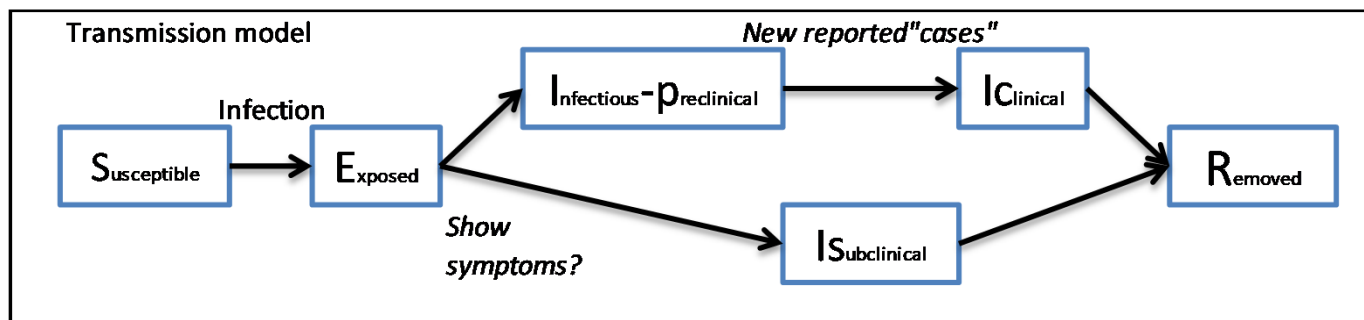
Do the results change?

Is the policy advice sensitive to its non-ABM assumptions?

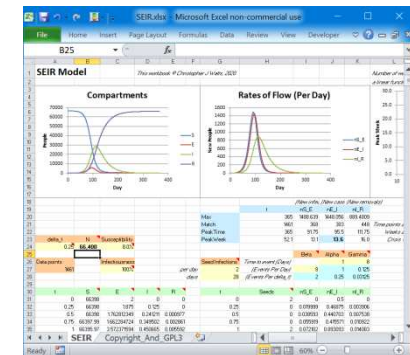
A brief introduction to

THE HYGIENE MODEL OF DAVIES ET AL. (2020)

Compartmental models: Quick-and-dirty solutions since 1927 (or the late 1970s)



- Susceptible-Infectious-Removed (SIR), SI, SEIR, SEIRD, etc.
 - Kermack & McKendrick (1927) SIR model
 - Popular among Epidemiologists since late 1970s (Anderson & May)
 - Drag-and-drop creation in SD software (e.g. Vensim)
 - Solve them in Excel
 - E.g. <https://github.com/innovative-simulator/PopScaleCompartmentModels/blob/master/xlsx/SEIR.xlsx>



The force of infection, λ

The rate at which infection events occur for a Susceptible person is a linear function of the Infectious proportion of the (sub-)Population

λ_i = The rate over time of infection events for a Susceptible in Age Group i

$$\lambda_i = u \cdot \sum_{j \in AG} C_{i,j,t} \cdot (Ip_{j,t} + Ic_{j,t} + f \cdot Is_{j,t}) / N_j$$

The weighted proportion of the entire population (in age group j) who are currently infectious

$N = S + E + Ip + Ic + Is + R$
NB: R includes the recovered, the isolating, and the dead!

Contacts per time step between people in age groups i and j

Contact matrices:

- Age Group i to Age Group j
- Home + School + Work + "Other"
- Empirical: E.g. the POLYMOD data
- Modify matrices to simulate interventions (e.g. School Closure)

Non-adaptive behaviour:

- Contact rates do not vary with disease state
- No re-allocation of contact time

The Universal Mixing Assumption:

I have C_{ij} contacts per day with people from group j and they may be **anyone** in that group, irrespective of their:

- Spatial distance
- Social network (Family? Friends? Co-workers?)
- Our disease states, e.g. whether or not they are dead

The Hygiene Model addresses the criticisms of the differential equation models

System Dynamics model / Differential Equations (SIR)	Individual-based dynamic stochastic model (Hygiene Model)
Deterministic; Variance around norm not represented	Stochastic flows between compartments $n_{S_{i,t} \rightarrow E_{i,t}} \sim \text{Bin}(S_{i,t}, 1 - e^{-\lambda_{i,t} \cdot \delta t})$
Homogenous population	Age-group sub-compartments
Universal mixing	Simulate counties, then sum to UK. Age-group-based contact matrices
No adaptive behaviour	Interventions simulated as time- based changes to contact matrices

The Hygiene Model addresses the criticisms of other stochastic / microsimulation models

Stochastic model	Individual-based dynamic stochastic model (Hygiene Model)
State transition times are Exponentially distributed = High variance	Sample next transition time (from a Gamma distribution) for each individual entering state

But this means a simulation of UK (Pop=66.4 million) needs hundreds of millions of random numbers per run

The county-level simulations need seed infections

County/Region-based model	Hygiene Model
No cross-border contacts; Counties simulations run independent of each other	Seed infections for each county

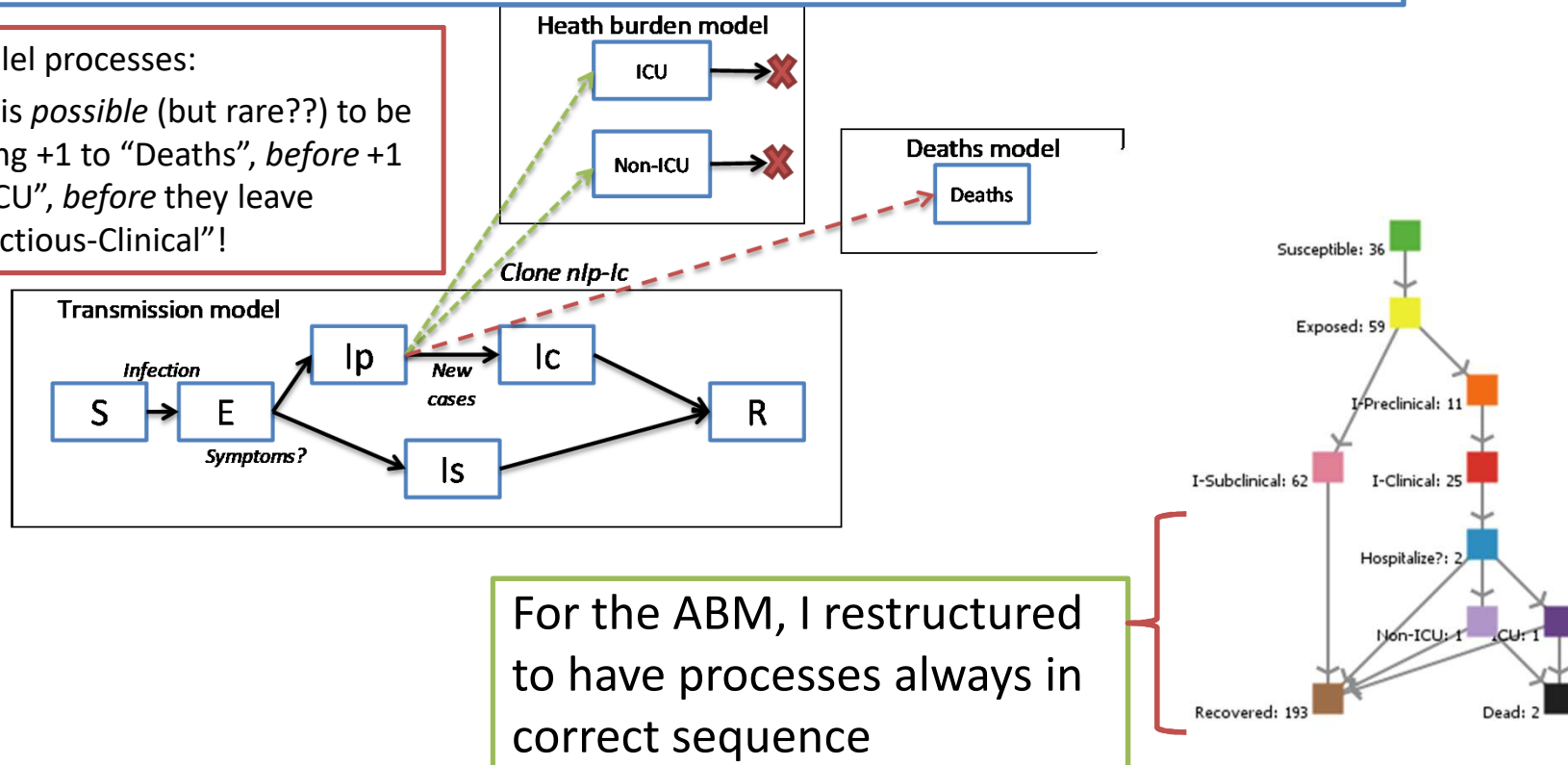
This raises a few problems...
(See later, and RofASSS article)

WHAT WE LEARNED FROM REPLICATING THE HYGIENE MODEL

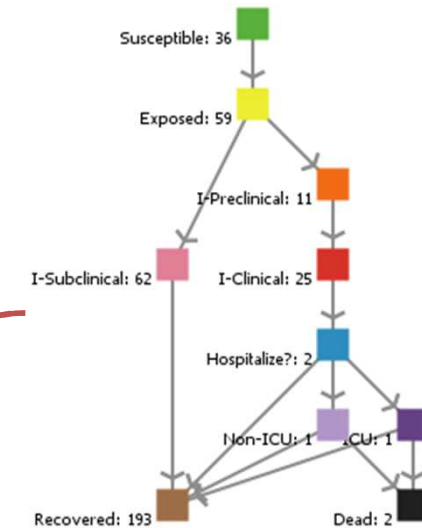
Hygiene Model *splits* models for Health Burden and Deaths from Transmission model!

Parallel processes:

So it is *possible* (but rare??) to be adding +1 to “Deaths”, *before* +1 to “ICU”, *before* they leave “Infectious-Clinical”!



For the ABM, I restructured to have processes always in correct sequence



Size does not matter (usually)

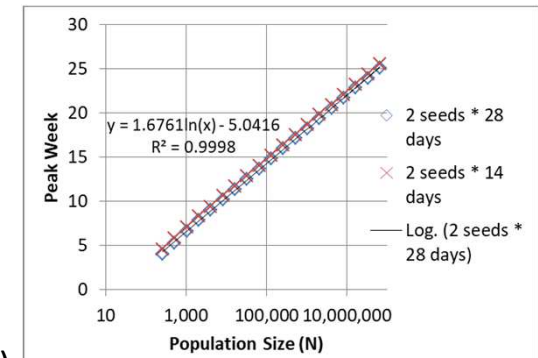
- A compartmental model produces Peak or Total Cases of $X\%$ of N people
 - Same $X\%$, whatever the N
- So small models ok. E.g. model 1000 agents. Then multiply by 66,400 to forecast the UK
- But population size does affect the *time* to peak cases and *time* to all ICU beds full
- Need to re-scale seed infections to $Y\%$ of N
- Very small models still a problem. 1 seed infectious person might remove too quick to spark epidemic

How to seed infections?

See <https://rofasss.org/2020/08/14/role-population-scale/>

Hygiene model:

- Same seed schedule used for every county: 2 per day for 28 days
- Time to peak (and critical levels of hospital beds) tends to linear function of $\text{Log}(\text{Population Size} / \text{Number of Seeds})$
- So smallest county (population=2200) peaks earlier than biggest (2.9 million)
- Why is this a problem?
 - Timing is important to the interventions' impact
 - UK aggregate of out-of-sync county peaks will be smaller
- Hard problem to solve:
 - Re-scale seeding for county size? – But small seeding might not generate an epidemic. Biggest county get 1300x seeding of smallest.
 - Model UK directly? – But universal mixing absurd
 - Model UK social network? – But you don't yet know what this looks like!



Lack of justification

- Although the Hygiene model *can* be fit to empirical data (using MCMC)...
 - E.g. posteriors for rate parameters , R_0 / Susceptibility
- It makes lots of under-justified assumptions
 - Seed infections schedule
 - Interventions altering the contact matrices
- There is very little sensitivity testing (No time?)
 - R_0
 - Intervention launch delays
 - But *not* the age-based rates for symptomatic rate, hospitalisation, fatality, etc.

Replication revealed a bug in their model!

- We found a bug in the LSHTM model!
 - Intervention scenarios were being run with the *default seeding* (=1 person on day 1)
 - *Not* the seeding schedule used for *Base scenario* (=2 per day for 28 days)
- The Lancet paper will have to be corrected!
 - (When Davies et al. have time...)
 - They massively overestimated the impact of interventions
 - Apparently this affects the paper, but *not* their advice to UK government. (Well, that's their story...)

Were the complications worth the effort?

Complications

- UK-size population (66.4 million)
- UK = Sum of Counties
- Age-group sub-compartments + parameters
- Stochastic flows
- Individual-based (i.e. sampled maturity times)

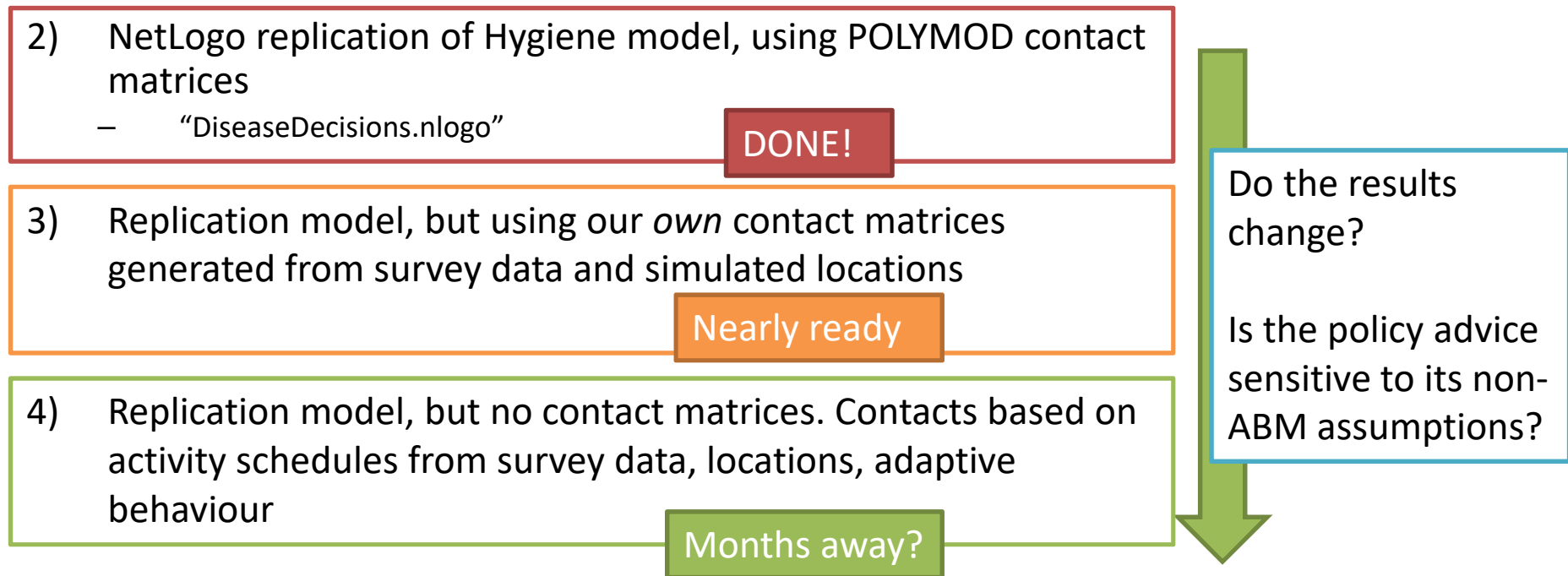
Results

- Too much code to check adequately
- Longer to run experiments (hours to days)
- Verification and validation focus on UK aggregate level, not county-level runs
- Modelling at county level led to seeding scale problem, and helped hide a bad bug
- More reliable policy modelling?
 - Not if you have to retract the paper!

Conclusions

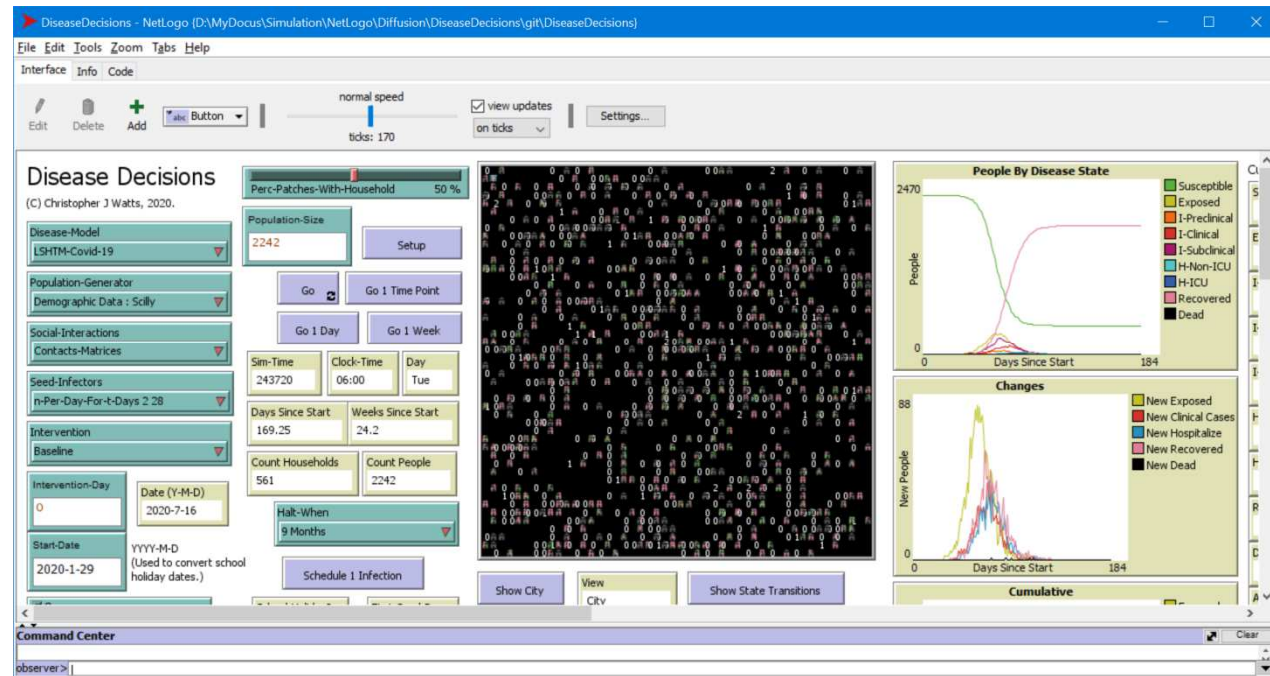
- Advantage of “individual-based, dynamic stochastic model” over quick-and-dirty differential equations / SD is not clear. Advantage over ABM??
- Advantage of county-level runs is not obvious: it still involves Universal Mixing
- AB modellers note: Small models good; Complexity must be easy to check

Nigel's strategy: Progress so far



DiseaseDecisions.nlogo

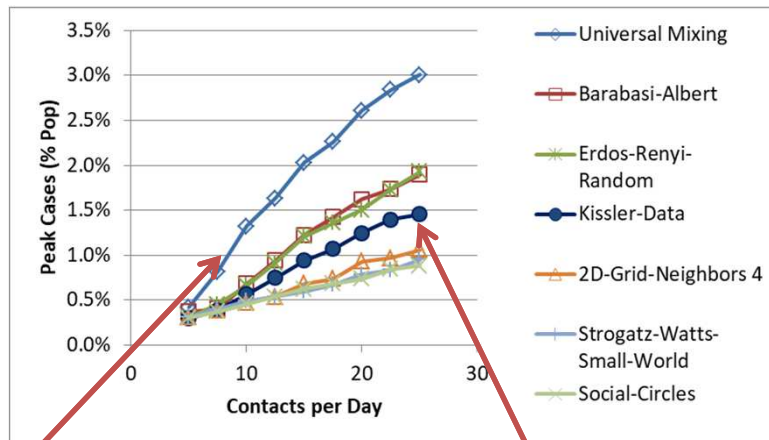
- Replicates Hygiene Model
- Can replace contact matrices with agents' activity schedules and locations
- Uses discrete-event simulation



Not universal mixing, but...

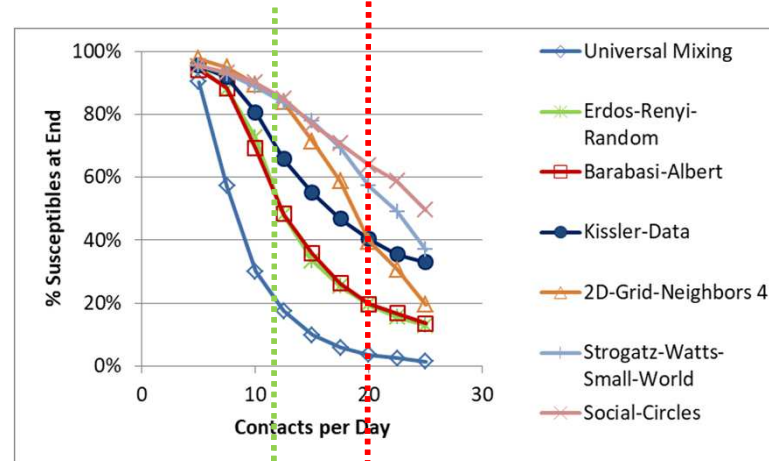
IS THE FUTURE NETWORKS OR REAL CONTACT DATA?

Universal mixing compares badly with network structure...
...but none of them look like **real contact data**!



Universal Mixing overestimates Peak Cases

Network derived from Kissler's real-world contact data behaves unlike any library model



Lockdown Effect

Network structure may make a big difference to impact of Interventions

Realistic contact data from Stephen Kissler (BBC Pandemic project, 2018)

- 3 days of contact data from small town (Haslemere) using mobile apps
- Filtered to distances $\leq 2m$
- 392 nodes, largest component 381 nodes, Network Density = 1.4%

What capabilities do the modelling approaches offer?

Requirement / Feature	System Dynamics / Differential Equations (SIR)	Individual-based Dynamic Stochastic Model (Hygiene Model)	Agent-Based Model (Your model here?)
Computation	Very Easy	Hours – Days for UK	Forget about the UK!
Heterogeneous Agents (Age, Location, Timing, Contact Rate)	Only with sub-compartments: e.g. Young-Old, City-Rural, etc.		Easy
Time to state transitions	Fixed	Exponential distribution, or Any if individual-based	Sample from any distribution
Contact Neighbours	Universal Mixing	Universal Mixing	Social Networks
Adaptive Agent Behaviour / Learning	No	No	Quite Easy
Fit to data	MCMC	MCMC	No likelihood! Anyone used ABC??
Estimating Risks	Averages only	Stochastic variation	Complexity

References

- The Hygiene Model
 - Davies et al. (2020) “Effects of non-pharmaceutical interventions on COVID-19 cases, deaths, and demand for hospital services in the UK: a modelling study”. The Lancet Public Health. Published online June 2, 2020 [https://doi.org/10.1016/S2468-2667\(20\)30133-X](https://doi.org/10.1016/S2468-2667(20)30133-X)
 - <https://github.com/cmmid/covid-UK>
- RofASSS article
 - <https://rofasss.org/2020/08/14/role-population-scale/>
- NetworksAndUniMix.nlogo (and **these slides**)
 - <https://github.com/innovative-simulator/NetworksAndUniMix>
- Kissler’s contact data
 - <https://github.com/skissler/haslemere>
- Email: cjwatts1007@gmail.com