policy analytics

LAMSADE
UMR CNRS 7243
laboratoire d'analyse et modélisation de systèmes pour l'aide à la décision

Une modélisation sociale et qualitative en épidémiologie pour s'intégrer à l'aide à la décision

27 mai 2021

Juliette Rouchier
(LAMSADE – CNRS, PSL, Université Paris-Dauphine)

cnrs

Principes de l'aide à la décision

- Aide à la décision comme domaine de recherche

  Formalisation du problème : classements, indicateurs composés, vote >> *propriétés mathématiques*

  Accompagnement : expression des préférences / valeurs, création d'alternatives de choix >> *comparabilité des options*

  Arbitrage : par le décideur rationnellement informé >> *multi-critères*

- Spécificités de l'aide à la décision publique

  Légitimité, applicabilité - « acceptabilité »

  « Accountability » - responsabilité / transparente

Modèles épidémiologiques durant la crise

- Modèles SIR individus centrés (ABM) / SIS

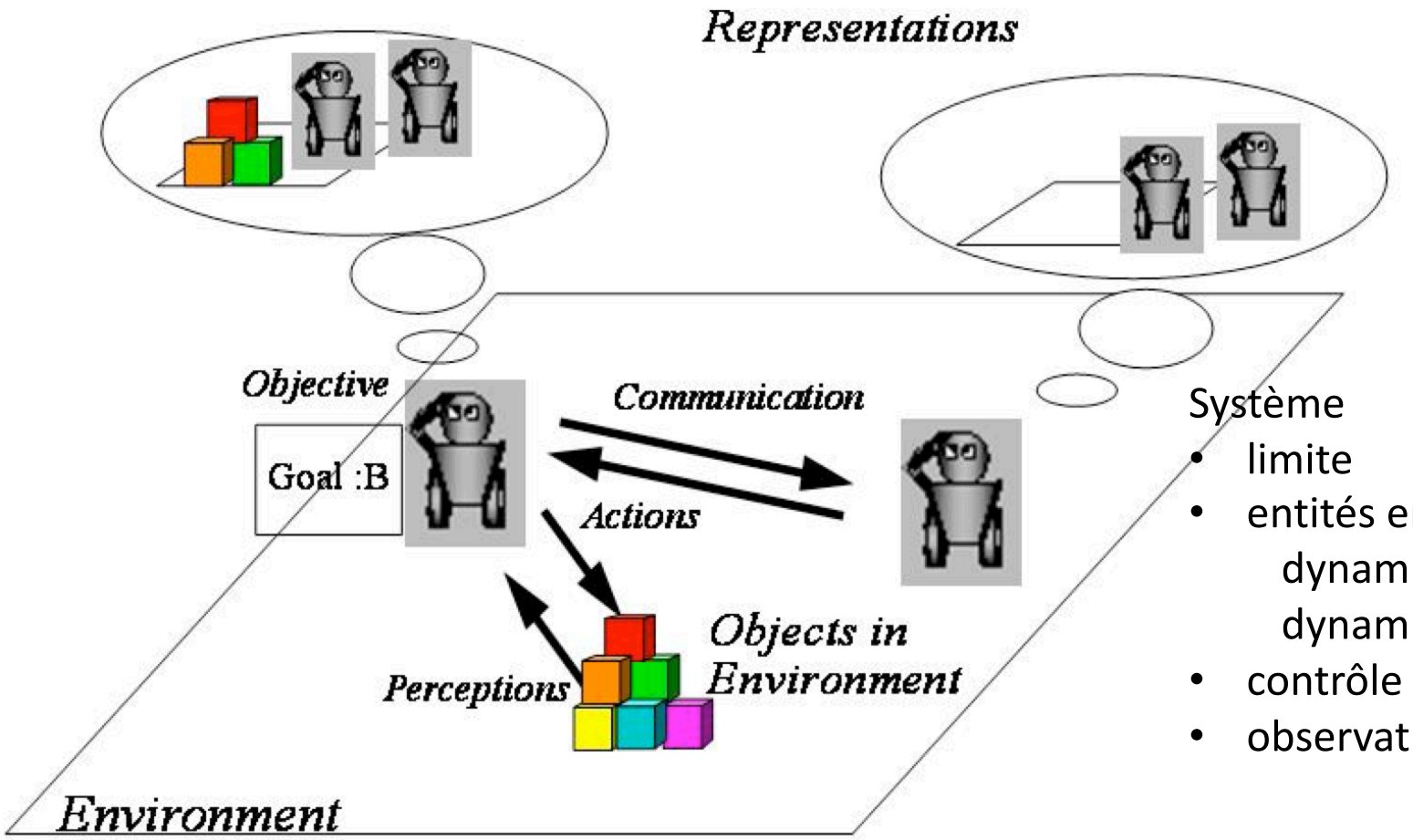
>> *« bottom-up »* - potentiellement explicatifs, travail possible sur les contrefactuels, évolution des hypothèses

- Modèles SIR « à compartiments »

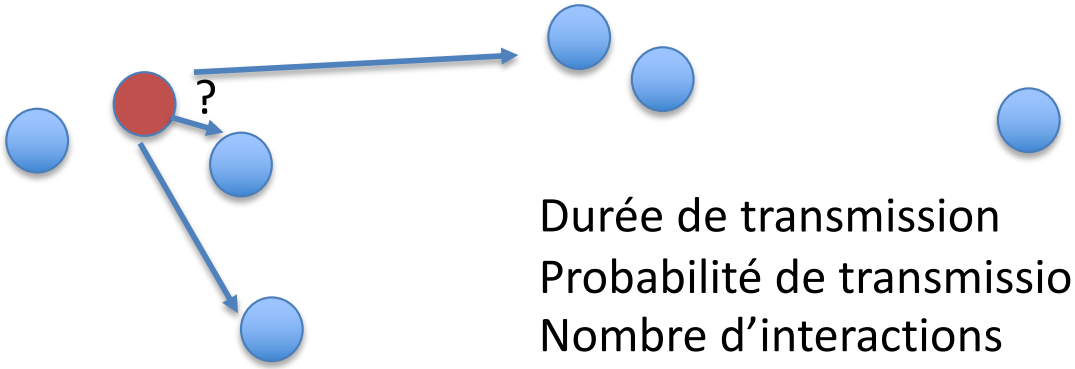
>> *« agent représentatif »* - « prédiction »

- Modèles « phénoménologiques »

>> *extrapolation* sans explication ni intégration de connaissances, peu de paramètres, « fit »



Système
- limite
- entités en interaction
  dynamique des entités
  dynamique des interaction
- contrôle / feedback
- observation



Durée de transmission
Probabilité de transmission
Nombre d'interactions

Les modélisations ABM durant la crise

- Quantitative / prédictive :
  - ABM présente dès l'origine des recommandations scientifiques (Ferguson, 2020) - prévoir le nombre d'agents à l'hôpital / en réanimation / morts
  - Mais finalement méthode peu reprise par la suite pour la prédiction (Hoertel, 2020)
  - Manzo et van de Rijt, 2020 (JASSS)

- Qualitative / explicative : décrire les émergences possibles

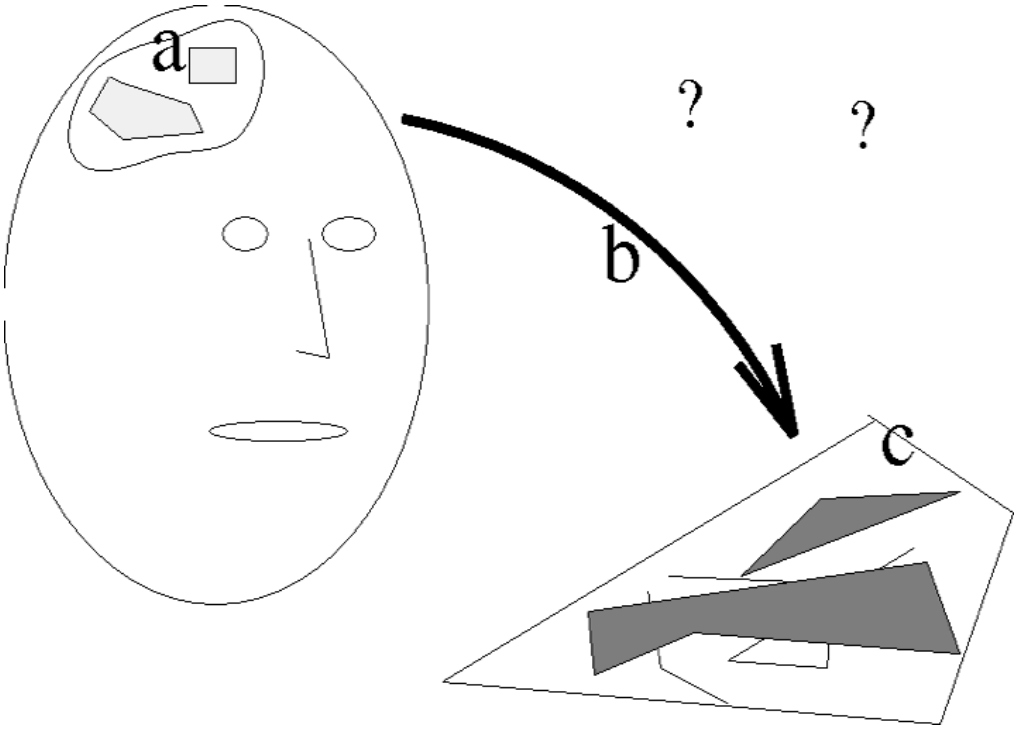
Un choix pédagogique durant le premier confinement :

https://covprehension.org/

- Système de questions,
- Discussions sur les thèmes
- Partage d'informations, articles
- Corrections croisées dans la rédaction des modèles

- Vision KISS de la modélisation : simple pour faire apparaître des effets compréhensibles
- « Styles de modélisation » et d'interprétation

Emergence



!!!!

Toujours des valeurs numériques

Pour trouver le taux de contagiosité, on fait des la « rétro-ingénierie » à partir du R0 estimé (entre 2 et 3 à l'état naturel)

Qualitative : le résultat est une *propriété du système - moins coûteux en données*

Quantitative : le résultat est une *valeur numérique*

peu recommandé en général - voir « Pourquoi la société ne se laisse pas mettre en équation » (Pablo Jensen)

Modèles qualitatifs « faits stylisés »

Régularités non-évidentes (tautologiques)

- les prophéties auto-réalisatrices (marchés financiers, racisme) - en sortir (*Heinke et al 2013*) ;
- La congestion (marchés, ressources) (Rouchier *et al,* 2001) ;
- La polarisation - le simple processus d'imitation d'opinion conduit à ce que l'imitation devienne impossible (Axelrod 1997 ; Deffuant *et al* 2002) ;
- Le processus de renforcement, à plusieurs échelles
  - L'amplification sur le désir de similarité (Schelling 1970),
  - Décision individuelle renforce des hiérarchies sans que personne ne souhaite les renforcer (Rouchier *et al* 2001),
  - Situations sous-optimales par imitation de comportement -des lock-in ou blocages ce qui est un résultat important concernant les marchés en particulier (Vriend 2000, Arthur 1988).

Les critiques concernant les prédictions avec ABM

+++ : On peut apprécier un élément *absent* des autres modélisations prédictives : rend possible une réflexion réelle sur la *DIFFUSION* comme processus basé sur les *interactions* (aspect social)

- rôle du réseau (forme, densité, clustering, hétérogénéité)
- taille de la population – pas d'extrapolation possible des résultats
- path-dependence plus fort que dans d'autres modèles (= dépendance au chemin – une même situation initiale peut avoir plusieurs conséquences très distinctes du fait de hasards historiques)

--- : Mais deux problèmes classiques :

- Grand jeu de paramètres et pas assez de données pour fitter
- Un problème général lors de l'utilisation des modèles de diffusion sur réseaux : la STRUCTURE du réseau doit être connue a minima…

Thiriot 2007

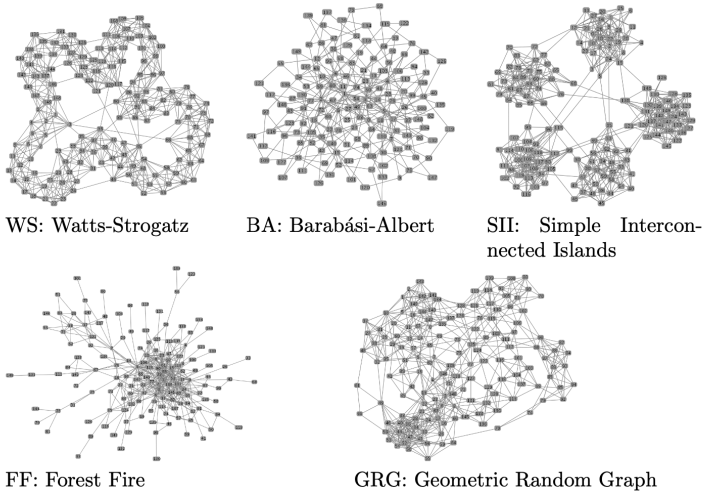


Figure 4: Examples of small-world networks created by each generator used in this paper.

« smallworld is not enough »(Thiriot, 2007) https://arxiv.org/abs/2002.03717

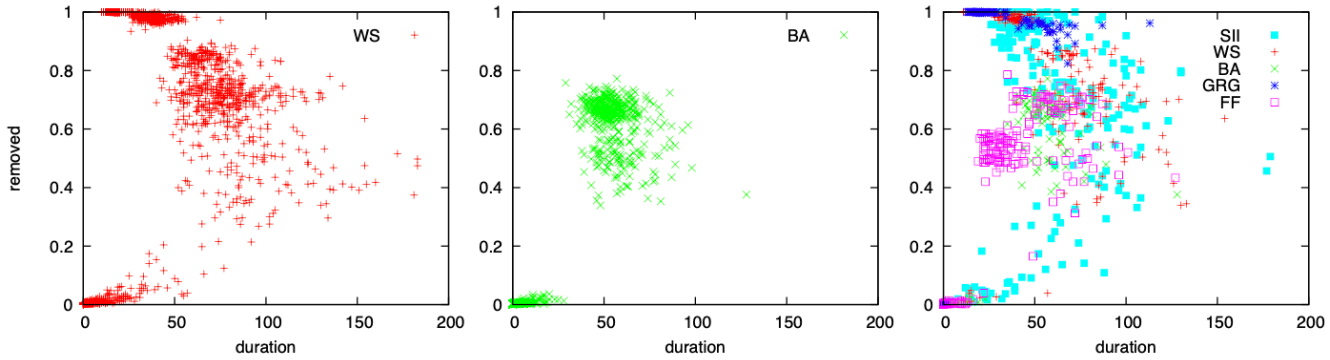


Figure 6: Dynamics $\mathcal{D}_{\mathrm{SIR}}$ of the SIR model supported by *(left)* Watts-Strogatz networks, *(center)* Barabási-Albert networks and *(right)* various small-world $\mathcal{X}^{qWS}$. Each dot in these figures represents the result of one simulation.

nature medicine

Explore Content · Journal Information · Publish With Us

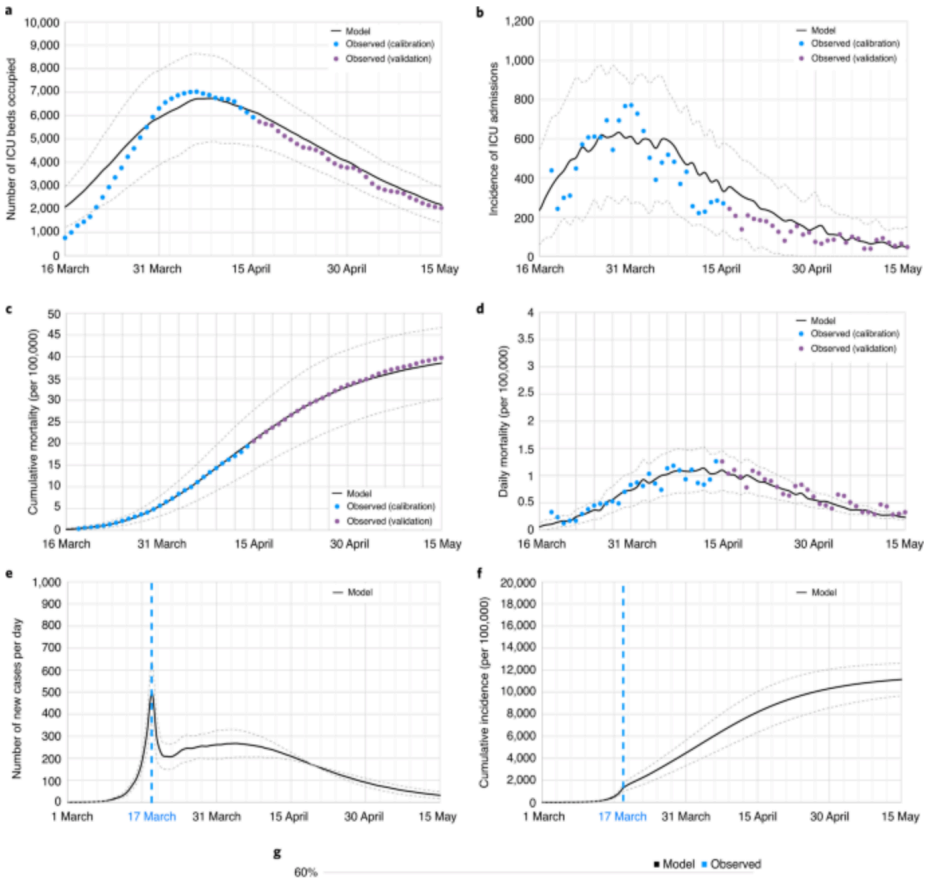
nature › nature medicine › letters › article

Letter | Published: 14 July 2020

A stochastic agent-based model of the SARS-CoV-2 epidemic in France

Nicolas Hoertel, Martin Blachier, Carlos Blanco, Mark Olfson, Marc Massetti, Marina Sánchez Rico, Frédéric Limosin & Henri Leleu

Fig. 1: Model-predicted and observed ICU-bed occupancy, daily ICU admissions, cumulative mortality, daily mortality and age distribution of deceased people, and model-predicted daily number of new cases and cumulative incidence.

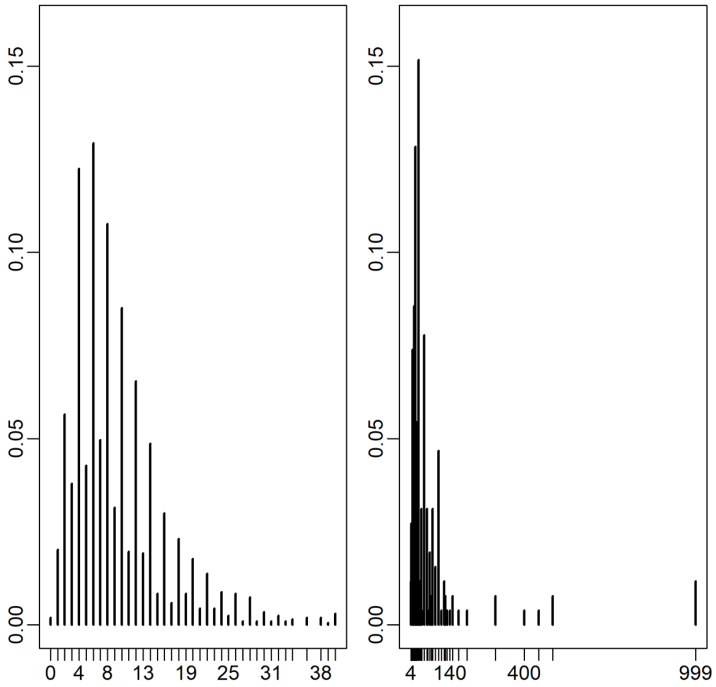


Paramètres et données

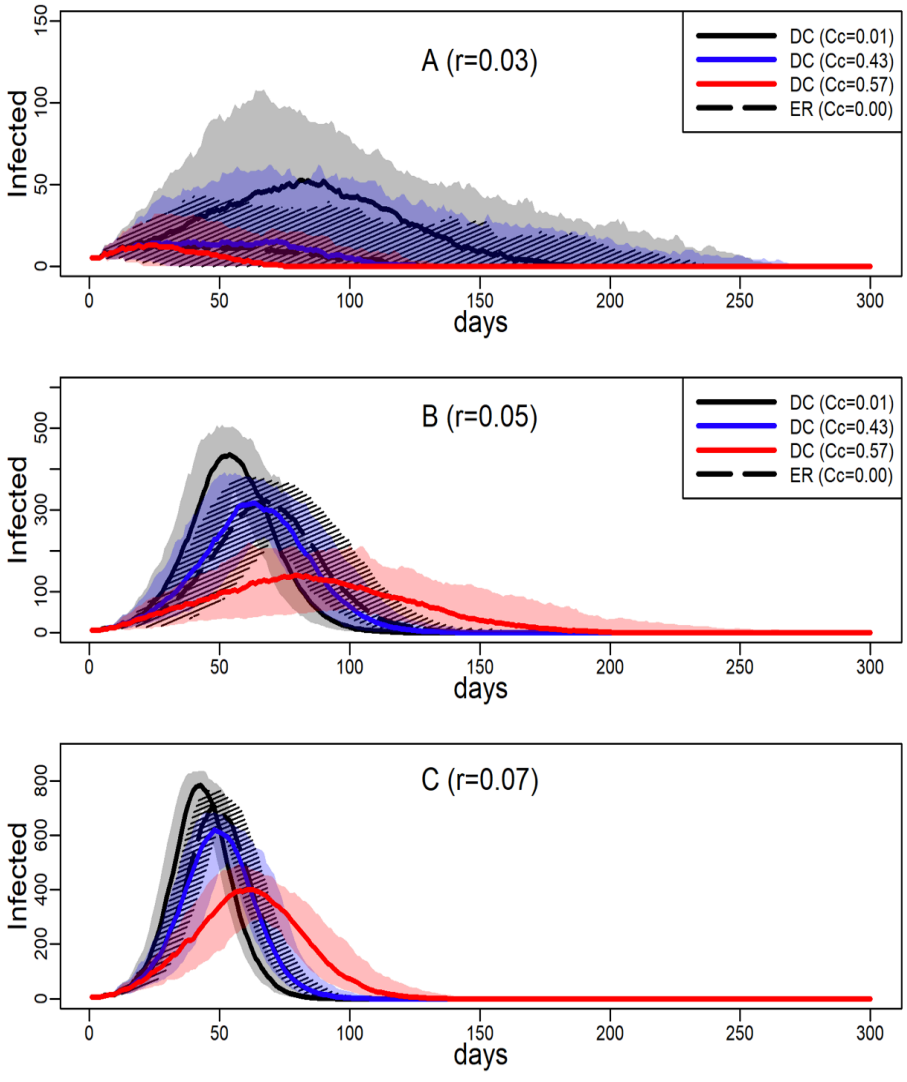
Of 194 model parameters, *192 were estimated from the literature (n = 170) or by assumption (n = 22), and 2 parameters, contamination risk (per min per m² of contact) and proportion of undiagnosed cases, could not be estimated and were therefore calibrated.* (…)
The stochastic ABM was run for 360 d on 500,000 individuals. The results were based on an average of 200 simulations. Analyses were performed on 17 May using data for model parameters until 15 April. Results were extrapolated to the French population of 67 million people. We provided uncertainty measures by using *200 bootstrap samples based on the random variation of all non-calibrated* parameters simultaneously, either within a 95% confidence interval for parameters estimated from the literature or within a ±20% interval if the parameter was assumed.

Représentation confinement :
70% des liens supprimés
1,2 fois au magasin où on rencontre 5 personnes en moyenne

Social contacts		
School class size (average)	30	Assumption
Proportion of small companies (<10 employees)	18.40%	INSEE, 2018
Number of colleagues in small companies (average)	2	Assumption
Number of colleagues in bigger companies (average)	10	Assumption
Employment rate (for people aged 20 to 65 years)	92%	INSEE, 2020
Shopping density (per 100,000 inhabitants)	29.3	Direction Générales des Entreprises (DGE), 2018
Number of shopping trips (average per week)	1.2	Assumption
Number of people met per shopping trip (average)	5	Assumption
Social network distance	22	Gilbert et al., 2010 [22]
Frequency of meeting friends (average per week)	1	Assumption
Event participations, i.e., museum, cinema, music and sport events (average per year)	5.4	Ministère de la Culture, 2015
Close encounters per event participation (average)	5	Assumption
Round trips with public transport (average per week) for workers	5	Assumption
Round trips with public transport (average per week) for non-workers	1.7	Assumption
Close encounters in public transport	3 to 5	Assumption, with work-related trips assumed to happen at peak times with more encounters
International contamination (average, per week)	1.8	Based on imported cases observed in France initially (Santé Publique France)



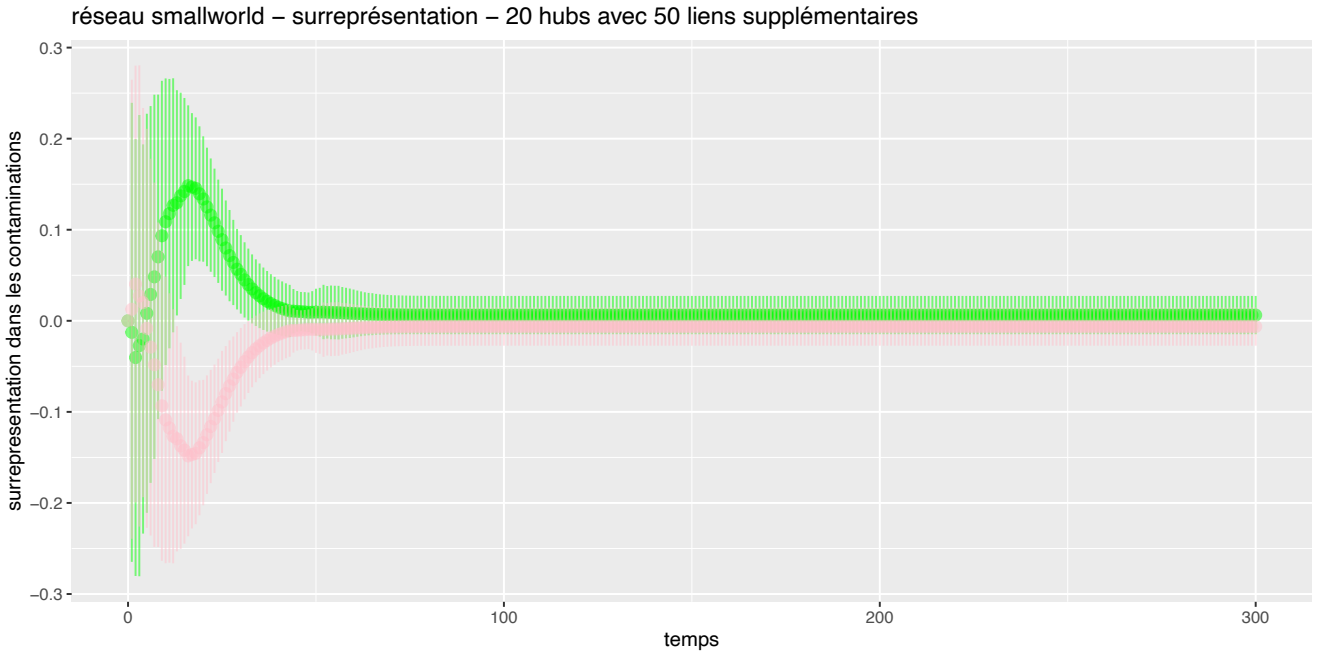
Gianluca Manzo et Arnout van de Rijt(2020)
http://jasss.soc.surrey.ac.uk/23/4/10.html

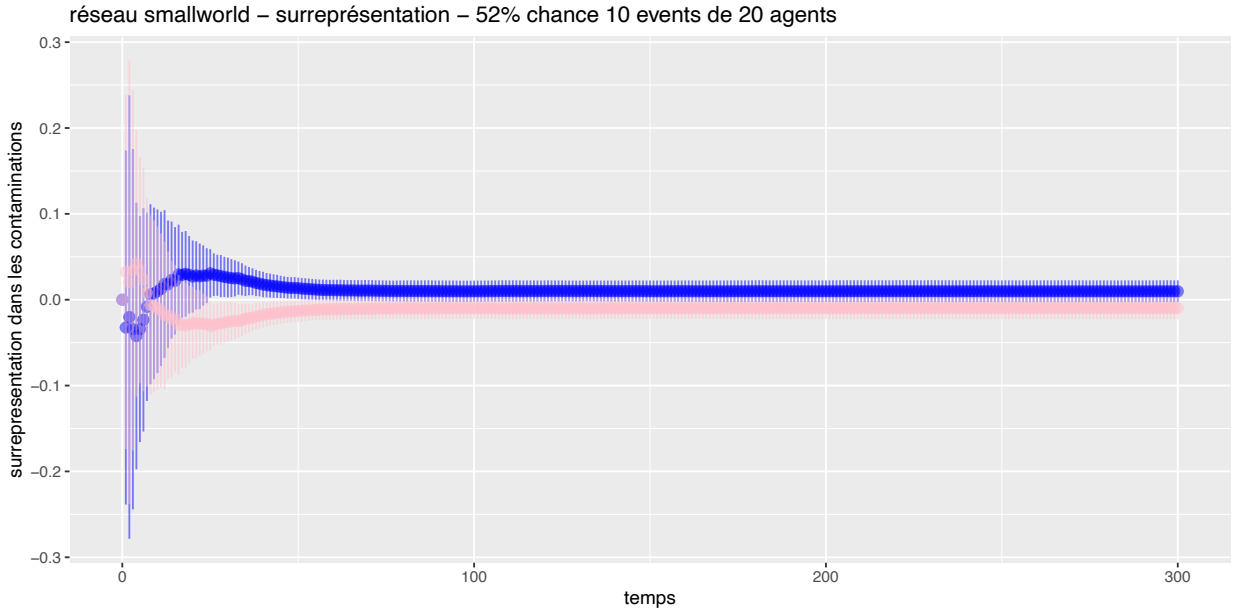


Importance de l'hétérogénéité des nombres de liens !!!!
« super-spreaders »

Conclusion sur « immunisation ciblée »
Bémol « efficacité »

Représenter / mesurer la contribution relatives des liens

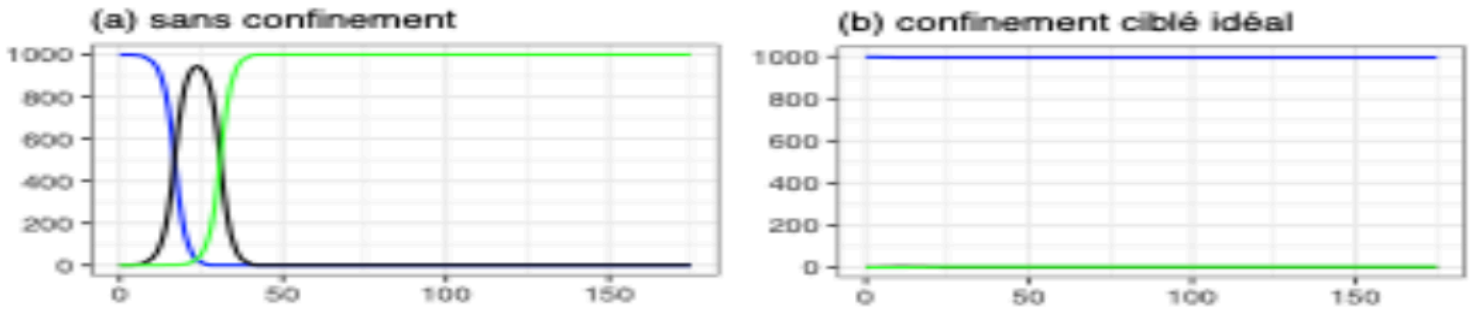




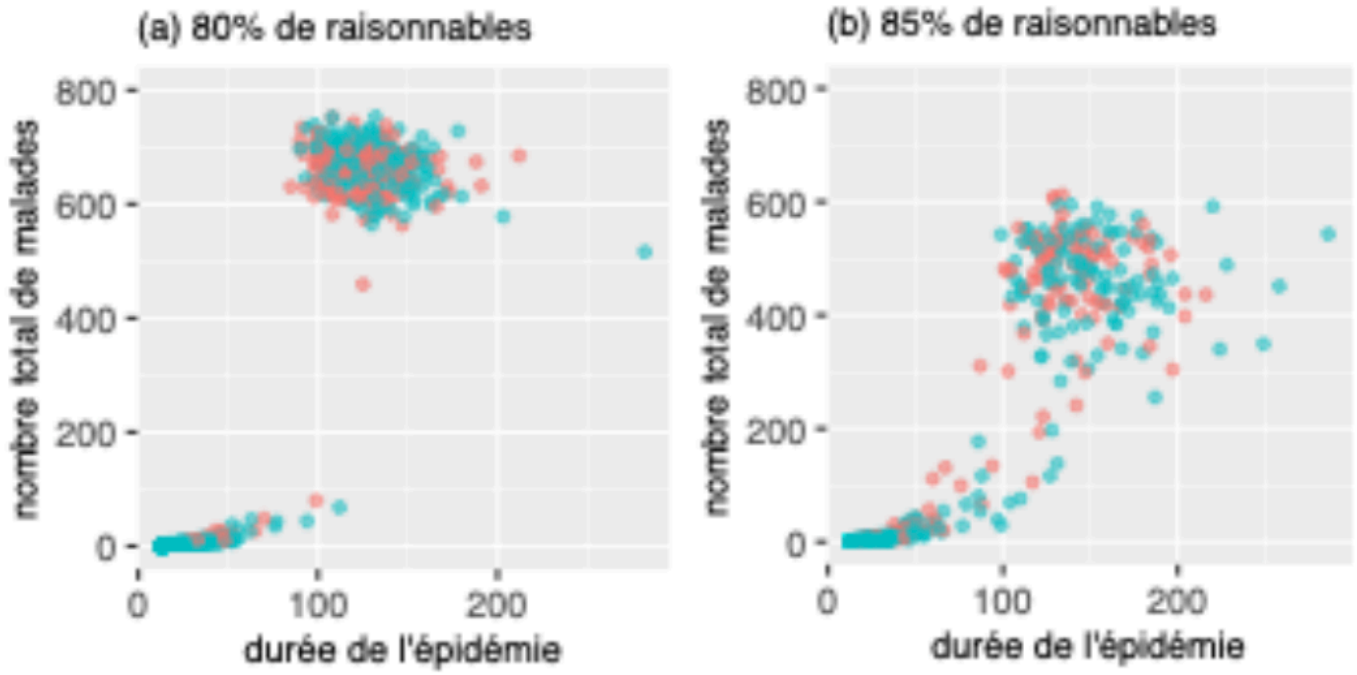
Faire un arbitrage : la nécessité de mobiliser plusieurs indicateurs

Modèle Suède

- Multi-critère minimaliste : le nombre de non-confinés comme une mesure d'activité économique
- Les malades guérissent – le choix de ne pas parler des morts est volontaire
- Probabilité de transmission : 20%.
- Durée de maladie et contagion 14 pas.
- Auto-confinement des malades (***Vulnérabilité du système***)
  - Réduction du réseau à 3 autres agents « à la maison »
  - % raisonnables [se confine + respecte les gestes barrière > division par 2 du risque]
  - Durée avant de détecter qu'on est malade

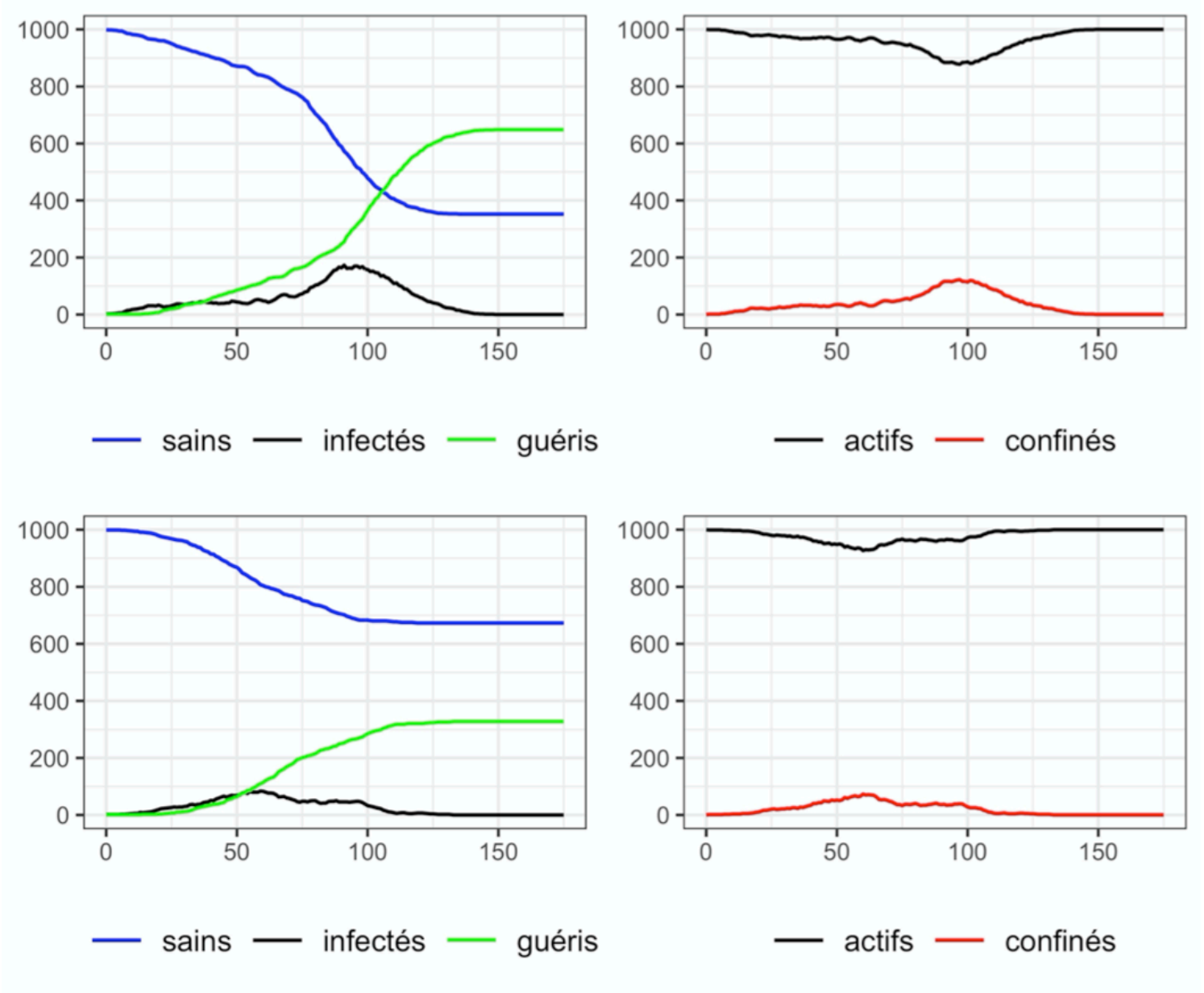


sains infectés guéris



agent zéro : non raisonnable raisonnable

Information parfaite



80

85

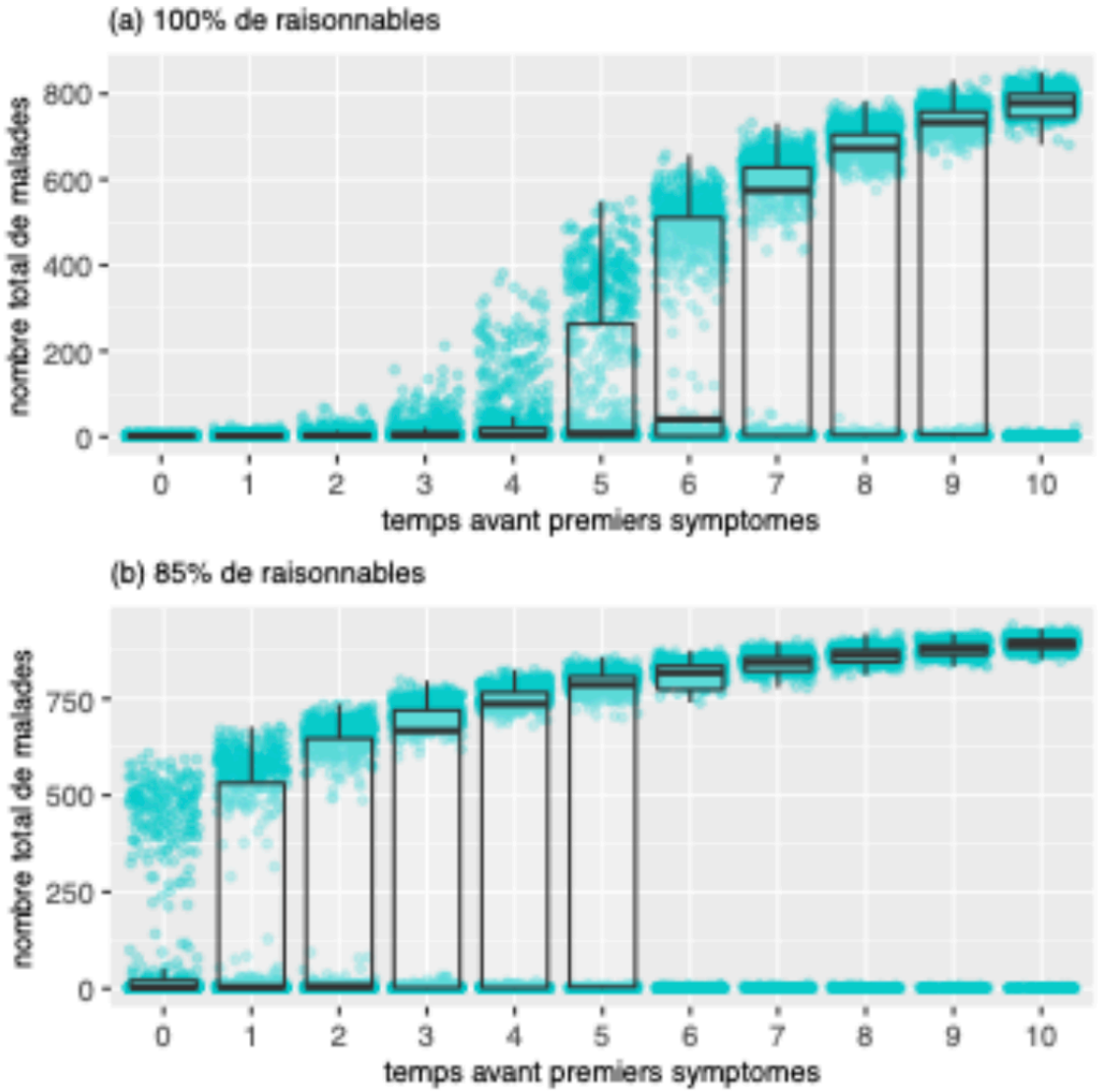
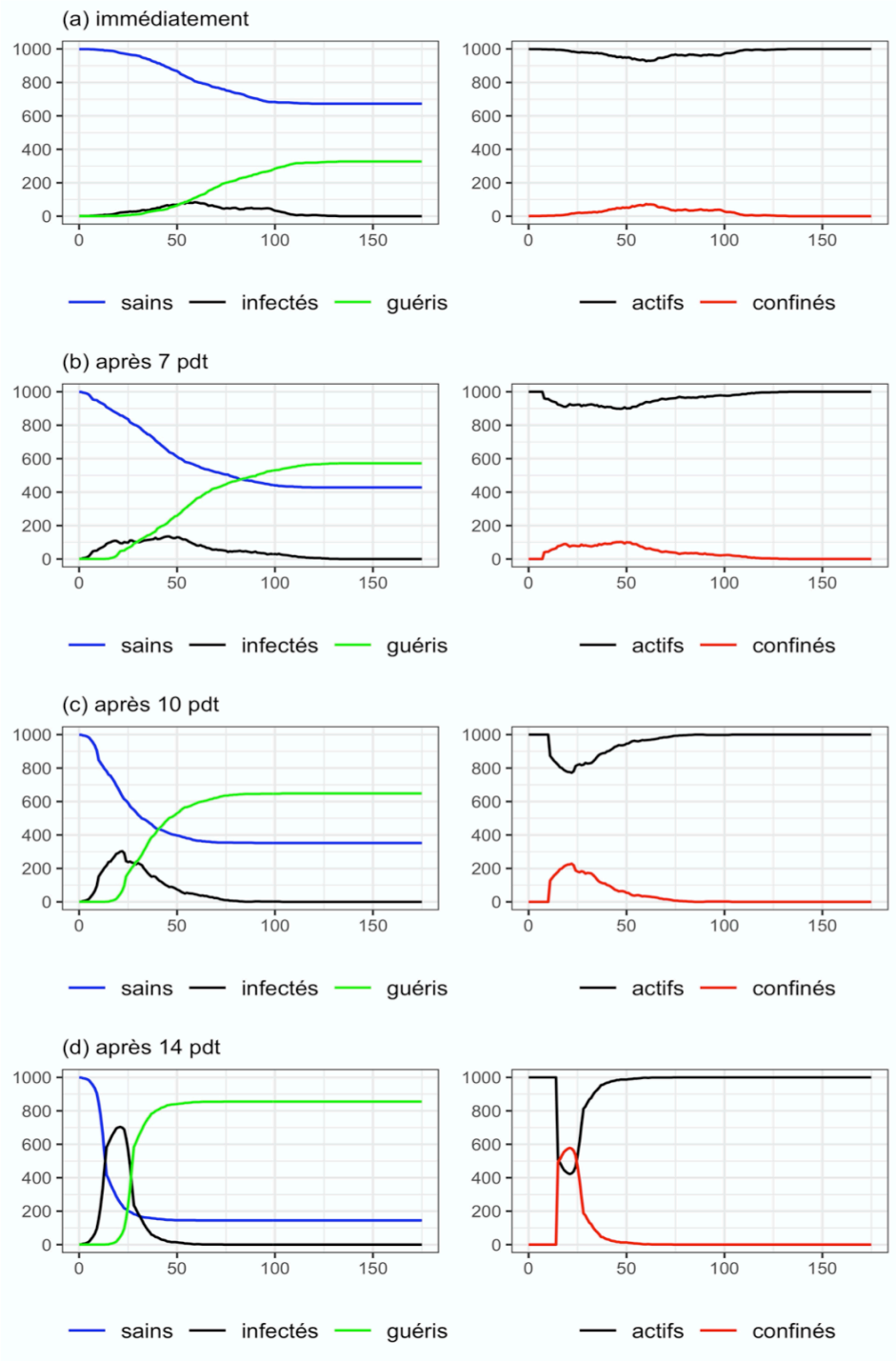


Figure 5 : *Pour 1000 agents, répartition du nombre total d'agents malades en fin d'épidémie en fonction du temps avant l'apparition des premiers symptômes (1000 simulations à chaque fois). En haut, avec 100% d'agents raisonnables, en bas, avec 85% d'agents raisonnables.*

Délai du démarrage du confinement ciblé



Les guéris ressortent ?

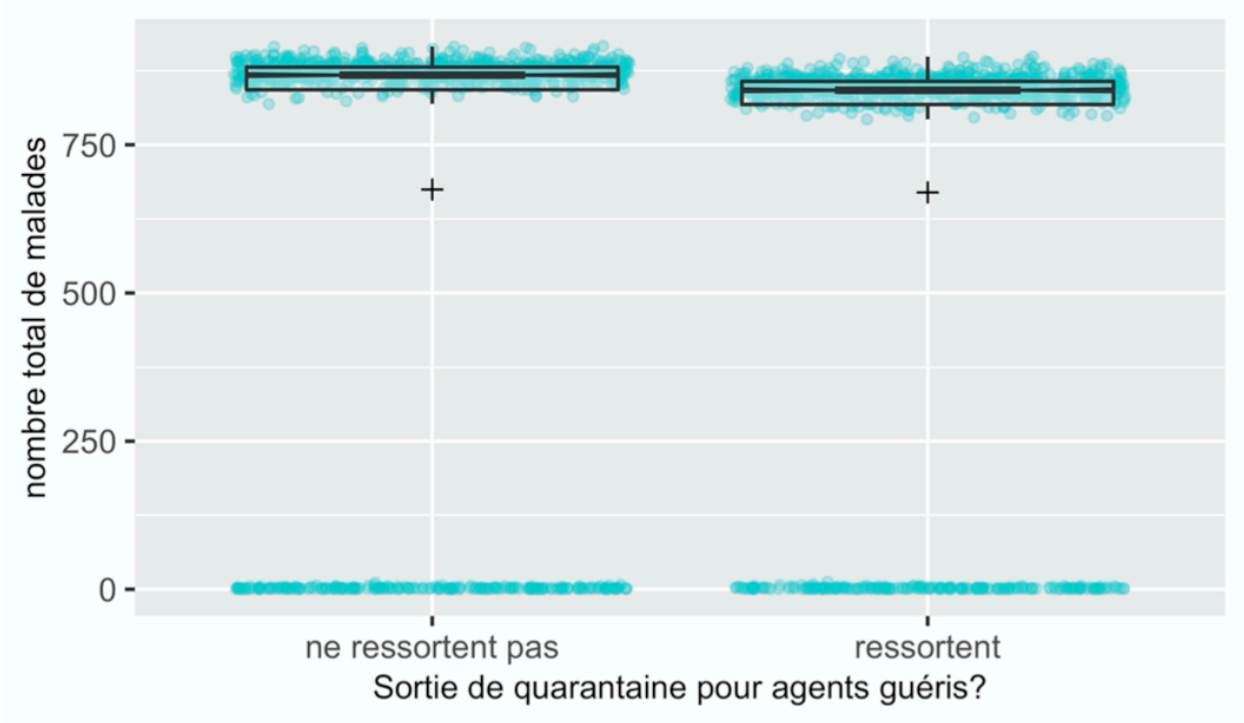


Figure 7 : *Répartition du nombre total d'agents malades en fin d'épidémie selon que les agents guéris restent confinés ou non - 1000 simulations à chaque fois. Les autres paramètres étaient fixés à 1000 agents, avec 100% d'agents raisonnables et 7 jours de retard dans l'apparition des symptômes. Les croix (+) représentent les moyennes de chaque série et les traits gras dans les "boites" leurs médianes.*

« Effet confinement » controversé

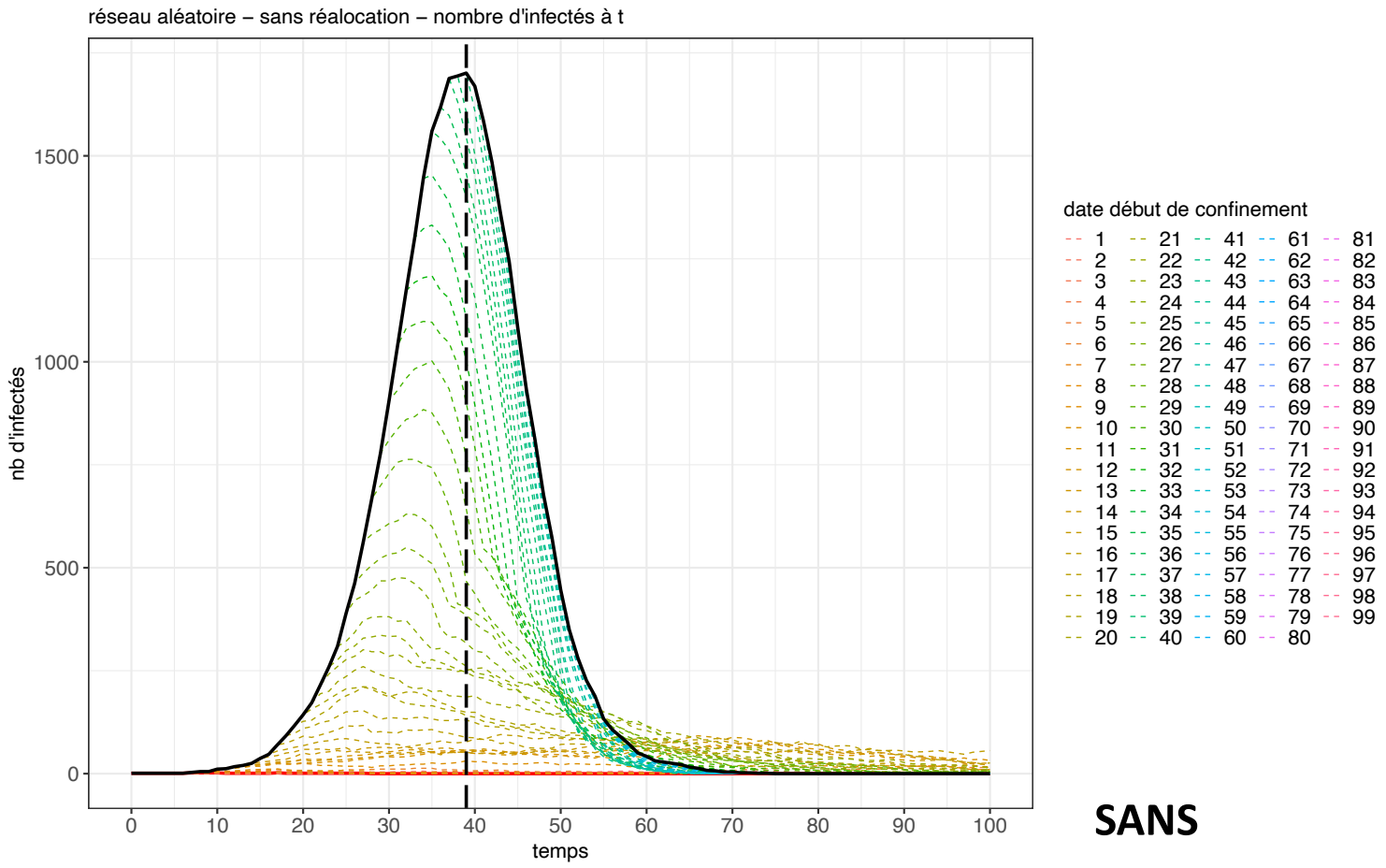
« Effet confinement »

Très discuté à l'international sous plusieurs formes de démonstration (controverse en cours)

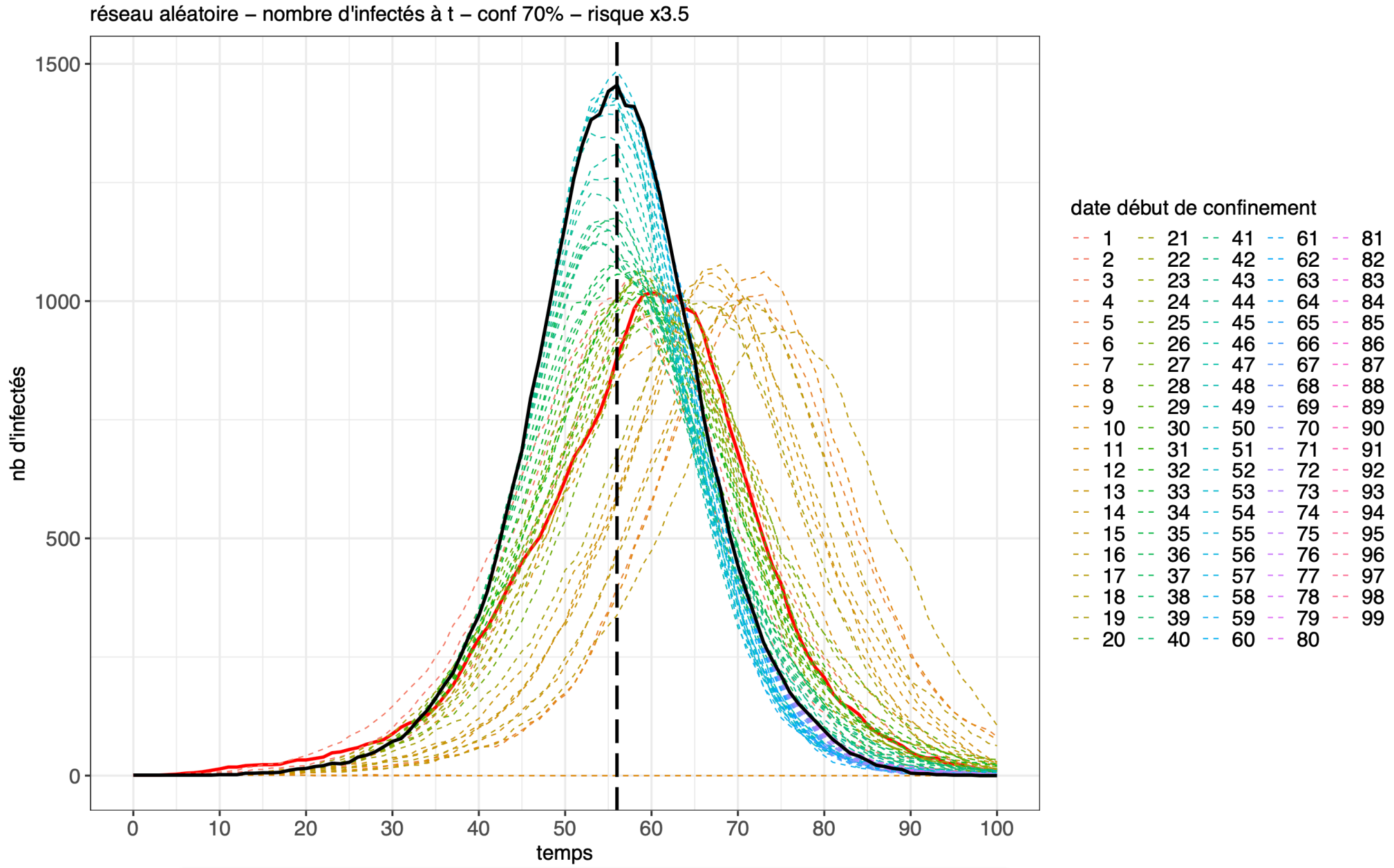
- Macro : comparaison internationale en fonction de la sévérité du confinement (Suède, « les deux Dakota »)
- Démonstration « ad hoc » -
    - « il y a n% de morts en moins que prévu par nos modèles »
    - ?? – modèles jamais validés par expérience naturelle, aucun fit.. ??
- Arbitrage de santé publique « années de vie perdues / gagnées »

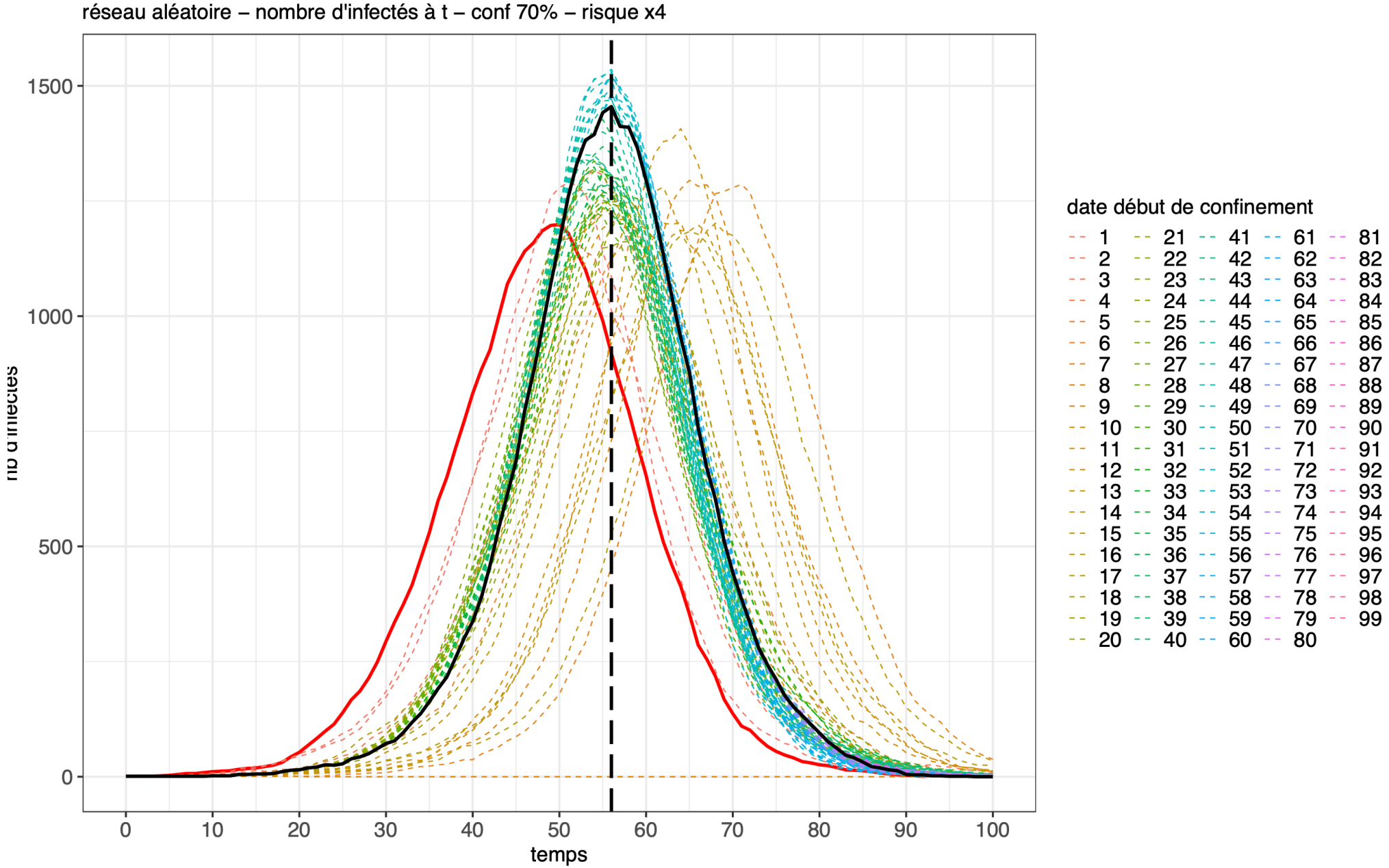
Quid du confinement trop tard ?

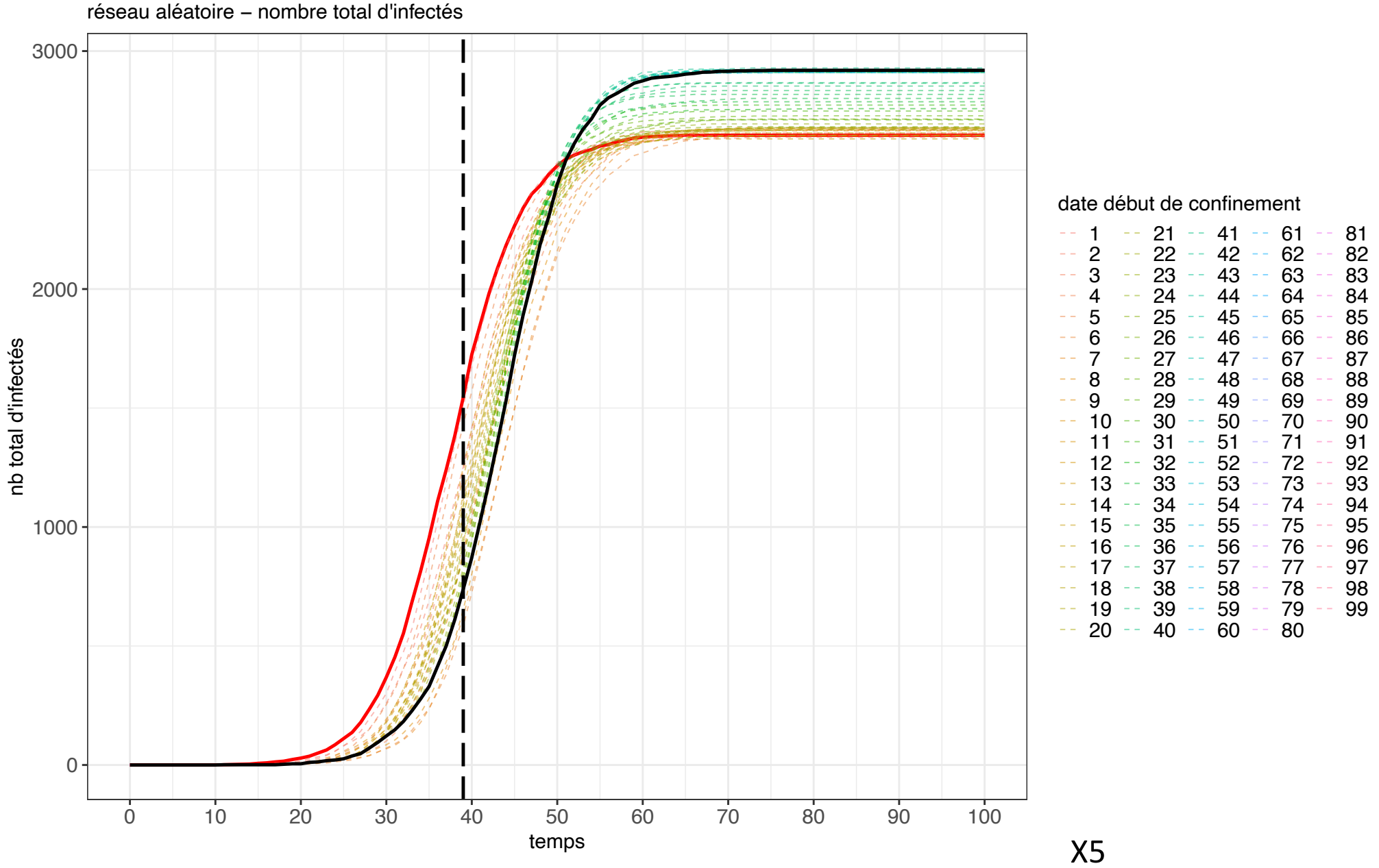
3700 agents - 10 liens en moyenne par agents - 0% d'immunisés
2.5 % de chance de transmission par interactions
10 pdt de contagiosité
70% de réduction de liens – avec ou sans augmentation d'intensité dans les relations conservées (maison, famille, travail)



SANS


réseau aléatoire – nombre d'infectés à t – conf 70% – risque x3.5
nb d'infectés
temps
0
500
1000
1500
0
10
20
30
40
50
60
70
80
90
100
date début de confinement
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20
21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40
41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60
61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80
81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99

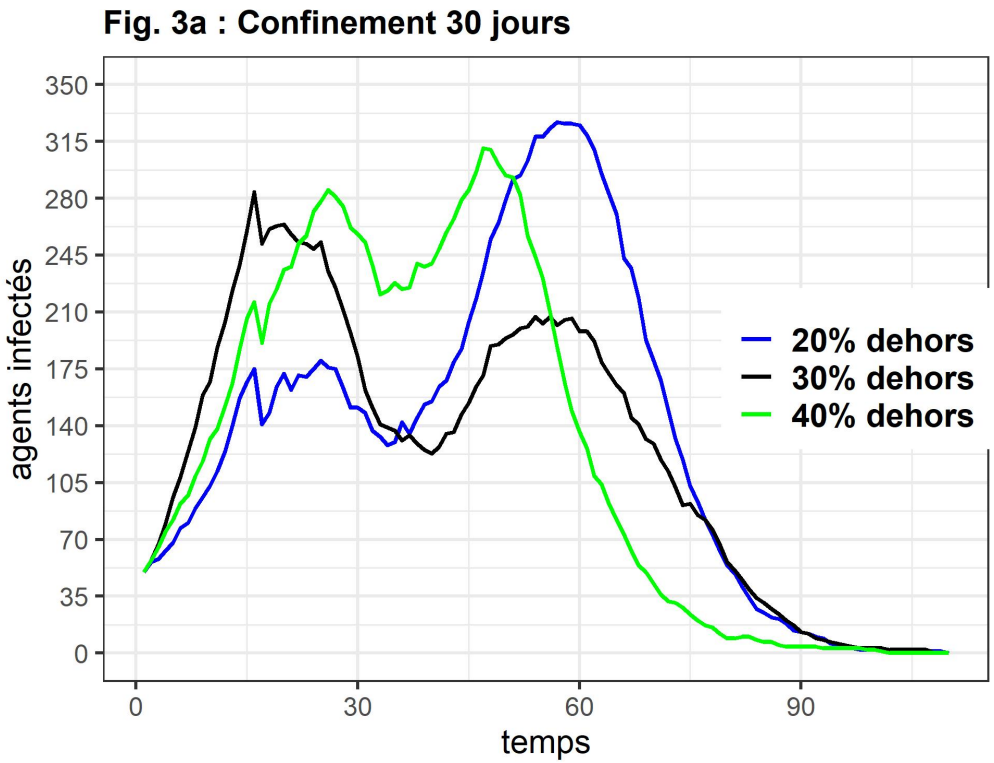

réseau aléatoire – nombre d'infectés à t – conf 70% – risque x4
nb d'infectés
1500
1000
500
0
0
10
20
30
40
50
60
70
80
90
100
temps
date début de confinement
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99


réseau aléatoire – nombre total d'infectés
nb total d'infectés
temps
3000
2000
1000
0
0
10
20
30
40
50
60
70
80
90
100
date début de confinement
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20
21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40
41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60
61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80
81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99

X5

Covid 19 : « Années gagnées sur la mort contre années perdues pour la vie »
Gaspard Koenig, Les échos.

Des vagues ou des cloches ?



Le rebond semble un artefact de modélisation : pourquoi les modèles n'ont pas été révisés après l'évidence qu'il y a une dynamique « biologique » ?
Notion de « vague » = métaphore adéquate ? (diffusion ! – « feu »)

Que faire des simulations qualitatives ?

Méthode : au-delà de l'évidence

- Modèle tout d'abord calé sur « l'évidence » : ce qu'on peut déduire logiquement (hausse contagiosité > hausse vitesse de diffusion)
- Être sûr de son modèle en variant les différentes dimensions et avoir des résultats logiques

PUIS : **créer les contrefactuels** – tirer les paramètres hors des zones de production de l'évidence

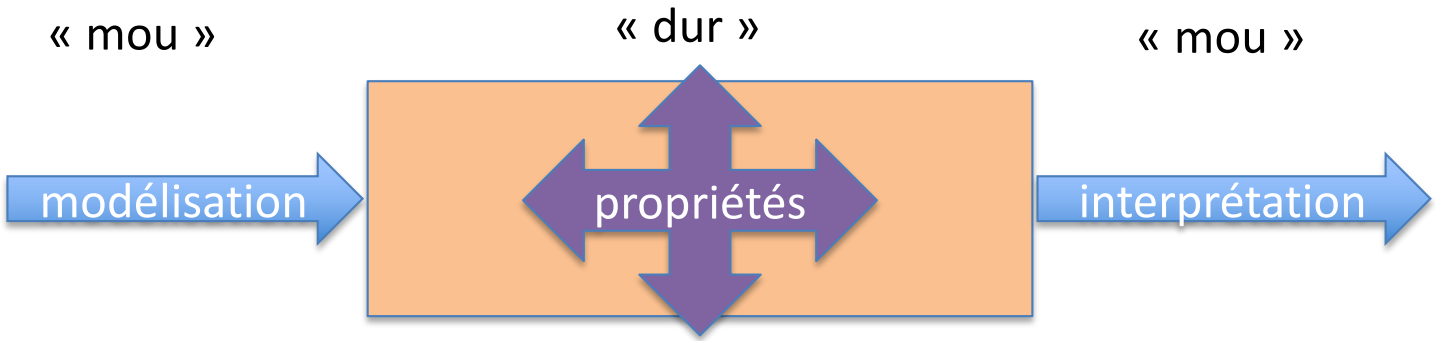
- Chercher des phénomènes surprenants ou non attendus
- Les interpréter, chercher plusieurs méthodes pour les produire, expliquer les processus….

Méthode : ressemblance / pertinence

Reconnaître la vérité en Science revient à « être convaincu »

- Nécessité d'un accord sur la **pertinence** d'un modèle – reconnaître qu'on peut déduire des choses du monde en l'utilisant / agir sur le monde grâce à lui
- La **ressemblance** relative peut-être un argument pour la « justesse » du modèle, qui semble très importante dans les rapports science-société

Ces deux dimensions résultent d'habitudes sociales, d'accords au sein des communautés de pratique, **pas de la vérité !**



Modèles qualitatifs : accompagner

- Bien préciser les questions et les éléments nécessaires à produire le résultat
  - Parfois c'est un trade-off entre plusieurs paramètres qu'on met en avant (Suède)
  - Parfois on peut chercher les hypothèses nécessaires pour obtenir un résultat (notre rebond)
  - Parfois on peut se poser des questions sur l'observation des phénomènes (super-spreading)
- « KISS » – en tentant de ne pas multiplier les paramètres
  - Contagiosité calée sur un R0 théorique entre 2 et 3 sans confinement
  - En fonction de la taille de la population : implique une contagiosité variable
- Un travail qui n'aurait eu de sens décisionnel *qu'intégré dans un dialogue* (qui n'a pas existé) – démonstration des possibles de la logique, des tensions entre paramètres, mais pas prédiction exacte … à doubler d'autres méthodes

Questions démocratiques

<u>Composition du conseil scientifique</u>

Économie ? Aide à la décision ? Statisticien ?

<u>Questions organisationnelles</u>

- Barbara Stiegler > nudge (BVA) – Mc Kinsey
  - Fabrique du consentement / manipulation des foules
- « Covid-19 : une crise organisationnelle » (Bergeron et al., octobre 2020)

<u>Plus généralement</u>

Dialogue scientifique polarisé et en berne (exemple : « origines »)

- Isabelle Stengers (esprit) > science simpliste qui interdit les discussions

  « Industrielle » - RCT, modèles, méta-analyses

Controverses multiples invisibilisées

Dénigrement / censure d'un nombre croissant de scientifiques de la recherche publique par des opérateurs privés (journalistes, GAFA)

Relation au risque et à l'incertitude

- Calculateurs personnels
  - « prévisions locales » https://cloudapps.france-bioinformatique.fr/covidici/
  - https://www.mediapart.fr/journal/france/120521/vaccins-covid-calculez-votre-balance-benefice-risque-personnelle
- RRR vs RRA (risque relatif vs risque absolu)
  - Papier original Pfizer : comparaison par groupes (vaccin / placebo)
  - Nombre d'infectés après n semaines
    - Deux variables aléatoires sont observées en une seule donnée (rencontre du virus + virus attrapé)
    - 95% - Résultat ne donne pas « la probabilité individuelle de ne pas être infecté » mais « la taille de population pour laquelle le vaccin est efficace »
    - Prévalence vers 1% > Réduction de risque individuel est vers 0,95%

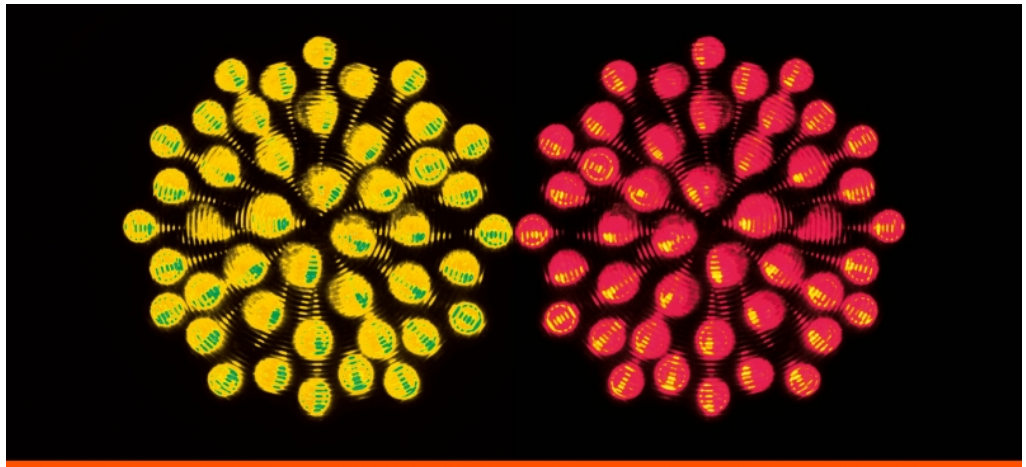
- Probabilités conditionnelles (faux-négatifs et faux-positifs)
- RRR vs RRA

Principes de l'aide à la décision

- Aide à la décision comme domaine de recherche
  - Formalisation : classements, indicateurs composés, vote >> *propriétés mathématiques*
  - Accompagnement : expression des préférences / valeurs, création d'alternatives de choix >> ***comparabilité des options***
  - Arbitrage : par le décideur rationnellement informé >> ***multi-critères***
- **Spécificités de l'aide à la décision publique**
  - **Légitimité, applicabilité - « acceptabilité »**
  - **« Accountability » - responsabilité transparente**

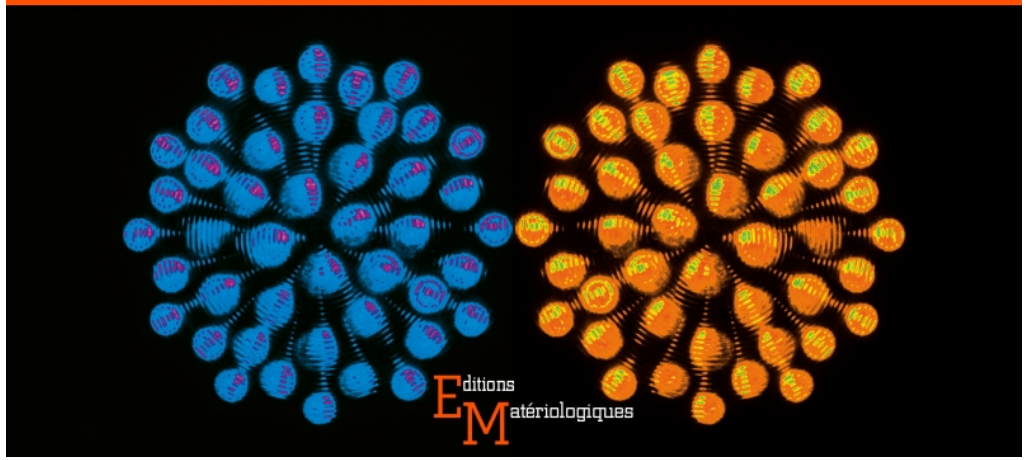
Conclusion

- Approche différente des modèles « phénoménologiques » (hospitalisation > hospitalisation)
  - Expression d'hypothèses, construction de contrefactuels, indicateurs de sortie multiples
  - Observer les « possibles » puis creuser les conditions d'émergence
- Approche *pédagogique* pour voir divers détails
  - questions classiques en ABM > effets déjà connus
  - plusieurs stratégies de modélisation > pas d'unicité de la pertinence / ressemblance
  - variation des hypothèses en fonction des connaissance car facilité de modifications - Cf Levitt ou Karl Friston, en modélisation rétrospective – révision de la susceptibilité (finalement moins d'infections que prévu – plusieurs explications)
- Création d'indicateurs / apparition d'effets contre-intuitifs / recherche d'explications / tests de contre-factuels / création d'indicateurs



Juliette **Rouchier** & Victorien **Barbet**

La diffusion de la Covid-19

Que peuvent les modèles ?

Préface de Franck Varenne