

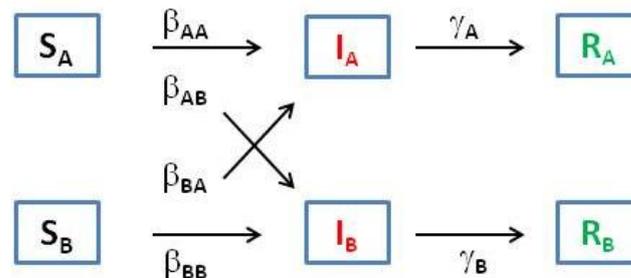


La sous-population invisible : un train peut en cacher un autre

<https://corona-circule.github.io/lettres/>

Nous avons, comme Santé Publique France, été surpris par la brutalité de l'arrivée de la seconde vague au début du mois d'octobre. Et nous avons formulé l'hypothèse que la contamination avait pu se répandre de façon très active dans une fraction de la population avec un faible taux de visibilité associé à la rareté des cas graves dans cette sous-population. Nous présentons ici un modèle simple à deux sous-populations basées sur les seules fractions S.I.R., qui permettra de tester la pertinence de cette hypothèse. Il nous restera à tenter d'analyser les données réelles de notre pays. Cet objectif plus ambitieux devra inclure un plus grand nombre de sous-populations (tranches d'âge par exemple) et inclure aussi, au minimum, la fraction décédés qui est la seule base statistique solide comme nous le répétons depuis le (presque) début de cette chronique.

Le schéma pertinent pour ce modèle simple tient compte d'une interaction possible entre les deux sous-populations au seul niveau de la contamination initiale. Le reste du parcours se fait sans interaction :



Le système d'équations différentielles couplées qui régit l'évolution des populations s'écrit alors :

$$dS_A/dt = - (\beta_{AA} I_A + \beta_{BA} I_B) S_A; \quad dI_A/dt = (\beta_{AA} I_A + \beta_{BA} I_B) S_A - \gamma_A I_A; \quad dR_A/dt = \gamma_A I_A$$

$$dS_B/dt = - (\beta_{BB} I_B + \beta_{AB} I_A) S_B; \quad dI_B/dt = (\beta_{BB} I_B + \beta_{AB} I_A) S_B - \gamma_B I_B; \quad dR_B/dt = \gamma_B I_B$$

La somme des dérivées de chaque sous-population est nulle, en accord avec

$S_A(t) + I_A(t) + R_A(t) = p_A$; $S_B(t) + I_B(t) + R_B(t) = p_B$ avec $p_A + p_B = 1$, où p_A, p_B sont les proportions de A et B dans la population totale.

Enfin, les cumuls infectés sont régis par $dIc_A/dt = (\beta_{AA} I_A + \beta_{BA} I_B) S_A$, $dIc_B/dt = (\beta_{BB} I_B + \beta_{AB} I_A) S_B$

Pour les matheux, ces équations peuvent s'écrire sous une forme matricielle qui aura l'avantage de rester compacte et élégante lorsqu'il faudra introduire un plus grand nombre de sous-populations. Le moment n'en est pas venu...

Sur cette base nous avons créé une version « SIR double » de notre fichier de travail Excel. Les caractéristiques des sous-populations A (active) et B (basique) seront choisies arbitrairement pour des simulations qui ne visent qu'à établir la pertinence du mécanisme.

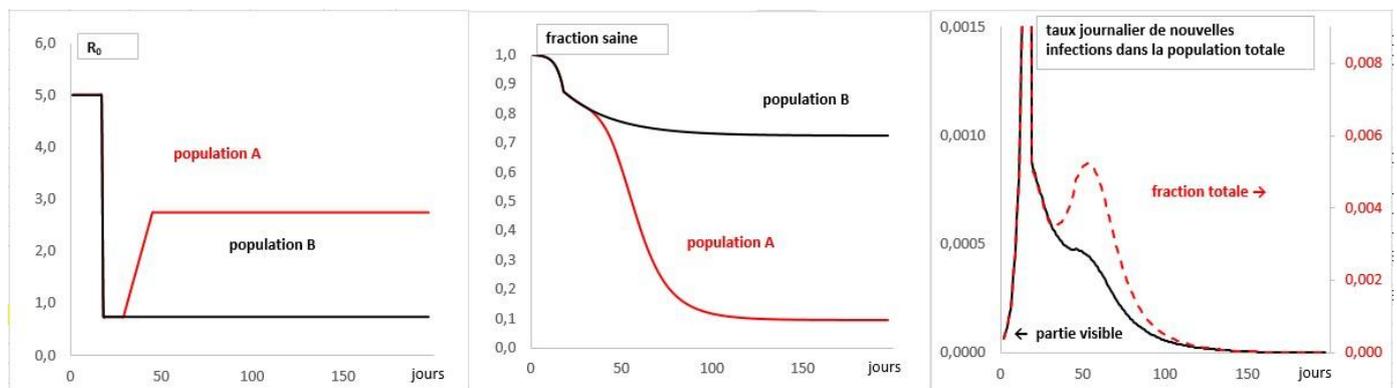
La dynamique choisie pour réaliser ces simulations a été simplifiée au maximum et comporte trois phases :

- 1- Démarrage exponentiel de l'épidémie avec les mêmes paramètres caractéristiques pour toute la population
- 2- Coup d'arrêt brutal dans l'ensemble de la population, toujours supposés uniforme
- 3- Reprise de la circulation du virus dans la sous-population active.

Pour le premier essai, nous avons supposé l'absence totale d'interactions entre les sous-populations. Le jeu de paramètres est résumé ci-dessous :

Sous-population	R_0 (1)	R_{eff} (2)	R_{eff} (3)	proportion	Taux de visibilité
A	5	0.75	2.75	0.2	0.05
B	5	0.75	0.75	0.8	0.2

Nous avons préféré mentionner le taux initial de reproduction $R_0 = \beta/\gamma$ plutôt que le taux effectif $R_{eff} = S \times R_0$. C'est ce premier taux qui représente la fréquence des contacts entre personnes, et c'est donc lui qu'il faut associer aux changements individuels de comportement.



La figure de gauche montre que nous avons choisi une évolution progressive entre les phases 2 et 3 de l'évolution de R_0 dans la sous-population active. Cette approche douce évite de voir apparaître des discontinuités appréciables dans les courbes de la figure de droite. *En effet, la dérivée dlc/dt est sensible aux problèmes numériques découlant de la discrétisation du temps dans le processus de résolution des équations différentielles; ces discontinuités déjà évoquées dans la lettre n°3 peuvent, par ailleurs, être réduites autant que nécessaire en abaissant la valeur de l'incrément de temps, prise à 1 jour par commodité.*

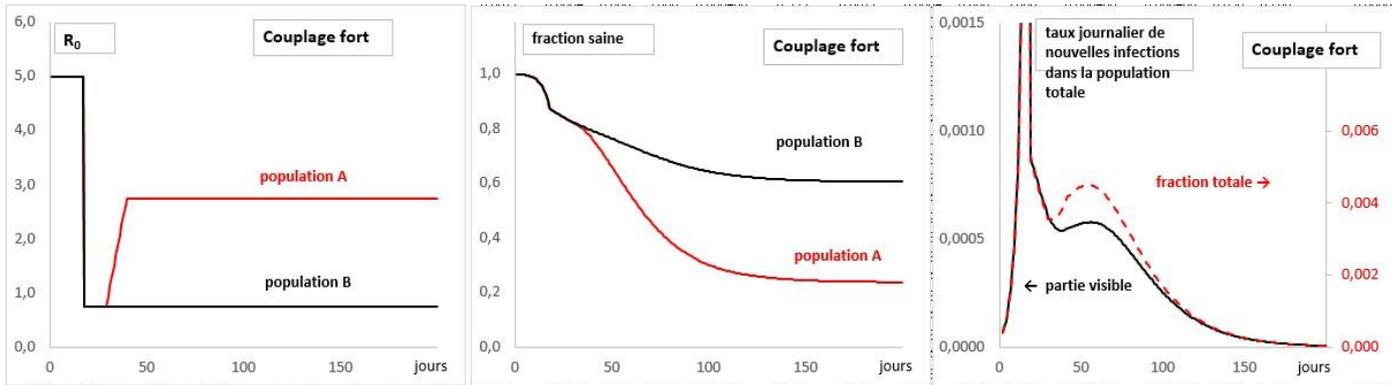
La figure centrale montre que, comme attendu, le changement de comportement de la sous-population A induit une reprise de l'épidémie dans cette même sous-population, mais pas dans l'autre.

La figure de droite tient compte des taux de visibilité différents des deux sous-populations. La nette reprise de l'épidémie (courbe pointillée rouge) est à peine visible sur l'ensemble de la population. C'est la combinaison de deux particularités : la sous-population active est à la fois minoritaire et sous-représentée.

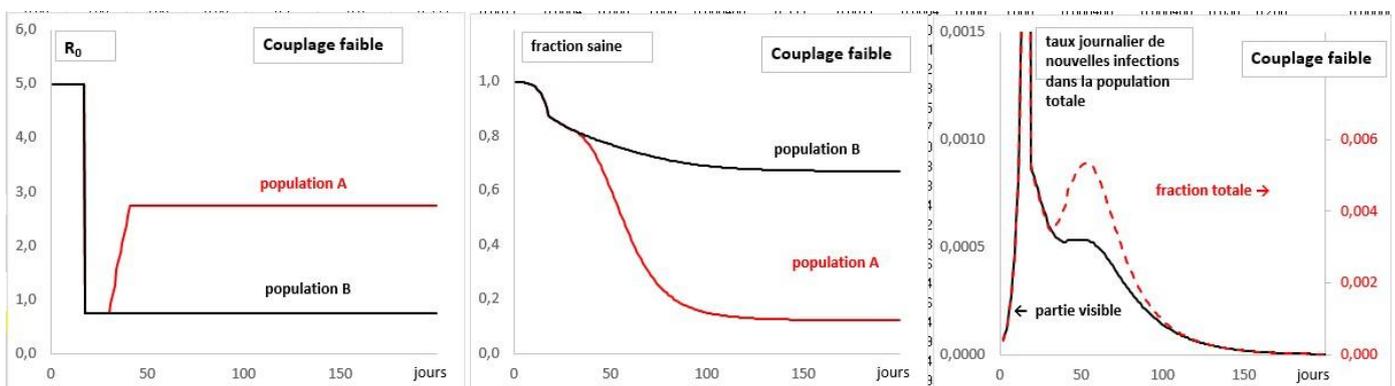
Les essais suivants ont pris en compte des interactions entre ces deux sous-populations dans la phase 3 de l'évolution. Par simplicité nous avons alors systématiquement pris $\gamma_{AB} / \gamma_{AA} = \gamma_{BA} / \gamma_{BB}$. Et, pour faciliter la comparaison aux résultats de l'essai sans interactions : $\gamma_{AA} + \gamma_{AB} = \gamma_{AA}$ (essai 1) ; $\gamma_{BB} + \gamma_{BA} = \gamma_{BB}$ (essai 1) ; et c'est la somme de ces paramètres d'interaction, $\gamma_{AA} + \gamma_{AB}$ et $\gamma_{BB} + \gamma_{BA} = \gamma_{BB}$ que nous avons reportée sur les figures de gauche qui seront présentées.

Nous avons considéré deux valeurs possibles du rapport des interactions :

$\gamma_{AB} / \gamma_{AA} = 1/2$, situation que nous pourrions qualifier de « **couplage fort** »:



$\gamma_{AB} / \gamma_{AA} = 1/10$, qui sera qualifié de « **couplage faible** »



En couplage fort, la contamination de la sous-population A entraîne, avec retard, la contamination du reste de la population, qui redevient donc visible. Avec un couplage plus faible, ces effets de contamination croisée sont atténués et le rebond cesse d'être visible. La faible visibilité du rebond a donc nécessité que la sous-population active, minoritaire et peu visible, ait eu relativement peu d'interactions avec le reste de la population.

Nous avons donc démontré la pertinence d'un mécanisme qui a pu retarder la prise de conscience de la gravité de la seconde vague. Nous ignorons pour l'instant dans quelle direction chercher cette (ou ces) sous-populations actives. S'agit-il d'une ou plusieurs régions en avance dans cette seconde flambée ? Ou bien d'une (ou plusieurs) tranches d'âge, par exemple impliquées dans le soutien de l'activité économique ? Ou peu concernées par la lutte contre la pandémie ? Nous allons nous adresser aux données du site GEODES, qui sont effectivement déclinées par tranches d'âge, ou par régions et départements et devraient permettre de conclure au bout d'un long travail de recherche. En effet, il faudra passer à un autre niveau de complexité du modèle, pour lequel l'ajustement des nombreux paramètres $\beta(t)$ passera par des méthodes statistiques telles que le recuit simulé (Méthode de Metropolis). Nous espérons pouvoir conclure avant l'arrivée massive des vaccins... à suivre donc.

Portez-vous bien, en respectant au mieux le confinement, les gestes barrière, en particulier le masque partout où il est requis. Et naturellement n'oubliez pas le nettoyage fréquent de vos mains.

Et continuons à réfléchir à une stratégie prudente pour les fêtes de fin d'année. L'ouverture à tous des tests PCR serait une bonne nouvelle.

François VARRET, Physicien, Professeur Emérite à l'Université de Versailles Saint-Quentin

Mathilde VARRET, Chargée de Recherche INSERM (Génétique, Biologie) Hôpital Bichat.