Analyse de l'article

Hydroxychloroquine and azithromycin as a treatment of COVID-19: results of an open-label non-randomized clinical trial

Remarques sur le protocole

Le protocole de cette étude a été déposé sur le registre des essais cliniques de l'Agence européenne du médicament (EU clinical trials register) le 10 mars 2020.
Les éléments de protocole disponible sur ce site font état

- d'un essai ouvert monocentrique prévoyant d'inclure 25 patients atteint d'une infection respiratoire à SARS-COV2 et traités par hydroxychloroquine 600 mg par jour
- ayant eu une autorisation ANSM le 5 mars et pas de date d'autorisation du CPP (l'article dit que le CPP a donné son accord le 6 mars 2020)
- dont l'objectif principal est de raccourcir la période de contagion ('To shorten the period of virus carrying and thus contagion')
- et les objectifs secondaires sont d'évaluer l'efficacité clinique sur le délai jusqu'à la disparition de la fièvre, la normalisation du rythme respiratoire, la durée de séjour à l'hôpital et la mortalité ('To evaluate the clinical effectiveness of treatment on time to apyrexia, normalization of respiratory rate, and average length of hospital stay and mortality').
- Le critère primaire de jugement est le résultat de la détection du virus ('Results of SARS-COV2 virus detection') avec des évaluations à Jour 1, 4, 7 et 14

L'article décrit une étude légèrement différente puisqu'il y a des patients non traités (témoins) dont on dit que ce sont les patients ayant des critères de non-inclusion (intolérance au traitement de l'essai, grossesse, allaitement)- par définition de mauvais témoins-, ou ayant refusé l'inclusion dans l'étude ou traités dans d'autres centres cliniques qu'à Marseille en plus des personnes incluses dans l'essai. Les patients traités ont été inclus dans l'essai entre début mars -avec une ambiguïté sur quand précisément- et le 16 mars. On ne connait pas les dates pour les témoins.

Dans l'article on parle de prélèvements journaliers (et non à J1, 4, 7 et 14) et on définit le critère primaire de jugement comme étant à J6 post inclusion (donc une ambiguïté sur savoir si cela correspond ou non à J7).

Concernant le calcul du nombre de sujets nécessaires, il est conduit comme si on avait une étude randomisée avec 2 bras (traité et non traité) et les éléments fournis ne permettent pas de reproduire ce calcul.
Le paragraphe est le suivant :
Assuming a 50% efficacy of hydroxychloroquine in reducing the viral load at day 7, a 85% power, a type I error rate of 5% and 10% loss to follow-up, we calculated that a total of 48 COVID-19 patients (ie, 24 cases in the hydroxychloroquine group and 24 in the control group) would be required for the analysis(Fleiss with CC).
Remarquons que le paragraphe en question parle, comme le protocole, d'une évaluation à 7 jours (et non 6) maintenant l'ambiguïté sur ce point.

On comprend que le calcul est fait en utilisant une correction de continuité (Fleiss CC) avec un risque de 1ere espèce de 5 % et une puissance de 85 %. On ne sait pas est si le calcul est fait pour un test uni- ou bilatéral. L'information de 50 % d'efficacité est imprécise. Il aurait fallu avoir le taux de de charge virale non détectable dans le groupe non traité et celui dans le groupe traité pour faire le calcul. Dire 50 % d'efficacité peut vouloir dire :

- 2 fois plus dans le groupe traité que dans le groupe non traité
- ou 50 % versus 0 %
- ou une différence de 50 % entre les 2 taux de succès.

Voici ce qu'on obtient avec différentes hypothèses d'efficacité dans les 2 groupes et en supposant un test bilatéral, un risque de 1ere espèce de 5 %, une puissance de 85 % en utilisant nQuery :

% de succès Non Traités	0,10	0,20	0,30	0,35	0,10	0,10
% de succès Traités	0,20	0,40	0,60	0,70	0,50	0,60
N par groupe	247	103	55	41	26	18
N avec 10% de perdus de vue	272	113	60	45	29	20

Remarques sur les résultats

Dans la section résultats de l'article, on mélange les personnes incluses dans l'essai avec les témoins en commençant par dire que 36 des 42 personnes incluses, qui ont le recul de 6 jours (toujours l'ambiguïté sur la date précise de l'évaluation), sont incluses dans l'analyse réalisée.
Il y a 16 non traités, sans que l'on sache à quelle catégorie de non traités ils appartiennent (critère de non inclusion, refus, autres centres).
Sur les 26 traités, on nous explique ensuite que 6 patients ont été perdus de vue, dont 3 pour passage en réanimation à J2, J3 et J4, 1 pour décès à J2, 1 qui est sortie de l'hôpital (décision du patient) à J3 et un arrêt de traitement pour évènement indésirable (nausée) à J3. Ces 6 patients sont exclus de l'analyse et leur caractéristiques ne sont pas décrites.
Cela pose un problème sévère, une analyse rigoureuse aurait dû considérer tout ou partie de ces cas comme des échecs. C'est par exemple ainsi qu'on analyse les essais dans le domaine du VIH ou le critère primaire est aussi un taux de charge virale en dessous du seuil de détection, toute donnée manquante, quel que soit la cause étant analysée comme un échec.

L'article ne porte donc que sur 36 personnes dont 20 des 26 inclus dans l'essai (sur un nombre attendu de 25 selon le registre des essais européens) et seules ces personnes sont décrites.

A J6 le nombre de personnes avec une charge virale en dessous du seuil de détection est de 14/20 chez les traités analysés et de 2/16 chez les non traités analysés avec un test exact de Fisher avec un degré de signification de 0,001.
Mais en fait ces résultats, ceux présentés par les auteurs, ne sont pas conformes à ce qui est été réalisé selon la suplementary Table 1. En effet dans le groupe non traité 11 personnes sur 16 ont été prélevées et 5 ne l'ont pas été et sont comptées comme positives ! En revanche dans le groupe traité, un participant n'a pas été prélevé et 19 l'ont été et le patient non prélevé a été compté comme négatif. Une analyse destinée à favoriser le groupe traitement.
Si on exclut les patients non prélevés alors 2/11 versus 13/19 p=0,021 et dans l'hypothèse la pire, inverse de celle choisie par les auteurs (manquant = négatif chez les non traités et manquant = positif chez les traités) cela devient 7/16 versus 13/20 p= 0,313.

Les résultats portant sur l'ajout d'azithromycine, essentiellement sur des patients avec pneumonie (4 des 6 recevant ce traitement), sont encore plus difficile à interpréter, d'autant que peu de personnes non traitées sont atteintes d'une pneumonie

S'agissant d'une étude non randomisée dont le groupe témoin est composé en proportion inconnue de personnes ayant des critères de non-inclusion ou ayant refusé de participer à l'essai ou suivi dans d'autres centres, il est impossible de juger si les 2 groupes sont ou non comparables et notamment sur le niveau de charge virale initiale puisque beaucoup des non traités sont évalués par uen technique qualitative et par conséquent d'apprécier si les éventuelles différences observées peuvent être attribuées ou non à l'exposition au traitement et ceci d'autant plus que les cas les plus sévères cliniquement (au moins 4 des 6 perdus de vue -3 passage en réanimation et un décès-) ont été exclus de la description et de l'analyse. Une complication supplémentaire émane du fait qu'il n'est pas certain que les évaluations virologiques aient été conduites de façon centralisé, ce qui est suggéré par la simple mention de positif et non d'un résultat quantitatif pour une grande proportion des témoins, alors que ce n'est jamais le cas pour les personnes traitées, ce qui pourrait entrainer des différences liés à la mesure et non au fait d'être ou non traités. Cela entraine aussi qu'on ne connait pas le niveau de charge virale dans les 2 groupes et bien sur il serait capital de savoir s'il est comparable ou non. L'analyse a été conduite en vue de favoriser le bras traité, probablement en utilisant la technique LOCF (last observation carried forward) totalement inadéquate dans le contexte.

De plus, si la prevention de la transmission est un critère important, dans la situation de crise actuelle le critère pertinent dans l'évaluation d'un traitement chez le spersonnes hospitalisées est plutôt un critère clinique tel que la nécessité de recourir à un respirateur ou la mortalité, ce qui a été le cas pour 4/26 patients du groupe traité dans cette étude et 0/20 du groupe non traité.

Au total, l'étude est conduite, décrite et analysée de façon non rigoureuse avec des imprécisions et des ambiguïtés et il s'agit d'une étude à fort risque de biais selon les standards internationaux. Dans ce contexte, il est donc impossible d'interpréter l'effet décrit comme étant attribuable au traitement par chloroquine. L'étude n'apporte juste aucune information du tout sur un éventuel effet de l'hydroxychloroquine.

Dominique COSTAGLIOLA

Membre de l'Académie des sciences
DRCE INSERM
Vice-Doyenne Déléguée Recherche de la Faculté de Médecine, Sorbonne Université

Directrice adjointe de l'Institut Pierre Louis d'Epidémiologie et de Santé Publique, Sorbonne Université, INSERM