Does Madagascar's herbal remedy, COVID-Organics work?

Article · June 2020

CITATIONS
0

READS
3,471

1 author:

Rick Sheridan
EMSKE Phytochem
6 PUBLICATIONS **0** CITATIONS

SEE PROFILE

Does Madagascar's herbal remedy, COVID-Organics work?

towardsdatascience.com/does-madagascars-covid-organics-work-74f577eb415b

January 5, 2021

A machine-learning approach to in silico drug docking sheds light

So over here in eastern Africa, a lot of news has been made of Madagascar's COVID-Organics extract made from a local plant, introduced by that country's Institut Malgache de Recherches Appliquées (Malagasy Institute of Applied Research) IMRA, and promulgated especially by the president of Madagascar himself, Andry Rajoelina.

COVID-Organics as extracted from Artemisia Annua (Credit: LHS: EMSKE Phytochem; RHS: Photo by Mathias Katz on Unsplash)

(You can find a less technical treatment of the topic **here**)

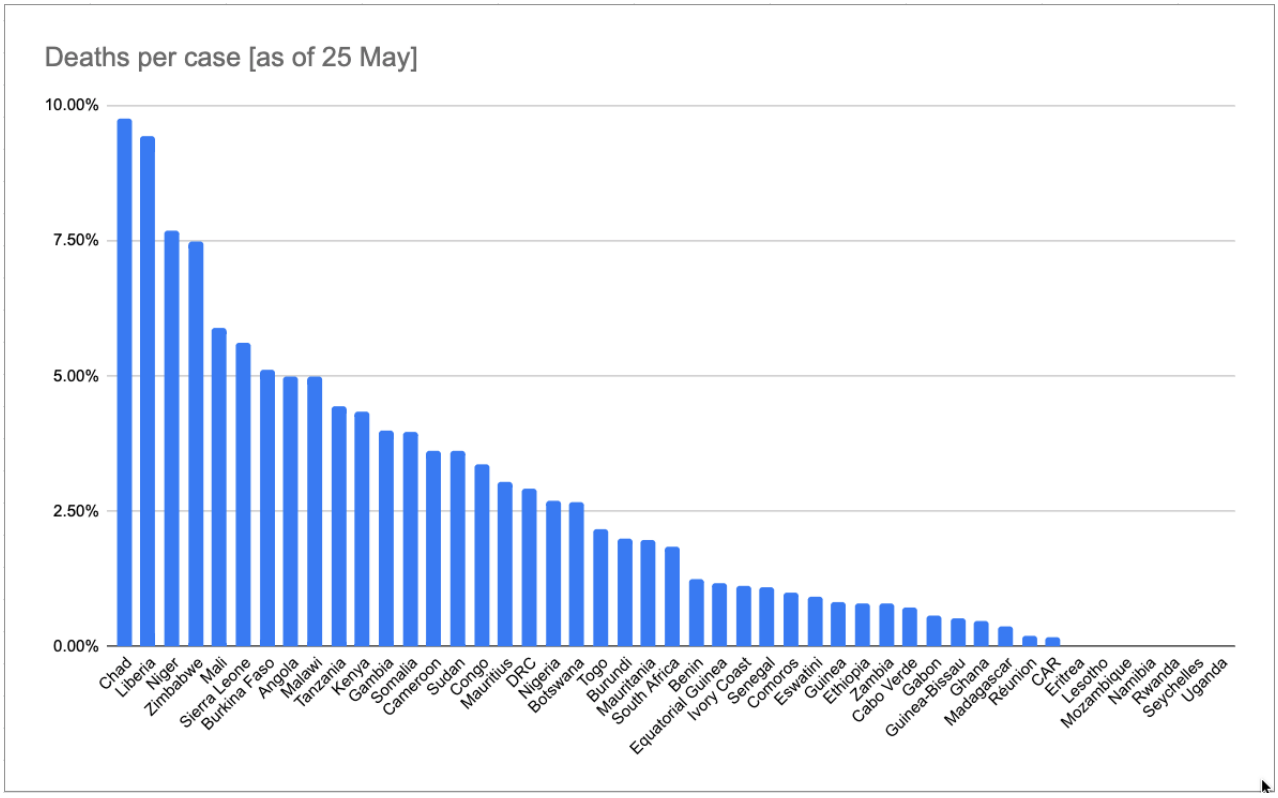
My experience is that articles on the topic particularly written for audiences outside Africa, like this one from the WSJ, tend to have a condescending tone. In contrast, representatives of many African countries have made a pilgrimage to ecologically-rich Madagascar to pick up a shipment of COVID-Organics, citing their own nations' histories of relying on traditional medicine.

Meanwhile, on the heels of my last two Medium articles on plant medicinals and participation in MIT's "Africa Takes on COVID" venture competition, a team has formed around investigating hundreds of candidate plant medicinals for inhibitory utility against the protease of the virus that causes COVID-19, SARS-CoV-2.

So while we've been proving out our toolchain and getting a UK patent filed on these screening endeavours and the candidate plant-derived compounds they are yielding, we would be remiss not to assess another plant-derived medicinal like COVID-Organics with the same level of rigor as we do our own drug candidates.

Sub-Saharan African Epidemiology

First, some epidemiology — the proof needs to be in the pudding, no matter what. So we collected the data from Worldometer on sub-Saharan Africa countries' case statistics. We isolated the number of fatalities and the number of cases for each country, took the ratio, and arrived at this chart:



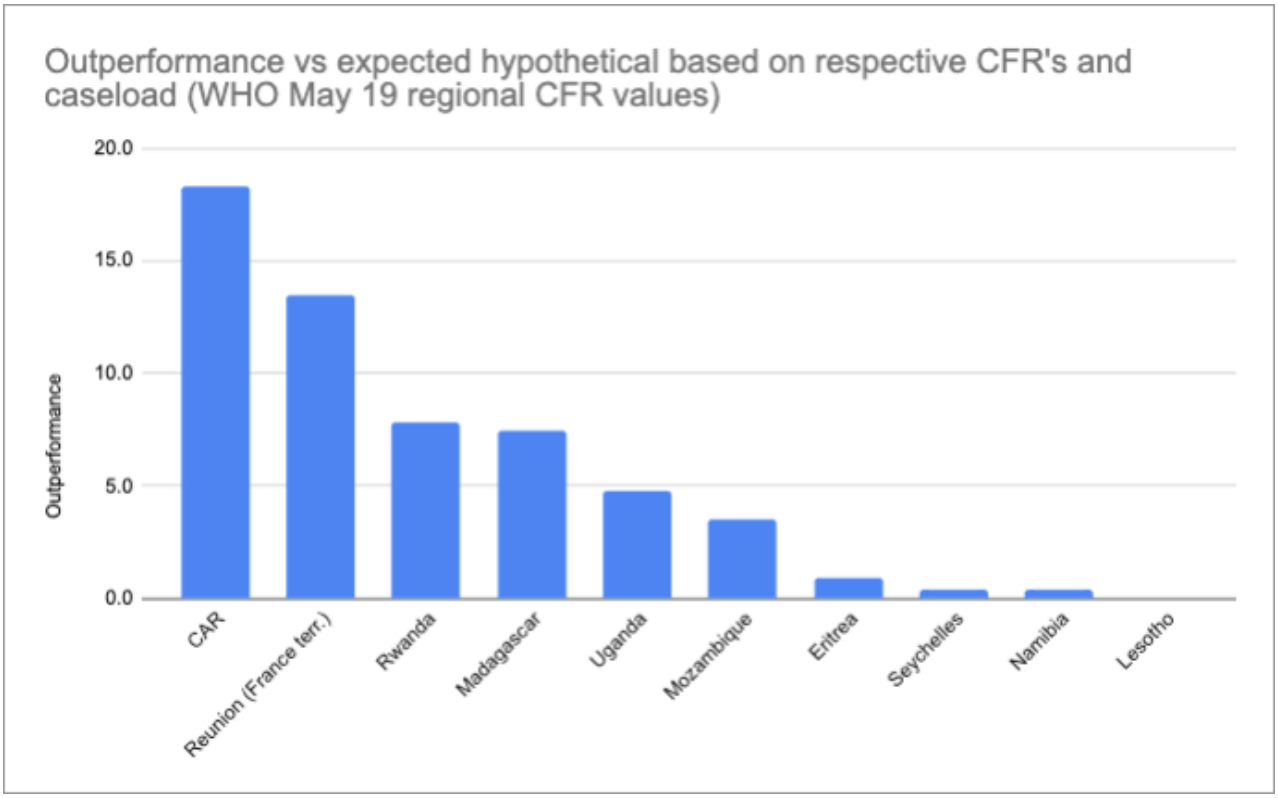
Recorded fatalities per recorded case, sub-Saharan African countries as of 25 May

Why did we select the fatalities-per-case ratio? First, because it features a degree of tolerance to each countries' unique patient data collection and recording regime: A country might overcount or undercount cases based on its health system's unique challenges and incentives. But they are less likely to misclassify or hide a coronvirus-

induced fatality from among those already registered into 'the system'. So we see it as a fairer comparison metric of how each country's health system treats cases that are accepted into it.

Second, it's important to keep in mind that COVID-Organics has been incorporated as *the* go-to therapeutic for Madagascar's caseload. Then benefiting from the fatalities-per-case dataset, we would expect to see signal in this particular ratio vis-a-vis other countries that don't incorporate this therapeutic (assuming all other factors held equal of course).

Therefore, note that Madagascar sits in the top-10 of the 47 countries in the data collection. Not a bad showing to be sure, but not strong enough to be compelling. So let's take a deeper look into those countries that seem to be doing better than Madagascar.



Countries with fatalities per case metrics ahead of Madagascar

Pictured is a measure we call 'Outperformance'. It is simply the number of recorded cases in each country, multiplied by the WHO's recorded average Case Fatality Rate (CFR) for that region of Africa, minus the actual number of fatalities of that country. In short, it tells us how many fatalities that country had vs. how many they *should* have had if they had the same average CFR as the sub-region of sub-Saharan Africa in which they are located.

(RegionalCFR) x (# of cases) - (# of fatalities) = Outperformance

Now a different story starts to emerge. A lot of the countries that show no fatalities also have very few cases(less than 50 even from Eritrea on down). Based on a typical CFR, we hardly expect any deaths at such low numbers, with departures from the regional CFR being overcome by noise. Madagascar now sits in 4th place, neck-and-neck with Rwanda in 3rd, with France's territorial island of Reunion in 2nd, with Central African Republic in 1st place with 600 cases and just 1 fatality.

We have now zoomed in to a level that it's worth looking at special cases of each country. Rwanda is famous for the strength of its public-sector. Its highways are the best kept in East Africa and its capital city is a showpiece capital of government project spending. To its credit, it should come as no surprise they have a strong public health sector. Reunion is a French territory, and has the benefit of France's similarly famous public health system's resources. So while Reunion's performance deserves kudos in its own right, it's not surprising given the France resource influence.

Central African Republic (CAR) however is known for none of these public sector tailwinds. It's a poor, landlocked country with a GDP per capita of $400 *per year* — one of the lowest in the world. And it's been wracked with civil war since 2012. Not the kind of place you'd expect to be doing well in a global pandemic of a highly infectious, life-threatening illness as COVID-19. And indeed, CAR has over 600 cases to its name.

But there is one thing notable about Central African Republic. They are the only one of the countries on the shortlist who took volume delivery of COVID-Organics from Madagascar by May 8th. It's not inappropriate to assume those volumes reach the hands (and mouths) of that country's caseload of COVID patients as intended. CAR has had 17 days for artemisia administerings to take place and start showing up in the epidemiological record as a lack of the fatalities that everyone would otherwise expect to be there.

So we realize now we're not just looking at a chart of Madagascar vs. all the other countries here, we are looking at a chart of Madagascar *& Central African Republic* vs. the other countries. Now the chart is delivering an unanticipated message indeed. CAR is top of the list, bar none.

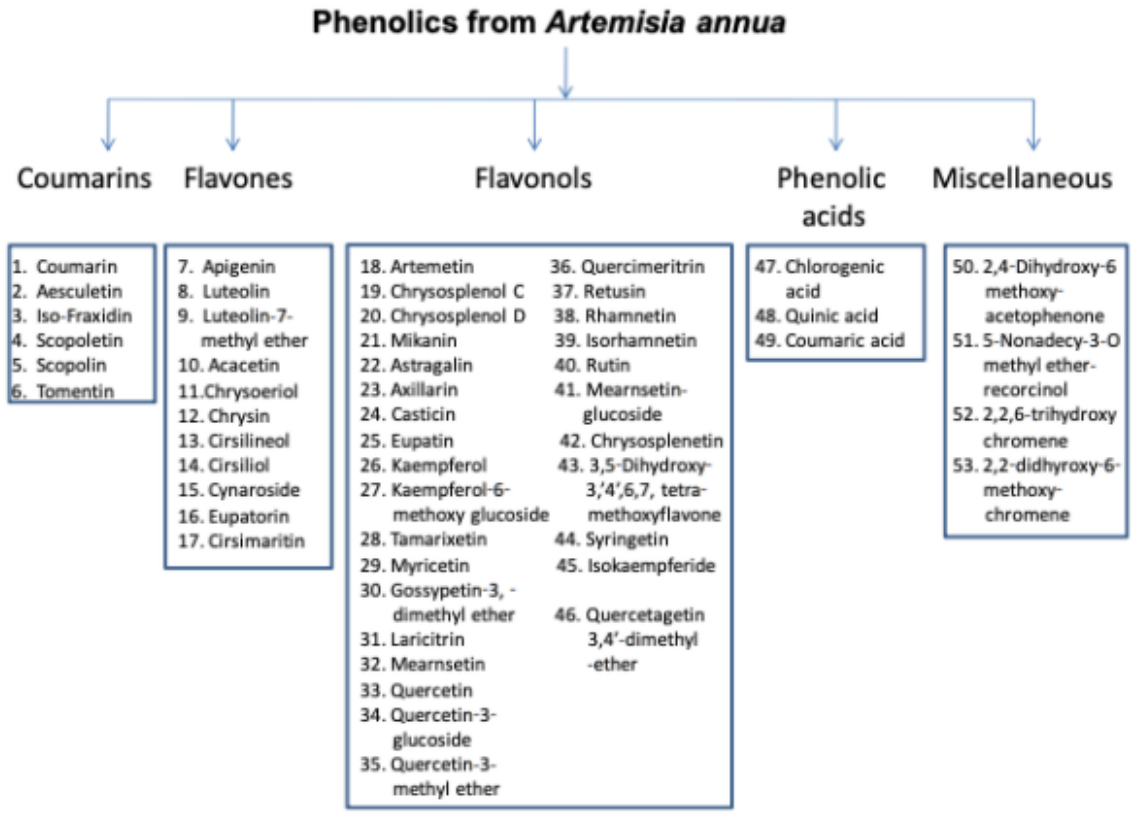
So what does this all mean? We say it's time to pay extremely close epidemiological and clinical attention to COVID-Organics' efficacy on COVID-19.

What could be going on in this novel potion? Let's peer into the cauldron and find out:

Artemisia Annua

COVID-Organics is quite literally an herbal 'tea' — a hot water extract of the plant *Artemisia Annua*. A. annua is a famous plant in phytochemical circles because it contains artemisinin. Artemisinin is the groundbreaking compound renowned in pharma circles for its inhibitory efficacy on *Plasmodium falciparum*, the parasite that causes malaria. Discovered by Chinese western-cum-traditional pharmacologist Tu Youyou, artemisinin earned its nature-derived place in the mainstream pharmacopoeia as *the* first-line, WHO-anointed, medication for malaria infection.

But the extract has a lot of other compounds too. Let's take a look:



Compounds in A. Annua (from Ferreira et al., "Flavonoids from Artemisia Annua . . . ")

In all, 53 identified compounds, plus the famous artemisinin. So we looked at which of these could feature the same kind of inhibitory efficacy on SARS-CoV-2's protease as we ordinarily screen for in the in silico assay we use in our main-line drug screening. How does inhibition work? Here is a really helpful animation to illustrate. Like most viruses, SARS-CoV-2 has an enzyme called its 'main protease'. After hijacking its way into the human cell, the protease acts like the virus' "laser-cutter", as you might find in a mechanical prototyping lab. It cuts ('cleaves') peptide chains to shape in order for new copies of the virus to get made. These newly-minted virions will go on to infect more cells.

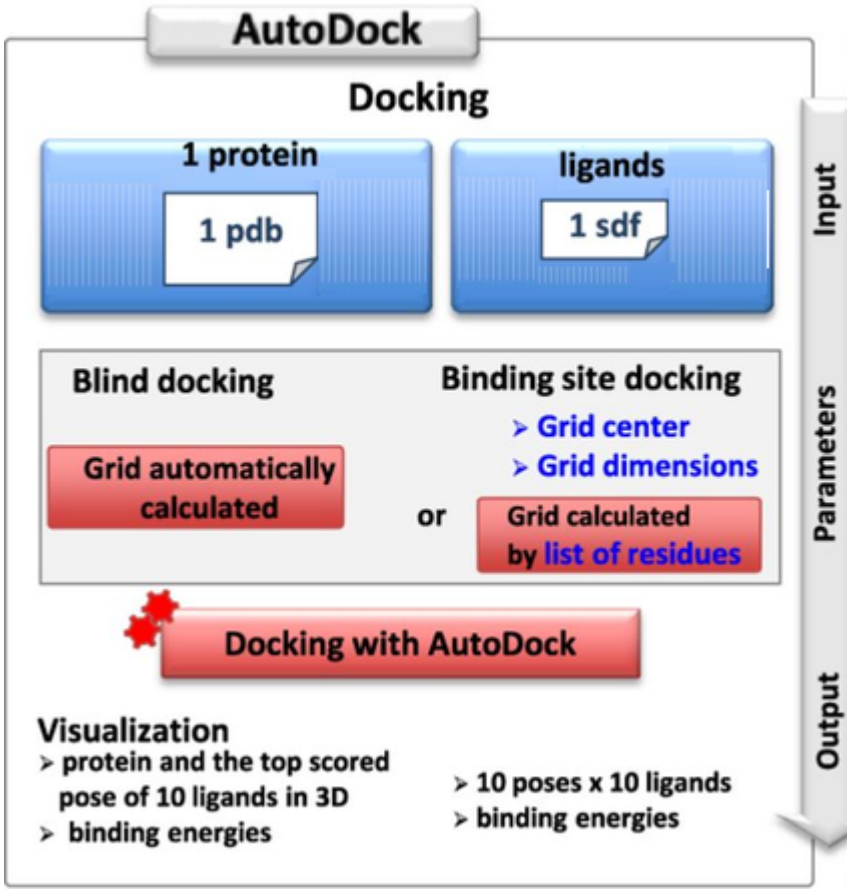
How does candidate drug screening work?

All about drug docking

(Note: you can skip this section if you want to go straight to the results)

Our screen rests atop a drug docking tool called Autodock Vina. Autodock and Autodock Vina are open-source tools put out by Scripps Research Institute. Rooted in machine-learning, it uses gradient-descent to find conforming, polarity-attractive "poses" of candidate compounds to conform onto any of a protein's receptor sites that are under study. The Autodock family of docking software has been extensively applied for screening by life sciences researchers across the body of published biochemistry literature

for over a decade. Vina is a cost-function and scoring refinement of Autodock which improves the hit-miss accuracy and significantly improves computation speed compared with the original Autodock. Here is the general workflow of autodocking tools:



Flow chart for Autodock (Source: MTIOpenScreen, (CC BY 4.0) modified for generalized Autodock)

Essentially, you download the (larger) pdb files from the Protein Database. The compounds (called 'ligands') you download in 3D SDF format typically from PubChem. For Autodock Vina to work with them, you need to convert the compounds and ligands into .pdbqt format. Converting is accomplished through the MGLTools download page from the Scripps website (if you're a Mac user, you'll want to make sure to run that on an un-upgraded 32-bit Mac OS version).
1.Use MGLTools' prepare_receptor4.py to convert the protein pdb file to .pdbqt
2. We prefer using the combination of obabel (obabel is its own installation effort, unfortunately) and prepare_ligand4.py to convert the .sdf file into .pdb and subsequently to .pdbqt respectively.

Now you're almost ready to do a docking run! The remaining item to sort out is the search box. Vina does its gradient-descent search by iterating across the many-dimensional space of 3D location, orientation, and ligand rotational bond conformations. It searches within a three-dimensional volume search box that the user specifies in units of Angstroms. The user also specifies a center for the search. Together, these parameters allow the selection of either a 'blind' search, which is a volume encompassing the entire protein (we use 150 Angstroms^3), or for targeted docking to active sites, a narrower search volume (we use 30 Angstroms^3). Using a 3D visualization program like PyMol, it is

easy to determine the coordinates for blind docking (typically the estimated geometric center of the protein). For a targeted site, just open up the human-readable pdbqt file itself and find the target site's amino acids of interest. The active site's amino acids are ascertained from published literature for the protein of interest. For the SARS-CoV main protease, we usually look at the active site defined by A-histidine-41 & A-cysteine-145. ('A' refers to one of the two chains of amino acid that combine to form the totality of the protein).

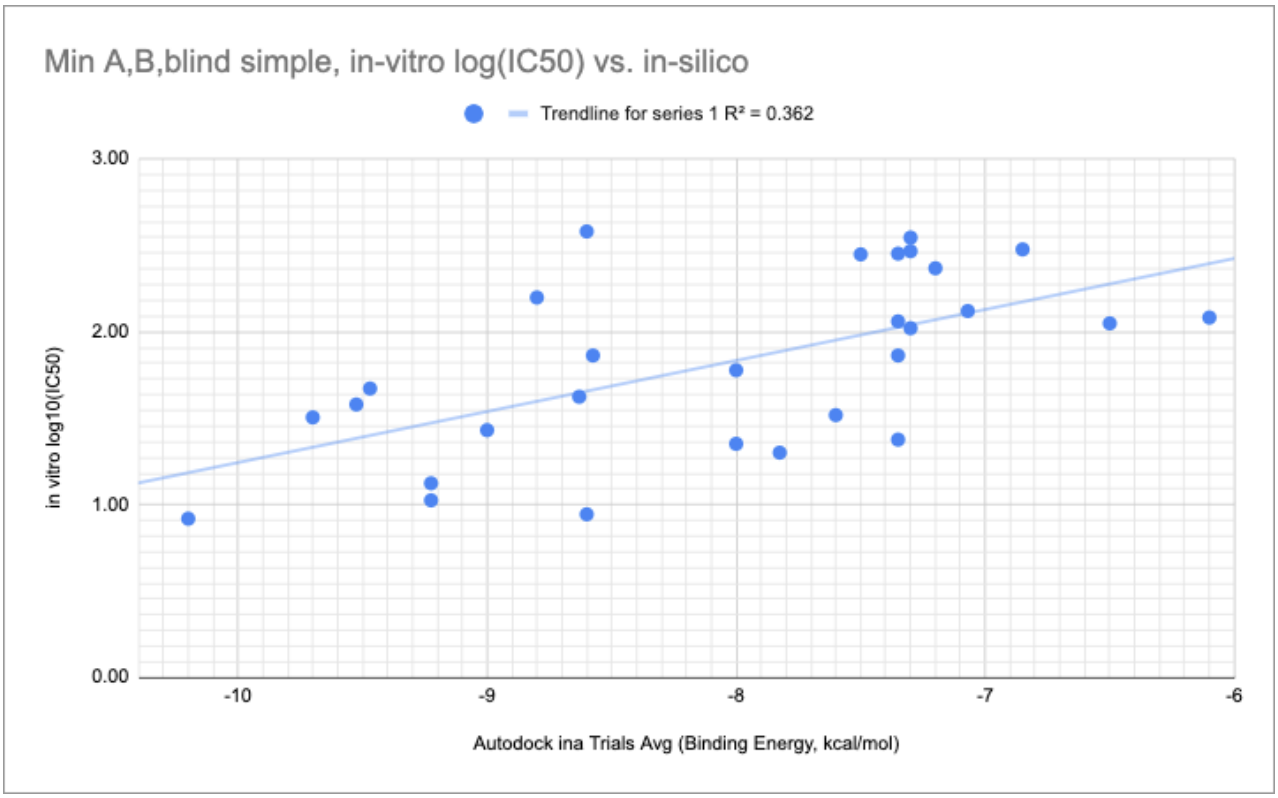
Under gradient-descent with a cost function defined by the ligand-protein complex' energy level, Vina iterates the ligand's position, rotation, and structural pose vis-a-vis the protein's surface. Critically, it takes into account the attraction of opposite charges on the ligand molecule and the protein receptor surface. Vina generates results along a -4 to -12 scale (formally: in units of kilocalories per mole). In other words, docking software like Autodock Vina thinks in terms of minimizing the "energy level" of a set of chemical bonds determined by a compound's pose against a protein receptor site. Our application of Vina essentially maintains this scale but simply applies some combination logic and averaging over an appropriate number of trials. That way our results (where the more negative number is more likely to be more inhibitory) may better reflect a compound's propensity to defeat the virus' ability to replicate itself, as illustrated next.

How do we effectively apply this tool to the coronavirus' protease?

When the earlier SARS virus first struck in 2003, there was sufficient time and funding in follow-on years for researchers to do in vitro inhibitory studies of the virus' main protease. Researchers were studying to determine which compounds would inhibit this proteases' growth in a petri dish (referred to as *in vitro*). Why target a protease? Because it's a one-shot-kill for viruses. Proteases' efficacy as a target was well-established especially thanks to the HIV protease inhibitors that came out in the 1990s.

Critically, the SARS 2003 coronavirus is very, very similar to the current 2019 coronavirus. The two virus' main proteases specifically are 95% identical in amino acid sequence, and differ primarily in their three-dimensional folding structure.

Treating in vitro results as our "source of truth", we compared the results of our Autodock studies on the 2003 SARS main protease with researchers' in vitro results over the 16 years since that epidemic. However it is impossible to know which target site in real life the in vitro researchers managed to bind their inhibitory compounds to. And just as a real compound can bind to different sites with the potential to inhibit, so too can we target autodocking at different sites. Therefore we must explore different combinations of the autodock targeting results. By applying results from blind, A, and B docking studies as #1 in isolation and #2 in different combinations with each other, we arrive at the following resulting combination that features the highest correlation with in vitro lab inhibitory results:



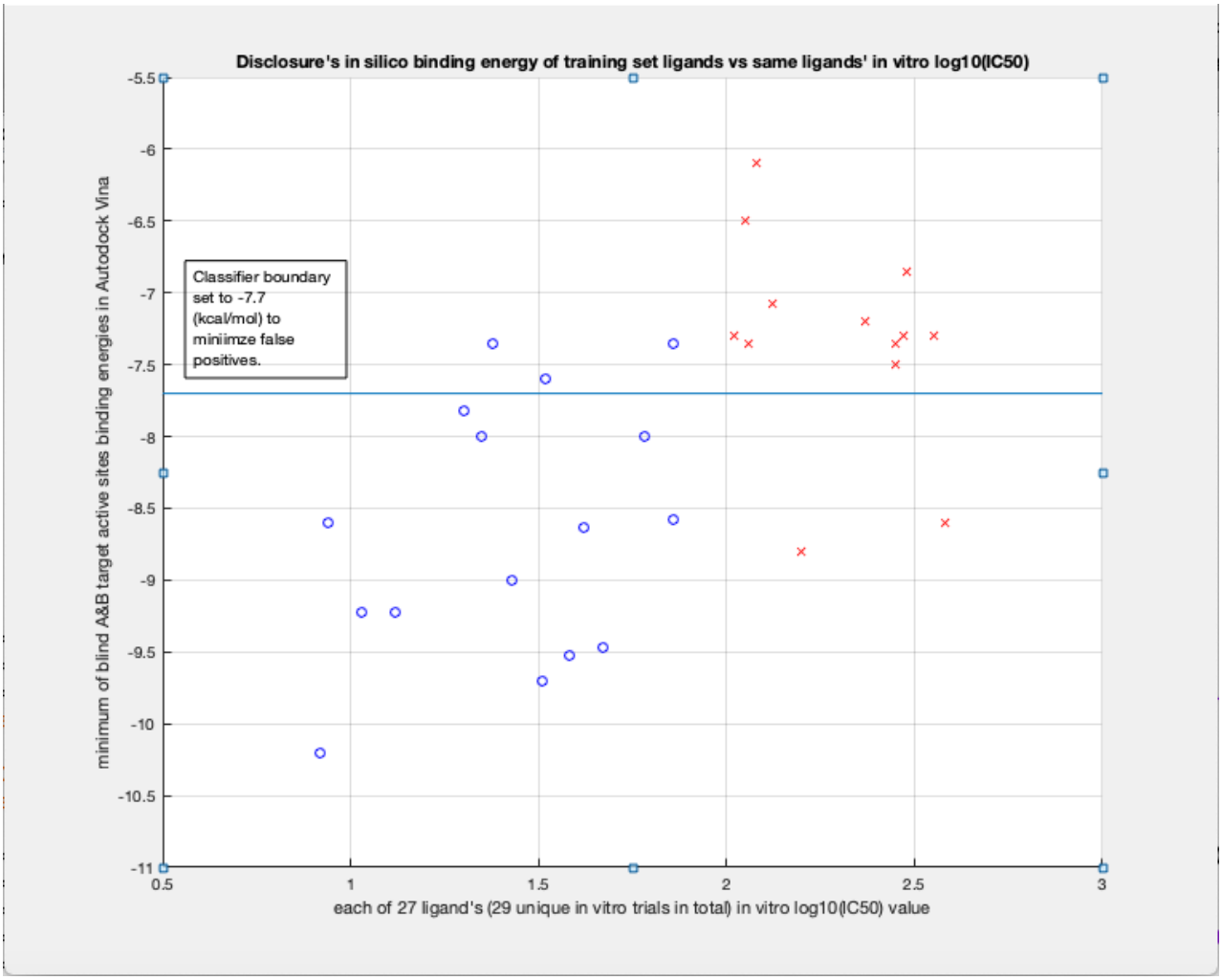
Correlation of processed Autodock results with in vitro study

We achieve an R^2 of 0.362, a value that is not atypical for biochemical virology investigation.

Time to form the classifier

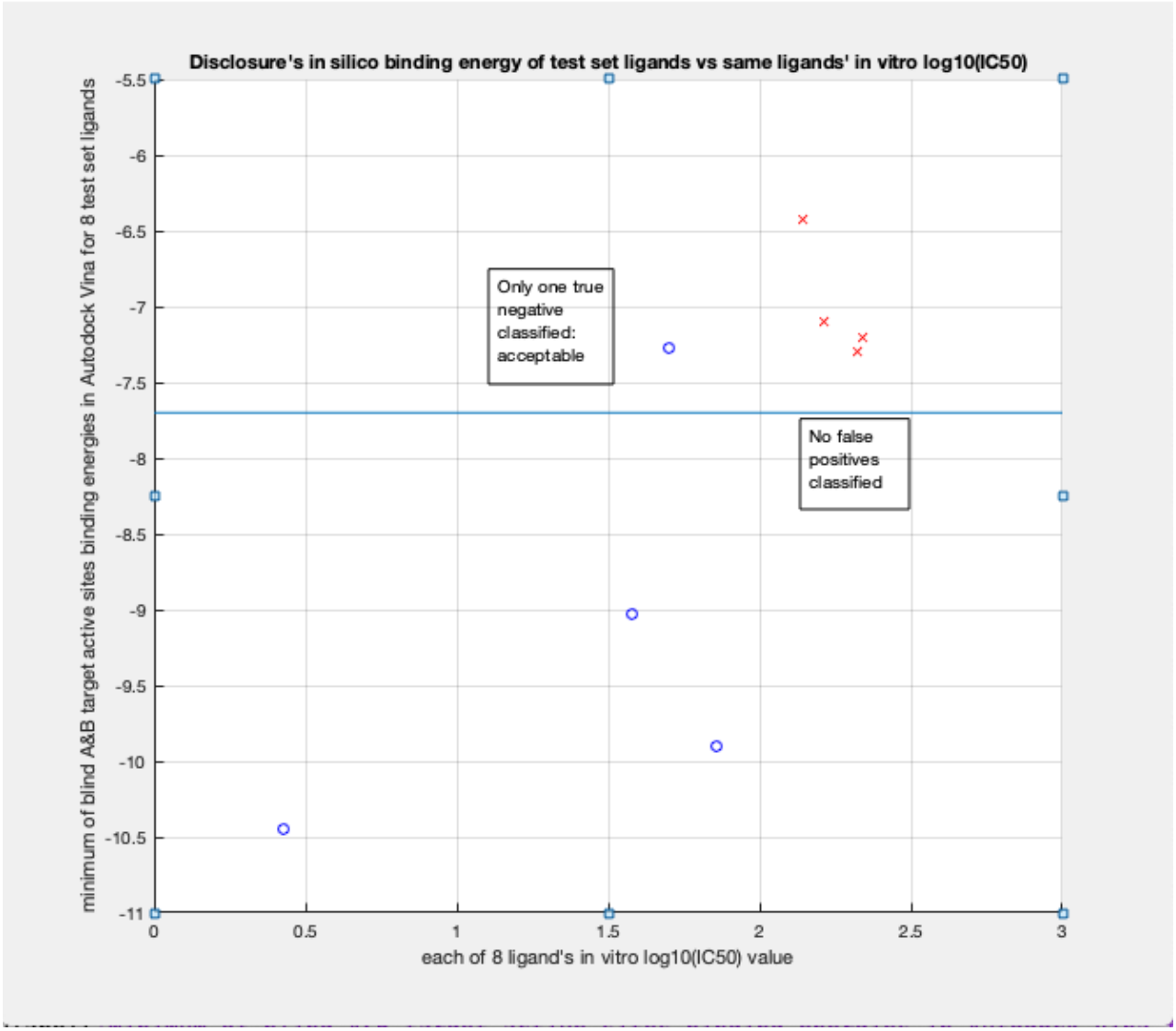
Correlation in hand, we can now develop a classifier for separating out the inhibitory 'hit' ligands from the worthless 'miss' ligands.

In vitro researchers treat any compound that inhibits with a concentration of 100 uM (micromolar) or less to be the low-side cutoff for an in vitro 'hit'. That corresponds to a log10 value of 2.0 . Therefore we trained our classifier by comparing Autodock Vina results of 27 compounds as simulated against multiple active sites of the similar, original SARS (2003) coronavirus main protease with those actual, in vitro lab inhibitory results. Our best-fit-to-reality formulation of Vina results across several active sites yields the following hit-miss chart:



The 27-compound training set for our ML classifier of Autodock Vina results applied to SARS-2003 compounds

Benefiting from this training set, we assayed our formulation of Vina results against an additional 8-compound test set. The classifier demonstrated that anything with a **-7.7 kcal/mol score** or above should be classified as a 'hit'.



Results of our ML classifier's test set of 8 additional compounds, also based on SARS-2003.

Based on the test set, our classifier demonstrates a classification accuracy of **87.5%**.

Artemisia's in silico assay results

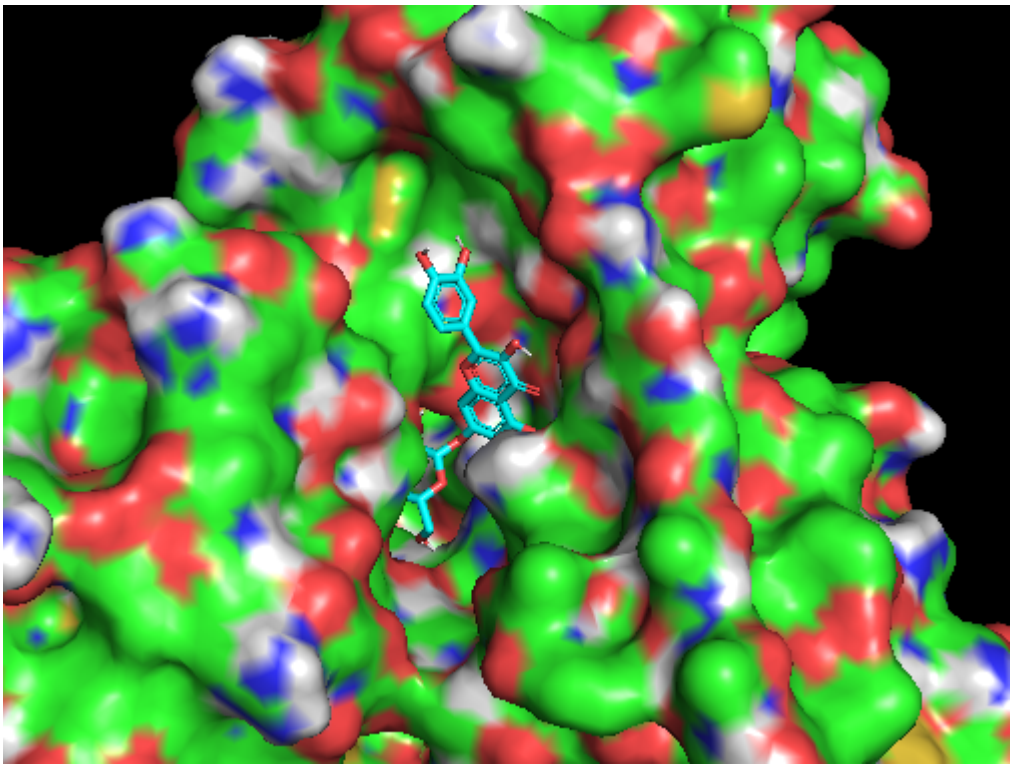
On the -4 to -12 scale and with the -7.7 kcal/mol threshold, here are the compounds of *A. Annua* yielded from our assay as strong performers:

Compound (of qty46 studied)	min(blind, A,B) binding energy kcal/mol:
5_7_4_Trihydroxy_3_methoxyflavone	-7.75
Astragalin	-7.95
Laricitrin	-7.90
Luteolin	-7.98
Mearnsetin	-7.90
Quercetagetin_3_4_dimethyl_ether	-8.13
Quercetin-7-glucoside	-9.15
Tamarixetin	-7.98
isoquercitrin	-8.40
rutin	-8.875

Hits of -7.7 kcal/mol and above *Source: EMSKE Phytochem*

First, it's worth noting that artemisinin doesn't even rate for purposes of COVID-19 virus' protease. Instead, the majority of these "placers" are common flavonoids and their glycosides. What's more, while all of these 10 compounds show baseline levels of potency, quercetin-7-glucoside (aka "quercimeritrin") actually presents as rather strong (!). In our screening, we consider -9 and above a 'strong hit'. As you can see from the SARS in vitro correlation data, you can generate models that show how a stronger hit translates to greater potency (defined as dosage required relative to body weight).

Here is the top pose determined by Autodock Vina of what quercetin-7-glucoside looks like in it's lowest energy (read: highest potency) binding state to the protease:



Pictured is quercimeritrin's best pose binding into the 'tunnel' structure of the protease. (Source EMSKE Phytochem, application: PyMol)

In short, our venture EMSKE Phytochem comes out very much in support of COVID-Organics, and we're happy to help put a loudspeaker on that. But you don't have to go all the way to Madagascar to get the benefits of this extract. Most of what we see as demonstrating efficacy in the extract is in a lot of common everyday fruits & vegetables as well. Most prevalently among them, red onions. (Of course, it would take consuming an entire onion or two per day to achieve the expected inhibitory dose required, so a more isolated / extracted approach might be preferred).

We also feel it's important to recognize the role that traditional medicine has played here. Madagascar's IMRA health research institute was the originator of trying Artemisia Annua extract on COVID patients. It was a much more efficient route to identifying drug 'hit' candidates than the formal Phase I-III clinical trials route espoused by the FDA in the US. While it's not without side effects, such as headaches, this plant has been known to (Chinese) traditional medicine to be safe for administering by way of ingestion. When the dust settles, the world may end up owing a great debt to Madagascar, IMRA, and the paradigm of traditional medicine development models as conducted through an investigatory (but systematic and safe) scientific lens for its treatments.

There's more in the research pipeline in terms of compounds which we are identifying based on similar efficacy studies as this. We'll be unveiling these results in the weeks ahead.

Update 3 July: The Max Planck Institute in Germany has been carrying out in vitro tests of different formulations of the artemisia extract as well as refined artemisinin. Their results are that the artemisia extract is indeed showing efficacy on SARS-CoV-2

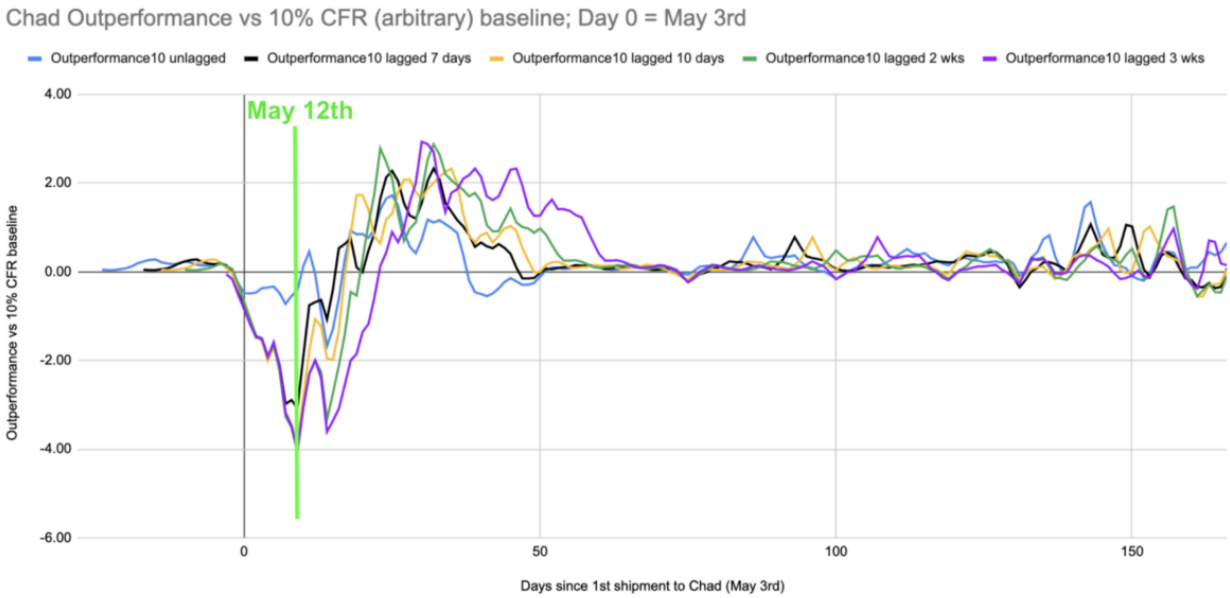
inhibition, and the refined artemisinin has no effect. In other words, very much in agreement so far with the above.

Update 26 Dec 2020: Improving the analysis, availing of Chad media reports as supported by Covid-Organics Africa-wide shipment data (obtained offline), the following findings present themselves:

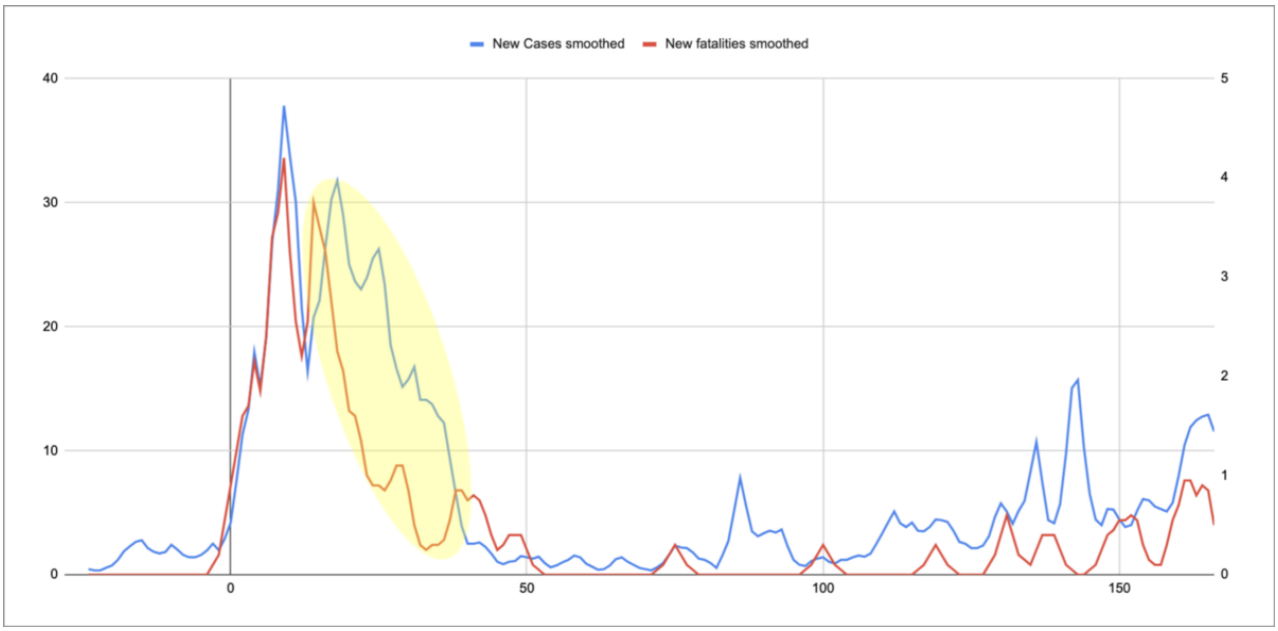
- Covid-Organics turns out to contain extracts primarily of two plants — Artemisia Annua and Ravintsara.
- From the audience gained through the above post, I was able to obtain confidential shipment data of CVO to African countries during the pandemic period. (Unfortunately I can't share that data for confidentiality reasons). Suffice to say that the above list of countries who received CVO was incomplete (and perhaps incorrect regarding Tanzania specifically), and of course lacked the important nuance of how many doses were shipped to each country.
- It was found through the shipment data that the nation of Chad was the only country to maintain sustained and significant shipment volumes of Covid-Organics relative to caseload, and was known through local media reports to have administered it with intention to its covid patients.
- That assertion turned out to be backed up by relevant media reports from Chad (in French). In chronological order they are 12-May: 1, 4-June: B, 11-June: C, and 16-June D. (I've made inquiries through a WHO program manager contact to Chad's Ministry of Health; as anywhere, covid-related data is heavily politicized, so I have only learned that they haven't produced any formal data for public consumption other than the media reports).
- From article C, we know that 34 'high-risk' covid patients were treated exclusively with Covid-Organics by 4-June. All were brought back to health again. Given Chad's high initial fatality rate at the start of their outbreak, this is notable.
- Article C implies that additional patients were treated with covid-organics alongside other medications. However Article D goes on to mention that covid-organics should not be used with hydroxychloroquine. (This isn't surprising as similar flavonoid glycosides are known to competitively inhibit two liver enzymes known for processing many common chronic medications — and so the possibility of hydroxychloroquine overdoses in patients so treated exists — such unfortunate events would explain their media's non-reporting of that patient segment)
- One of their government ministers is quoted saying (Google translated), "We are pleased today to reiterate what our Minister of Health has already said. Covid-Organics has been a positive experience in Chad"

But is that conclusion supported by data?

- Given that there is typically a lagtime between a reported case and any fatality associated with it, the Outperformance metric described earlier is brought out in terms of zero lag, 1-week lag, 10-day lag, 2-wk lag, and 3-wk lag.
- The outperformance data for Chad accounting for case-fatality lag, (set relative to an arbitrary 10% baseline) is:



- What we see is the case fatality rate initially doing poorly; Then from May 12th we start to see an improvement in Chad's case fatality ratio performance (across all lag metrics).
- This can be compared to the raw Case vs. Fatality charts for Chad from Worldometer below — note the gap of the case trace (blue) over the fatality trace (red) in days 20–40. So the above chart is essentially highlighting the gap seen in yellow below:



Day 0 = May 3rd. Blue: Chad covid cases (left side scale), Red: Chad covid fatalities (right side scale). Yellow highlighted is anticipateded fatalities that (thankfully) never materialized.

- Indeed, 12-May (as per Article A) is when the first Covid-Organics (aka Tambavy CVO) shipment is confirmed to have arrived in Chad. We don't know precisely how soon after 12-May they start administering to patients, but we presume soon afterward as by 4-June they report having already administered to patients and seeing encouraging results.

- The right-side and left-side scales have been proportionally aligned to each other for easy comparison. Reviewing the yellow highlight, we offer that it's *possible* that Chad saved up to 20–25 lives with their CVO intervention.
- Might any lives have been *lost* due to the intervention? It's actually not outside the realm of possibility. Article D seems to acknowledge the possibility when they warn about co-administration of CVO with chloroquine. This drug interaction makes sense, as flavonoids competitively inhibit liver enzymes that are relied upon to process xenobiotics such as chloroquine. (The resulting overdose causes heart arrhythmia which could prove difficult to manage in a low capability health care setting).
- Net-net of any such complications, it looks like Chad and its patient caseload came out the better for the intervention. In the face of a pandemic outbreak with essentially nil treatment options at the time (remember this is all before dexamethasone's efficacy was identified), it's hard not to appreciate the unique actions that their health ministry took during their May outbreak.

Formally what we're showing here is just a correlation. There can be other causes for Chad's improvement in performance from the primary outbreak — perhaps improved patient detection and handling protocols as healthcare workers became experienced with the novel pandemic, or additional ventilators made available, etc. But coupled with the reports from their government we believe we are seeing a little more causality than just strict correlation.

To stay apprised of our latest in silico studies results and clinical trials readiness, stay tuned on Twitter @EMSKEPhyto