

Du bon (ou du mauvais) choix des statistiques à la lumière de la pandémie

Hervé Seitz (herve.seitz@igh.cnrs.fr)

IGH (UMR 9002 CNRS et univ. de Montpellier)

13 mars 2021

Ce diaporama est accessible à :

<http://www.igh.cnrs.fr/equip/Seitz/CZLR.pdf>

Statistiques et sciences expérimentales

Rumeurs pseudo-scientifiques sur le Covid-19

Sélection et invention de données

Tests d'hypothèses multiples

Publication de données non finalisées

Catégorisation fautive

Significativité, ampleur de l'effet, et effectifs

Coincidence attendue

Données non représentatives

Résultats tronqués

Pas de patient fractionnaire

Biais de composition des cohortes

Interprétation orientée

Le deuxième pic épidémique

Variants et vitesse de propagation

Conclusion

Statistiques et sciences expérimentales

Du bon (ou du mauvais) choix des statistiques à la lumière de la pandémie

Statistiques et sciences expérimentales

Rumeurs pseudo-scientifiques sur le Covid-19

Sélection et invention de données

Tests d'hypothèses multiples

Publication de données non finalisées

Catégorisation fautive

Significativité, ampleur de l'effet, et effectifs

Coincidence attendue

Données non représentatives

Résultats tronqués

Pas de patient fractionnaire

Biais de composition des cohortes

Interprétation orientée

Le deuxième pic épidémique

Variants et vitesse de propagation

Conclusion

Statistiques et sciences expérimentales

Irreproductibilité de la mesure : si on effectue deux fois la même mesure (même sur le même objet), les valeurs ne seront pas parfaitement identiques, avec tous leurs chiffres après la virgule.

Du bon (ou du mauvais) choix des statistiques à la lumière de la pandémie

Statistiques et sciences expérimentales

Rumeurs pseudo-scientifiques sur le Covid-19

Sélection et invention de données

Tests d'hypothèses multiples

Publication de données non finalisées

Catégorisation fautive

Significativité, ampleur de l'effet, et effectifs

Coincidence attendue

Données non représentatives

Résultats tronqués

Pas de patient fractionnaire

Biais de composition des cohortes

Interprétation orientée

Le deuxième pic épidémique

Variants et vitesse de propagation

Conclusion

Statistiques et sciences expérimentales

Irreproductibilité de la mesure : si on effectue deux fois la même mesure (même sur le même objet), les valeurs ne seront pas parfaitement identiques, avec tous leurs chiffres après la virgule.

Origine : imprécisions dans l'expérience ; variabilité des objets mesurés.

Statistiques et sciences expérimentales

Rumeurs pseudo-scientifiques sur le Covid-19

Sélection et invention de données

Tests d'hypothèses multiples

Publication de données non finalisées

Catégorisation fautive

Significativité, ampleur de l'effet, et effectifs

Coincidence attendue

Données non représentatives

Résultats tronqués

Pas de patient fractionnaire

Biais de composition des cohortes

Interprétation orientée

Le deuxième pic épidémique

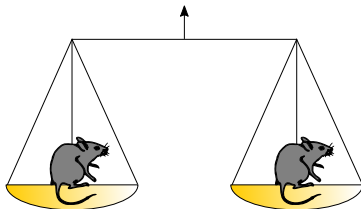
Variants et vitesse de propagation

Conclusion

Statistiques et sciences expérimentales

Irreproductibilité de la mesure : si on effectue deux fois la même mesure (même sur le même objet), les valeurs ne seront pas parfaitement identiques, avec tous leurs chiffres après la virgule.

Origine : imprécisions dans l'expérience ; variabilité des objets mesurés.



Statistiques et sciences expérimentales

Rumeurs pseudo-scientifiques sur le Covid-19

Sélection et invention de données

Tests d'hypothèses multiples

Publication de données non finalisées

Catégorisation fautive

Significativité, ampleur de l'effet, et effectifs

Coincidence attendue

Données non représentatives

Résultats tronqués

Pas de patient fractionnaire

Biais de composition des cohortes

Interprétation orientée

Le deuxième pic épidémique

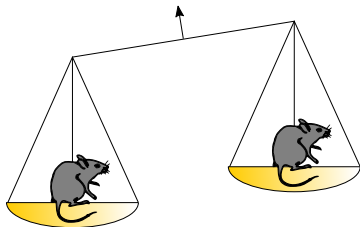
Variants et vitesse de propagation

Conclusion

Statistiques et sciences expérimentales

Irreproductibilité de la mesure : si on effectue deux fois la même mesure (même sur le même objet), les valeurs ne seront pas parfaitement identiques, avec tous leurs chiffres après la virgule.

Origine : imprécisions dans l'expérience ; variabilité des objets mesurés.



Statistiques et sciences expérimentales

Rumeurs pseudo-scientifiques sur le Covid-19

Sélection et invention de données

Tests d'hypothèses multiples

Publication de données non finalisées

Catégorisation fautive

Significativité, ampleur de l'effet, et effectifs

Coincidence attendue

Données non représentatives

Résultats tronqués

Pas de patient fractionnaire

Biais de composition des cohortes

Interprétation orientée

Le deuxième pic épidémique

Variants et vitesse de propagation

Conclusion

Statistiques et sciences expérimentales

Irreproductibilité de la mesure : si on effectue deux fois la même mesure (même sur le même objet), les valeurs ne seront pas parfaitement identiques, avec tous leurs chiffres après la virgule.

Origine : imprécisions dans l'expérience ; variabilité des objets mesurés.

—> Comparer deux conditions expérimentales : je sais d'avance que les résultats seront différents !

Statistiques et sciences expérimentales

Rumeurs pseudo-scientifiques sur le Covid-19

Sélection et invention de données

Tests d'hypothèses multiples

Publication de données non finalisées

Catégorisation fautive

Significativité, ampleur de l'effet, et effectifs

Coincidence attendue

Données non représentatives

Résultats tronqués

Pas de patient fractionnaire

Biais de composition des cohortes

Interprétation orientée

Le deuxième pic épidémique

Variants et vitesse de propagation

Conclusion

Statistiques et sciences expérimentales

Principe de l'analyse : faire plusieurs mesures dans chaque condition, et voir à quel point la variabilité inter-groupe est supérieure à la variabilité intra-groupe.

Du bon (ou du mauvais) choix des statistiques à la lumière de la pandémie

Statistiques et sciences expérimentales

Rumeurs pseudo-scientifiques sur le Covid-19

Sélection et invention de données

Tests d'hypothèses multiples

Publication de données non finalisées

Catégorisation fautive

Significativité, ampleur de l'effet, et effectifs

Coincidence attendue

Données non représentatives

Résultats tronqués

Pas de patient fractionnaire

Biais de composition des cohortes

Interprétation orientée

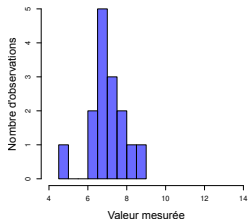
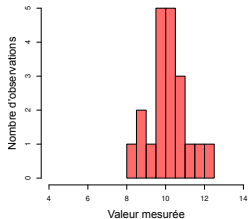
Le deuxième pic épidémique

Variants et vitesse de propagation

Conclusion

Statistiques et sciences expérimentales

Du bon (ou du mauvais) choix des statistiques à la lumière de la pandémie



Statistiques et sciences expérimentales

Rumeurs pseudo-scientifiques sur le Covid-19

Sélection et invention de données

Tests d'hypothèses multiples

Publication de données non finalisées

Catégorisation fautive

Significativité, ampleur de l'effet, et effectifs

Coincidence attendue

Données non représentatives

Résultats tronqués

Pas de patient fractionnaire

Biais de composition des cohortes

Interprétation orientée

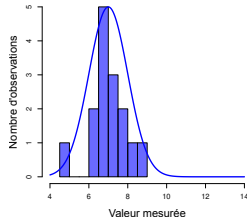
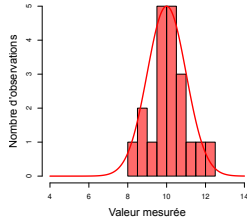
Le deuxième pic épidémique

Variants et vitesse de propagation

Conclusion

Statistiques et sciences expérimentales

Du bon (ou du mauvais) choix des statistiques à la lumière de la pandémie



Statistiques et sciences expérimentales

Rumeurs pseudo-scientifiques sur le Covid-19

Sélection et invention de données

Tests d'hypothèses multiples

Publication de données non finalisées

Catégorisation fautive

Significativité, ampleur de l'effet, et effectifs

Coincidence attendue

Données non représentatives

Résultats tronqués

Pas de patient fractionnaire

Biais de composition des cohortes

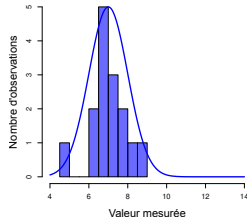
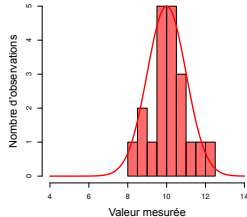
Interprétation orientée

Le deuxième pic épidémique

Variants et vitesse de propagation

Conclusion

Statistiques et sciences expérimentales



Probabilité que les populations infinies échantillonnées aient la même moyenne : $\approx 0,000000000004$.

Du bon (ou du mauvais) choix des statistiques à la lumière de la pandémie

Statistiques et sciences expérimentales

Rumeurs pseudo-scientifiques sur le Covid-19

Sélection et invention de données

Tests d'hypothèses multiples

Publication de données non finalisées

Catégorisation fautive

Significativité, ampleur de l'effet, et effectifs

Coincidence attendue

Données non représentatives

Résultats tronqués

Pas de patient fractionnaire

Biais de composition des cohortes

Interprétation orientée

Le deuxième pic épidémique

Variants et vitesse de propagation

Conclusion

Statistiques et sciences expérimentales

La « significativité » d'une différence (\neq amplitude) : mesure la confiance qu'on a dans cette différence mesurée.

Du bon (ou du mauvais) choix des statistiques à la lumière de la pandémie

Statistiques et sciences expérimentales

Rumeurs pseudo-scientifiques sur le Covid-19

Sélection et invention de données

Tests d'hypothèses multiples

Publication de données non finalisées

Catégorisation fautive

Significativité, ampleur de l'effet, et effectifs

Coincidence attendue

Données non représentatives

Résultats tronqués

Pas de patient fractionnaire

Biais de composition des cohortes

Interprétation orientée

Le deuxième pic épidémique

Variants et vitesse de propagation

Conclusion

Statistiques et sciences expérimentales

La « significativité » d'une différence (\neq amplitude) : mesure la confiance qu'on a dans cette différence mesurée.

Condition d'applicabilité du *t-test* : les données sont échantillonnées à partir d'une loi normale.

Statistiques et sciences expérimentales

Rumeurs pseudo-scientifiques sur le Covid-19

Sélection et invention de données

Tests d'hypothèses multiples

Publication de données non finalisées

Catégorisation fautive

Significativité, ampleur de l'effet, et effectifs

Coincidence attendue

Données non représentatives

Résultats tronqués

Pas de patient fractionnaire

Biais de composition des cohortes

Interprétation orientée

Le deuxième pic épidémique

Variants et vitesse de propagation

Conclusion

Statistiques et sciences expérimentales

La « significativité » d'une différence (\neq amplitude) : mesure la confiance qu'on a dans cette différence mesurée.

Condition d'applicabilité du *t-test* : les données sont échantillonnées à partir d'une loi normale.

Variantes : mesures appariées

Statistiques et sciences expérimentales

Rumeurs pseudo-scientifiques sur le Covid-19

Sélection et invention de données

Tests d'hypothèses multiples

Publication de données non finalisées

Catégorisation fautive

Significativité, ampleur de l'effet, et effectifs

Coincidence attendue

Données non représentatives

Résultats tronqués

Pas de patient fractionnaire

Biais de composition des cohortes

Interprétation orientée

Le deuxième pic épidémique

Variants et vitesse de propagation

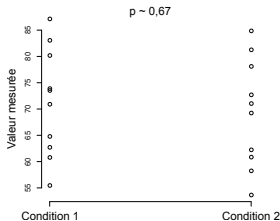
Conclusion

Statistiques et sciences expérimentales

La « significativité » d'une différence (\neq amplitude) : mesure la confiance qu'on a dans cette différence mesurée.

Condition d'applicabilité du *t-test* : les données sont échantillonnées à partir d'une loi normale.

Variantes : mesures appariées



Statistiques et sciences expérimentales

Rumeurs pseudo-scientifiques sur le Covid-19

Sélection et invention de données

Tests d'hypothèses multiples

Publication de données non finalisées

Catégorisation fautive

Significativité, ampleur de l'effet, et effectifs

Coincidence attendue

Données non représentatives

Résultats tronqués

Pas de patient fractionnaire

Biais de composition des cohortes

Interprétation orientée

Le deuxième pic épidémique

Variants et vitesse de propagation

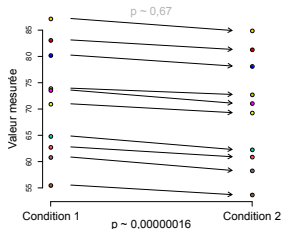
Conclusion

Statistiques et sciences expérimentales

La « significativité » d'une différence (\neq amplitude) : mesure la confiance qu'on a dans cette différence mesurée.

Condition d'applicabilité du *t-test* : les données sont échantillonnées à partir d'une loi normale.

Variantes : mesures appariées



Statistiques et sciences expérimentales

Rumeurs pseudo-scientifiques sur le Covid-19

Sélection et invention de données

Tests d'hypothèses multiples

Publication de données non finalisées

Catégorisation fautive

Significativité, ampleur de l'effet, et effectifs

Coincidence attendue

Données non représentatives

Résultats tronqués

Pas de patient fractionnaire

Biais de composition des cohortes

Interprétation orientée

Le deuxième pic épidémique

Variants et vitesse de propagation

Conclusion

Statistiques et sciences expérimentales

La « significativité » d'une différence (\neq amplitude) : mesure la confiance qu'on a dans cette différence mesurée.

Condition d'applicabilité du *t-test* : les données sont échantillonnées à partir d'une loi normale.

Variantes : mesures appariées ; comparaison de plus de deux groupes ; valeurs mesurées : toujours entières ; ...

Statistiques et sciences expérimentales

Rumeurs pseudo-scientifiques sur le Covid-19

Sélection et invention de données

Tests d'hypothèses multiples

Publication de données non finalisées

Catégorisation fautive

Significativité, ampleur de l'effet, et effectifs

Coincidence attendue

Données non représentatives

Résultats tronqués

Pas de patient fractionnaire

Biais de composition des cohortes

Interprétation orientée

Le deuxième pic épidémique

Variants et vitesse de propagation

Conclusion

Statistiques et sciences expérimentales

La « significativité » d'une différence (\neq amplitude) : mesure la confiance qu'on a dans cette différence mesurée.

Condition d'applicabilité du *t-test* : les données sont échantillonnées à partir d'une loi normale.

Variantes : mesures appariées ; comparaison de plus de deux groupes ; valeurs mesurées : toujours entières ; ...

→ Évaluer la situation avant de choisir un test statistique (le choix du test peut profondément changer le résultat).

Statistiques et sciences expérimentales

Rumeurs pseudo-scientifiques sur le Covid-19

Sélection et invention de données

Tests d'hypothèses multiples

Publication de données non finalisées

Catégorisation fautive

Significativité, ampleur de l'effet, et effectifs

Coincidence attendue

Données non représentatives

Résultats tronqués

Pas de patient fractionnaire

Biais de composition des cohortes

Interprétation orientée

Le deuxième pic épidémique

Variants et vitesse de propagation

Conclusion

Rumeurs pseudo-scientifiques sur le Covid-19

Du bon (ou du mauvais) choix des statistiques à la lumière de la pandémie

Statistiques et sciences expérimentales

Rumeurs pseudo-scientifiques sur le Covid-19

Sélection et invention de données

Tests d'hypothèses multiples

Publication de données non finalisées

Catégorisation fautive

Significativité, ampleur de l'effet, et effectifs

Coincidence attendue

Données non représentatives

Résultats tronqués

Pas de patient fractionnaire

Biais de composition des cohortes

Interprétation orientée

Le deuxième pic épidémique

Variants et vitesse de propagation

Conclusion

Rumeurs pseudo-scientifiques sur le Covid-19

Janvier 2020 : l'épidémie est détectée en France (touriste chinois à Paris ; premier décès hors-Asie le 14 février [Source]).

Du bon (ou du mauvais) choix des statistiques à la lumière de la pandémie

Statistiques et sciences expérimentales

Rumeurs pseudo-scientifiques sur le Covid-19

Sélection et invention de données

Tests d'hypothèses multiples

Publication de données non finalisées

Catégorisation fautive

Significativité, ampleur de l'effet, et effectifs

Coincidence attendue

Données non représentatives

Résultats tronqués

Pas de patient fractionnaire

Biais de composition des cohortes

Interprétation orientée

Le deuxième pic épidémique

Variants et vitesse de propagation

Conclusion

Rumeurs pseudo-scientifiques sur le Covid-19

Janvier 2020 : l'épidémie est détectée en France (touriste chinois à Paris ; premier décès hors-Asie le 14 février [Source]).

En quelques semaines : du mépris à l'affolement.

Du bon (ou du mauvais) choix des statistiques à la lumière de la pandémie

Statistiques et sciences expérimentales

Rumeurs pseudo-scientifiques sur le Covid-19

Sélection et invention de données

Tests d'hypothèses multiples

Publication de données non finalisées

Catégorisation fautive

Significativité, ampleur de l'effet, et effectifs

Coincidence attendue

Données non représentatives

Résultats tronqués

Pas de patient fractionnaire

Biais de composition des cohortes

Interprétation orientée

Le deuxième pic épidémique

Variants et vitesse de propagation

Conclusion

Rumeurs pseudo-scientifiques sur le Covid-19

Janvier 2020 : l'épidémie est détectée en France (touriste chinois à Paris ; premier décès hors-Asie le 14 février [Source]).

En quelques semaines : du mépris à l'affolement. Des comportements irrationnels, surprenants ...

Du bon (ou du mauvais) choix des statistiques à la lumière de la pandémie

Statistiques et sciences expérimentales

Rumeurs pseudo-scientifiques sur le Covid-19

Sélection et invention de données

Tests d'hypothèses multiples

Publication de données non finalisées

Catégorisation fautive

Significativité, ampleur de l'effet, et effectifs

Coincidence attendue

Données non représentatives

Résultats tronqués

Pas de patient fractionnaire

Biais de composition des cohortes

Interprétation orientée

Le deuxième pic épidémique

Variants et vitesse de propagation

Conclusion

Rumeurs pseudo-scientifiques sur le Covid-19

Janvier 2020 : l'épidémie est détectée en France (touriste chinois à Paris ; premier décès hors-Asie le 14 février [Source]).

En quelques semaines : du mépris à l'affolement. Des comportements irrationnels, surprenants ... notamment : la population entière se prend subitement de passion pour la biologie moléculaire.

Statistiques et sciences expérimentales

Rumeurs pseudo-scientifiques sur le Covid-19

Sélection et invention de données

Tests d'hypothèses multiples

Publication de données non finalisées

Catégorisation fautive

Significativité, ampleur de l'effet, et effectifs

Coincidence attendue

Données non représentatives

Résultats tronqués

Pas de patient fractionnaire

Biais de composition des cohortes

Interprétation orientée

Le deuxième pic épidémique

Variants et vitesse de propagation

Conclusion

Rumeurs pseudo-scientifiques sur le Covid-19

Janvier 2020 : l'épidémie est détectée en France (touriste chinois à Paris ; premier décès hors-Asie le 14 février [Source]).

En quelques semaines : du mépris à l'affolement. Des comportements irrationnels, surprenants ... notamment : la population entière se prend subitement de passion pour la biologie moléculaire.

Un public affolé, passionné, mais peu instruit scientifiquement : terreau fertile pour les rumeurs.

Statistiques et sciences expérimentales

Rumeurs pseudo-scientifiques sur le Covid-19

Sélection et invention de données

Tests d'hypothèses multiples

Publication de données non finalisées

Catégorisation fautive

Significativité, ampleur de l'effet, et effectifs

Coincidence attendue

Données non représentatives

Résultats tronqués

Pas de patient fractionnaire

Biais de composition des cohortes

Interprétation orientée

Le deuxième pic épidémique

Variants et vitesse de propagation

Conclusion

Rumeurs pseudo-scientifiques sur le Covid-19

Janvier 2020 : l'épidémie est détectée en France (touriste chinois à Paris ; premier décès hors-Asie le 14 février [Source]).

En quelques semaines : du mépris à l'affolement. Des comportements irrationnels, surprenants ... notamment : la population entière se prend subitement de passion pour la biologie moléculaire.

Un public affolé, passionné, mais peu instruit scientifiquement : terreau fertile pour les rumeurs.

Tout ce qui suit : scripts et données publiquement accessibles sur https://github.com/HKeyHKey/Covid19/tree/master/CZLR_March_2021.

Statistiques et sciences expérimentales

Rumeurs pseudo-scientifiques sur le Covid-19

Sélection et invention de données

Tests d'hypothèses multiples

Publication de données non finalisées

Catégorisation fautive

Significativité, ampleur de l'effet, et effectifs

Coincidence attendue

Données non représentatives

Résultats tronqués

Pas de patient fractionnaire

Biais de composition des cohortes

Interprétation orientée

Le deuxième pic épidémique

Variants et vitesse de propagation

Conclusion

Rumeurs pseudo-scientifiques sur le Covid-19

Du bon (ou du mauvais) choix des statistiques à la lumière de la pandémie

The screenshot shows a GitHub repository page for 'Covid19/CZLR_March_2021' at the master branch. The repository is owned by 'HKeyHKey'. The commit history shows several files added via upload, including folders for 'Biais_de_composition_des_cohortes', 'Coincidence_attendue', 'Donnees_non_representatives', 'Pas_de_patient_fractionnaire', 'Significativite_ampleur_de_l_effet_et_effectifs', 'Statistiques_et_sciences_experimentales', and 'Tests_d_hypotheses_multiples'. The README.md file was created 5 days ago. The commit message is 'HKeyHKey Add files via upload'.

File	Action	Time
...
Biais_de_composition_des_cohortes	Add files via upload	2 days ago
Coincidence_attendue	Update README.md	10 minutes ago
Donnees_non_representatives	Add files via upload	2 days ago
Pas_de_patient_fractionnaire	Add files via upload	4 minutes ago
Significativite_ampleur_de_l_effet_et_effectifs	Add files via upload	1 hour ago
Statistiques_et_sciences_experimentales	Add files via upload	now
Tests_d_hypotheses_multiples	Add files via upload	21 minutes ago
README.md	Create README.md	5 days ago

Statistiques et sciences expérimentales

Rumeurs pseudo-scientifiques sur le Covid-19

Sélection et invention de données

Tests d'hypothèses multiples

Publication de données non finalisées

Catégorisation fautive

Significativité, ampleur de l'effet, et effectifs

Coincidence attendue

Données non représentatives

Résultats tronqués

Pas de patient fractionnaire

Biais de composition des cohortes

Interprétation orientée

Le deuxième pic épidémique

Variants et vitesse de propagation

Conclusion

Sélection et invention de données

Source : article de mars 2020 par Gautret *et al.* [[Lien](#)]

Statistiques et sciences expérimentales

Rumeurs pseudo-scientifiques sur le Covid-19

Sélection et invention de données

Tests d'hypothèses multiples

Publication de données non finalisées

Catégorisation fautive

Significativité, ampleur de l'effet, et effectifs

Coincidence attendue

Données non représentatives

Résultats tronqués

Pas de patient fractionnaire

Biais de composition des cohortes

Interprétation orientée

Le deuxième pic épidémique

Variants et vitesse de propagation

Conclusion

Sélection et invention de données

Source : article de mars 2020 par Gautret *et al.* [[Lien](#)]



26 patients traités HCQ
(Marseille)



16 patients non-traités HCQ
(Marseille, Nice, Avignon et Briançon)

Du bon (ou du mauvais) choix des statistiques à la lumière de la pandémie

Statistiques et sciences expérimentales

Rumeurs pseudo-scientifiques sur le Covid-19

Sélection et invention de données

Tests d'hypothèses multiples

Publication de données non finalisées

Catégorisation fautive

Significativité, ampleur de l'effet, et effectifs

Coincidence attendue

Données non représentatives

Résultats tronqués

Pas de patient fractionnaire

Biais de composition des cohortes

Interprétation orientée

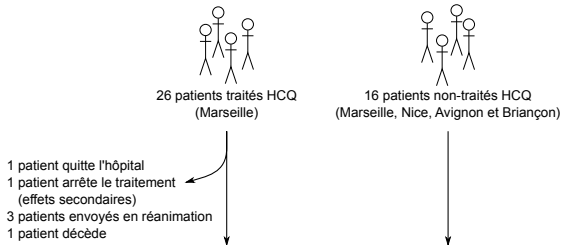
Le deuxième pic épidémique

Variants et vitesse de propagation

Conclusion

Sélection et invention de données

Source : article de mars 2020 par Gautret *et al.* [Lien]



Statistiques et sciences expérimentales

Rumeurs pseudo-scientifiques sur le Covid-19

Sélection et invention de données

Tests d'hypothèses multiples

Publication de données non finalisées

Catégorisation fautive

Significativité, ampleur de l'effet, et effectifs

Coincidence attendue

Données non représentatives

Résultats tronqués

Pas de patient fractionnaire

Biais de composition des cohortes

Interprétation orientée

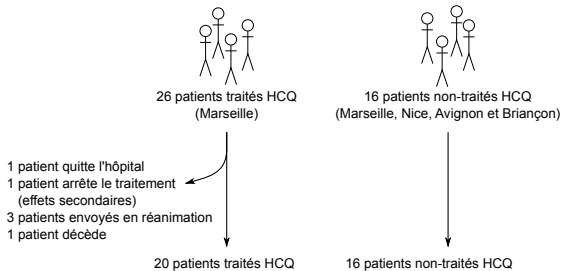
Le deuxième pic épidémique

Variants et vitesse de propagation

Conclusion

Sélection et invention de données

Source : article de mars 2020 par Gautret *et al.* [[Lien](#)]



Statistiques et sciences expérimentales

Rumeurs pseudo-scientifiques sur le Covid-19

Sélection et invention de données

Tests d'hypothèses multiples

Publication de données non finalisées

Catégorisation fautive

Significativité, ampleur de l'effet, et effectifs

Coincidence attendue

Données non représentatives

Résultats tronqués

Pas de patient fractionnaire

Biais de composition des cohortes

Interprétation orientée

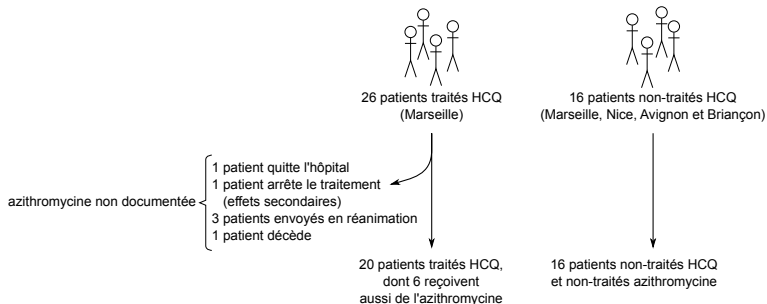
Le deuxième pic épidémique

Variants et vitesse de propagation

Conclusion

Sélection et invention de données

Source : article de mars 2020 par Gautret *et al.* [[Lien](#)]



Statistiques et sciences expérimentales

Rumeurs pseudo-scientifiques sur le Covid-19

Sélection et invention de données

Tests d'hypothèses multiples

Publication de données non finalisées

Catégorisation fautive

Significativité, ampleur de l'effet, et effectifs

Coincidence attendue

Données non représentatives

Résultats tronqués

Pas de patient fractionnaire

Biais de composition des cohortes

Interprétation orientée

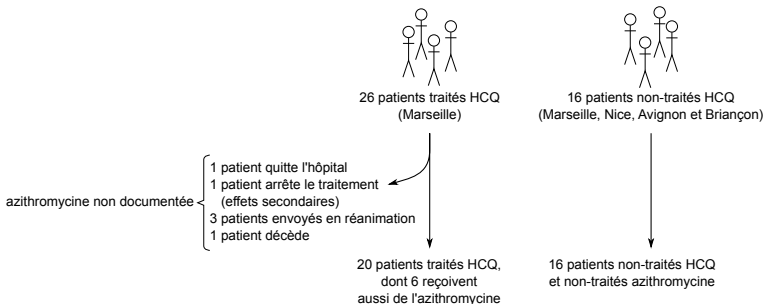
Le deuxième pic épidémique

Variants et vitesse de propagation

Conclusion

Sélection et invention de données

Source : article de mars 2020 par Gautret *et al.* [Lien]



Sélection de données (exclusion des patients traités dont l'état s'est aggravé).

Statistiques et sciences expérimentales

Rumeurs pseudo-scientifiques sur le Covid-19

Sélection et invention de données

Tests d'hypothèses multiples

Publication de données non finalisées

Catégorisation fautive

Significativité, ampleur de l'effet, et effectifs

Coincidence attendue

Données non représentatives

Résultats tronqués

Pas de patient fractionnaire

Biais de composition des cohortes

Interprétation orientée

Le deuxième pic épidémique

Variants et vitesse de propagation

Conclusion

Sélection et invention de données

Table supplémentaire de l'article [Source] :

D0	D1	D2	D3	D4	D5	D6
31	NEG	NEG	NEG	NEG	NEG	NEG
26	ND	33	34	NEG	34	NEG
26	31	23	22	27	NEG	26
24	NEG	33	33	NEG	NEG	32
24	24	24	27	NEG	31	29
POS	ND	POS	ND	POS	ND	POS
28	ND	ND	ND	26	ND	30
POS	ND	POS	ND	POS	POS	POS
POS	ND	POS	ND	POS	ND	POS
POS	ND	POS	ND	ND	ND	POS
POS	ND	POS	ND	POS	ND	ND
ND	ND	POS	ND	POS	ND	ND
POS	ND	POS	ND	POS	ND	POS
POS	ND	POS	ND	ND	POS	ND
ND	ND	ND	POS	ND	POS	ND
POS	ND	ND	ND	ND	POS	ND
30	ND	29	26	32	26	31
29	NEG	NEG	NEG	NEG	NEG	NEG
23	25	28	25	NEG	NEG	NEG
30	NEG	NEG	NEG	NEG	NEG	NEG
34	27	19	16	34	24	22
28	NEG	32	34	NEG	NEG	NEG
22	21	30	NEG	32	28	NEG
17	21	23	21	26	24	24
22	ND	28	21	15	20	17

Statistiques et sciences expérimentales

Rumeurs pseudo-scientifiques sur le Covid-19

Sélection et invention de données

Tests d'hypothèses multiples

Publication de données non finalisées

Catégorisation fautive

Significativité, ampleur de l'effet, et effectifs

Coincidence attendue

Données non représentatives

Résultats tronqués

Pas de patient fractionnaire

Biais de composition des cohortes

Interprétation orientée

Le deuxième pic épidémique

Variants et vitesse de propagation

Conclusion

Sélection et invention de données

Table supplémentaire de l'article [Source] :

D0	D1	D2	D3	D4	D5	D6
31	NEG	NEG	NEG	NEG	NEG	NEG
26	ND	33	34	NEG	34	NEG
26	31	23	22	27	NEG	26
24	NEG	33	33	NEG	NEG	32
24	24	24	27	NEG	31	29
POS	ND	POS	ND	POS	ND	POS
28	ND	ND	ND	26	ND	30
POS	ND	POS	ND	POS	POS	POS
POS	ND	POS	ND	POS	ND	POS
POS	ND	POS	ND	ND	ND	POS
POS	ND	POS	ND	POS	ND	ND
ND	ND	POS	ND	POS	ND	ND
POS	ND	POS	ND	POS	ND	POS
POS	ND	POS	ND	ND	POS	ND
ND	ND	ND	POS	ND	POS	ND
POS	ND	ND	ND	ND	POS	ND
30	ND	29	26	32	26	31
29	NEG	NEG	NEG	NEG	NEG	NEG
23	25	28	25	NEG	NEG	NEG
30	NEG	NEG	NEG	NEG	NEG	NEG
34	27	19	16	34	24	22
28	NEG	32	34	NEG	NEG	NEG
22	21	30	NEG	32	28	NEG
17	21	23	21	26	24	24
22	ND	28	21	15	20	17

Statistiques et sciences expérimentales

Rumeurs pseudo-scientifiques sur le Covid-19

Sélection et invention de données

Tests d'hypothèses multiples

Publication de données non finalisées

Catégorisation fautive

Significativité, ampleur de l'effet, et effectifs

Coincidence attendue

Données non représentatives

Résultats tronqués

Pas de patient fractionnaire

Biais de composition des cohortes

Interprétation orientée

Le deuxième pic épidémique

Variants et vitesse de propagation

Conclusion

Sélection et invention de données

Table supplémentaire de l'article [Source] :

D0	D1	D2	D3	D4	D5	D6
31	NEG	NEG	NEG	NEG	NEG	NEG
26	ND	33	34	NEG	34	NEG
26	31	23	22	27	NEG	26
24	NEG	33	33	NEG	NEG	32
24	24	24	27	NEG	31	29
POS	ND	POS	ND	POS	ND	POS
28	ND	ND	ND	26	ND	30
POS	ND	POS	ND	POS	POS	POS
POS	ND	POS	ND	POS	ND	POS
POS	ND	POS	ND	ND	ND	POS
POS	ND	POS	ND	POS	ND	ND
ND	ND	POS	ND	POS	ND	ND
POS	ND	POS	ND	POS	ND	POS
POS	ND	POS	ND	ND	POS	ND
ND	ND	ND	POS	ND	POS	ND
POS	ND	ND	ND	ND	POS	ND
30	ND	29	26	32	26	31
29	NEG	NEG	NEG	NEG	NEG	NEG
23	25	28	25	NEG	NEG	NEG
30	NEG	NEG	NEG	NEG	NEG	NEG
34	27	19	16	34	24	22
28	NEG	32	34	NEG	NEG	NEG
22	21	30	NEG	32	28	NEG
17	21	23	21	26	24	24
22	ND	28	21	15	20	17

Erreurs de mesures (classique en qPCR).

Statistiques et sciences expérimentales

Rumeurs pseudo-scientifiques sur le Covid-19

Sélection et invention de données

Tests d'hypothèses multiples

Publication de données non finalisées

Catégorisation fautive

Significativité, ampleur de l'effet, et effectifs

Coincidence attendue

Données non représentatives

Résultats tronqués

Pas de patient fractionnaire

Biais de composition des cohortes

Interprétation orientée

Le deuxième pic épidémique

Variants et vitesse de propagation

Conclusion

Sélection et invention de données

Table supplémentaire de l'article [Source] :

D0	D1	D2	D3	D4	D5	D6
31	NEG	NEG	NEG	NEG	NEG	NEG
26	ND	33	34	NEG	34	NEG
26	31	23	22	27	NEG	26
24	NEG	33	33	NEG	NEG	32
24	24	24	27	NEG	31	29
POS	ND	POS	ND	POS	ND	POS
28	ND	ND	ND	26	ND	30
POS	ND	POS	ND	POS	POS	POS
POS	ND	POS	ND	POS	ND	POS
POS	ND	POS	ND	ND	ND	POS
POS	ND	POS	ND	POS	ND	ND
ND	ND	POS	ND	POS	ND	ND
POS	ND	POS	ND	POS	ND	POS
POS	ND	POS	ND	ND	POS	ND
ND	ND	ND	POS	ND	POS	ND
POS	ND	ND	ND	ND	POS	ND
30	ND	29	26	32	26	31
29	NEG	NEG	NEG	NEG	NEG	NEG
23	25	28	25	NEG	NEG	NEG
30	NEG	NEG	NEG	NEG	NEG	NEG
34	27	19	16	34	24	22
28	NEG	32	34	NEG	NEG	NEG
22	21	30	NEG	32	28	NEG
17	21	23	21	26	24	24
22	ND	28	21	15	20	17

Erreurs de mesures (classique en qPCR).

Statistiques et sciences expérimentales

Rumeurs pseudo-scientifiques sur le Covid-19

Sélection et invention de données

Tests d'hypothèses multiples

Publication de données non finalisées

Catégorisation fautive

Significativité, ampleur de l'effet, et effectifs

Coincidence attendue

Données non représentatives

Résultats tronqués

Pas de patient fractionnaire

Biais de composition des cohortes

Interprétation orientée

Le deuxième pic épidémique

Variants et vitesse de propagation

Conclusion

Sélection et invention de données

Table supplémentaire de l'article [Source] :

D0	D1	D2	D3	D4	D5	D6
31	NEG	NEG	NEG	NEG	NEG	NEG
26	ND	33	34	NEG	34	NEG
26	31	23	22	27	NEG	26
24	NEG	33	33	NEG	NEG	32
24	24	24	27	NEG	31	29
POS	ND	POS	ND	POS	ND	POS
28	ND	ND	ND	26	ND	30
POS	ND	POS	ND	POS	POS	POS
POS	ND	POS	ND	POS	ND	POS
POS	ND	POS	ND	ND	ND	POS
POS	ND	POS	ND	POS	ND	ND
ND	ND	POS	ND	POS	ND	ND
POS	ND	POS	ND	POS	ND	POS
POS	ND	POS	ND	ND	POS	ND
ND	ND	ND	POS	ND	POS	ND
POS	ND	ND	ND	ND	POS	ND
30	ND	29	26	32	26	31
29	NEG	NEG	NEG	NEG	NEG	NEG
23	25	28	25	NEG	NEG	NEG
30	NEG	NEG	NEG	NEG	NEG	NEG
34	27	19	16	34	24	22
28	NEG	32	34	NEG	NEG	NEG
22	21	30	NEG	32	28	NEG
17	21	23	21	26	24	24
22	ND	28	21	15	20	17

Erreurs de mesures (classique en qPCR). Invention de données.

Statistiques et sciences expérimentales

Rumeurs pseudo-scientifiques sur le Covid-19

Sélection et invention de données

Tests d'hypothèses multiples

Publication de données non finalisées

Catégorisation fautive

Significativité, ampleur de l'effet, et effectifs

Coincidence attendue

Données non représentatives

Résultats tronqués

Pas de patient fractionnaire

Biais de composition des cohortes

Interprétation orientée

Le deuxième pic épidémique

Variants et vitesse de propagation

Conclusion

Tests d'hypothèses multiples

Du bon (ou du mauvais) choix des statistiques à la lumière de la pandémie

Statistiques et sciences expérimentales

Rumeurs pseudo-scientifiques sur le Covid-19

Sélection et invention de données

Tests d'hypothèses multiples

Publication de données non finalisées

Catégorisation fautive

Significativité, ampleur de l'effet, et effectifs

Coincidence attendue

Données non représentatives

Résultats tronqués

Pas de patient fractionnaire

Biais de composition des cohortes

Interprétation orientée

Le deuxième pic épidémique

Variants et vitesse de propagation

Conclusion

Tests d'hypothèses multiples

Dissimulation de données : les résultats des mesures décrites pour l'essai clinique sont beaucoup plus riches que les résultats présentés dans l'article [[Source](#)] :

E.5 End points		
E.5.1	Primary end point(s)	Results of SARS-COV2 virus detection
E.5.1.1	Timepoint(s) of evaluation of this end point	Day 1, Day 4, Day 7 and Day 14
E.5.2	Secondary end point(s)	Apyrexia, normalization of respiratory rate, and average length of hospital stay and mortality.
E.5.2.1	Timepoint(s) of evaluation of this end point	Day 1, Day 4, Day 7, Day 14 and Days of hospital discharge

Statistiques et sciences expérimentales

Rumeurs pseudo-scientifiques sur le Covid-19

Sélection et invention de données

Tests d'hypothèses multiples

Publication de données non finalisées

Catégorisation fautive

Significativité, ampleur de l'effet, et effectifs

Coincidence attendue

Données non représentatives

Résultats tronqués

Pas de patient fractionnaire

Biais de composition des cohortes

Interprétation orientée

Le deuxième pic épidémique

Variants et vitesse de propagation

Conclusion

Tests d'hypothèses multiples

Dissimulation de données : les résultats des mesures décrites pour l'essai clinique sont beaucoup plus riches que les résultats présentés dans l'article [[Source](#)] :

E.5 End points		
E.5.1	Primary end point(s)	Results of SARS-COV2 virus detection
E.5.1.1	Timepoint(s) of evaluation of this end point	Day 1, Day 4, Day 7 and Day 14
E.5.2	Secondary end point(s)	Apyrexia, normalization of respiratory rate, and average length of hospital stay and mortality.
E.5.2.1	Timepoint(s) of evaluation of this end point	Day 1, Day 4, Day 7, Day 14 and Days of hospital discharge

Détection de l'ARN viral : pas de différence significative (test exact de Fisher, jour 1 : $p = 1$; jour 4 : $p = 0,15$) [[Code source](#)].

Statistiques et sciences expérimentales

Rumeurs pseudo-scientifiques sur le Covid-19

Sélection et invention de données

Tests d'hypothèses multiples

Publication de données non finalisées

Catégorisation fautive

Significativité, ampleur de l'effet, et effectifs

Coincidence attendue

Données non représentatives

Résultats tronqués

Pas de patient fractionnaire

Biais de composition des cohortes

Interprétation orientée

Le deuxième pic épidémique

Variants et vitesse de propagation

Conclusion

Tests d'hypothèses multiples

Dissimulation de données : les résultats des mesures décrites pour l'essai clinique sont beaucoup plus riches que les résultats présentés dans l'article [[Source](#)] :

E.5 End points		
E.5.1	Primary end point(s)	Results of SARS-COV2 virus detection
E.5.1.1	Timepoint(s) of evaluation of this end point	Day 1, Day 4, Day 7 and Day 14
E.5.2	Secondary end point(s)	Apyrexia, normalization of respiratory rate, and average length of hospital stay and mortality.
E.5.2.1	Timepoint(s) of evaluation of this end point	Day 1, Day 4, Day 7, Day 14 and Days of hospital discharge

Détection de l'ARN viral : pas de différence significative (test exact de Fisher, jour 1 : $p = 1$; jour 4 : $p = 0,15$) [[Code source](#)].

Problème classique du test d'hypothèses multiples : plus on teste, plus on prend le risque de tomber sur un faux positif [[illustration humoristique](#)].

Statistiques et sciences expérimentales

Rumeurs pseudo-scientifiques sur le Covid-19

Sélection et invention de données

Tests d'hypothèses multiples

Publication de données non finalisées

Catégorisation fautive

Significativité, ampleur de l'effet, et effectifs

Coincidence attendue

Données non représentatives

Résultats tronqués

Pas de patient fractionnaire

Biais de composition des cohortes

Interprétation orientée

Le deuxième pic épidémique

Variants et vitesse de propagation

Conclusion

Tests d'hypothèses multiples

Dissimulation de données : les résultats des mesures décrites pour l'essai clinique sont beaucoup plus riches que les résultats présentés dans l'article [[Source](#)] :

E.5 End points		
E.5.1	Primary end point(s)	Results of SARS-COV2 virus detection
E.5.1.1	Timepoint(s) of evaluation of this end point	Day 1, Day 4, Day 7 and Day 14
E.5.2	Secondary end point(s)	Apyrexia, normalization of respiratory rate, and average length of hospital stay and mortality.
E.5.2.1	Timepoint(s) of evaluation of this end point	Day 1, Day 4, Day 7, Day 14 and Days of hospital discharge

Détection de l'ARN viral : pas de différence significative (test exact de Fisher, jour 1 : $p = 1$; jour 4 : $p = 0,15$) [[Code source](#)].

Problème classique du test d'hypothèses multiples : plus on teste, plus on prend le risque de tomber sur un faux positif [[illustration humoristique](#)].

► Comment je ferais

Statistiques et sciences expérimentales

Rumeurs pseudo-scientifiques sur le Covid-19

Sélection et invention de données

Tests d'hypothèses multiples

Publication de données non finalisées

Catégorisation fautive

Significativité, ampleur de l'effet, et effectifs

Coincidence attendue

Données non représentatives

Résultats tronqués

Pas de patient fractionnaire

Biais de composition des cohortes

Interprétation orientée

Le deuxième pic épidémique

Variants et vitesse de propagation

Conclusion

Tests d'hypothèses multiples

Janvier 2021 : publication d'un correctif ([Source](#)).

Du bon (ou du mauvais) choix des statistiques à la lumière de la pandémie

Statistiques et sciences expérimentales

Rumeurs pseudo-scientifiques sur le Covid-19

Sélection et invention de données

Tests d'hypothèses multiples

Publication de données non finalisées

Catégorisation fautive

Significativité, ampleur de l'effet, et effectifs

Coincidence attendue

Données non représentatives

Résultats tronqués

Pas de patient fractionnaire

Biais de composition des cohortes

Interprétation orientée

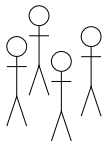
Le deuxième pic épidémique

Variants et vitesse de propagation

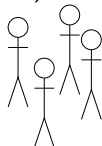
Conclusion

Tests d'hypothèses multiples

Janvier 2021 : publication d'un correctif ([Source](#)).



24 patients traités HCQ



18 patients non-traités HCQ

Dont **16** sans azithromycine

Dont **8** reçoivent aussi de l'azithromycine

1 décès, **2** transferts en réanimation

1 décès, **2** transferts en réanimation

0 décès, **1** transfert en réanimation

Statistiques et sciences expérimentales

Rumeurs pseudo-scientifiques sur le Covid-19

Sélection et invention de données

Tests d'hypothèses multiples

Publication de données non finalisées

Catégorisation fautive

Significativité, ampleur de l'effet, et effectifs

Coincidence attendue

Données non représentatives

Résultats tronqués

Pas de patient fractionnaire

Biais de composition des cohortes

Interprétation orientée

Le deuxième pic épidémique

Variants et vitesse de propagation

Conclusion

Publication de données non finalisées

Historique :

- ▶ 9 avril 2020 : pendant une visite du président de la République, l'IHU Marseille annonce un taux de létalité de 0,47% dans sa cohorte traitée [[Source](#)]

Statistiques et sciences expérimentales

Rumeurs pseudo-scientifiques sur le Covid-19

Sélection et invention de données

Tests d'hypothèses multiples

Publication de données non finalisées

Catégorisation fautive

Significativité, ampleur de l'effet, et effectifs

Coincidence attendue

Données non représentatives

Résultats tronqués

Pas de patient fractionnaire

Biais de composition des cohortes

Interprétation orientée

Le deuxième pic épidémique

Variants et vitesse de propagation

Conclusion

Publication de données non finalisées

Historique :

- ▶ 9 avril 2020 : pendant une visite du président de la République, l'IHU Marseille annonce un taux de létalité de 0,47% dans sa cohorte traitée [[Source](#)]
- ▶ 5 mai 2020 : l'article est publié en ligne [[Lien](#)], le texte principal annonce 0,75% de létalité ;

Statistiques et sciences expérimentales

Rumeurs pseudo-scientifiques sur le Covid-19

Sélection et invention de données

Tests d'hypothèses multiples

Publication de données non finalisées

Catégorisation fautive

Significativité, ampleur de l'effet, et effectifs

Coincidence attendue

Données non représentatives

Résultats tronqués

Pas de patient fractionnaire

Biais de composition des cohortes

Interprétation orientée

Le deuxième pic épidémique

Variants et vitesse de propagation

Conclusion

Publication de données non finalisées

Historique :

- ▶ 9 avril 2020 : pendant une visite du président de la République, l'IHU Marseille annonce un taux de létalité de 0,47% dans sa cohorte traitée [[Source](#)]
- ▶ 5 mai 2020 : l'article est publié en ligne [[Lien](#)], le texte principal annonce 0,75% de létalité ;
- ▶ ... au détour d'une note de bas de page, il annonce 0,9% suite à deux décès non comptabilisés dans le texte principal ...

Statistiques et sciences expérimentales

Rumeurs pseudo-scientifiques sur le Covid-19

Sélection et invention de données

Tests d'hypothèses multiples

Publication de données non finalisées

Catégorisation fautive

Significativité, ampleur de l'effet, et effectifs

Coincidence attendue

Données non représentatives

Résultats tronqués

Pas de patient fractionnaire

Biais de composition des cohortes

Interprétation orientée

Le deuxième pic épidémique

Variants et vitesse de propagation

Conclusion

Publication de données non finalisées

Historique :

- ▶ 9 avril 2020 : pendant une visite du président de la République, l'IHU Marseille annonce un taux de létalité de 0,47% dans sa cohorte traitée [[Source](#)]
- ▶ 5 mai 2020 : l'article est publié en ligne [[Lien](#)], le texte principal annonce 0,75% de létalité ;
- ▶ ... au détour d'une note de bas de page, il annonce 0,9% suite à deux décès non comptabilisés dans le texte principal ...
- ▶ ... sachant que 4 patients y sont encore en sursis (le taux de létalité final s'établira donc entre 0,9 et 1,2%, jamais publié) ; mensonge entretenu par D. Raoult le 26 mai 2020 [[Source](#)].

Statistiques et sciences expérimentales

Rumeurs pseudo-scientifiques sur le Covid-19

Sélection et invention de données

Tests d'hypothèses multiples

Publication de données non finalisées

Catégorisation fautive

Significativité, ampleur de l'effet, et effectifs

Coincidence attendue

Données non représentatives

Résultats tronqués

Pas de patient fractionnaire

Biais de composition des cohortes

Interprétation orientée

Le deuxième pic épidémique

Variants et vitesse de propagation

Conclusion

Publication de données non finalisées

Historique :

- ▶ 9 avril 2020 : pendant une visite du président de la République, l'IHU Marseille annonce un taux de létalité de 0,47% dans sa cohorte traitée [[Source](#)]
- ▶ 5 mai 2020 : l'article est publié en ligne [[Lien](#)], le texte principal annonce 0,75% de létalité ;
- ▶ ... au détour d'une note de bas de page, il annonce 0,9% suite à deux décès non comptabilisés dans le texte principal ...
- ▶ ... sachant que 4 patients y sont encore en sursis (le taux de létalité final s'établira donc entre 0,9 et 1,2%, jamais publié) ; mensonge entretenu par D. Raoult le 26 mai 2020 [[Source](#)].

En France au printemps 2020, le taux de létalité du Covid-19 est estimé à 0,7 [[Source](#)] ou 0,8% [[Source](#)].

Statistiques et sciences expérimentales

Rumeurs pseudo-scientifiques sur le Covid-19

Sélection et invention de données

Tests d'hypothèses multiples

Publication de données non finalisées

Catégorisation fautive

Significativité, ampleur de l'effet, et effectifs

Coincidence attendue

Données non représentatives

Résultats tronqués

Pas de patient fractionnaire

Biais de composition des cohortes

Interprétation orientée

Le deuxième pic épidémique

Variants et vitesse de propagation

Conclusion

Catégorisation fautive

Source : manuscrit mis en ligne le 11 mai 2020 par une équipe de l'hôpital de Garches (professeurs Davido, Perronne, ...) [[Lien](#)].

Statistiques et sciences expérimentales

Rumeurs pseudo-scientifiques sur le Covid-19

Sélection et invention de données

Tests d'hypothèses multiples

Publication de données non finalisées

Catégorisation fautive

Significativité, ampleur de l'effet, et effectifs

Coincidence attendue

Données non représentatives

Résultats tronqués

Pas de patient fractionnaire

Biais de composition des cohortes

Interprétation orientée

Le deuxième pic épidémique

Variants et vitesse de propagation

Conclusion

Catégorisation fautive

Source : manuscrit mis en ligne le 11 mai 2020 par une équipe de l'hôpital de Garches (professeurs Davido, Perronne, ...) [[Lien](#)].

45 patients traités hydroxychloroquine+azithromycine comparés à 87 patients traités autrement.

Statistiques et sciences expérimentales

Rumeurs pseudo-scientifiques sur le Covid-19

Sélection et invention de données

Tests d'hypothèses multiples

Publication de données non finalisées

Catégorisation fautive

Significativité, ampleur de l'effet, et effectifs

Coincidence attendue

Données non représentatives

Résultats tronqués

Pas de patient fractionnaire

Biais de composition des cohortes

Interprétation orientée

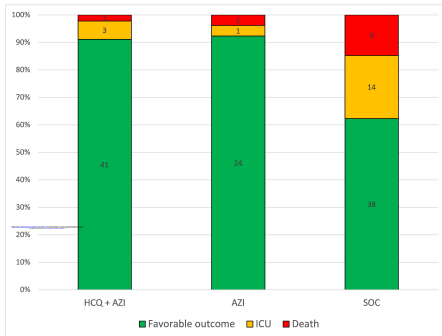
Le deuxième pic épidémique

Variants et vitesse de propagation

Conclusion

Catégorisation fautive

Source : manuscrit mis en ligne le 11 mai 2020 par une équipe de l'hôpital de Garches (professeurs Davido, Perronne, ...) [[Lien](#)].



[Source]

Statistiques et sciences expérimentales

Rumeurs pseudo-scientifiques sur le Covid-19

Sélection et invention de données

Tests d'hypothèses multiples

Publication de données non finalisées

Catégorisation fautive

Significativité, ampleur de l'effet, et effectifs

Coincidence attendue

Données non représentatives

Résultats tronqués

Pas de patient fractionnaire

Biais de composition des cohortes

Interprétation orientée

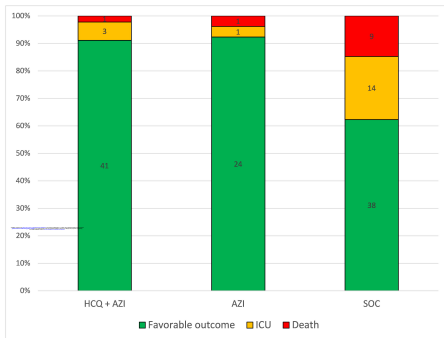
Le deuxième pic épidémique

Variants et vitesse de propagation

Conclusion

Catégorisation fautive

Source : manuscrit mis en ligne le 11 mai 2020 par une équipe de l'hôpital de Garches (professeurs Davido, Perronne, ...) [[Lien](#)].



[[Source](#)]

9 patients HCQ+AZI, décédés ou envoyés en réanimation en moins de 48h, comptés comme « autre traitement ».

Statistiques et sciences expérimentales

Rumeurs pseudo-scientifiques sur le Covid-19

Sélection et invention de données

Tests d'hypothèses multiples

Publication de données non finalisées

Catégorisation fautive

Significativité, ampleur de l'effet, et effectifs

Coincidence attendue

Données non représentatives

Résultats tronqués

Pas de patient fractionnaire

Biais de composition des cohortes

Interprétation orientée

Le deuxième pic épidémique

Variants et vitesse de propagation

Conclusion

Catégorisation fautive

Tentative de rétractation [[Source](#)], la publication finale [[Source](#)] ne cherche plus à comparer la bithérapie aux autres traitements.

Statistiques et sciences expérimentales

Rumeurs pseudo-scientifiques sur le Covid-19

Sélection et invention de données

Tests d'hypothèses multiples

Publication de données non finalisées

Catégorisation fautive

Significativité, ampleur de l'effet, et effectifs

Coincidence attendue

Données non représentatives

Résultats tronqués

Pas de patient fractionnaire

Biais de composition des cohortes

Interprétation orientée

Le deuxième pic épidémique

Variants et vitesse de propagation

Conclusion

Significativité, ampleur de l'effet, et effectifs

Source : D. Raoult dans *Marianne* : « plus l'échantillon d'un test clinique est faible, plus ses résultats sont significatifs » [Source], répété devant la commission d'enquête de l'Assemblée nationale [Source], repris par C. Perronne [Source].

Statistiques et sciences expérimentales

Rumeurs pseudo-scientifiques sur le Covid-19

Sélection et invention de données

Tests d'hypothèses multiples

Publication de données non finalisées

Catégorisation fautive

Significativité, ampleur de l'effet, et effectifs

Coincidence attendue

Données non représentatives

Résultats tronqués

Pas de patient fractionnaire

Biais de composition des cohortes

Interprétation orientée

Le deuxième pic épidémique

Variants et vitesse de propagation

Conclusion

Significativité, ampleur de l'effet, et effectifs

Source : D. Raoult dans *Marianne* : « plus l'échantillon d'un test clinique est faible, plus ses résultats sont significatifs » [Source], répété devant la commission d'enquête de l'Assemblée nationale [Source], repris par C. Perronne [Source].

C'est faux : les petits effectifs ne détecteront « significatifs » que des effets amples, mais les gros effectifs les détecteraient aussi (et même mieux) !

Statistiques et sciences expérimentales

Rumeurs pseudo-scientifiques sur le Covid-19

Sélection et invention de données

Tests d'hypothèses multiples

Publication de données non finalisées

Catégorisation fautive

Significativité, ampleur de l'effet, et effectifs

Coincidence attendue

Données non représentatives

Résultats tronqués

Pas de patient fractionnaire

Biais de composition des cohortes

Interprétation orientée

Le deuxième pic épidémique

Variants et vitesse de propagation

Conclusion

Significativité, ampleur de l'effet, et effectifs

Source : D. Raoult dans *Marianne* : « plus l'échantillon d'un test clinique est faible, plus ses résultats sont significatifs » [Source], répété devant la commission d'enquête de l'Assemblée nationale [Source], repris par C. Perronne [Source].

C'est faux : les petits effectifs ne détecteront « significatifs » que des effets amples, mais les gros effectifs les détecteraient aussi (et même mieux) !

Cond. 1	Cond. 2
6	3
14	17

$$p \approx 0,45$$

Cond. 1	Cond. 2
60	30
140	170

$$p \approx 0,00047$$

(test exact de Fisher)

Coïncidence attendue

Source : manuscrit décrivant des similitudes de séquence entre les génomes viraux HIV et SARS-CoV-2 [[Source](#)], le 31 janvier 2020.

Statistiques et sciences expérimentales

Rumeurs pseudo-scientifiques sur le Covid-19

Sélection et invention de données

Tests d'hypothèses multiples

Publication de données non finalisées

Catégorisation fautive

Significativité, ampleur de l'effet, et effectifs

Coïncidence attendue

Données non représentatives

Résultats tronqués

Pas de patient fractionnaire

Biais de composition des cohortes

Interprétation orientée

Le deuxième pic épidémique

Variants et vitesse de propagation

Conclusion

Coïncidence attendue

Source : manuscrit décrivant des similitudes de séquence entre les génomes viraux HIV et SARS-CoV-2 [[Source](#)], le 31 janvier 2020.

Des segments très courts (6 à 12 acides aminés après traduction), en tolérant des insertions, en comparant de nombreux variants de longs génomes viraux.

Statistiques et sciences expérimentales

Rumeurs pseudo-scientifiques sur le Covid-19

Sélection et invention de données

Tests d'hypothèses multiples

Publication de données non finalisées

Catégorisation fautive

Significativité, ampleur de l'effet, et effectifs

Coïncidence attendue

Données non représentatives

Résultats tronqués

Pas de patient fractionnaire

Biais de composition des cohortes

Interprétation orientée

Le deuxième pic épidémique

Variants et vitesse de propagation

Conclusion

Coïncidence attendue

Source : manuscrit décrivant des similitudes de séquence entre les génomes viraux HIV et SARS-CoV-2 [[Source](#)], le 31 janvier 2020.

Des segments très courts (6 à 12 acides aminés après traduction), en tolérant des insertions, en comparant de nombreux variants de longs génomes viraux.

→ ressemblance fortuite ; manuscrit rétracté le 2 février 2020 [[Source](#)].

Statistiques et sciences expérimentales

Rumeurs pseudo-scientifiques sur le Covid-19

Sélection et invention de données

Tests d'hypothèses multiples

Publication de données non finalisées

Catégorisation fautive

Significativité, ampleur de l'effet, et effectifs

Coïncidence attendue

Données non représentatives

Résultats tronqués

Pas de patient fractionnaire

Biais de composition des cohortes

Interprétation orientée

Le deuxième pic épidémique

Variants et vitesse de propagation

Conclusion

Coïncidence attendue

Source : manuscrit décrivant des similitudes de séquence entre les génomes viraux HIV et SARS-CoV-2 [[Source](#)], le 31 janvier 2020.

Des segments très courts (6 à 12 acides aminés après traduction), en tolérant des insertions, en comparant de nombreux variants de longs génomes viraux.

→ ressemblance fortuite ; manuscrit rétracté le 2 février 2020 [[Source](#)].

En France, le Prof. Montagnier reprend cette affirmation dans les médias, pour en conclure que le SARS-CoV-2 a été créé par l'Homme [[Source](#)].

Statistiques et sciences expérimentales

Rumeurs pseudo-scientifiques sur le Covid-19

Sélection et invention de données

Tests d'hypothèses multiples

Publication de données non finalisées

Catégorisation fautive

Significativité, ampleur de l'effet, et effectifs

Coïncidence attendue

Données non représentatives

Résultats tronqués

Pas de patient fractionnaire

Biais de composition des cohortes

Interprétation orientée

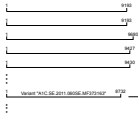
Le deuxième pic épidémique

Variants et vitesse de propagation

Conclusion

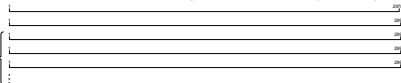
Coïncidence attendue

HIV databases (<https://www.hiv.lanl.gov/content/index>) :
12501 variants du HIV :



9943 variants
21 nucléotides
identiques

GISAID (<https://platform.gisaid.org>) :
10184 variants correctement séquencés du SARS-CoV-2 (22/04/2020) :



Statistiques et sciences expérimentales

Rumeurs pseudo-scientifiques sur le Covid-19

Sélection et invention de données

Tests d'hypothèses multiples

Publication de données non finalisées

Catégorisation fautive

Significativité, ampleur de l'effet, et effectifs

Coïncidence attendue

Données non représentatives

Résultats tronqués

Pas de patient fractionnaire

Biais de composition des cohortes

Interprétation orientée

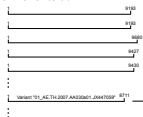
Le deuxième pic épidémique

Variants et vitesse de propagation

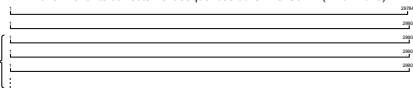
Conclusion

Coïncidence attendue

HIV databases (<https://www.hiv.lanl.gov/content/index>) :
12501 variants du HIV :



GISAID (<https://platform.gisaid.org>) :
10184 variants correctement séquencés du SARS-CoV-2 (22/04/2020) :



2 variants
21 nucléotides identiques

...AACTACTGTGTGTATTAACAGTCTCTC...
...TAACTACTGTGTGTATTAACAGTCTCTC...

Statistiques et sciences expérimentales

Rumeurs pseudo-scientifiques sur le Covid-19

Sélection et invention de données

Tests d'hypothèses multiples

Publication de données non finalisées

Catégorisation fautive

Significativité, ampleur de l'effet, et effectifs

Coïncidence attendue

Données non représentatives

Résultats tronqués

Pas de patient fractionnaire

Biais de composition des cohortes

Interprétation orientée

Le deuxième pic épidémique

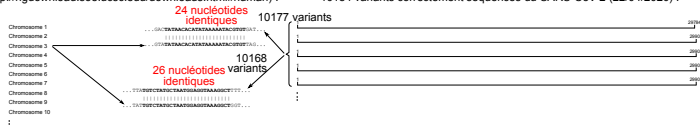
Variants et vitesse de propagation

Conclusion

Coïncidence attendue

Génome humain de référence
(<http://hgdownload.soe.ucsc.edu/downloads.html#human>) :

GISAID (<https://platform.gisaid.org>) :
10184 variants correctement séquencés du SARS-CoV-2 (22/04/2020) :



Statistiques et sciences expérimentales

Rumeurs pseudo-scientifiques sur le Covid-19

Sélection et invention de données

Tests d'hypothèses multiples

Publication de données non finalisées

Catégorisation fautive

Significativité, ampleur de l'effet, et effectifs

Coïncidence attendue

Données non représentatives

Résultats tronqués

Pas de patient fractionnaire

Biais de composition des cohortes

Interprétation orientée

Le deuxième pic épidémique

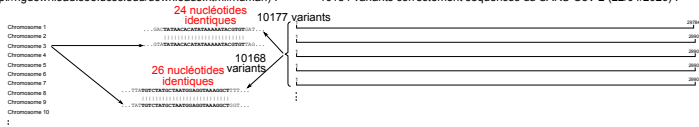
Variants et vitesse de propagation

Conclusion

Coïncidence attendue

Génome humain de référence
(<http://hgdownload.soe.ucsc.edu/downloads.html#human>) :

GISAIID (<https://platform.gisaid.org>) :
10184 variants correctement séquencés du SARS-CoV-2 (22/04/2020) :



Dans la collection de variants HIV et SARS-CoV-2 connue fin avril : $\approx 0,25$ similarités aussi belles sont attendues par hasard.

Statistiques et sciences expérimentales

Rumeurs pseudo-scientifiques sur le Covid-19

Sélection et invention de données

Tests d'hypothèses multiples

Publication de données non finalisées

Catégorisation fautive

Significativité, ampleur de l'effet, et effectifs

Coïncidence attendue

Données non représentatives

Résultats tronqués

Pas de patient fractionnaire

Biais de composition des cohortes

Interprétation orientée

Le deuxième pic épidémique

Variants et vitesse de propagation

Conclusion

Données non représentatives

- ▶ « On meurt moins à Marseille » (É. Chabrière, IHU Marseille, le 27 mars 2020 [[Source](#)])

Statistiques et sciences expérimentales

Rumeurs pseudo-scientifiques sur le Covid-19

Sélection et invention de données

Tests d'hypothèses multiples

Publication de données non finalisées

Catégorisation fautive

Significativité, ampleur de l'effet, et effectifs

Coincidence attendue

Données non représentatives

Résultats tronqués

Pas de patient fractionnaire

Biais de composition des cohortes

Interprétation orientée

Le deuxième pic épidémique

Variants et vitesse de propagation

Conclusion

Données non représentatives

- ▶ « On meurt moins à Marseille » (É. Chabrière, IHU Marseille, le 27 mars 2020 [[Source](#)])
- ▶ « 0,5% de mortalité, est-ce que ce n'est pas le taux de mortalité général ? » / « Vous rigolez ? Vous avez vu à Paris, ce que c'est la mortalité ? » (D. Pujadas et D. Raoult, le 26 mai 2020 [[Source](#)])

Statistiques et sciences expérimentales

Rumeurs pseudo-scientifiques sur le Covid-19

Sélection et invention de données

Tests d'hypothèses multiples

Publication de données non finalisées

Catégorisation fautive

Significativité, ampleur de l'effet, et effectifs

Coincidence attendue

Données non représentatives

Résultats tronqués

Pas de patient fractionnaire

Biais de composition des cohortes

Interprétation orientée

Le deuxième pic épidémique

Variants et vitesse de propagation

Conclusion

Données non représentatives

- ▶ « On meurt moins à Marseille » (É. Chabrière, IHU Marseille, le 27 mars 2020 [[Source](#)])
- ▶ « 0,5% de mortalité, est-ce que ce n'est pas le taux de mortalité général ? » / « Vous rigolez ? Vous avez vu à Paris, ce que c'est la mortalité ? » (D. Pujadas et D. Raoult, le 26 mai 2020 [[Source](#)])
- ▶ « Vous verrez la mortalité hospitalière par région, et puis vous verrez celle que l'on aura » (D. Raoult, IHU Marseille, le 27 octobre 2020 ; [[Source](#)])

Statistiques et sciences expérimentales

Rumeurs pseudo-scientifiques sur le Covid-19

Sélection et invention de données

Tests d'hypothèses multiples

Publication de données non finalisées

Catégorisation fautive

Significativité, ampleur de l'effet, et effectifs

Coincidence attendue

Données non représentatives

Résultats tronqués

Pas de patient fractionnaire

Biais de composition des cohortes

Interprétation orientée

Le deuxième pic épidémique

Variants et vitesse de propagation

Conclusion

Données non représentatives

- ▶ « On meurt moins à Marseille » (É. Chabrière, IHU Marseille, le 27 mars 2020 [[Source](#)])
- ▶ « 0,5% de mortalité, est-ce que ce n'est pas le taux de mortalité général ? » / « Vous rigolez ? Vous avez vu à Paris, ce que c'est la mortalité ? » (D. Pujadas et D. Raoult, le 26 mai 2020 [[Source](#)])
- ▶ « Vous verrez la mortalité hospitalière par région, et puis vous verrez celle que l'on aura » (D. Raoult, IHU Marseille, le 27 octobre 2020 ; [[Source](#)])

Statistiques et sciences expérimentales

Rumeurs pseudo-scientifiques sur le Covid-19

Sélection et invention de données

Tests d'hypothèses multiples

Publication de données non finalisées

Catégorisation fautive

Significativité, ampleur de l'effet, et effectifs

Coincidence attendue

Données non représentatives

Résultats tronqués

Pas de patient fractionnaire

Biais de composition des cohortes

Interprétation orientée

Le deuxième pic épidémique

Variants et vitesse de propagation

Conclusion

Données non représentatives

- ▶ « On meurt moins à Marseille » (É. Chabrière, IHU Marseille, le 27 mars 2020 [[Source](#)])
- ▶ « 0,5% de mortalité, est-ce que ce n'est pas le taux de mortalité général ? » / « Vous rigolez ? Vous avez vu à Paris, ce que c'est la mortalité ? » (D. Pujadas et D. Raoult, le 26 mai 2020 [[Source](#)])
- ▶ « Vous verrez la mortalité hospitalière par région, et puis vous verrez celle que l'on aura » (D. Raoult, IHU Marseille, le 27 octobre 2020 ; [[Source](#)])

Comparaison de Marseille, ou des Bouches-du-Rhône, uniquement avec les villes et départements les plus sévèrement touchés.

Statistiques et sciences expérimentales

Rumeurs pseudo-scientifiques sur le Covid-19

Sélection et invention de données

Tests d'hypothèses multiples

Publication de données non finalisées

Catégorisation fautive

Significativité, ampleur de l'effet, et effectifs

Coincidence attendue

Données non représentatives

Résultats tronqués

Pas de patient fractionnaire

Biais de composition des cohortes

Interprétation orientée

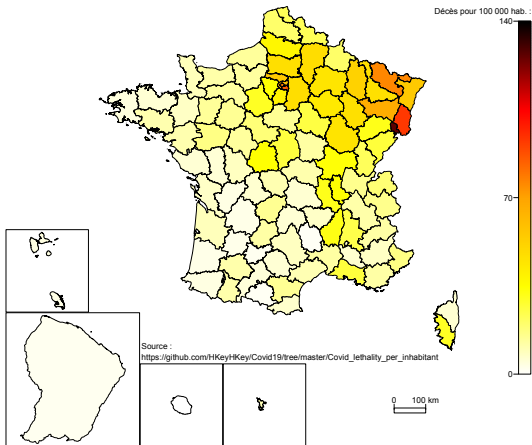
Le deuxième pic épidémique

Variants et vitesse de propagation

Conclusion

Données non représentatives

Décès hospitaliers pour Covid-19 entre le 18 mars et le 20 juin 2020 :



(contours des départements : © les contributeurs d'OpenStreetMap sous licence ODbL)

Du bon (ou du mauvais) choix des statistiques à la lumière de la pandémie

Statistiques et sciences expérimentales

Rumeurs pseudo-scientifiques sur le Covid-19

Sélection et invention de données

Tests d'hypothèses multiples

Publication de données non finalisées

Catégorisation fautive

Significativité, ampleur de l'effet, et effectifs

Coincidence attendue

Données non représentatives

Résultats tronqués

Pas de patient fractionnaire

Biais de composition des cohortes

Interprétation orientée

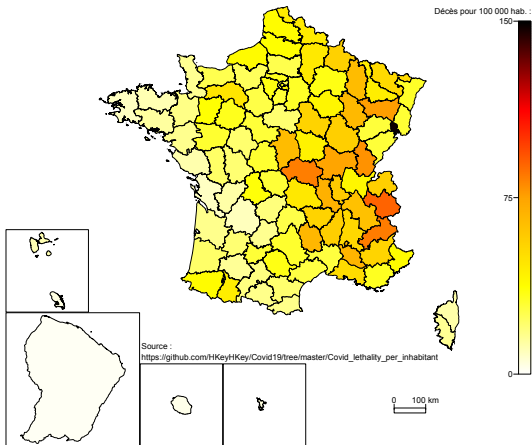
Le deuxième pic épidémique

Variants et vitesse de propagation

Conclusion

Données non représentatives

Décès hospitaliers pour Covid-19 entre le 20 octobre 2020 et le 8 janvier 2021 :



(contours des départements : © les contributeurs d'OpenStreetMap sous licence ODbL)

Du bon (ou du mauvais) choix des statistiques à la lumière de la pandémie

Statistiques et sciences expérimentales

Rumeurs pseudo-scientifiques sur le Covid-19

Sélection et invention de données

Tests d'hypothèses multiples

Publication de données non finalisées

Catégorisation fautive

Significativité, ampleur de l'effet, et effectifs

Coincidence attendue

Données non représentatives

Résultats tronqués

Pas de patient fractionnaire

Biais de composition des cohortes

Interprétation orientée

Le deuxième pic épidémique

Variants et vitesse de propagation

Conclusion

Données non représentatives

Mortalité parmi les patients Covid-19 hospitalisés :

- ▶ Première vague (18 mars – 20 juin) : 25 départements ont des mortalités inférieures à celle des Bouches-du-Rhône ;
- ▶ Deuxième vague (20 octobre – 8 janvier) : 13 départements.

Statistiques et sciences expérimentales

Rumeurs pseudo-scientifiques sur le Covid-19

Sélection et invention de données

Tests d'hypothèses multiples

Publication de données non finalisées

Catégorisation fautive

Significativité, ampleur de l'effet, et effectifs

Coincidence attendue

Données non représentatives

Résultats tronqués

Pas de patient fractionnaire

Biais de composition des cohortes

Interprétation orientée

Le deuxième pic épidémique

Variants et vitesse de propagation

Conclusion

Données non représentatives

Mortalité parmi les patients Covid-19 hospitalisés :

- ▶ Première vague (18 mars – 20 juin) : 25 départements ont des mortalités inférieures à celle des Bouches-du-Rhône ;
- ▶ Deuxième vague (20 octobre – 8 janvier) : 13 départements.

Pas d'hydroxychloroquine pour le traitement du Covid-19 dans la 2ème vague [[Source](#)].

Statistiques et sciences expérimentales

Rumeurs pseudo-scientifiques sur le Covid-19

Sélection et invention de données

Tests d'hypothèses multiples

Publication de données non finalisées

Catégorisation fautive

Significativité, ampleur de l'effet, et effectifs

Coincidence attendue

Données non représentatives

Résultats tronqués

Pas de patient fractionnaire

Biais de composition des cohortes

Interprétation orientée

Le deuxième pic épidémique

Variants et vitesse de propagation

Conclusion

Résultats tronqués

Source : une **pré-publication** (mise en ligne le 10 avril, ensuite **publiée dans un journal** le 21 août) évalue la toxicité de l'hydroxychloroquine (avec ou sans azithromycine) sur de larges cohortes, en étude rétrospective.

Statistiques et sciences expérimentales

Rumeurs pseudo-scientifiques sur le Covid-19

Sélection et invention de données

Tests d'hypothèses multiples

Publication de données non finalisées

Catégorisation fautive

Significativité, ampleur de l'effet, et effectifs

Coincidence attendue

Données non représentatives

Résultats tronqués

Pas de patient fractionnaire

Biais de composition des cohortes

Interprétation orientée

Le deuxième pic épidémique

Variants et vitesse de propagation

Conclusion

Résultats tronqués

Source : une **pré-publication** (mise en ligne le 10 avril, ensuite **publiée dans un journal** le 21 août) évalue la toxicité de l'hydroxychloroquine (avec ou sans azithromycine) sur de larges cohortes, en étude rétrospective.

Plus de 900 000 patients traités à l'hydroxychloroquine (référence : plus de 300 000 patients traités à la sulfasalazine) : pas de toxicité observée sur des traitements à court terme.

Statistiques et sciences expérimentales

Rumeurs pseudo-scientifiques sur le Covid-19

Sélection et invention de données

Tests d'hypothèses multiples

Publication de données non finalisées

Catégorisation fautive

Significativité, ampleur de l'effet, et effectifs

Coincidence attendue

Données non représentatives

Résultats tronqués

Pas de patient fractionnaire

Biais de composition des cohortes

Interprétation orientée

Le deuxième pic épidémique

Variants et vitesse de propagation

Conclusion

Résultats tronqués

Source : une **pré-publication** (mise en ligne le 10 avril, ensuite **publiée dans un journal** le 21 août) évalue la toxicité de l'hydroxychloroquine (avec ou sans azithromycine) sur de larges cohortes, en étude rétrospective.

Plus de 900 000 patients traités à l'hydroxychloroquine (référence : plus de 300 000 patients traités à la sulfasalazine) : pas de toxicité observée sur des traitements à court terme.

Plus de 300 000 patients traités à l'hydroxychloroquine+azithromycine (référence : plus de 350 000 patients traités à l'hydroxychloroquine+amoxicilline) : doublement du risque de mort par arrêt cardiaque sous 30 jours.

Statistiques et sciences expérimentales

Rumeurs pseudo-scientifiques sur le Covid-19

Sélection et invention de données

Tests d'hypothèses multiples

Publication de données non finalisées

Catégorisation fautive

Significativité, ampleur de l'effet, et effectifs

Coincidence attendue

Données non représentatives

Résultats tronqués

Pas de patient fractionnaire

Biais de composition des cohortes

Interprétation orientée

Le deuxième pic épidémique

Variants et vitesse de propagation

Conclusion

Résultats tronqués

D. Raoult cite cet article (il l'a donc lu, même évalué) pour affirmer que l'hydroxychloroquine n'augmente pas le risque cardiaque ... en passant sous silence l'association avec l'azithromycine (qui, elle, l'augmente) [[Source](#)].

Statistiques et sciences expérimentales

Rumeurs pseudo-scientifiques sur le Covid-19

Sélection et invention de données

Tests d'hypothèses multiples

Publication de données non finalisées

Catégorisation fautive

Significativité, ampleur de l'effet, et effectifs

Coincidence attendue

Données non représentatives

Résultats tronqués

Pas de patient fractionnaire

Biais de composition des cohortes

Interprétation orientée

Le deuxième pic épidémique

Variants et vitesse de propagation

Conclusion

Pas de patient fractionnaire

Source : vidéo de l'IHU Marseille du 19/05/2020 [\[Lien\]](#)

Argumentaire : en Île-de-France et Grand-Est, de même qu'en Chine en début d'épidémie, des patients jeunes sont morts, mais pas dans la cohorte traitée à l'IHU de Marseille. Donc nous avons sauvé des jeunes patients.

Statistiques et sciences expérimentales

Rumeurs pseudo-scientifiques sur le Covid-19

Sélection et invention de données

Tests d'hypothèses multiples

Publication de données non finalisées

Catégorisation fautive

Significativité, ampleur de l'effet, et effectifs

Coincidence attendue

Données non représentatives

Résultats tronqués

Pas de patient fractionnaire

Biais de composition des cohortes

Interprétation orientée

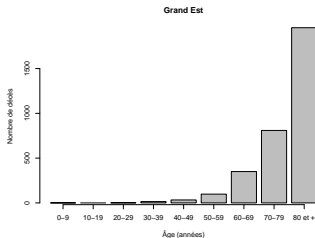
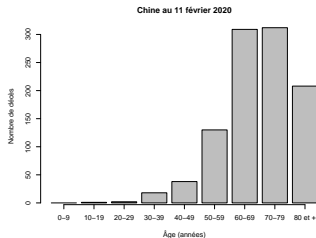
Le deuxième pic épidémique

Variants et vitesse de propagation

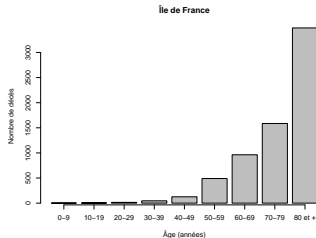
Conclusion

Pas de patient fractionnaire

80-89 et 90+↓



80-89 et 90+↓



Statistiques et sciences expérimentales

Rumeurs pseudo-scientifiques sur le Covid-19

Sélection et invention de données

Tests d'hypothèses multiples

Publication de données non finalisées

Catégorisation fautive

Significativité, ampleur de l'effet, et effectifs

Coincidence attendue

Données non représentatives

Résultats tronqués

Pas de patient fractionnaire

Biais de composition des cohortes

Interprétation orientée

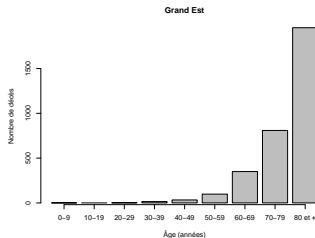
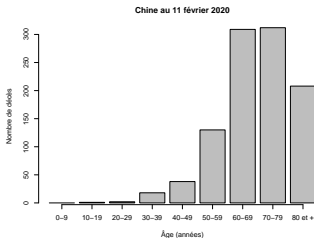
Le deuxième pic épidémique

Variants et vitesse de propagation

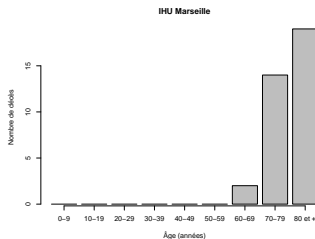
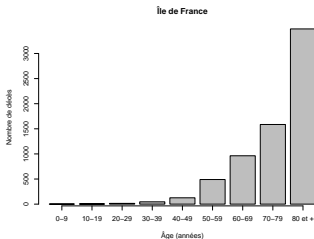
Conclusion

Pas de patient fractionnaire

80-89 et 90+↓



80-89 et 90+↓



Statistiques et sciences expérimentales

Rumeurs pseudo-scientifiques sur le Covid-19

Sélection et invention de données

Tests d'hypothèses multiples

Publication de données non finalisées

Catégorisation fautive

Significativité, ampleur de l'effet, et effectifs

Coincidence attendue

Données non représentatives

Résultats tronqués

Pas de patient fractionnaire

Biais de composition des cohortes

Interprétation orientée

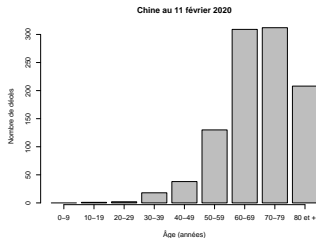
Le deuxième pic épidémique

Variants et vitesse de propagation

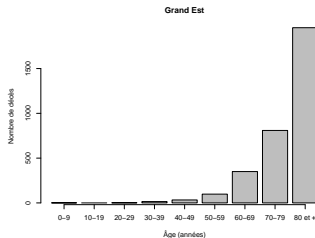
Conclusion

Du bon (ou du mauvais) choix des statistiques à la lumière de la pandémie

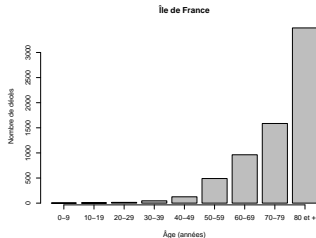
Pas de patient fractionnaire



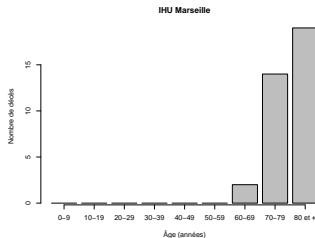
80-89 et 90+↓



80-89 et 90+↓



IHU Marseille



Statistiques et sciences expérimentales

Rumeurs pseudo-scientifiques sur le Covid-19

Sélection et invention de données

Tests d'hypothèses multiples

Publication de données non finalisées

Catégorisation fautive

Significativité, ampleur de l'effet, et effectifs

Coincidence attendue

Données non représentatives

Résultats tronqués

Pas de patient fractionnaire

Biais de composition des cohortes

Interprétation orientée

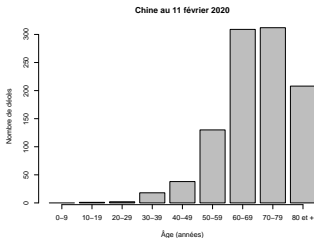
Le deuxième pic épidémique

Variants et vitesse de propagation

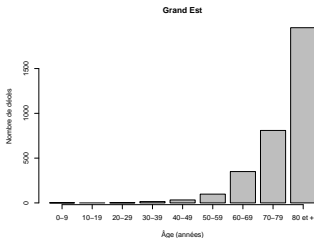
Conclusion

Du bon (ou du mauvais) choix des statistiques à la lumière de la pandémie

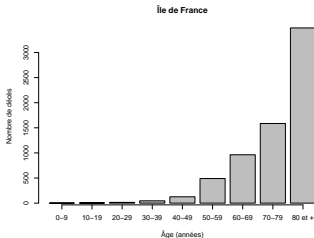
Pas de patient fractionnaire



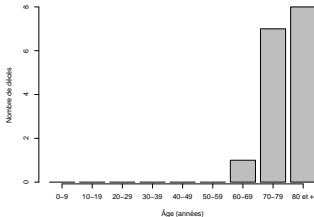
80-89 et 90+↓



80-89 et 90+↓



IHU Marseille traités bithérapie



Statistiques et sciences expérimentales

Rumeurs pseudo-scientifiques sur le Covid-19

Sélection et invention de données

Tests d'hypothèses multiples

Publication de données non finalisées

Catégorisation fautive

Significativité, ampleur de l'effet, et effectifs

Coincidence attendue

Données non représentatives

Résultats tronqués

Pas de patient fractionnaire

Biais de composition des cohortes

Interprétation orientée

Le deuxième pic épidémique

Variants et vitesse de propagation

Conclusion

Pas de patient fractionnaire

Un problème de distribution (nombres entiers, répartis entre catégories dont la répartition sera comparée entre différentes conditions)

Statistiques et sciences expérimentales

Rumeurs pseudo-scientifiques sur le Covid-19

Sélection et invention de données

Tests d'hypothèses multiples

Publication de données non finalisées

Catégorisation fautive

Significativité, ampleur de l'effet, et effectifs

Coincidence attendue

Données non représentatives

Résultats tronqués

Pas de patient fractionnaire

Biais de composition des cohortes

Interprétation orientée

Le deuxième pic épidémique

Variants et vitesse de propagation

Conclusion

Pas de patient fractionnaire

Un problème de distribution (nombres entiers, répartis entre catégories dont la répartition sera comparée entre différentes conditions) \implies **Test du χ^2 ou, de préférence : test exact de Fisher.**

Statistiques et sciences expérimentales

Rumeurs pseudo-scientifiques sur le Covid-19

Sélection et invention de données

Tests d'hypothèses multiples

Publication de données non finalisées

Catégorisation fautive

Significativité, ampleur de l'effet, et effectifs

Coincidence attendue

Données non représentatives

Résultats tronqués

Pas de patient fractionnaire

Biais de composition des cohortes

Interprétation orientée

Le deuxième pic épidémique

Variants et vitesse de propagation

Conclusion

Pas de patient fractionnaire

Un problème de distribution (nombres entiers, répartis entre catégories dont la répartition sera comparée entre différentes conditions) \implies **Test du χ^2 ou, de préférence : test exact de Fisher.**

Grand-Est contre IHU : $p = 0,46$

Grand-Est contre IHU bithérapie : $p = 0,56$

Île-de-France contre IHU : $p = 0,21$

Île-de-France contre IHU bithérapie : $p = 0,52$

Statistiques et sciences expérimentales

Rumeurs pseudo-scientifiques sur le Covid-19

Sélection et invention de données

Tests d'hypothèses multiples

Publication de données non finalisées

Catégorisation fautive

Significativité, ampleur de l'effet, et effectifs

Coincidence attendue

Données non représentatives

Résultats tronqués

Pas de patient fractionnaire

Biais de composition des cohortes

Interprétation orientée

Le deuxième pic épidémique

Variants et vitesse de propagation

Conclusion

Pas de patient fractionnaire

Un problème de distribution (nombres entiers, répartis entre catégories dont la répartition sera comparée entre différentes conditions) \implies **Test du χ^2 ou, de préférence : test exact de Fisher.**

Grand-Est contre IHU : $p = 0,46$

Grand-Est contre IHU bithérapie : $p = 0,56$

Île-de-France contre IHU : $p = 0,21$

Île-de-France contre IHU bithérapie : $p = 0,52$

\implies Pas de différence significative. S'il n'y a pas de décès de patient jeune à l'IHU Marseille, c'est plus probablement parce que la cohorte était trop petite pour les observer.

Statistiques et sciences expérimentales

Rumeurs pseudo-scientifiques sur le Covid-19

Sélection et invention de données

Tests d'hypothèses multiples

Publication de données non finalisées

Catégorisation fautive

Significativité, ampleur de l'effet, et effectifs

Coincidence attendue

Données non représentatives

Résultats tronqués

Pas de patient fractionnaire

Biais de composition des cohortes

Interprétation orientée

Le deuxième pic épidémique

Variants et vitesse de propagation

Conclusion

Biais de composition des cohortes

Source : articles de l'IHU Marseille présentant les résultats du traitement à l'hydroxychloroquine et azithromycine [[Source](#), [Source](#)].

Statistiques et sciences expérimentales

Rumeurs pseudo-scientifiques sur le Covid-19

Sélection et invention de données

Tests d'hypothèses multiples

Publication de données non finalisées

Catégorisation fautive

Significativité, ampleur de l'effet, et effectifs

Coincidence attendue

Données non représentatives

Résultats tronqués

Pas de patient fractionnaire

Biais de composition des cohortes

Interprétation orientée

Le deuxième pic épidémique

Variants et vitesse de propagation

Conclusion

Biais de composition des cohortes

Source : articles de l'IHU Marseille présentant les résultats du traitement à l'hydroxychloroquine et azithromycine [[Source](#), [Source](#)].

Puisque ce traitement peut provoquer des troubles cardiaques, des examens cardiaques sont nécessaires au cours du traitement. Patients à risque : interrompent le traitement.

Statistiques et sciences expérimentales

Rumeurs pseudo-scientifiques sur le Covid-19

Sélection et invention de données

Tests d'hypothèses multiples

Publication de données non finalisées

Catégorisation fautive

Significativité, ampleur de l'effet, et effectifs

Coincidence attendue

Données non représentatives

Résultats tronqués

Pas de patient fractionnaire

Biais de composition des cohortes

Interprétation orientée

Le deuxième pic épidémique

Variants et vitesse de propagation

Conclusion

Biais de composition des cohortes

Source : articles de l'IHU Marseille présentant les résultats du traitement à l'hydroxychloroquine et azithromycine [[Source](#), [Source](#)].

Puisque ce traitement peut provoquer des troubles cardiaques, des examens cardiaques sont nécessaires au cours du traitement. Patients à risque : interrompent le traitement.

Les patients qui interrompent le traitement dans les 3 premiers jours sont reversés dans la cohorte contrôle, donc enrichie en patients plus âgés, et avec des comorbidités.

Statistiques et sciences expérimentales

Rumeurs pseudo-scientifiques sur le Covid-19

Sélection et invention de données

Tests d'hypothèses multiples

Publication de données non finalisées

Catégorisation fautive

Significativité, ampleur de l'effet, et effectifs

Coincidence attendue

Données non représentatives

Résultats tronqués

Pas de patient fractionnaire

Biais de composition des cohortes

Interprétation orientée

Le deuxième pic épidémique

Variants et vitesse de propagation

Conclusion

Biais de composition des cohortes

Source : article initial du labo Raoult sur ce traitement ([Gautret *et al.*, 2020](#)) et son correctif [[Source](#)].

Du bon (ou du mauvais) choix des statistiques à la lumière de la pandémie

Statistiques et sciences expérimentales

Rumeurs pseudo-scientifiques sur le Covid-19

Sélection et invention de données

Tests d'hypothèses multiples

Publication de données non finalisées

Catégorisation fautive

Significativité, ampleur de l'effet, et effectifs

Coincidence attendue

Données non représentatives

Résultats tronqués

Pas de patient fractionnaire

Biais de composition des cohortes

Interprétation orientée

Le deuxième pic épidémique

Variants et vitesse de propagation

Conclusion

Biais de composition des cohortes

Source : article initial du labo Raoult sur ce traitement ([Gautret *et al.*, 2020](#)) et son correctif [[Source](#)].

Cohortes de petite taille, constituées sans randomisation.

Statistiques et sciences expérimentales

Rumeurs pseudo-scientifiques sur le Covid-19

Sélection et invention de données

Tests d'hypothèses multiples

Publication de données non finalisées

Catégorisation fautive

Significativité, ampleur de l'effet, et effectifs

Coincidence attendue

Données non représentatives

Résultats tronqués

Pas de patient fractionnaire

Biais de composition des cohortes

Interprétation orientée

Le deuxième pic épidémique

Variants et vitesse de propagation

Conclusion

Biais de composition des cohortes

Source : article initial du labo Raoult sur ce traitement ([Gautret et al., 2020](#)) et son correctif [[Source](#)].

Cohortes de petite taille, constituées sans randomisation.

Cohorte « contrôle » : malades depuis moins longtemps ; cohortes traitées : davantage de patients dont l'infection est descendue dans les poumons.

Statistiques et sciences expérimentales

Rumeurs pseudo-scientifiques sur le Covid-19

Sélection et invention de données

Tests d'hypothèses multiples

Publication de données non finalisées

Catégorisation fautive

Significativité, ampleur de l'effet, et effectifs

Coincidence attendue

Données non représentatives

Résultats tronqués

Pas de patient fractionnaire

Biais de composition des cohortes

Interprétation orientée

Le deuxième pic épidémique

Variants et vitesse de propagation

Conclusion

Biais de composition des cohortes

Source : article initial du labo Raoult sur ce traitement ([Gautret et al., 2020](#)) et son correctif [[Source](#)].

Cohortes de petite taille, constituées sans randomisation.

Cohorte « contrôle » : malades depuis moins longtemps ; cohortes traitées : davantage de patients dont l'infection est descendue dans les poumons.

Le test PCR ne détecte plus l'ARN viral quand l'infection est suffisamment ancienne.

Statistiques et sciences expérimentales

Rumeurs pseudo-scientifiques sur le Covid-19

Sélection et invention de données

Tests d'hypothèses multiples

Publication de données non finalisées

Catégorisation fautive

Significativité, ampleur de l'effet, et effectifs

Coincidence attendue

Données non représentatives

Résultats tronqués

Pas de patient fractionnaire

Biais de composition des cohortes

Interprétation orientée

Le deuxième pic épidémique

Variants et vitesse de propagation

Conclusion

Biais de composition des cohortes

Source : article initial du labo Raoult sur ce traitement ([Gautret et al., 2020](#)) et son correctif [[Source](#)].

Cohortes de petite taille, constituées sans randomisation.

Cohorte « contrôle » : malades depuis moins longtemps ; cohortes traitées : davantage de patients dont l'infection est descendue dans les poumons.

Le test PCR ne détecte plus l'ARN viral quand l'infection est suffisamment ancienne.

Nécessité de tenir compte de ces variables confondantes (et de respecter les données, et d'éviter les tests d'hypothèses multiples).

Statistiques et sciences expérimentales

Rumeurs pseudo-scientifiques sur le Covid-19

Sélection et invention de données

Tests d'hypothèses multiples

Publication de données non finalisées

Catégorisation fautive

Significativité, ampleur de l'effet, et effectifs

Coincidence attendue

Données non représentatives

Résultats tronqués

Pas de patient fractionnaire

Biais de composition des cohortes

Interprétation orientée

Le deuxième pic épidémique

Variants et vitesse de propagation

Conclusion

Biais de composition des cohortes

Source : article initial du labo Raoult sur ce traitement ([Gautret et al., 2020](#)) et son correctif [[Source](#)].

Cohortes de petite taille, constituées sans randomisation.

Cohorte « contrôle » : malades depuis moins longtemps ; cohortes traitées : davantage de patients dont l'infection est descendue dans les poumons.

Le test PCR ne détecte plus l'ARN viral quand l'infection est suffisamment ancienne.

Nécessité de tenir compte de ces variables confondantes (et de respecter les données, et d'éviter les tests d'hypothèses multiples). Une telle analyse ne détecte pas d'effet du traitement [[Lien](#)].

Statistiques et sciences expérimentales

Rumeurs pseudo-scientifiques sur le Covid-19

Sélection et invention de données

Tests d'hypothèses multiples

Publication de données non finalisées

Catégorisation fautive

Significativité, ampleur de l'effet, et effectifs

Coincidence attendue

Données non représentatives

Résultats tronqués

Pas de patient fractionnaire

Biais de composition des cohortes

Interprétation orientée

Le deuxième pic épidémique

Variants et vitesse de propagation

Conclusion

Interprétation orientée

Source : le 15 juillet 2020, un article signé par des membres du collectif « Laissons les médecins prescrire » (V. Guérin, M. Wonner, ...) compare des cohortes traitées à l'azithromycine (avec ou sans hydroxychloroquine) à une cohorte traitée autrement [[Lien](#)].

Statistiques et sciences expérimentales

Rumeurs pseudo-scientifiques sur le Covid-19

Sélection et invention de données

Tests d'hypothèses multiples

Publication de données non finalisées

Catégorisation fautive

Significativité, ampleur de l'effet, et effectifs

Coincidence attendue

Données non représentatives

Résultats tronqués

Pas de patient fractionnaire

Biais de composition des cohortes

Interprétation orientée

Le deuxième pic épidémique

Variants et vitesse de propagation

Conclusion

Interprétation orientée

Source : le 15 juillet 2020, un article signé par des membres du collectif « Laissons les médecins prescrire » (V. Guérin, M. Wonner, ...) compare des cohortes traitées à l'azithromycine (avec ou sans hydroxychloroquine) à une cohorte traitée autrement [[Lien](#)].

Patients : médecins eux-mêmes, ou leur entourage ; 51 sur 88 sont confirmés Covid-19 par PCR (les 37 autres : suspicion sur la base des symptômes) ; mesure du délai avant « guérison » (non décrite : apparemment, auto-évaluation).

Statistiques et sciences expérimentales

Rumeurs pseudo-scientifiques sur le Covid-19

Sélection et invention de données

Tests d'hypothèses multiples

Publication de données non finalisées

Catégorisation fautive

Significativité, ampleur de l'effet, et effectifs

Coincidence attendue

Données non représentatives

Résultats tronqués

Pas de patient fractionnaire

Biais de composition des cohortes

Interprétation orientée

Le deuxième pic épidémique

Variants et vitesse de propagation

Conclusion

Interprétation orientée

Source : le 15 juillet 2020, un article signé par des membres du collectif « Laissons les médecins prescrire » (V. Guérin, M. Wonner, ...) compare des cohortes traitées à l'azithromycine (avec ou sans hydroxychloroquine) à une cohorte traitée autrement [[Lien](#)].

Patients : médecins eux-mêmes, ou leur entourage ; 51 sur 88 sont confirmés Covid-19 par PCR (les 37 autres : suspicion sur la base des symptômes) ; mesure du délai avant « guérison » (non décrite : apparemment, auto-évaluation).

AZM	AZM+HCQ	Autre traitement	Aucun traitement
34	20	23	11

(AZM : azithromycine, ou (pour 2 patients) clarithromycine ; HCQ : hydroxychloroquine)

Statistiques et sciences expérimentales

Rumeurs pseudo-scientifiques sur le Covid-19

Sélection et invention de données

Tests d'hypothèses multiples

Publication de données non finalisées

Catégorisation fautive

Significativité, ampleur de l'effet, et effectifs

Coincidence attendue

Données non représentatives

Résultats tronqués

Pas de patient fractionnaire

Biais de composition des cohortes

Interprétation orientée

Le deuxième pic épidémique

Variants et vitesse de propagation

Conclusion

Interprétation orientée

Résultats : différence significative entre « AZM » et « autres/aucun traitement », mais pas entre « AZM » et « AZM+HCQ ».

Statistiques et sciences expérimentales

Rumeurs pseudo-scientifiques sur le Covid-19

Sélection et invention de données

Tests d'hypothèses multiples

Publication de données non finalisées

Catégorisation fautive

Significativité, ampleur de l'effet, et effectifs

Coincidence attendue

Données non représentatives

Résultats tronqués

Pas de patient fractionnaire

Biais de composition des cohortes

Interprétation orientée

Le deuxième pic épidémique

Variants et vitesse de propagation

Conclusion

Interprétation orientée

Résultats : différence significative entre « AZM » et « autres/aucun traitement », mais pas entre « AZM » et « AZM+HCQ ».

—> montre un effet bénéfique de l'antibiotique (attendu en cas de surinfection bactérienne), mais pas d'effet de la bithérapie avec hydroxychloroquine.

Statistiques et sciences expérimentales

Rumeurs pseudo-scientifiques sur le Covid-19

Sélection et invention de données

Tests d'hypothèses multiples

Publication de données non finalisées

Catégorisation fautive

Significativité, ampleur de l'effet, et effectifs

Coincidence attendue

Données non représentatives

Résultats tronqués

Pas de patient fractionnaire

Biais de composition des cohortes

Interprétation orientée

Le deuxième pic épidémique

Variants et vitesse de propagation

Conclusion

Interprétation orientée

Résultats : différence significative entre « AZM » et « autres/aucun traitement », mais pas entre « AZM » et « AZM+HCQ ».

—> montre un effet bénéfique de l'antibiotique (attendu en cas de surinfection bactérienne), mais pas d'effet de la bithérapie avec hydroxychloroquine.

Pourtant les auteurs titrent sur l'efficacité de la bithérapie : *Azithromycin and hydroxychloroquine accelerate recovery of outpatients with mild/moderate COVID-19.*

Statistiques et sciences expérimentales

Rumeurs pseudo-scientifiques sur le Covid-19

Sélection et invention de données

Tests d'hypothèses multiples

Publication de données non finalisées

Catégorisation fautive

Significativité, ampleur de l'effet, et effectifs

Coincidence attendue

Données non représentatives

Résultats tronqués

Pas de patient fractionnaire

Biais de composition des cohortes

Interprétation orientée

Le deuxième pic épidémique

Variants et vitesse de propagation

Conclusion

Interprétation orientée

Résultats : différence significative entre « AZM » et « autres/aucun traitement », mais pas entre « AZM » et « AZM+HCQ ».

→ montre un effet bénéfique de l'antibiotique (attendu en cas de surinfection bactérienne), mais pas d'effet de la bithérapie avec hydroxychloroquine.

Pourtant les auteurs titrent sur l'efficacité de la bithérapie : *Azithromycin and hydroxychloroquine accelerate recovery of outpatients with mild/moderate COVID-19.*

Illustration de la pollution de la littérature scientifique par des revues prédatrices (ici, *Asian journal of medicine and health*, comme plaisamment démontré par la publication d'un article-canular ; [[Lien](#) ; avant rétractation : [Lien](#)]).

Statistiques et sciences expérimentales

Rumeurs pseudo-scientifiques sur le Covid-19

Sélection et invention de données

Tests d'hypothèses multiples

Publication de données non finalisées

Catégorisation fautive

Significativité, ampleur de l'effet, et effectifs

Coincidence attendue

Données non représentatives

Résultats tronqués

Pas de patient fractionnaire

Biais de composition des cohortes

Interprétation orientée

Le deuxième pic épidémique

Variants et vitesse de propagation

Conclusion

Le deuxième pic épidémique

Source : J.-F. Toussaint dément la dangerosité de la deuxième vague au motif que les tests positifs augmentent plus vite que les décès [[Lien](#)].

Statistiques et sciences expérimentales

Rumeurs pseudo-scientifiques sur le Covid-19

Sélection et invention de données

Tests d'hypothèses multiples

Publication de données non finalisées

Catégorisation fautive

Significativité, ampleur de l'effet, et effectifs

Coincidence attendue

Données non représentatives

Résultats tronqués

Pas de patient fractionnaire

Biais de composition des cohortes

Interprétation orientée

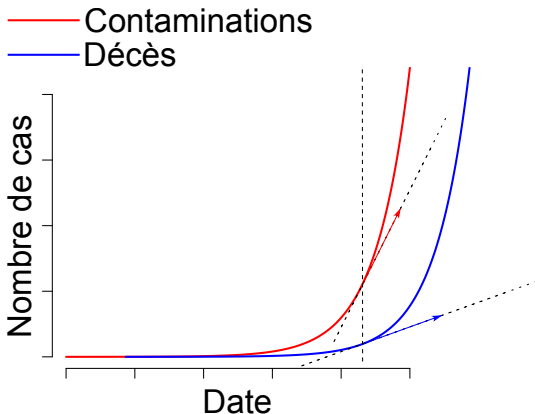
Le deuxième pic épidémique

Variants et vitesse de propagation

Conclusion

Le deuxième pic épidémique

Source : J.-F. Toussaint dément la dangerosité de la deuxième vague au motif que les tests positifs augmentent plus vite que les décès [[Lien](#)].



Statistiques et sciences expérimentales

Rumeurs pseudo-scientifiques sur le Covid-19

Sélection et invention de données

Tests d'hypothèses multiples

Publication de données non finalisées

Catégorisation fautive

Significativité, ampleur de l'effet, et effectifs

Coincidence attendue

Données non représentatives

Résultats tronqués

Pas de patient fractionnaire

Biais de composition des cohortes

Interprétation orientée

Le deuxième pic épidémique

Variants et vitesse de propagation

Conclusion

Variants et vitesse de propagation

Source : D. Raoult estime que le 2ème pic épidémique, à l'automne, est dû à l'émergence d'un nouveau variant (une « nouvelle épidémie ») [[Lien](#)].

Statistiques et sciences expérimentales

Rumeurs pseudo-scientifiques sur le Covid-19

Sélection et invention de données

Tests d'hypothèses multiples

Publication de données non finalisées

Catégorisation fautive

Significativité, ampleur de l'effet, et effectifs

Coincidence attendue

Données non représentatives

Résultats tronqués

Pas de patient fractionnaire

Biais de composition des cohortes

Interprétation orientée

Le deuxième pic épidémique

Variants et vitesse de propagation

Conclusion

Variants et vitesse de propagation

Source : D. Raoult estime que le 2ème pic épidémique, à l'automne, est dû à l'émergence d'un nouveau variant (une « nouvelle épidémie ») [[Lien](#)].

Corrélation n'est pas causalité. Le pic épidémique se fait avec le variant du moment, mais ne lui est dû que si le variant est lui-même plus contagieux (\neq variants dits « anglais », « sud-africain », « brésilien/japonais » de l'hiver).

Statistiques et sciences expérimentales

Rumeurs pseudo-scientifiques sur le Covid-19

Sélection et invention de données

Tests d'hypothèses multiples

Publication de données non finalisées

Catégorisation fautive

Significativité, ampleur de l'effet, et effectifs

Coincidence attendue

Données non représentatives

Résultats tronqués

Pas de patient fractionnaire

Biais de composition des cohortes

Interprétation orientée

Le deuxième pic épidémique

Variants et vitesse de propagation

Conclusion

Variants et vitesse de propagation

Source : D. Raoult estime que le 2ème pic épidémique, à l'automne, est dû à l'émergence d'un nouveau variant (une « nouvelle épidémie ») [[Lien](#)].

Corrélation n'est pas causalité. Le pic épidémique se fait avec le variant du moment, mais ne lui est dû que si le variant est lui-même plus contagieux (\neq variants dits « anglais », « sud-africain », « brésilien/japonais » de l'hiver).

Exemple : [[Lien](#)].

Statistiques et sciences expérimentales

Rumeurs pseudo-scientifiques sur le Covid-19

Sélection et invention de données

Tests d'hypothèses multiples

Publication de données non finalisées

Catégorisation fautive

Significativité, ampleur de l'effet, et effectifs

Coincidence attendue

Données non représentatives

Résultats tronqués

Pas de patient fractionnaire

Biais de composition des cohortes

Interprétation orientée

Le deuxième pic épidémique

Variants et vitesse de propagation

Conclusion

Conclusion

Du bon (ou du mauvais) choix des statistiques à la lumière de la pandémie

Statistiques et sciences expérimentales

Rumeurs pseudo-scientifiques sur le Covid-19

Sélection et invention de données

Tests d'hypothèses multiples

Publication de données non finalisées

Catégorisation fautive

Significativité, ampleur de l'effet, et effectifs

Coincidence attendue

Données non représentatives

Résultats tronqués

Pas de patient fractionnaire

Biais de composition des cohortes

Interprétation orientée

Le deuxième pic épidémique

Variants et vitesse de propagation

Conclusion

Conclusion

Les statistiques sont indispensables en sciences expérimentales.

Du bon (ou du mauvais) choix des statistiques à la lumière de la pandémie

Statistiques et sciences expérimentales

Rumeurs pseudo-scientifiques sur le Covid-19

Sélection et invention de données

Tests d'hypothèses multiples

Publication de données non finalisées

Catégorisation fautive

Significativité, ampleur de l'effet, et effectifs

Coincidence attendue

Données non représentatives

Résultats tronqués

Pas de patient fractionnaire

Biais de composition des cohortes

Interprétation orientée

Le deuxième pic épidémique

Variants et vitesse de propagation

Conclusion

Conclusion

Les statistiques sont indispensables en sciences expérimentales.

Le meilleur outil ne donnera rien s'il est mal utilisé.

Statistiques et sciences expérimentales

Rumeurs pseudo-scientifiques sur le Covid-19

Sélection et invention de données

Tests d'hypothèses multiples

Publication de données non finalisées

Catégorisation fautive

Significativité, ampleur de l'effet, et effectifs

Coincidence attendue

Données non représentatives

Résultats tronqués

Pas de patient fractionnaire

Biais de composition des cohortes

Interprétation orientée

Le deuxième pic épidémique

Variants et vitesse de propagation

Conclusion

Conclusion

Les statistiques sont indispensables en sciences expérimentales.

Le meilleur outil ne donnera rien s'il est mal utilisé.

Un contexte spécial, qui a révélé incompétence et malhonnêteté chez des responsables médico-scientifiques puissants.

Statistiques et sciences expérimentales

Rumeurs pseudo-scientifiques sur le Covid-19

Sélection et invention de données

Tests d'hypothèses multiples

Publication de données non finalisées

Catégorisation fautive

Significativité, ampleur de l'effet, et effectifs

Coincidence attendue

Données non représentatives

Résultats tronqués

Pas de patient fractionnaire

Biais de composition des cohortes

Interprétation orientée

Le deuxième pic épidémique

Variants et vitesse de propagation

Conclusion

Conclusion

Les statistiques sont indispensables en sciences expérimentales.

Le meilleur outil ne donnera rien s'il est mal utilisé.

Un contexte spécial, qui a révélé incompétence et malhonnêteté chez des responsables médico-scientifiques puissants.

- ▶ Cas isolés ?
- ▶ Un problème systémique ?

Statistiques et sciences expérimentales

Rumeurs pseudo-scientifiques sur le Covid-19

Sélection et invention de données

Tests d'hypothèses multiples

Publication de données non finalisées

Catégorisation fautive

Significativité, ampleur de l'effet, et effectifs

Coincidence attendue

Données non représentatives

Résultats tronqués

Pas de patient fractionnaire

Biais de composition des cohortes

Interprétation orientée

Le deuxième pic épidémique

Variants et vitesse de propagation

Conclusion

Conclusion

Les statistiques sont indispensables en sciences expérimentales.

Le meilleur outil ne donnera rien s'il est mal utilisé.

Un contexte spécial, qui a révélé incompétence et malhonnêteté chez des responsables médico-scientifiques puissants.

- ▶ Cas isolés ?
- ▶ Un problème systémique ?

Au printemps 2020, l'esprit critique semblait manquer chez la plupart des experts et politiciens appelés à commenter [exemple, exemple, exemple, ...].

Statistiques et sciences expérimentales

Rumeurs pseudo-scientifiques sur le Covid-19

Sélection et invention de données

Tests d'hypothèses multiples

Publication de données non finalisées

Catégorisation fautive

Significativité, ampleur de l'effet, et effectifs

Coincidence attendue

Données non représentatives

Résultats tronqués

Pas de patient fractionnaire

Biais de composition des cohortes

Interprétation orientée

Le deuxième pic épidémique

Variants et vitesse de propagation

Conclusion

Conclusion

Les statistiques sont indispensables en sciences expérimentales.

Le meilleur outil ne donnera rien s'il est mal utilisé.

Un contexte spécial, qui a révélé incompétence et malhonnêteté chez des responsables médico-scientifiques puissants.

- ▶ Cas isolés ?
- ▶ Un problème systémique ?

Au printemps 2020, l'esprit critique semblait manquer chez la plupart des experts et politiciens appelés à commenter [exemple, exemple, exemple, ...].

Importance de la culture scientifique dans le grand public (à améliorer avant la prochaine pandémie ...).

Statistiques et sciences expérimentales

Rumeurs pseudo-scientifiques sur le Covid-19

Sélection et invention de données

Tests d'hypothèses multiples

Publication de données non finalisées

Catégorisation fautive

Significativité, ampleur de l'effet, et effectifs

Coincidence attendue

Données non représentatives

Résultats tronqués

Pas de patient fractionnaire

Biais de composition des cohortes

Interprétation orientée

Le deuxième pic épidémique

Variants et vitesse de propagation

Conclusion