



CRAIG LASSIG/REUTERS

CORONAVÍRUS

Quais as bandeiras vermelhas do estudo sobre a hidroxicloroquina na *Lancet*?

Professor do Instituto Superior Técnico explica ao PÚBLICO que sinais vermelhos encontra no estudo publicado na revista *The Lancet* sobre a hidroxicloroquina que levantou suspeitas de fraude. Autor principal do artigo científico pede desculpa por não ter garantido que a fonte dos dados era apropriada.

Filipa Almeida Mendes
5 de Junho de 2020, 16:02

Não tardou até que a comunidade científica levantasse questões sobre o estudo publicado, a 22 de Maio, pela prestigiada revista *The Lancet* sobre a hidroxicloroquina. A fiabilidade dos dados analisados no estudo, fornecidos pela Surgisphere (que se apresenta como uma empresa de análise de dados em saúde com sede nos Estados Unidos), foi posta em causa e vários meios de comunicação e especialistas começaram a investigar a própria empresa, ao passo que iam surgindo cada vez mais incongruências, levantando-se suspeitas de que tudo não terá passado de uma fraude.

O estudo referia ter analisado dados de cerca de 96 mil pessoas com covid-19 internadas em 671 hospitais por todo o mundo. Na última quinta-feira, três dos quatro autores pediram que o estudo fosse retirado da revista *The Lancet*.

Mas, afinal, quais são as bandeiras vermelhas em relação ao estudo? Mário Figueiredo, professor catedrático no Instituto Superior Técnico (Lisboa) e coordenador do mestrado em Engenharia e Ciência de Dados, destaca ao PÚBLICO que este é um estudo que inclui mais de 90 mil pacientes de hospitais de vários países, que começou logo por revelar “uma série de discrepâncias com os números”, nomeadamente em relação às vítimas mortais na Austrália. “O número de mortos na Austrália que eles reportavam, com base nos dados dos hospitais, ultrapassava o número de mortos de que havia [registo] na Austrália e houve uma série de pistas que levava a crer que havia problemas nos dados. Essa foi a principal causa de levantar suspeitas”, explica Mário Figueiredo.

Posteriormente, as atenções viraram-se para a Surgisphere. “Descobriu-se que era uma empresa muito pequenina, que tem muito pouca gente e que as pessoas que aparecem como funcionários não têm formação nessas áreas e isso pôs em causa a credibilidade técnica da empresa”, acrescenta o investigador português.

Saltou à vista o facto de este estudo ter “uma quantidade muito grande de dados”, o que poderá facilitar a sua subamostragem. “Posso pegar num conjunto de dados para mostrar o que eu quiser e posso tirar a conclusão que eu quiser, porque não sabemos como é que foi feita a selecção final”, nota Mário Figueiredo. “Em relação à hidroxicloroquina, provavelmente o critério para dar esse remédio foi diferente em diferentes hospitais. Basta que tenham escolhido hospitais em que esse medicamento só fosse dado a doentes que já estavam em pior condição, o que vai aumentar logo a probabilidade destes que o tomam morrerem.”



O investigador Mário Figueiredo DÉBORA RODRIGUES/IST



Anonimização dos dados

Mário Figueiredo, um dos académicos mais citados a nível mundial na investigação em *machine learning*, destaca que os autores “não explicitaram como é que retiraram os dados” nem partilharam a informação, pelo que “não é possível replicar as análises que fizeram para ver se essas análises e as conclusões estão correctas”, o que faz com que o estudo tenha “muito pouca validade”.

A Surgisphere recusou-se a partilhar o conjunto completo de dados, alegando que isso “iria violar os acordos com os clientes e requisitos de confidencialidade”. Mário Figueiredo refere que “esse é um problema muito crítico em análise estatística”, nomeadamente quando os dados envolvem pessoas.

“A anonimização dos dados é muito complicada e essa é também uma das críticas a fazer”, afirma, destacando que “não é óbvio” como é que “uma equipa tão pequenina” conseguiu aparentemente contactar tantos hospitais de vários países, “todos eles com leis de anonimização de dados muito diferentes”. É, portanto, questionável como é que conseguiram “verificar e satisfazer todos os critérios de anonimização”.

“Podem estar a jogar muito à defesa, porque provavelmente os dados que recolheram não estão devidamente anonimizados e isto não significa simplesmente retirar o nome das pessoas”, afirma. O professor do Instituto Superior Técnico nota ainda que “estes estudos que envolvem pessoas têm de ser sempre pré-validados por comissões éticas” de determinados hospitais e universidades, “o que demora muito tempo”. “É tudo muito delicado e isto foi aparentemente uma coisa feita muito à pressa”, conclui.

“A validação em ciência é crucial”

De forma a garantir a anonimização e a partilha dos dados, a solução poderá passar por “fornecer os dados agregados”, por exemplo por hospital ou mesmo por país. “Quanto maior for o agregado, maior é a garantia de anonimidade perante os pacientes”, nota Mário Figueiredo, sublinhando, porém, que os autores do estudo em questão “não explicaram como é que fizeram nada”, o que impossibilitou a validação do artigo. No entanto, “a validação em ciência é crucial”.

Questionado pelo jornal *The Guardian* sobre a forma como conseguiu a colaboração de tantos hospitais em todo o mundo, Sapan Desai, o director executivo da Surgisphere e co-autor do artigo na *The Lancet*, revelou que a empresa “serve como um agregador de dados e realiza análises de dados”. “Não somos responsáveis pela fonte dos dados, pelo que a tarefa trabalhosa de exportar os dados dos registos electrónicos de saúde, convertê-los para o formato exigido pelo nosso dicionário de dados [coleção de metadados] e remover completamente a identificação dos dados são realizados pelo parceiro de saúde”, respondeu ao jornal britânico.

Mário Figueiredo explica que é comum a exportação, conversão e anonimização dos dados estar a cabo dos parceiros e que a Surgisphere pode ter alguns “contratos com vários



hospitais e até com empresas que trabalham com hospitais”. Porém, destaca que a própria Surgisphere, uma empresa da qual “nunca tinha ouvido falar”, descarta a sua responsabilidade (nomeadamente em relação à admissão dos dados e aos registos electrónicos de saúde), mas depois fornece esses mesmos dados para um estudo científico “sem conseguir ela própria garantir a sua qualidade”.

“Tentam desresponsabilizar-se mas, por outro lado, publicam na *The Lancet* — que é uma revista de alta credibilidade — um *paper* que vai ter um impacto enorme” e para o qual todos os olhos estarão virados, afirma.

A fiabilidade dos dados, assim como a sua devida recolha, processamento e análise é, por isso, “completamente crucial”. “Os estudos clínicos são coisas muito delicadas, que têm de ser feitas com muito cuidado. Têm aspectos de anonimização, aspectos éticos de vários níveis e a qualidade dos dados é crucial. Se os dados forem maus, as conclusões são inválidas”, resume Mário Figueiredo.

Um pedido de desculpa

Mandeep R. Mehra, director executivo do Centro Avançado para as Doenças Cardíacas do Brigham and Women’s Hospital, em Boston, e principal autor do estudo garante, por sua vez, numa resposta enviada por *e-mail* ao PÚBLICO, que sua intenção era “contribuir para a discussão científica” e assegurar que a medicina é praticada com base “nas melhores provas disponíveis”.

O investigador explica que, através de um dos seus co-autores, foi apresentado a Sapan Desai e à Surgisphere, “uma empresa privada que alegava ter dados de hospitais de todo o mundo, que poderiam ser aproveitados para responder a questões de saúde pública importantes” que o próprio colocou perante a pandemia da covid-19. “Para responder a essas questões, o Dr. Desai, que foi co-autor e cuja equipa mantinha esta base de dados observacionais, conduziu várias análises. Enquanto primeiro autor, elas foram-me fornecidas e, com base nessas análises, foram publicados dois artigos revistos pelos pares, um na *New England Journal of Medicine* (NEJM) e outro na *The Lancet*”, afirma Mandeep R. Mehra.

O investigador sublinha que “quando começaram a surgir discrepâncias nos dados”, ele e os restantes co-autores “solicitaram imediatamente uma reanálise [por parte] da Surgisphere” e contrataram o Medical Technology & Practice Patterns Institute (uma organização não governamental) para realizar uma auditoria, mas a Surgisphere recusou-se a partilhar os dados. Perante a incapacidade “de verificar a fonte primária”, Mandeep R. Mehra destaca que perdeu a confiança “na origem e veracidade dos dados ou nos resultados a que eles levaram”.

“É agora claro para mim que, na minha esperança de contribuir com esta investigação durante um período de grande necessidade, não fiz o suficiente para garantir que a fonte dos dados era apropriada para este uso. Por isso, e por todas as perturbações — directa e indirectamente — peço desculpa”, conclui.