Fundamental Research xxx (xxxx) xxx

Contents lists available at ScienceDirect

Fundamental Research

journal homepage: http://www.keaipublishing.com/en/journals/fundamental-research/

Article

Tailoring combinational therapy with Monte Carlo method-based regression modeling

Boqian Wang[a,1], Shuofeng Yuan[b,c,1], Chris Chun-Yiu Chan[b,c], Jessica Oi-Ling Tsang[b,c], Yiwu He[d], Kwok-Yung Yuen[b,c,e], Xianting Ding[a,*], Jasper Fuk-Woo Chan[b,c,e,*]

[a] State Key Laboratory of Oncogenes and Related Genes, Institute for Personalized Medicine, School of Biomedical Engineering, Shanghai Jiao Tong University, Shanghai 200030, China
[b] State Key Laboratory of Emerging Infectious Diseases, Department of Microbiology, School of Clinical Medicine, Li Ka Shing Faculty of Medicine, The University of Hong Kong, Pokfulam, Hong Kong 999077, China
[c] Department of Infectious Disease and Microbiology, The University of Hong Kong-Shenzhen Hospital, Shenzhen 518048, China
[d] Technology Transfer Office, The University of Hong Kong, Pokfulam, Hong Kong 999077, China
[e] Hainan Medical University-The University of Hong Kong Joint Laboratory of Tropical Infectious Diseases, The University of Hong Kong, Pokfulam, Hong Kong 999077, China

ARTICLE INFO

Article history:
Received 13 August 2022
Received in revised form 5 December 2022
Accepted 17 March 2023
Available online xxx

Keywords:
Combinational therapy
Regression modeling
Dose optimization
Monte Carlo method
SARS-CoV-2

ABSTRACT

Combinatorial drug therapies are generally more effective than monotherapies in treating viral infections. However, it is critical for dose optimization to maximize the efficacy and minimize side effects. Although various strategies have been devised to accelerate the optimization process, their efficiencies were limited by the high noises and suboptimal reproducibility of biological assays. With conventional methods, variances among the replications are used to evaluate the errors of the readouts alone rather than actively participating in the optimization. Herein, we present the Regression Modeling Enabled by Monte Carlo Method (ReMEMC) algorithm for rapid identification of effective combinational therapies. ReMEMC transforms the sample variations into probability distributions of the regression coefficients and predictions. *In silico* simulations revealed that ReMEMC outperformed conventional regression methods in benchmark problems, and demonstrated its superior robustness against experimental noises. Using COVID-19 as a model disease, ReMEMC successfully identified an optimal 3-drug combination among 10 anti-SARS-CoV-2 drug compounds within two rounds of experiments. The optimal combination showed 2-log and 3-log higher load reduction than non-optimized combinations and monotherapy, respectively. Further workflow refinement allowed identification of personalized drug combinational therapies within 5 days. The strategy may serve as an efficient and universal tool for dose combination optimization.

1. Introduction

Combinatorial drug therapy, also known as cocktail therapy, has been widely used in the treatment of different diseases including cancers and infections. To maximize the efficacy and minimize potential toxicity of combinatorial drug therapies, precise optimization of the dose regimen is essential [1–4]. Various methodologies have been devised to achieve rapid and reliable dose optimization of drug combinations and to expedite new drug development and repositioning [5–7]. Among these methods, artificial intelligence (AI) featuring model regression has been increasingly used due to its high adaptability and requirement of relatively little data from experiments [8,9]. Polynomial regression-based AI methods, such as the Phenotypic Response Surface (PRS) platform, have been applied to optimize combinatorial drug therapies for cancers [10], infections [11,12], and transplant rejections [13]. Nevertheless, the efficiency of these methods is limited by the high noises and poor reproducibility of the biological assays involved in the optimization process, either *in vitro* or *in vivo* [14]. Conventional methods are prone to model bias caused by such uncertainties, which may limit further translational applications. To cancel the influence of noises on the modeling accuracy of conventional methods, it requires multiples times more experimental efforts. In the previous study [15], we introduced a projection distance-based scoring method named STRICT for evaluating drug efficacies and drug interactions, which showed excellent

* Corresponding authors.
E-mail addresses: dingxianting@sjtu.edu.cn (X. Ding), jfwchan@hku.hk (J.F.-W. Chan).
[1] These authors contributed equally to this work.

https://doi.org/10.1016/j.fmre.2023.03.008

Please cite this article as: B. Wang, S. Yuan, C.C.-Y. Chan et al., Tailoring combinational therapy with Monte Carlo method-based regression modeling, Fundamental Research, https://doi.org/10.1016/j.fmre.2023.03.008

robustness against experimental noises. However, the scoring algorithm could not be straightforwardly used to predict the optimal combinations.

To overcome these limitations, we introduced the Monte Carlo (MC) method, a simulation approach via random sampling [16,17], into the regression modeling process (ReMEMC). Conventional modeling methods commonly adopt the mean values of repeated measurements as the only information for modeling, while the errors, which represent the uncertainties of the information, are generally not utilized. To make maximal use of all available information generated in the experiments, ReMEMC uses both the mean values and sample standard deviations (SSDs) for modeling. The method features multiple parallel modeling calculations. Briefly, in each calculation, a normally distributed estimation of the true experimental outcome is generated to conduct dose-response modeling, replacing the commonly used mean value. With numerous simulations performed, the MC method provides the probability distribution, namely the confidence interval (CI), of every coefficient in the model and every prediction of efficacy. In this way, rather than generating a definite fixed model which real biological studies usually lack, ReMEMC maximally considers variances acquired from experiments and generates a dynamic model involving such uncertainties. *In silico* simulations suggested that the more informative and efficient ReMEMC outperforms conventional methods in optimizing benchmark problems.

To demonstrate the usefulness of ReMEMC in optimizing combinatorial drug therapy for major health threats, we exploited this novel modeling method to optimize combinational therapy for coronavirus disease 2019 (COVID-19) caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). To mimic the scenario of applying ReMEMC to quickly optimize drug compounds that were available in the early phase of an emerging disease outbreak, we purposely selected drug compounds that were identified in a previous study published during the first year of the pandemic in which we evaluated 22 potential antiviral drug compounds against SARS-CoV-2, including 17 small molecule drug compounds and 5 recombinant IFNs [18]. In addition to remdesivir, lopinavir, and chloroquine, we identified the *in vitro* anti-SARS-CoV-2 activity of types I and II recombinant IFNs, 25-hydroxycholesterol, and the sterol regulatory element-binding protein (SREBP) modulator AM580 [19]. Based on these findings, in this study, we selected 5 small molecule drug compounds and 5 recombinant IFNs for further optimization of combinational therapy using our novel ReMEMC algorithm.

2. Material and methods/experiment

2.1. Virus, cell lines, and drug compounds

SARS-CoV-2 strain HKU-001a (GenBank accession number: MT230904) was prepared as we described previously [18,20]. VeroE6 (ATCC® CRL-1586™) cells were cultured in Dulbecco's modified eagle medium (DMEM, Gibco, CA, USA) with 10% fetal bovine serum, 50 U/mL penicillin, and 50 μg/mL streptomycin. All the experiments involving live SARS-CoV-2 were conducted in the Biosafety Level 3 facility at the Department of Microbiology, the University of Hong Kong, following the approved standard operating procedures as previously described [21,22]. The recombinant IFNs were obtained from the following sources: Pegasys (Roche, Basel, Switzerland), Avonex (UCB, Brussels, Belgium), Rebif (EMD Serono, Inc. Rockland, MA, UA), Betaferon (Bayer Schering Pharma, Berlin, Germany), and Immukin (Boehringer Ingelheim, Ingelheim am Rhein, Germany). All other drug compounds were purchased from MedChemExpress (Monmouth Junction, NJ, USA).

2.2. 5SARS-CoV-2 viral load reduction assay

Viral load reduction assay was performed as described previously with modifications [18,23]. Briefly, VeroE6 cells were seeded at the density of 2×10^4 cells/well in 96-well plates 24 h before the experiment. On the following day, SARS-CoV-2 was used to infect VeroE6 cells at the MOI 0.01, then incubated at 37 ˚C / 5% CO_2 for one hour. The infected cells were washed with PBS once and replaced with DMEM medium containing compound cocktails, and then incubated at 37 ˚C / 5% CO_2 for 48 h. Supernatant viral copies were determined by qRT-PCR. A total of 40μl of culture supernatant was lysed with 160 μl of AVL buffer. The viral RNA was subsequently extracted from the mixture with the QIAamp viral RNA mini kit (Qiagen, Hilden, Germany). qRT-PCR was performed using the QuantiNova Probe RT-PCR kit (Qiagen) with a LightCycler 480 Real-Time PCR System (Roche). The primers and probe sequences were targeting the RNA-dependent RNA polymerase/Helicase (RdRP/Hel) gene region of SARS-CoV-2: Forward primer: 5′ CGCATACAGTCTTRCAGGCT-3′; Reverse primer: 5′-GTGTGATGTTGAWATGACATGGTC-3′; specific probe: 5′-FAM TTAAGATGTGGTGCTTGCATACGTAGAC-IABkFQ-3′.

2.3. Algorithms and benchmarks

Data analysis and the algorithms of ReMEMC was realized with the built-in function "lasso" and self-written code in MATLAB©. In each independent modeling process, Lasso regression was conducted with the randomly sampled input and output to fit a quadratic polynomial model. The model was then used to predict the efficacy of every possible dose combination. For every ReMEMC model, the modeling process was repeated 10,000 times in parallel. For the data of the ten-drug anti-SARS-CoV-2 combinations, the whole modeling process took about 2.5 h with a hexa-core processer. By integrating the results in all the 10,000 independent calculation, the MC method provided the probability distribution of efficacy prediction of every possible combination. The design of experiment (DOE) for the first round of experiment was generated according to uniform design. Then the DOEs of subsequent rounds were generated by ReMEMC.

The benchmarks were also conducted with self-written codes in MATLAB©. The stepwise regression and Lasso regression were performed with the built-in function "stepwiselm" and "lasso" in MATLAB©. For each set of the method and benchmark condition, the optimization process was repeated 50 times with randomized starting DOEs to cancel the randomness of optimization results. The DOEs of the subsequent rounds were generated by the tested algorithm as part of the output. Each iteration contained 48 combinations. The curves in the figures were smoothed using the built-in "smoothdata" function in MATLAB©.

2.4. Code and data availability

The ReMEMC algorithm with MATLAB© implementation as well as the SARS-CoV-2 dataset and benchmark functions is available at Github repository https://github.com/JackW-SJTU/ReMEMC.

3. Results and discussion

3.1. Adopting Monte Carlo method in regression modeling

When modeling drug responses for dose optimization for combinational therapies, there are two critical factors that determine the efficiency and accuracy of the outcomes. The first factor is a proper model that fits the drug response curves well and the other is the cancelation of model bias to the greatest extent. Previous studies have revealed that polynomial models, especially quadratic model integrated with an iterative feedback scheme, could be an ideal solution for mechanism-free modeling in optimizing combinations of 5 to 15 drugs. However, a satisfying solution to the problem of model bias is still lacking. As the model bias is mainly caused by the high noise rate and poor reproducibility of biological experiment outcomes, a straightforward solution is to increase the number of experiments, either by repeating the same

experiments to cancel the errors in experiments, or by adding more data points to cancel the errors in modeling. Nevertheless, this is usually not a feasible option as it greatly increases the cost and labor-intensiveness of the study. Particularly, the improvement in model accuracy is not linear to the increase of the number of experiments.

Another strategy is to increase the utility of the available data. Biological experiments, *in vitro* or *in vivo*, are usually repeated multiple times to reduce randomness. However, conventional regression methods use only the mean values of the replicates for modeling. The information underlying the distribution of the replicates is not fully utilized. Considering an experiment with n replicates, in most cases, the outcome Y follows a Gaussian distribution $N(\hat{Y}, \sigma^2)$, where $\hat{Y}$ is the ideally unbiased outcome of the experiment. Therefore, the mean $\bar{Y}$ of the n replicates also follow a Gaussian distribution of

$$\overline{Y} \sim N\left(\hat{Y}, \frac{\sigma^2}{n}\right) \tag{1}$$

Reversely, when $\hat{Y}$ and σ^2 are unknown, they can be estimated by the sample mean $\bar{Y}$ and the sample variance s^2, as the probability distribution:

$$\hat{Y} \sim N\left(\bar{Y}, \frac{s^2}{n}\right) \tag{2}$$

The probability distribution represents the uncertainty of estimation of $\hat{Y}$, which is often neglected by conventional modeling methods. In contrast, ReMEMC utilizes such information of the uncertainties in the modeling process. Considering the fact that it is difficult to directly evaluate the effect of the uncertainties on the model and prediction, Monte Carlo simulation is adopted to solve the problem through repeated random sampling. In each simulation, a random estimation of $\hat{Y}$ is generated for every combination with the distribution of Eq. 2, composing the dependent variable in modeling. The drug doses as the independent variables are also added with normally distributed noises, representing the errors in performing the experiments. Then, the regression is performed with these randomly sampled values, resulting in a set of coefficients of the variables in the model. By repeating the random sampling and regression process numerous times, results from all the simulations jointly make up the probability distributions of the coefficients and predictions (Fig. 1a).

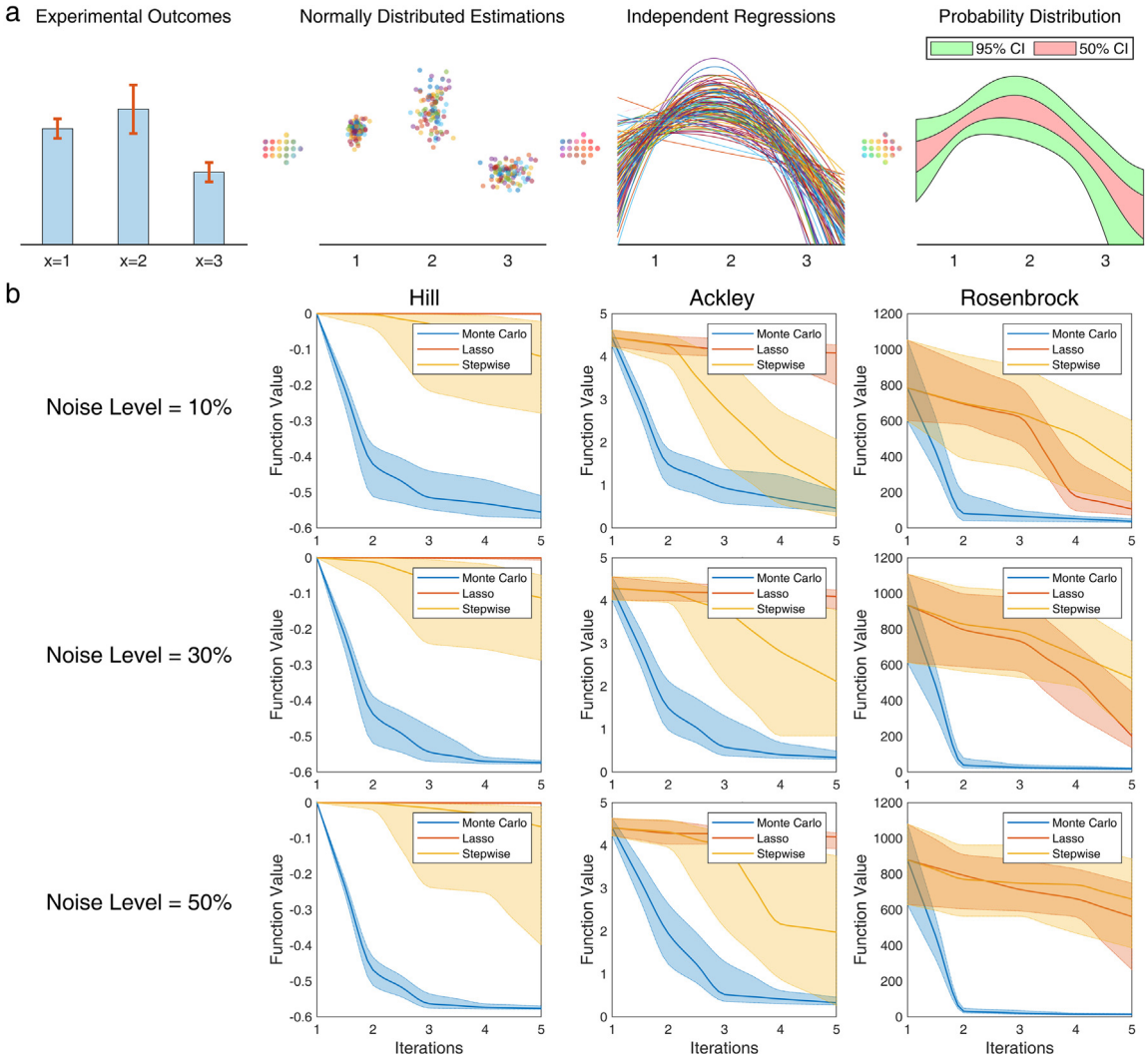


Fig. 1. The Monte Carlo (MC) regression and *in silico* validation. (a) The schematic of the MC regression. Normally distributed estimations of the true values are generated with the sample means and sample variances. Each set of estimations go through an independent regression process, with tens of thousands regressions conducted in parallel. Predictions from all the models jointly form the probability distribution at every input value. (b) The ReMEMC method (blue) was compared with Lasso regression (orange) and stepwise regression (yellow) in optimizing three benchmark functions at different levels of synthetic noise. Dimension of the benchmarks functions was set as 10. The solid lines and the translucent areas respectively indicate the median performance and the 50% CI of 50 parallel optimization processes with random initial DOE.

Of note, conventional methods are also able to calculate CIs for coefficients and predictions. However, the implications of CIs are different between conventional methods and ReMEMC. As conventional methods use definite values (the averages of replicates) for modeling, CIs represent only the magnitude of the model bias. For ReMEMC, CIs are determined not only by the biases in modeling but also the variance of every input, therefore making it more practical and reliable for prediction. Therefore, as a modeling algorithm, our method does not require any given function of dose-effect relationship. It generated a series of polynomial functions as "sampling" to depict the potential dose-effect relationships, based on a very limited available experimental data set.

Conventional methods simply select the best predicted combinations without further interrogating CIs. With the additional information of CIs, there are two distinct strategies to select combinations for the next iteration of experiments. The first strategy is called "exploitation" which is to choose the ascertained best predictions (minimal CI). The second strategy is called "exploration" which is to choose combinations with great potentials yet also with higher uncertainties (large CIs). Exploration may result in more effective combinations, but also could possibly end up with waste of efforts. Therefore, an upper confidence bound (UCB) acquisition function is adopted to balance the weights of the median and the CI for maximal optimization efficiency, denoted as (assuming the optimal efficacy refers to the maximal prediction)

$$UCB = M(x) - kCI_{50}(x) = M(x) - k\left|P_{75}(x) - P_{25}(x)\right| \quad (3)$$

where M, P_{25}, P_{75} are respectively the median, 25th percentile and 75th percentile of the MC predictions; k is an adjustable parameter to adjust the weights. In this study, k was set 1 to give the median and the CI equal weight. Combinations with the highest UCB are selected as the DOE for the next iteration. The iteration terminated when no combination was predicted to have larger UCB value than the current largest one. Although this criterion design does not guarantee the optimal combination identified is the global optimum, it maximizes the cost-effectiveness by quickly identifying an effective drug combination with minimal allowed experimental materials. As a matter of fact, one cannot justify whether the optimal combination identified is the real global optimum unless all possibilities are enumeratively tested.

3.2. In silico validation of the Monte Carlo regression

To testify the performance of ReMEMC in drug combination optimization, we adopted three benchmark functions to compare its optimization efficiency *in silico* with two most commonly used regression methods, stepwise regression [24] and Lasso regression [25]. The Ackley function [26] and Rosenbrock function [27] are two classic evaluation functions for optimization algorithms, representing mono-modal and multi-modal problems, respectively. The Hill function is adapted from the Hill equation that characterizes the combinatorial drug response curves, which is expected to be more comparable to real-world problems of combinatorial drug optimization [28]. The details of these benchmark functions are listed in Table S1.

Of note, to maximally simulate real-world dose optimization problems, we designed the benchmark process different from regular benchmarks. Since the typical application of optimization algorithms is *in silico*, for example identifying the maximum of a function which cannot be directly obtained with derivative, the benchmarks usually run the algorithms for thousands of iterations, with only one combination tested in each iteration. We herein reduced the number of iterations to 5, and included 48 combinations in each iteration, which are about the magnitude of typical biological experiments.

The results of the *in silico* simulations showed that ReMEMC outperformed stepwise regression and Lasso regression on the three benchmark functions in most scenarios of dimensions and noise levels (Figs. 1b, S1, S2). Notably, the optimization performance of ReMEMC was robust despite the increase of synthetic noises level, while stepwise regression and Lasso regression were both negatively affected. These results highlighted ReMEMC as a promising method for dose optimization of combinatorial drug therapies.

3.3. Optimization of combinations against SARS-CoV-2 from a ten-drug library

To further demonstrate its ability to solve practical problems, ReMEMC was exploited for optimizing a ten-drug combination against SARS-CoV-2. In our previous study, the ten compounds identified to exhibit potent anti-SARS-CoV-2 activity were individually tested, but their interactions and optimal dose ratios in combinational therapies remained unknown. If the dose of every drug was chosen from 3 levels (0, low, and high), there are about 59,000 possible combinations, which is beyond the capability of most research laboratories using conventional trial-and-error methods, and with one more dose level added, the number will dramatically increase to over 1000,000. Moreover, when the viral load was inhibited to about 10^7 copies/ml, the data became highly noised. The standard deviations of repeated experiments reached about 50% on average. These characteristics made the optimization of anti-SARS-CoV-2 drug combinations a perfect scenario to apply our novel ReMEMC algorithm.

In the first iteration, three doses were selected for each drug (0, low, and median, see Table S2). For non-IFN compounds, high doses were determined as 1/20 of their corresponding CC_{50} concentrations, and low doses were equivalent to 1/5 of the high doses. For IFN compounds, high doses were determined based on the observed 100-fold viral load reduction concentration compared to PBS control, and low doses were equivalent to 1/10 of the high doses. The DOE containing 36 experiments for the first iteration was generated according to uniform design so as to ensure that the combinations were evenly distributed in the combinatorial space. The concept of accumulative dose was induced to compare the doses of the drug combinations, which is directly related to toxicity. Since the therapeutic potencies of the drugs are different, doses within a combination cannot be simply summed up to compare the accumulative dose. We therefore normalized the dose of each drug in a logarithmic way, where the low dose of each drug was set as 1, and for each dilution fold increased, the normalized dose +1. The dilution folds were determined by single-drug response curves, which were determined in the previous study [20]. Though this accumulative dose might not be a perfect evaluation of combinatorial drug toxicity, it provided an objective and logical solution to conduct the optimization. Of note, the algorithm is compatible with any quantification of toxicity as one of the optimization targets.

The 36 combinations were tested on VeroE6 cells against SARS-CoV-2. The results were measured by qRT-PCR with a specific probe recognizing viral RNA. The viral loads of the 36 combinations varied from 7.2×10^5 to 8.2×10^6 copies/ml (Fig. 2a), corresponding to about 30,00- to 30,000-fold viral load reduction. ReMEMC was then used to build an MC model with these results. In each simulation of the MC method, an interaction model (linear terms + interaction terms) was regressed. Totally 10,000 independent simulations were conducted to generate the probability distribution of the coefficients (Fig. S3). The prediction of the optimal combinations was made afterward, and combinations to be tested in the next iteration were selected using the previously introduced UCB function. Additional criteria applied to the selection included (1) the number of drugs in the combination was limited to no more than 4 in consideration of clinical practice to avoid side effects and/or antagonisms, and (2) no more than one IFN drug should be included in each combination to avoid over-activation of the host immune response [29]. As a result, eight drugs (AM580, Intron A, Avonex, Rebif, Betaferon, Immunkin, chloroquine, and remdesivir) were selected to compose totally 36 ternary and quaternary combinations for the next iterations. Moreover, as the number of drugs reduced in the combinations, a higher dose option was added for every drug.

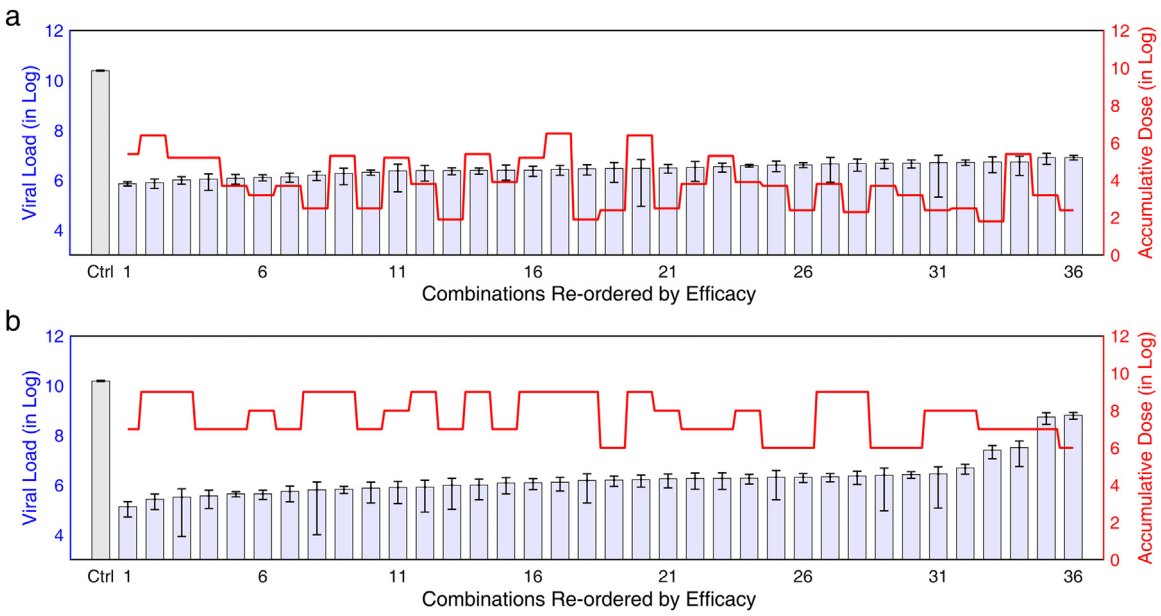


Fig. 2. The efficacy and accumulative doses of the tested combinations. (a) The 36 combinations in the 1st iteration. (b) The 36 combinations in the 2nd iteration. The bar plots refer to the logarithm transformed viral loads. The red lines refer to the sum of the logarithm transformed doses of all the drugs in every combination. The sequences of the combinations were rearranged by the efficacy from high to low. The optimal combination was identified in the 2nd iteration, marked as #1 in (b), which was composed of AM580, Betaferon, and chloroquine.

These combinations were tested with the same assays. The viral loads varied from 1.3×10^5 to 6.4×10^8 copies/ml, corresponding to about 24- to 110,000-fold viral load reduction (Fig. 2b). The results were then used to regress a new MC model (Fig. 3). Prediction of optimal combinations was performed as before with the same criterions. However, none of the top combinations selected by the UCB function was predicted to have better efficacy (median prediction) than the best combination tested in 2nd iteration. This indicated the endpoint of the optimization process as there was no combination requiring further exploration. Therefore, the best combination in the 2nd iteration was the optimal combination identified by ReMEMC, which was composed of 5.1 μM AM580, 300 UI/mL Betaferon, and 10 μM chloroquine.

3.4. *Response curves of the optimal ternary combination*

Apart from prediction of discrete doses for combinational therapies, response surfaces can also be generated from the MC model, which helps to take an intuitive look of the interactions in the combinations. Thus, a series of response surfaces of the optimal ternary combination have been plotted (Fig. 4a-c). Different from regular response surfaces, the MC model additionally provided the CI for each point on the surfaces to show the uncertainty of prediction.

Some clues of the interaction profiles can be found from the response surfaces. The anti-SARS-CoV-2 activity of AM580 alone is not obvious. When Betaferon is absent or at low concentration, AM580 showed little anti-SARS-CoV-2 effect (Fig. 4a). This is consistent with our previous finding that AM580 did not show efficacy until its dose was at or above 20 μM, [18]. In the present study, the maximal concentration was only 10 μM. However, strong synergy between Betaferon and AM580 could be observed (Fig. 4a). When Betaferon was absent, AM580 itself did not show obvious anti-viral effect, and neither did Betaferon without AM580. However, when treated with Betaferon and AM580 jointly, the viral load decreased dramatically. Additivity was observed between AM580 and chloroquine (Fig. 4a and 4b), as the surfaces were close to planes. In addition, the dose of AM580 unexpectedly changed the shape of the response surface of Betaferon and chloroquine, suggesting that AM580 might be able to induce the synergy of the other two drugs in the ternary combination (Fig. 4c). This explained why AM580 was

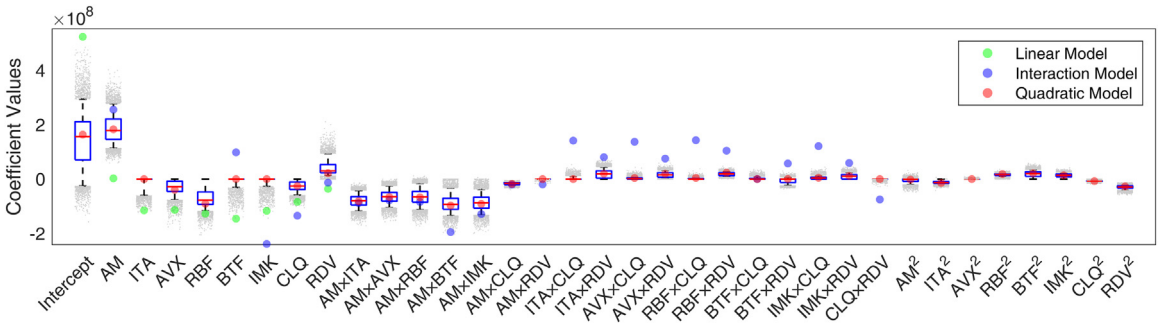


Fig. 3. The probability distributions of the coefficients in the ReMEMC model. For each coefficient, the red horizontal line, blue box, and the whiskers respectively indicate the median, 50% CI, and 90% CI of the probability distribution. The small light gray scatters are considered as outliers. For comparison, conventional modeling was also conducted with Lasso regression. The colored dots indicate the coefficient values when a linear (green), interaction (blue), or quadratic polynomial model (red) was built.

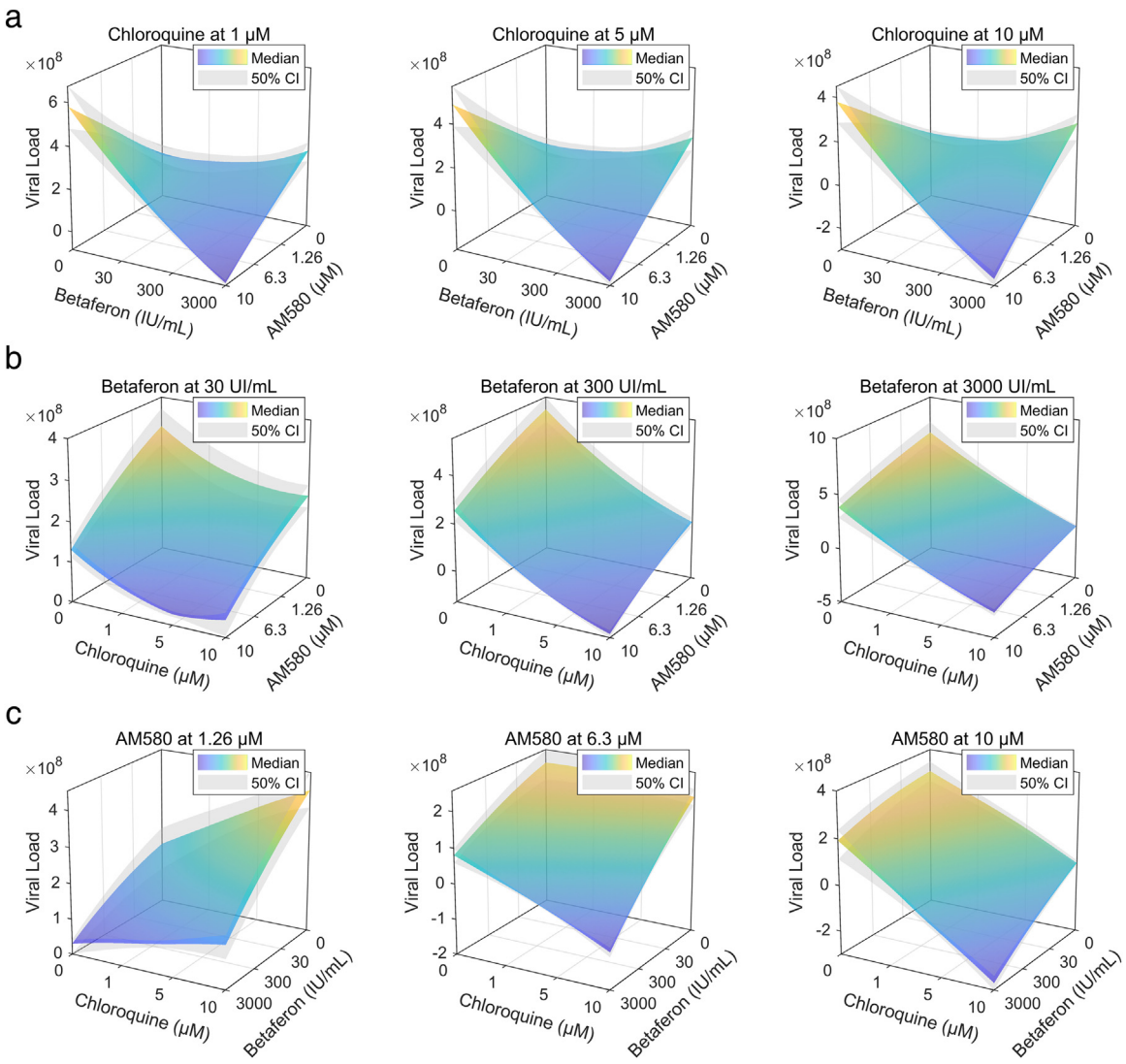


Fig. 4. Response surfaces of the optimal ternary combination. For visualization, the dose of one drug in the combination was fixed in each row of subplots: (a) Chloroquine; (b) Betaferon; (c) AM580. Three different fix doses were taken for each drug (low, median, and high, details shown in Table S2). The curved surfaces indicate the change in viral load under the change of the other two drugs. The medians of the MC predictions are shown as the colored surfaces, while the light gray translucent surfaces indicate the 50% CIs.

selected in the optimal combination. These response surfaces also revealed that increasing the dose of the three drugs in the optimal combination would likely further enhance its antiviral effects. If the combination was applied in clinical trials, these findings might indicate that an increased dosage could be an option for treating patients with severe infection as long as side effects were tolerable.

3.5. Proposed workflow for rapid identification of combinatorial drug therapies in personalized precision medicine

Through the *in silico* simulations and *in vitro* optimization of the anti-SARS-CoV-2 combination, ReMEMC has demonstrated its efficiency and reliability in optimizing drug combinations. However, because of heterogeneity of individual patients and pathogens, drug combination requires personalized optimization to achieve maximal treatment efficiency. Even for the same disease, the optimal treatment of one patient may be less effective or more toxic to another due to the mutation of pathogens and differences in physical conditions.

The efficiency in screening and optimizing combinational therapies for individualized patients quickly is the key to its successful application in personalized medicine. In this study, the whole experimental and data analysis process of every iteration took about 78 h at minimum with regular scheduling, as it was the first time the strategy was used for optimizing combinational therapy with a large number of drug compounds (Fig. 5a). However, by simply rescheduling the processes with multi-tasking, the time for each iteration can be shortened to 54 h (Fig. 5b). Considering that it generally takes 2 ~ 3 iteration on average, the total time required could be less than 5 days. Thus, the optimized drug choices and dosages could become available after 5 days for COVID-19 patients who fail to improve with standard treatment protocols initially which would be shorter than the median duration from onset of symptoms to ICU admission (9.5 days) [30]. The time window could be further shortened if the incubation time could be reduced without affecting the drug reactions. The schedule may also vary for treating different diseases and using different disease models. For example, the incubation time of bacterial infections can be as short as 12 h [15]; 24 to 48 h for cancer cells [31]; and 5 to 7 days for mini patient-derived xenograft (miniPDX) [32].

Therefore, the implementation of the novel MC regression algorithm, along with advances in drug screening technologies, signifies the

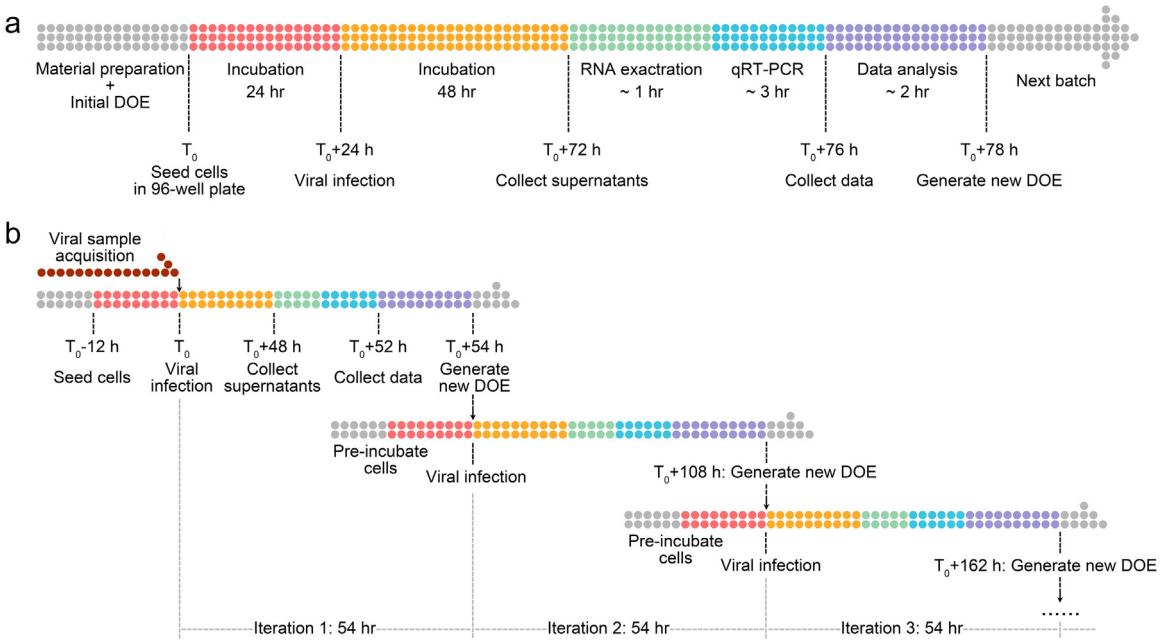


Fig. 5. The timeline of the optimization strategy for anti-SARS-CoV-2 combinatorial drugs. (a) Regular scheduling takes about 78 h for each iteration of experiments and data analysis, which requires about 1 ~ 2 skilled technicians and 1 data analyst. (b) An optimized multi-task scheduling reduces the time of each iteration to 54 h, with just 1 ~ 2 more technicians required. Ideally the optimization process finishes within 2 ~ 3 iterations, which takes 5 ~ 7 days immediately after the separation of viral sample.

realization of individualized combinatorial drug therapy, with dynamically-optimized drug repurposing and novel combination therapy development.

4. Conclusion

In this study, we exploited our novel ReMEMC algorithm to rapidly optimize combinational therapy for COVID-19. The aim of optimization was set to minimize the SARS-CoV-2 viral load while controlling the accumulative dose of the combination in a reasonable range that would not cause major adverse effects. Within just two rounds of experiments and a total of 72 combinations tested, the ReMEMC algorithm identified the optimal combination. The ternary combination managed to reduce the viral load to $1.3 \pm 0.8 \times 10^5$, which is about 100,000-fold less than the control group. The novel ternary drug combination was composed of Betaferon (recombinant IFN-β1b), AM580, and chloroquine. We have previously identified Betaferon as the most potent recombinant IFN against SARS-CoV-2 and MERS-CoV *in vitro* [18,33]. Treatment of MERS-CoV-infected common marmosets with Betaferon significantly improved the clinical, virological, and histopathological parameters [33,34]. AM580 is a selective retinoic acid receptor-α agonist that we recently found to exhibit broad-spectrum antiviral activities against coronaviruses and other emerging and circulating viruses via inhibition of SREBP [19,35]. Our data predicted that the combination of these three drugs which target different steps in the viral replication cycle should provide synergistic or additive effects against SARS-CoV-2.

The optimization of the anti-SARS-CoV-2 drug combination demonstrated the capability of the ReMEMC algorithm dealing with highly noised data. The modeling process of ReMEMC is phenotype-based and mechanism-free, designed for maximal optimization efficiency of drug combinations in an iterative way. Compared with conventional modeling and other artificial intelligence approaches, ReMEMC requires minimal experimental efforts, and maximally exploits the information from biological experiments.

Our novel optimization strategy could potentially serve as a universal and tool for rapid and reliable optimization of drug doses in combinational therapies, and help to realize personalized precision medicine. Our proposed workflow of the strategy allows individualized optimization of combinational therapies within 5 days. This would be clinically relevant even for acute infections like COVID-19 and more so for chronic infections like tuberculosis and HIV infection which require prolonged use of multi-drug regimens.

Declaration of competing interest

The authors declare that they have no conflicts of interest in this work.

Acknowledgments

This study was supported by the Health and Medical Research Fund (20190572), the Food and Health Bureau, The Government of the Hong Kong Special Administrative Region; General Research Fund (17122322 and 17126919), the Research Grants Council of the Hong Kong Special Administrative Region Government; NSFC Projects (T2122002, 22077079, 81871448), Ministry of Science and Technology of China Project (2022YFC2601700, 2022YFF0710202), Shanghai Municipal Science and Technology Project (22Z510202478), Shanghai Jiao Tong University Projects (YG2021ZD19), Shanghai Municipal Health Commission Project (2019CXJQ03); Sanming Project of Medicine in Shenzhen, China (SZSM201911014); the High Level-Hospital Program, Health Commission of Guangdong Province, China; the research project of Hainan Academician Innovation Platform (YSPTZX202004); the University of Hong Kong Outstanding Young Researcher Award; and the University of Hong Kong Research Output Prize (Li Ka Shing Faculty of Medicine). We thank Yiyang Li (Shanghai Jiao Tong University) for assistance in the experiments.

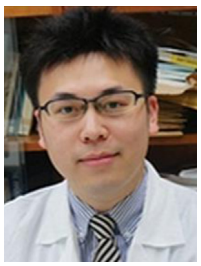
Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.fmre.2023.03.008.

References

[1] M.B.M.A. Rashid, T.B. Toh, L. Hooi, et al., Optimizing drug combinations against multiple myeloma using a quadratic phenotypic optimization platform (QPOP), Sci. Transl. Med. 10 (453) (2018) eaan0941.
[2] P.A. Ott, F.S. Hodi, H. Kaufman, et al., Combination immunotherapy: a road map, J. Immunother. Cancer 5 (1) (2017) 16.
[3] A.R. Brochado, A. Telzerow, J. Bobonis, et al., Species-specific activity of antibacterial drug combinations, Nature 559 (7713) (2018) 259–263.
[4] G. Middleton, P. Fletcher, S. Popat, et al., The National Lung Matrix Trial of personalized therapy in lung cancer, Nature 583 (7818) (2020) 807–812.
[5] G. Adam, L. Rampášek, Z. Safikhani, et al., Machine learning approaches to drug response prediction: challenges and recent progress, npj Precis. Oncol. 4 (1) (2020) 19.
[6] D. Ho, S.R. Quake, E.R.B. McCabe, et al., Enabling technologies for personalized and precision medicine, Trends Biotechnol. 38 (5) (2020) 497–518.
[7] Y.-C. Lo, S.E. Rensi, W. Torng, et al., Machine learning in chemoinformatics and drug discovery, Drug Discov. Today 23 (8) (2018) 1538–1546.
[8] D. Ho, Artificial intelligence in cancer therapy, Science 367 (6481) (2020) 982–983.
[9] A. Abdulla, B. Wang, F. Qian, et al., Project IDentif.AI: harnessing artificial intelligence to rapidly optimize combination therapy development for infectious disease intervention, Adv. Therapeut. 3 (7) (2020) 2000034.
[10] X. Ding, V.H.S. Chang, Y. Li, et al., Harnessing an artificial intelligence platform to dynamically individualize combination therapy for treating colorectal carcinoma in a rat model, Adv. Therapeut. 3 (4) (2019) 1900127.
[11] Y. Shen, T. Liu, J. Chen, et al., Harnessing artificial intelligence to optimize long-term maintenance dosing for antiretroviral-naive adults with HIV-1 infection, Adv. Therapeut. 3 (4) (2019) 1900114.
[12] B.-Y. Lee, D.L. Clemens, A. Silva, et al., Drug regimens identified and optimized by output-driven platform markedly reduce tuberculosis treatment time, Nat. Commun. 8 (1) (2017) 14183.
[13] A. Zarrinpar, D.K. Lee, A. Silva, et al., Individualizing liver transplant immunosuppression using a phenotypic personalized medicine platform, Sci. Transl. Med. 8 (333) (2016) 333ra49.
[14] B. Voelkl, N.S. Altman, A. Forsman, et al., Reproducibility of animal research in light of biological variation, Nat. Rev. Neurosci. 21 (7) (2020) 384–393.
[15] J. Sun, B. Wang, A.R. Warden, et al., Overcoming multidrug-resistance in bacteria with a two-step process to repurpose and recombine established drugs, Anal. Chem. 91 (21) (2019) 13562–13569.
[16] W. Li, C. Li, L. Gao, et al., Risk-based design optimization under hybrid uncertainties, Eng. Comput. 38 (2020) 2037–2049.
[17] A. Naseri, M. Jamei, I. Ahmadianfar, et al., Nanofluids thermal conductivity prediction applying a novel hybrid data-driven model validated using Monte Carlo-based sensitivity analysis, Eng. Comput. 38 (2020) 815–839.
[18] S. Yuan, C.C.-Y. Chan, K.K.-H. Chik, et al., Broad-spectrum host-based antivirals targeting the interferon and lipogenesis pathways as potential treatment options for the pandemic coronavirus disease 2019 (COVID-19), Viruses 12 (6) (2020) 628.
[19] S. Yuan, H. Chu, J.F.-W. Chan, et al., SREBP-dependent lipidomic reprogramming as a broad-spectrum antiviral target, Nat. Commun. 10 (1) (2019) 120.
[20] H. Chu, J.F.-W. Chan, T.T.-T. Yuen, et al., Comparative tropism, replication kinetics, and cell damage profiling of SARS-CoV-2 and SARS-CoV with implications for clinical manifestations, transmissibility, and laboratory studies of COVID-19: an observational study, Lancet Microbe 1 (1) (2020) e14–e23.
[21] J.F.-W. Chan, C.C.-Y. Yip, K.K.-W. To, et al., Improved molecular diagnosis of COVID-19 by the novel, highly sensitive and specific COVID-19-RdRp/Hel real-time reverse transcription-PCR assay validated in vitro and with clinical specimens, J. Clin. Microbiol. 58 (5) (2020) e00310–e00320.
[22] J.F.-W. Chan, A.J. Zhang, S. Yuan, et al., Simulation of the clinical and pathological manifestations of coronavirus disease 2019 (COVID-19) in a golden Syrian hamster model: implications for disease pathogenesis and transmissibility, Clin. Infect. Dis. 71 (9) (2020) 2428–2446.
[23] S. Yuan, J.F.-W. Chan, K.K.-H. Chik, et al., Discovery of the FDA-approved drugs bexarotene, cetilistat, diiodohydroxyquinoline, and abiraterone as potential COVID-19 treatments with a robust two-tier screening system, Pharmacol. Res. 159 (2020) 104960.
[24] R.R. Hocking, A biometrics invited paper. The analysis and selection of variables in linear regression, Biometrics, 32 (1) (1976) 1.
[25] R. Tibshirani, The Lasso method for variable selection in the Cox model, Stat. Med. 16 (4) (1997) 385–395.
[26] D.H. Ackley, in: A Connectionist Machine for Genetic hillclimbing. The Kluwer International Series in Engineering and Computer Science, Kluwer Academic Publishers. XII, Boston, 1987, p. 260.
[27] H.H. Rosenbrock, An automatic method for finding the greatest or least value of a function, Comput. J. 3 (3) (1960) 175–184.
[28] A. Zimmer, I. Katzir, E. Dekel, et al., Prediction of multidimensional drug dose responses based on measurements of drug pairs, Proc. Natl. Acad. Sci. U. S. A., 113 (37) (2016) 10442–10447.
[29] N. Mangalmurti, C.A. Hunter, Cytokine Storms: understanding COVID-19, Immunity 53 (1) (2020) 19–25.
[30] X. Yang, Y. Yu, J. Xu, et al., Clinical course and outcomes of critically ill patients with SARS-CoV-2 pneumonia in Wuhan, China: a single-centered, retrospective, observational study, Lancet Respir. Med. 8 (5) (2020) 475–481.
[31] A. Wang, H. Xu, X. Ding, Simultaneous optimization of drug combination dose-ratio sequence with innovative design and active learning, Adv. Therapeut. 3 (4) (2020) 1900135.
[32] P. Zhao, H. Chen, D. Wen, et al., Personalized treatment based on mini patient-derived xenografts and WES/RNA sequencing in a patient with metastatic duodenal adenocarcinoma, Cancer Commun. 38 (1) (2018) 54.
[33] J.F.W. Chan, K.-H. Chan, R.Y.T. Kao, et al., Broad-spectrum antivirals for the emerging Middle East respiratory syndrome coronavirus, J. Infect. 67 (6) (2013) 606–616.
[34] J.F.-W. Chan, Y. Yao, M.-L. Yeung, et al., Treatment with lopinavir/ritonavir or interferon-beta1b improves outcome of MERS-CoV infection in a nonhuman primate model of common marmoset, J. Infect. Dis. 212 (12) (2015) 1904–1913.
[35] M. Gianni, M.L. Calzi, M. Terao, et al., AM580, a stable benzoic derivative of retinoic acid, has powerful and selective cyto-differentiating effects on acute promyelocytic leukemia cells, Blood 87 (4) (1996) 1520–1531.

Boqian Wang is a research assistant at the School of Biomedical Engineering, Institute of Personalized Medicine, Shanghai Jiao Tong University (SJTU), China. He received his Ph.D. degree in biomedical engineering in 2022 at SJTU. His-research interests include drug combination optimization and single-cell data analysis.



Jasper Fuk-Woo Chan is a tenured clinical associate professor at the Department of Microbiology, School of Clinical Medicine, Li Ka Shing Faculty of Medicine, The University of Hong Kong, Hong Kong, China. He also serves as an Honorary Consultant at Queen Mary Hospital (Hong Kong) and The University of Hong Kong-Shenzhen Hospital (Shenzhen). His research focuses on establishing and applying novel disease models to optimize the diagnosis, treatment, and control of emerging infectious diseases with pandemic potential and those with special relevance to Asia.