1

2

3

4   Drug combination sensitivity scoring facilitates the discovery of synergistic and efficacious drug

5   combinations in cancer

6

7

8   Alina Malyutina[1], Muntasir Mamun Majumder[1], Wenyu Wang[1], Alberto Pessia[1], Caroline A.

9   Heckman[1], Jing Tang[1,2*]

10

11

12

13   [1]Insitute for Molecular Medicine Finland (FIMM), Helsinki Institute of Life Science, University of
14   Helsinki, Helsinki, Finland

15
16   [2]Department of Mathematics and Statistics, University of Turku, Turku, Finland

17
18

19   * Corresponding author

20   E-mail: jing.tang@helsinki.fi (JT)

21

22

23

24

25

26

Abstract

27

28   High-throughput drug sensitivity screening has been utilized for facilitating the discovery of drug

29   combinations in cancer. Many existing studies adopted a dose-response matrix design, aiming for

30   the characterization of drug combination sensitivity and synergy. However, there is lack of

31   consensus on the definition of sensitivity and synergy, leading to the use of different mathematical

32   models that do not necessarily agree with each other. We proposed a cross design to enable a more

33   cost-effective testing of sensitivity and synergy for a drug pair. We developed a drug combination

34   sensitivity score (CSS) to summarize the drug combination dose-response curves. Using a high-

35   throughput drug combination dataset, we showed that the CSS is highly reproducible among the

36   replicates. With machine learning approaches such as Elastic Net, Random Forests and Support

37   Vector Machines, the CSS can also be predicted with high accuracy. Furthermore, we defined a

38   synergy score based on the difference between the drug combination and the single drug dose-

39   response curves. We showed that the CSS-based synergy score is able to detect true synergistic and

40   antagonistic drug combinations. The cross drug combination design coupled with the CSS scoring

41   facilitated the evaluation of drug combination sensitivity and synergy using the same scale, with

42   minimal experimental material that is required. Our approach could be utilized as an efficient

43   pipeline for improving the discovery rate in high-throughput drug combination screening. The R

44   scripts for calculating and predicting CSS are available at https://github.com/amalyutina/CSS.

45

Author summary

47   Being a complex disease, cancer is one of the main death causes worldwide. Although new

48   treatment strategies have been achieved with cancers, they still have limited efficacy. Even when

49   there is an initial treatment response, cancer cells can develop drug resistance thus cause disease

50   recurrence. To achieve more effective and safe therapies to treat cancer, patients critically need

51   multi-targeted drug combinations that will kill cancer cells at reduced dosages and thereby avoid

52   side effects that are often associated with the standard treatment. However, the increasing number

53   of possible drug combinations makes a pure experimental approach unfeasible, even with

54   automated drug screening instruments. Therefore, we have proposed a new experimental set up to

55   get the drug combination sensitivity data cost-efficiently and developed a score to quantify the

56   efficiency of the drug combination, called drug combination sensitivity score (CSS). Using public

57   datasets, we have shown that the CSS robustness and its highly predictive nature with an accuracy

58   comparable to the experimental replicates. We have also defined a CSS-based synergy score as a

59   metric of drug interaction and justified its relevance. Thus, we expect the proposed computational

60   techniques to be easily applicable and beneficial in the field of drug combination discovery.

61

62   **Introduction**

63   Despite the great advances in the understanding of cancer, there remains a major gap

64   between the vast knowledge of molecular biology and effective anti-cancer treatments. Next

65   generation sequencing has revealed the intrinsic heterogeneity in cancer survival pathways, which

66   partly explains why patients respond differently to the same therapy [1]. To reach effective and

67   durable clinical responses, cancer patients who become resistant to standard treatments need

68   multi-targeted drug combinations, which shall effectively inhibit the cancer cells and block the

69   emergence of drug resistance [2-4].

70   In order to predict novel drug combinations, high-throughput drug screening has been

71   applied on a large variety of cancer cell lines and more recently on patient-derived cancer samples

72   [5-6]. Ideally, a promising drug combination should achieve the therapeutic efficacy at reduced

73   dosages, and therefore also minimize the toxicity and other side effects associated with high doses

74   of single drugs [7-8]. Therefore, both the sensitivity and synergy of a drug combination need to be

75   considered when evaluating the high-throughput screening results for further validation [9].

76   Many high-throughput drug combination screens combine two drugs at a full matrix of

77   multiple doses, for which the cell viability or growth inhibition effects are measured [10-11]. The

78   drug combinations are usually ranked based on the degree of synergy, such that the drug

79   combinations that produce higher growth inhibition effects compared to the single drugs will be

80   prioritized. However, there have been multiple methods to score drug synergy, each of which relies

81   on a different mathematical model that do not fully agree with each other [12]. The lack of

82   consensus on the choice of synergy scoring methods may partly explain the difficulty to validate

83   drug combination discoveries in a high-throughput setting [13]. On the other hand, focusing only on

84   synergy but not on sensitivity may produce false positive drug combinations that do not necessarily

85   reach therapeutic efficacy, despite being synergistic [14]. However, unlike the sensitivity of single

86   drugs which can be directly derived from dose-response curves [15], the sensitivity of a drug

87   combination remains largely undefined, as the same sensitivity can be achieved using different dose

88   combinations. Furthermore, there is a lack of scoring approaches to fully capture the synergy and

89   sensitivity simultaneously.

90   On the other hand, the dose-response matrix utilizes a full factorial design to test multiple

91   dose combinations, and thus demands a relatively large amount of cancer cells. For patient-derived

92   cancer samples which are typically rare and restricted in volume, the full dose-response matrix

93   design may be infeasible for testing even a minimal number drug combinations. Furthermore,

94   cancer samples of different genetic profiles are known to respond differently to the same drug

95   combination. With the limited amount of drug combinations as the training data, it becomes a

4

96   daunting task for any machine learning approach to navigate the combinatorial space to pinpoint

97   the most promising drug combinations that are specific for individual cancer samples [16].

98   To overcome these challenges, we proposed a cost-effective experimental and

99   computational procedure to facilitate the prediction of drug combination synergy and sensitivity.

100   We utilized an experimental design where two drugs are crossed at their IC_{50} concentrations, and

101   either drug is allowed to span over multiple doses while the concentration of the other drug is fixed.

102   The resulting dose-response curves are utilized to defined a drug combination sensitivity score

103   (CSS). Using a large scale of drug combination study, referred to as the O'Neil data [17], we showed

104   that the CSS is highly reproducible, suggesting its robustness as a metric for characterizing drug

105   combination responses. Furthermore, we found that the CSS can be predicted at high accuracy using

106   chemical and pharmacological features of the drug combinations. Based on the difference between

107   the observed and expected CSS values, a drug synergy score can be determined straightforwardly.

108   We showed that such a CSS-based synergy score can also detect the true synergistic and

109   antagonistic drug combinations with high accuracy. Compared to the dose-response matrix design,

110   the cross design requires minimal amount of experimental materials, while it still maintains a

111   sufficient level of accuracy for capturing both synergy and sensitivity simultaneously. We foresee

112   that such an experimental design and its CSS scoring would facilitate the standardization of drug

113   combination analysis that is currently lacking in functional chemical screening, and would allow for

114   the scale up of drug combination testing eventually for personalized medicine.

115

116   **Results**

117   **CSS values are highly reproducible and robust**

118   We applied the CSS scoring on the O'Neil drug combination data, which consists of 22,737

119   drug combinations for 39 cancer cells [17]. We found that the CSS_1 and CSS_2 values calculated using

120   either drug fixed at its IC_{50} concentration are highly correlated (Pearson correlation = 0.82, p-value

121   = 2×10^{-16}; Fig 1A). Both CSS_1 and CSS_2 values ranged from 0 to 50, with a marginal absolute difference

122   of 5.62 (Fig 1B). Such a high level of consistency holds true for all the 39 cancer cell lines and the

123   majority of the 38 unique drugs, suggesting the robustness of the CSS scoring method (Figs 1C and

124   D, Fig S1). We also found a high correlation between the CSS value and those derived from single

125   replicates (minimal correlation = 0.97, Table S3). In order to check whether the CSS values are within

126   the range of the CSS replicates for each drug combination, we calculated the minimal and maximal

127   values over the CSS replicates for each drug combination and plotted them together with CSS values

128   over the standard deviation of the CSS replicates. For a better visualization, we applied a generalized

129   additive model to smoothen the CSS lines and obtain 95% pointwise confidence interval around the

130   mean (Fig S2). Only 4% of the drug combinations have the CSS values being out of the CSS replicate-

131   based limits, however this can be explained by the higher variance over the replicates.

132   **Fig 1. Robustness and replicability of CSS.** (A) The correlation of CSS_1 and CSS_2 over all the drug

133   combinations; (B) Density plot of the CSS_1 and CSS_2 distributions; (C) The correlation per cell line

134   colored according to the tissue type; and (D) The correlation per drug colored according to the drug

135   target class.

136   Notably, we found that drug combinations that involved bortezomib showed much lower

137   correlation (0.26) between the CSS_1 and CSS_2 values compared to other drug combinations. Since

138   the O'Neil data contains the replicates for single drug screening, we analyzed the coefficient of

139   variation (CV) of the cell viability readout for each drug in the replicates. As expected, we found that

140   bortezomib has the highest CV (0.28), suggesting a relative low quality of the drug sensitivity data

141   involving this drug (Fig S3). Therefore, the lower correlation between the CSS_1 and CSS_2 for a drug

142   combination may be attributed to a higher experimental variation, and thus were considered as low

143   quality data points. For the subsequent analysis, we selected only those drug combinations that

6

144   have the absolute difference between CSS_1 and CSS_2 lower than 10, resulting in a total of 18,905

145   drug combinations. After this filtering the correlation between CSS_1 and CSS_2 was further improved

146   (Pearson correlation = 0.93, p-value = 2×10^{-16}). Furthermore, the mean absolute difference between

147   CSS_1 and CSS_2 was 3.83, which became comparable to the variability determined from the technical

148   replicates of CSS_1 and CSS_2 (2.92 and 3.06 respectively), suggesting that the difference between CSS_1

149   and CSS_2 is similar to what is expected when repeating the experiment. Taken together, CSS_1 and

150   CSS_2 values are highly consistent and therefore supported their averaging as a summary for the drug

151   combination sensitivity score.

152   **CSS can be predicted using machine learning approaches**

153   Given that the CSS is highly reproducible as a summary of the overall sensitivity of a drug

154   combination, we explored whether CSS can be predicted using pharmacological and chemical

155   information of the drugs. We considered a drug combination as a combination of its drugs target

156   profiles as well as their chemical fingerprints, with which the machine learning approaches

157   illustrated in the previous section can be optimized by exploring the feature space using the training

158   data. We examined three major machine learning methods for predictions: Elastic Net, Random

159   Forests and Support Vector Machines.

160   We found that all of these machine learning approaches worked reasonably well, with the

161   Elastic Net consistently achieving the best performance, with the mean MAE of 4.01 which is

162   comparable to that (2.07) of a technical replicate (Table 1). Note that in our cross-validation setting

163   the drug combinations in the test data were not present in the training data, the machine learning

164   methods were still able to predict the CSS values for new drug combinations by exploring the feature

165   similarity in the drug targets and chemical fingerprints. The prediction performance thus validated

166   our hypothesis that a drug combination can be considered as a combination of their drug target

167   profiles and chemical-structural properties, with which the CSS score can be predicted with high

168   confidence using the state-of-the-art machine learning approaches.

169   **Table 1. The prediction performance for Elastic Net, Random Forests and Support Vector**

170   **Machines, as compared to the upper limit when selecting randomly one technical replicate as the**

171   **prediction.**

Method	RMSE	R2	COR	MAE
Elastic Net	5.20±1.11	0.80±0.06	0.90±0.03	4.01±0.86
Random Forests	6.30±1.18	0.71±0.07	0.85±0.04	4.75±0.9
Support Vector Machines	7.47±1.32	0.57±0.08	0.80±0.04	5.80±1.07
Technical replicate	2.87±0.59	0.93±0.04	0.97±0.02	2.07±0.45

172   All the values are mean+/-standard deviation.

173   Since both the drug target profiles and chemical fingerprints were considered as the drug

174   combination features, we next evaluated their prediction performances separately using the Elastic

175   Net method. For drug-target profiles we collected known targets that were experimentally validated

176   as well as the additional secondary targets that were predicted with high confidence using the SEA

177   method. For chemical fingerprints we used the MACCS fingerprint which contains 166 structural

178   features [18]. As expected, when combining all the features the model achieved the best

179   performance (Table 2). We found that in general drug target profiles were predictive of CSS,

180   especially when including the experimentally validated targets. The predicted targets using the SEA

181   method did not improve the prediction accuracy significantly, indicating that even though

182   secondary drug target interactions may occur, most likely they have minor functional impact that

183   may not lead to the changes in cancer cell viability and thus does not contribute to the prediction

184   of CSS. On the other hand, we found that chemical fingerprints were less predictive of CSS compared

185   to the drug-target profiles, suggesting that the use of MACCS might be suboptimal to capture the

186   relevant structural information for predicting the drug combination sensitivity. However, as the

187   focus of this study was to show the validity of using machine learning methods to predict the CSS

188   score, we decided to explore other chemical fingerprint features as a future step.

189   **Table 2. The prediction performances for drug-target features and chemical fingerprint features**

190   **using Elastic Net.**

Feature	RMSE	R2	COR	MAE
Validated targets	5.66±1.27	0.77±0.07	0.88±0.04	4.26±0.95
Validated + predicted targets	5.70±1.22	0.76±0.06	0.87±0.04	4.34±0.95
Fingerprints	6.27±1.19	0.71±0.06	0.85±0.04	4.87±0.94
Validated targets + fingerprints	5.30±1.14	0.79±0.06	0.89±0.03	4.07±0.88
All features	5.20±1.11	0.80±0.06	0.90±0.03	4.01±0.86

191   All the values are mean+/-standard deviation.

192   We considered the regression coefficients that were determined in the Elastic Net model as

193   an indication of their importance to contribute to the CSS prediction. We collected 67 features that

194   have their absolute coefficients greater than 3 for at least one cell line. Unsupervised hierarchical

195   clustering with the Manhattan metric was then applied to group the cell lines and the selected

196   features (Fig 2). We found that certain drug target features were present with high coefficients

197   across all the cell lines. For example, DNA topoisomerases including TOP1MT, TOP2A, TOP2B and

198   TOP1 were selected, with an average coefficients of 8.2, 2.7, 2.6 and 1.0 separately. Despite the

199   difference in the level of variable importance, all the DNA topoisomerases showed positive

200   coefficients in 38 of 39 cell lines, suggesting that targeting DNA topoisomerases were associated

201   with a higher CSS. DNA topoisomerases are known proteins which are essential for cell replication

202   and metabolism [19]. Including a topoisomerase inhibitor can thus enhance the drug combination

203   sensitivity in many cancer cell lines. On the other hand, the only cell line that showed negative

9

204   coefficients for TOP1MT was LNCAP (prostate cancer), which turned out to be the cell line that has

205   the smallest average CSS scores for drug combinations involving the TOP1MT inhibitor (topotecan)

206   (Fig S4).

207   **Fig. 2. The feature importance map by Elastic Net for each cell line.** Cell-line independent as well

208   as cancer subtype-specific features can be identified by evaluating the regression coefficients of the

209   Elastic Net model. Features such as TOP1MT, TOP1, TOP2A/B has shown consistently positive

210   coefficients as compared to features such as AKT1/2/3 which showed cancer subtype specificity in

211   breast cancer (indicated as arrows).

212   At the cell line level, we found that cell lines of the same tissue type did not necessarily

213   cluster together, indicating their distinctive drug combination response profiles. However, we found

214   that the breast cancer cell lines did form a major cluster including KPL1, ZR751, EFM192B, OCUBM

215   and T47D, while the only outlier was MDAMB436. Indeed, MDAMB436 is the only triple negative

216   breast cancer (TNBC) subtype, while the other cell lines are either ER positive (KPL1, ZR751 and

217   T47D), or HER2 positive (EFM192B and OCUBM). It has been known that TNBC respond anticancer

218   drugs differently from ER and HER2 positive breast cancers due to the distinctive disease

219   mechanisms [20]. The features selected for the CSS prediction separated these two distinctive

220   breast cancer subtypes, suggesting the validity of using CSS to cluster cancer of different subtypes.

221   Furthermore, we found that AKT targets (AKT1/2/3) were among the top ones that showed higher

222   importance in the non-TNBC group. A combination of an AKT inhibitor and TOP1MT inhibitor

223   therefore can be suggested to treat non-TNBC but not necessarily for TNBC breast cancers. On the

224   other hand, we found that CHEK1 and PARP3/4 targets were selected for MDAMB436 but not for

225   non-TNBC group, suggesting that a combination of a CHEK inhibitor and PARP inhibitor might be

226   tested for TNBC, but not for non-TNBC. The mechanisms of actions for the proposed drug

227   combination may prove interesting for experimental validations. Taken together, the features that

228　were selected from the CSS prediction may help to pinpoint the underlying target interactions,

229　which are of pivotal importance to identify the drug combination response biomarkers.

230

231　**The CSS-based synergy scores can predict the true synergy and antagonism**

232　　　Next, we defined the degree of drug synergy as the differences between the dose-response

233　curves of a drug combination and its single drugs. We derived three variants of the CSS-based

234　synergy score (S_{sum}, S_{max}, S_{mean}) and compared them with the HSA, Bliss, Loewe and ZIP synergy

235　scores that were determined using the full-dose response matrix. We evaluated whether the CSS-

236　based synergy scores using the cross design can capture the ground truth. Although determined

237　using only one row and one column from the dose-response matrix, all the CSS-based synergy scores

238　managed to obtain a good correlation with the synergy scores based on the dose-response matrix

239　design (Table 3).

240　**Table 3. Correlations of the CSS-based synergy scores with those derived using four reference**

241　**models that were calculated using the full dose-response matrices.**

Synergy Scores	HSA	Bliss	Loewe	ZIP
S_{sum}	0.72	0.72	0.46	0.55
S_{max}	0.71	0.65	0.51	0.49
S_{mean}	0.65	0.63	0.41	0.44

242

243　　　We found that the CSS-based synergy scores correlated relatively well with the HSA and Bliss

244　scores, while the correlation started to decrease when comparing to the Loewe and ZIP scores. Since

245　all the synergy scoring models utilized different assumptions for the reference of no synergy, we

246　therefore did not expect a perfect correlations in such pairwise comparisons. Of all the three

247　variations of CSS-based synergy score, we found that S_{sum} showed the best correlation with those

248   determined using the full dose-response matrix. As S_{sum} considers the additive effect of single drug

249   sensitivities as the expectation of no synergy, it thus can be considered a more conservative scoring

250   method compared to S_{max} and S_{mean}, where the maximal and average effect of single drugs were

251   considered as reference separately. To control the false discovery rate of detecting synergistic

252   combinations, we therefore proposed S_{sum} as a more appropriate scoring method for the cross drug

253   combination design.

254   Furthermore, we evaluated the predicting accuracy of the CSS-based synergy scores for true

255   synergistic and antagonistic drug combinations. We applied a stringent criteria to determine the

256   ground truth from the dose response matrix data, such that all the synergy scores (HSA, Bliss, Loewe

257   and ZIP) must be higher than 5, or lower than -5, to be classified as a true synergistic or antagonistic

258   drug combination, respectively. From the O'Neil data, we identified 3,716 true synergistic and 218

259   true antagonistic drug combinations. We then asked the question of whether the CSS-based synergy

260   scores that were determined using the cross design can predict the ground truth. We showed that

261   the CSS-based synergy scores managed to achieve the area under the ROC curves of 0.997 (S_{sum}),

262   0.996 (S_{max}) and 0.992 (S_{mean}) to detect the true synergistic and antagonistic combinations correctly

263   (Fig 3A). Note that in order to calculate the synergy score, only two vectors of the drug combination

264   responses are needed, rather than the full dose-response matrix. Therefore, the CSS-based synergy

265   score needs a substantially fewer measurements compared to the other well-established synergy

266   scores. Still, the CSS-based synergy scores can predict the most synergistic and antagonistic drug

267   combinations with high accuracy. On the other hand, the CSS-based synergy score and the CSS drug

268   combination sensitivity score were using the same unit as the percentage of the actual drug

269   response compared to the theoretical upper limit. Therefore, the synergy score can be interpreted

270   as the extra benefit of combining two drugs that can achieve an effect closer to the upper limit. We

271   summarized both CSS drug sensitivity scores and CSS-based synergy scores for all the drug

12

272   combinations as an S (sensitivity)-S (synergy) plot (Fig 3B; Table S4). By applying a threshold of the

273   3rd quantiles for CSS and S, we can clearly identify the most promising drug combinations that fulfill

274   both the sensitivity and synergy criteria, while avoiding the false positive drug combinations that

275   might be synergistic but do not achieve a sufficient high level of sensitivity. Taken together, the

276   combined use of CSS drug combination sensitivity score and its associated synergy score allows a

277   simultaneous evaluation of the sensitivity and synergy for a drug combination, which will facilitate

278   a more systematic analysis of high-throughput drug combination data with much less experimental

279   materials.

280   **Fig. 3. (A) The ROC curves for the CSS-based synergy scores to detect true synergistic and**

281   **antagonistic drug combinations. (B) The S-S plot for all the drug combinations.** The drug

282   combinations with the 75th percentile and above for both the CSS and the S scores were considered

283   as the prioritized hits for further experimental validation and highlighted in red.

284

285   **Discussion**

286   Drug combinations may potentially lead to more durable clinical responses by overcoming

287   intra-tumoral heterogeneity and drug resistance to monotherapies. Identifying drug combinations

288   that are tailored for personalized medicine is a challenge, as the number of possible combinations

289   may easily grow exponentially [21]. High-throughput drug combination screening has been

290   increasingly utilized for early detection of true synergistic and effective drug combinations.

291   However, systematic identification of drug combinations is difficult, as the concepts of synergistic

292   versus effective drug combinations are often intertwined and sometimes interchanged without

293   sufficient clarification. Furthermore, there is a lack of consensus on what the definition of synergy

294   is, which might contribute to the poor reproducibility of many drug combination studies. The

295   uncertainty about the endpoint measurement in drug combination screens brings additional

296   complexity for any machine learning approach to tackle the prediction problem.

297   We developed a novel scoring approach called CSS for drug combinations that can be

298   efficiently determined using a simple experimental design. We found that the CSS is highly replicable

299   and therefore can be considered as a robust metric to characterize drug combination sensitivity. To

300   leverage the drug combination CSS data, we also developed a testing platform to allow for a

301   systematic evaluation of the prediction accuracy of different machine learning methods. We found

302   that the target information for the compounds as well as their chemical fingerprints are highly

303   predictive of the CSS values, with an accuracy comparable to the experimental replicates. Therefore,

304   the rationale of considering a drug combination as a combination of their target and fingerprint

305   profiles can be justified. This would also allow the augmentation of single-drug screening and drug

306   combination screening data to train a machine learning model, as many drugs are multi-targeted

307   which are equivalent to a drug combination with the same target profile. In this study we focused

308   on drug combination prediction within the same cell line. In the future, we would include the

309   molecular features of the cell lines to improve the prediction accuracy as well as identify drug

310   combination specific biomarkers across different cell lines. On the other hand, as the focus of the

311   current study is to propose the new experimental design and to justify its associated drug

312   combination scoring methods, we tested the predictability of CSS using conventional machine

313   learning methods, and showed that CSS can be accurately inferred from the pharmacological

314   features of drug combinations. More advanced machine learning methods such as Deep Learning

315   [16] or network-based methods [22] may further improve the prediction accuracy, which will be

316   tested as future work.

317   A truly promising drug combinations shall reach therapeutic efficacy via a strong synergy.

318   While there have been multiple synergy scoring methods that can be applied to the full dose-

319 response matrix design, they do not always produce the consistent results. The truly synergistic and

320 antagonistic drug combinations may therefore be determined by finding the consensus across the

321 different scoring methods [12]. We developed a CSS-based synergy score to quantify the degree of

322 interactions in a drug pair, and showed that the CSS-based synergy score can identify the truly

323 synergistic and antagonistic drug combinations accurately. Therefore, the CSS-based synergy score

324 can be used for the prioritization of a primary drug combination screen using the cross design, after

325 which only the significant drug combinations warrant a confirmation screen using the full dose-

326 response matrix design. Furthermore, we proposed a novel S-S plot to visualize drug combination

327 sensitivity and synergy using the same scale, which enables an unbiased way to explore high-

328 throughput drug combination data more efficiently. On the other hand, the CSS is defined at the

329 IC_{50} concentrations of the background drugs. Therefore, a synergistic drug combination determined

330 by the CSS-based synergy score should be more therapeutically relevant than a drug combination

331 where the synergy is detected at a higher concentrations, which are often associated with unwanted

332 off-target effects and side-effects.

333   The advantage of the proposed cross design is that drug combination screens and single-

334 drug screens can be implemented in a sequential manner, which requires much less cells and

335 compound materials compared to a full dose-response matrix. With the introduction of CSS scoring,

336 we are foreseeing a lower technical barrier to carry out large scale drug combination studies with

337 minimal cellular materials. Although the drug combination data we explored here involves cancer

338 cell lines, the cross design and its CSS scoring can be readily applied for ex-vivo drug screening,

339 where the amount of patient-derived materials can be extremely limited and technically difficult to

340 obtain due to culture constraints [23]. With the help of CSS and its visualization tools, drug

341 combination discovery can be more quickly advance and may eventually lead to the validation of

342 personalized drug combinations in clinical trials.

343

Materials and Methods

The cross drug combination design

346   We proposed a cross design to test the synergy and sensitivity of a drug pair by first

347   introducing the concepts of background drug and foreground drug: background drug is the drug

348   fixed at its IC_{50} concentration while foreground drug is added into the background drug with multiple

349   concentrations. We allow that either drug in the pair to be the background drug, so that two vectors

350   of dose mixtures will be produced and intersected at the IC_{50} concentrations (Fig 4). The dose-

351   response curves for these two vectors will be determined using cell viability or toxicity assays, where

352   the inhibition percentages can be calculated using negative and positive controls. Note that the

353   cross design requires specifically the combinations at the IC_{50} concentrations, which need to be

354   determined based on single drug sensitivity screening or prior knowledge.

355   **Fig. 4. The cross design to determine the drug combination sensitivity score.** Compared to the full-

356   dose response matrix design (left panel), only the single row and single column that correspond to

357   the IC_{50} concentrations of the two drugs were utilized for the calculation of CSS (middle panel).

358   Either Drug1 or Drug2 can be considered as the background drug fixed at its IC_{50} concentration while

359   the other is considered as the foreground drug with multiple doses being titrated. The resulting two

360   dose-response curves will be summarized as the drug combination sensitivity score (CSS), from

361   which a synergy score can also be calculated as the deviation from the expected value when there

362   is no interaction.

Determination of the CSS drug combination sensitivity scores

364   With the drug combination dose-response curves determined in the cross design, the CSS

365   summarizes the area under the curve similar to the AUC and DSS (Drug Sensitivity Score) scoring

16

366   approaches [24-25]. Namely, a four-parameter log-logistic function is used to fit the dose-response

367   curve for a concentration x of the foreground drug according to:

368
$$y = y_{min} + \frac{y_{max} - y_{min}}{1 + 10^{\lambda(\log_{10}IC_{50} - \log_{10}x)}}, \quad (1)$$

369   where y_{min} and y_{max} are the minimal and maximal inhibition (the bottom and top asymptotes of the

370   curve, $0 \leq y, y_{min}, y_{max} \leq 1$; IC_{50} is the concentration of the foreground drug with which the drug

371   combination reaches 50% inhibition of the cell growth; λ is the slope of the dose-response curve.

372   The dose-response curve (1) is transformed by substituting x with $x' = \log_{10}(x)$ as:

373
$$y = y_{min} + \frac{y_{max} - y_{min}}{1 + 10^{\lambda(\log_{10}IC_{50} - x')}} \quad (2)$$

374   The area under the $\log_{10}$-scaled dose-response curve (AUC) is determined according to

375
$$AUC = \int_{c_1}^{c_2} y_{min} + \frac{y_{max} - y_{min}}{1 + 10^{\lambda(m - x')}} dx' = y_{min}(c_2 - c_1) + (y_{max} - y_{min})\frac{1}{\lambda}\log_{10}\left(\frac{1 + 10^{\lambda(c_2 - m)}}{1 + 10^{\lambda(c_1 - m)}}\right), \quad (3)$$

376   where $[c_1, c_2]$ is the $\log_{10}$ concentration range for the foreground drug tested in the experiment, and

377    $m = \log_{10}(IC_{50})$.

378   The AUC is further normalized as the proportion of its theoretical upper bound according to:

379
$$AUC' = \frac{AUC - t(c_2 - c_1)}{(1 - t)(c_2 - c_1)}, \quad (4)$$

380   where t is the minimum inhibition level that is considered meaningful (by default it is fixed at 10%,

381   assuming that the inhibition below 10% is experimental noise).

382   The CSS for the foreground drug is defined as a percentage and varies between 0 and 100:

383
$$CSS = 100 AUC' \quad (5)$$

384   As there are two drug combination dose-response curves depending on which drug is fixed

385   as the background drug, we refer to the results of Eq. (5) for either scenario as CSS_1 and CSS_2, and

386   consider them as two samples that are generated from the same random variable. We take the

387   average of CSS_1 and CSS_2 as the CSS for the drug pair, i.e.

17

388
$$\text{CSS} = \frac{(\text{CSS}_1 + \text{CSS}_2)}{2} \tag{6}$$

389

390 **The O'Neil drug combination data**

391   Dose-response was measured as percentage of cell viability and retrieved from the

392 supplementary material of [17], which includes 22,737 experiments for 583 drug pairs that involves

393 38 unique drugs in 39 cancer cell lines, representing 7 tissue types. At the first stage, single-drug

394 screening was done using 8 concentrations to determine the IC_{50} concentration for each drug with

395 six replicates. At the second stage, a 4 by 4 dose matrix was utilized to cover the span of IC_{50}

396 concentrations for a drug pair with four replicates. To utilize the cross design, we picked up only the

397 row and the column corresponding to the concentrations closest to the IC_{50} of the single drugs.

398 These two vectors thus allowed the fitting of drug combination dose-response curves with which

399 the CSS can be calculated. The cell viability percentage was first transformed to inhibition

400 percentage according to:

401
$$\%inhibition = 100 - \%viability \tag{7}$$

402   In our analysis, the average % inhibition of the four replicates was used to calculate the CSS.

403 The robustness of the CSS values was assessed using the Pearson correlation across the four

404 replicates.

405

406 **Predicting the CSS using machine learning approaches**

407   With the CSS being determined for each drug combination, we sought to evaluate the

408 prediction accuracy of multiple machine learning methods. We considered a drug combination as a

409 combination of their targets and chemical fingerprints. We collected the known targets that have

410 been experimentally validated for the 38 drugs from Drugbank [26] and ChEMBL [27]. Furthermore,

411 we also utilized the Similarity Ensemble Approach (SEA) to predict additional secondary targets

18

412   based on the chemical structures of the drugs [28]. The targets that were predicted with Z-score

413   higher than 20, Tanimoto coefficient higher than 0.4 and P-value smaller than 0.01 were included.

414   The MACCS fingerprints of the drugs were determined using the SMILES strings with the R package

415   *rcdk* [29]. The resulting feature set for a drug combination altogether included 398 validated and

416   predicted targets and 166 MACCS fingerprints (Table S1 and S2).

417   We compared three state-of-the-art machine learning methods for the CSS prediction:

418   Elastic Net [30], Random Forests [31] and Support Vector Machines [32]. Elastic Net is a

419   regularization and feature selection method that combines both ridge and lasso regression by

420   including the L_1 and L_2 penalty terms. This method depends heavily on its penalty term that is

421   regulated by hyper parameters α and λ. In our studies, α was selected from the interval [0.1, 1] and

422   λ was chosen to minimize the difference between predicted and actual CSS scores. Random Forests

423   is an ensemble learning method that constructs multiple decision trees. In our studies, we set the

424   number of randomly selected predictors that is used at each split of the decision tree equal to the

425   rounded down square root of the number of variables. For Support Vector Machines, the tuning

426   parameters are the cost parameter C that sets the penalty for misclassification of a training point

427   and a smoothing parameter σ, based on the accuracy of predictions in cross-validation.

428   We focused on the model performance for predicting new drug combinations within the

429   same cell line, as the set of drug combinations in the training data did not overlap with that in the

430   test data. For each cell line, we randomly sampled 70% of the drug combinations to train multiple

431   machine learning models using 10-fold cross-validation. The optimized models were then used for

432   predicting the CSS values for the remaining 30% of the novel drug combinations. The same

433   processes were repeated 20 times by randomly splitting the training and test data. Four metrics

434   including coefficient of determination (R2), root mean square error (RMSE), mean absolute error

435   (MAE) and Pearson correlation (COR) were utilized for model comparison. To benchmark the

19

436   performance of the machine learning methods, we utilized one randomly selected technical

437   replicate as the prediction to obtain the upper limit of the performance. All the methods were

438   implemented and evaluated using the R package *caret* [33].

439

440   **Determination of the CSS-based drug synergy scores**

441   The advantage of CSS is that it allows a direct comparison of the sensitivity between a drug

442   combination and its single drugs, and hence facilitates the quantification of drug synergy. The

443   degree of synergy is often calculated as the deviation of the observed drug combination effect from

444   the reference, which is defined as the expectation effect if the drugs are not interacting. We defined

445   three variants of CSS-based synergy scores (termed as S scores) as:

$$S_{sum} = CSS - sum(DSS_1, DSS_2), \tag{8}$$

$$S_{max} = CSS - max(DSS_1, DSS_2), \tag{9}$$

$$S_{mean} = CSS - mean(DSS_1, DSS_2), \tag{10}$$

449   where the expectation was determined as a summary statistics of the normalized AUC based on the

450   single drug dose-response curves (termed as DSS values as proposed in [25]). For a non-synergistic

451   drug pair, the S score is expected to be zero. To evaluate the prediction accuracy of the CSS-based

452   synergy scores, we defined a set of true synergistic and antagonistic drug combinations as the gold

453   standard, which were determined using the full dose-response matrix data. We utilized the R

454   package *synergyfinder* [34] to calculate multiple versions of synergy scores including the HSA

455   (Highest single agency, [35]), the Bliss [36], the Loewe [37] and the ZIP synergy scores [38]. The

456   principles of these four models were briefed as below:

457   Consider that drug 1 at concentration x_1 and drug 2 at concentration x_2 were combined to

458   produce the inhibition effect of y_c, while their respective single drug effects were $y_1(x_1)$ and $y_2(x_2)$.

20

459   The synergy score was calculated as the difference between y_c and the expected effect y_e if there is

460   no synergy. Each synergy scoring took a different model for y_e:

461   1. HSA: y_e is the maximal single drug effect, defining

462
$$S_{\text{HSA}} = y_c - max\left(y_1(x_1), y_2(x_2)\right) \tag{11}$$

463   2. Bliss: y_e is the expected effect of two drugs acting independently, defining

464
$$S_{\text{Bliss}} = y_c - \left(y_1(x_1) + y_2(x_2) - y_1(x_1)y_2(x_2)\right) \tag{12}$$

465   3. Loewe: y_e is the expected effect of a drug combined with itself, defining

466
$$S_{\text{Loewe}} = y_c - y_1(x_1 + x_2) = y_c - y_2(x_1 + x_2) \tag{13}$$

467   4. ZIP: y_e is the expected effect of two drugs that do not potentiate each other, defining

468
$$S_{\text{ZIP}} = y_c' - \left(y_1'(x_1) + y_2'(x_2) - y_1'(x_1)y_2'(x_2)\right), \tag{14}$$

469   where $y_c', y_1'(x_1)$ and $y_2'(x_2)$ are the fitted values based on the full-dose response matrix for the

470   combination and single drugs, respectively.

471   For each of the four models, the synergy scores were determined first for a given dose

472   combination and then were averaged over the full dose-response matrix. With the four synergy

473   scores determined for each drug combination, the true synergistic and antagonistic drug

474   combinations are those with all the four synergy scores consistently higher than 5 and lower than

475   5, respectively. The aim was then to use the CSS-based synergy score which was determined by the

476   cross design data to predict the ground truth determined by the full dose-response matrix design.

477   The area under the ROC curve was used for evaluating how well the CSS-based synergy scores can

478   predict the consensus drug combinations determined using the full dose-response matrix data.

479

480   **Acknowledgements**

486

References

488   1. Lawrence MS, Stojanov P, Mermel CH, Robinson JT, Garraway LA, Golub TR, Meyerson M, Gabriel

489   SB, Lander ES, Getz G. Discovery and saturation analysis of cancer genes across 21 tumour types.

490   Nature. 2014;505(7484):495.

491   2. Hanahan D. Rethinking the war on cancer. The Lancet. 2014;383(9916):558-563.

492   3. Crystal AS, Shaw AT, Sequist LV, Friboulet L, Niederst MJ, Lockerman EL, Frias RL, Gainor JF,

493   Amzallag A, Greninger P, Lee D. Patient-derived models of acquired resistance can identify effective

494   drug combinations for cancer. Science. 2014;346(6216):1480-1486.

495   4. Dry JR, Yang M, Saez-Rodriguez J. Looking beyond the cancer cell for effective drug combinations.

496   Genome Medicine. 2016;8(1):125.

497   5. Held MA, Langdon CG, Platt JT, Graham-Steed T, Liu Z, Chakraborty A, Bacchiocchi A, Koo A,

498   Haskins JW, Bosenberg MW, Stern DF. Genotype-selective combination therapies for melanoma

499   identified by high-throughput drug screening. Cancer Discovery. 2013;3(1):52-67.

500   6. Gianni M, Qin Y, Wenes G, Bandstra B, Conley AP, Subbiah V, Leibowitz-Amit R, Ekmekcioglu S,

501   Grimm EA, Roszik J. High-throughput architecture for discovering combination cancer therapeutics.

502   JCO Clinical Cancer Informatics. 2018;11;2:1-2.

503   7. Al-Lazikani B, Banerji U, Workman P. Combinatorial drug therapy for cancer in the post-genomic

504   era. Nature Biotechnology. 2012;30(7):679-692.

505   8. Gottesman MM, Lavi O, Hall MD, Gillet JP. Toward a better understanding of the complexity of

506   cancer drug resistance. Annual Review of Pharmacology and Toxicology. 2016;56:85-102.

507 9. Doroshow JH, Simon RM. On the design of combination cancer therapy. Cell. 2017;171(7):1476-

508 1478.

509 10. Griner LA, Guha R, Shinn P, Young RM, Keller JM, Liu D, Goldlust IS, Yasgar A, McKnight C, Boxer

510 MB, Duveau DY. High-throughput combinatorial screening identifies drugs that cooperate with

511 ibrutinib to kill activated B-cell–like diffuse large B-cell lymphoma cells. Proceedings of the National

512 Academy of Sciences. 2014;111(6):2349-2354.

513 11. Licciardello MP, Ringler A, Markt P, Klepsch F, Lardeau CH, Sdelci S, Schirghuber E, Müller AC,

514 Caldera M, Wagner A, Herzog R. A combinatorial screen of the CLOUD uncovers a synergy targeting

515 the androgen receptor. Nature Chemical Biology. 2017;13(7):771-778.

516 12. Tang J, Wennerberg K, Aittokallio T. What is synergy? The Saariselkä agreement revisited.

517 Frontiers in Pharmacology. 2015;6:181.

518 13. Haverty PM, Lin E, Tan J, Yu Y, Lam B, Lianoglou S, Neve RM, Martin S, Settleman J, Yauch RL,

519 Bourgon R. Reproducible pharmacogenomic profiling of cancer cell line panels. Nature.

520 2016;533(7603):333-337.

521 14. Palmer AC, Sorger PK. Combination cancer therapy can confer benefit via patient-to-patient

522 variability without drug additivity or synergy. Cell. 2017;171(7):1678-1691.

523 15. Smirnov P, Kofia V, Maru A, Freeman M, Ho C, El-Hachem N, Adam GA, Ba-alawi W, Safikhani Z,

524 Haibe-Kains B. PharmacoDB: an integrative database for mining in vitro anticancer drug screening

525 studies. Nucleic Acids Research. 2017;46(D1):D994-1002.

526 16. Preuer K, Lewis RP, Hochreiter S, Bender A, Bulusu KC, Klambauer G. DeepSynergy: predicting

527 anti-cancer drug synergy with Deep Learning. Bioinformatics. 2017;34(9):1538-1546.

528 17. O'Neil J, Benita Y, Feldman I, Chenard M, Roberts B, Liu Y, Li J, Kral A, Lejnine S, Loboda A, Arthur

529 W. An unbiased oncology compound screen to identify novel combination strategies. Molecular

530 Cancer Therapeutics. 2016;15(6):1155-1162.

531   18. Durant JL, Leland BA, Henry DR, Nourse JG. Reoptimization of MDL keys for use in drug discovery.

532   Journal of Chemical Information and Computer Sciences. 2002;42(6):1273-1280.

533   19. Pommier Y, Leo E, Zhang H, Marchand C. DNA topoisomerases and their poisoning by anticancer

534   and antibacterial drugs. Chemistry & Biology. 2010;17(5):421-433.

535   20. Bianchini G, Balko JM, Mayer IA, Sanders ME, Gianni L. Triple-negative breast cancer: challenges

536   and opportunities of a heterogeneous disease. Nature Reviews Clinical Oncology. 2016;13(11):674.

537   21. Iorio F, Knijnenburg TA, Vis DJ, Bignell GR, Menden MP, Schubert M, Aben N, Gonçalves E,

538   Barthorpe S, Lightfoot H, Cokelaer T. A landscape of pharmacogenomic interactions in cancer. Cell.

539   2016;166(3):740-754.

540   22. Li H, Li T, Quang D, Guan Y. Network propagation predicts drug synergy in cancers. Cancer

541   Research. 2018;78(18):5446-5457.

542   23. Letai A. Functional precision cancer medicine—moving beyond pure genomics. Nature

543   Medicine. 2017;23(9):1028-1035.

544   24. Yang W, Soares J, Greninger P, Edelman EJ, Lightfoot H, Forbes S, Bindal N, Beare D, Smith JA,

545   Thompson IR, Ramaswamy S. Genomics of Drug Sensitivity in Cancer (GDSC): a resource for

546   therapeutic biomarker discovery in cancer cells. Nucleic Acids Research. 2012;41(D1):D955-961.

547   25. Yadav B, Pemovska T, Szwajda A, Kulesskiy E, Kontro M, Karjalainen R, Majumder MM, Malani

548   D, Murumägi A, Knowles J, Porkka K. Quantitative scoring of differential drug sensitivity for

549   individually optimized anticancer therapies. Scientific Reports. 2014;4:5193.

550   26. Wishart DS, Feunang YD, Guo AC, Lo EJ, Marcu A, Grant JR, Sajed T, Johnson D, Li C, Sayeeda Z,

551   Assempour N. DrugBank 5.0: a major update to the DrugBank database for 2018. Nucleic Acids

552   Research. 2017;46(D1):D1074-1082.

553   27. Gaulton A, Hersey A, Nowotka M, Bento AP, Chambers J, Mendez D, Mutowo P, Atkinson F, Bellis

554   LJ, Cibrián-Uhalte E, Davies M. The ChEMBL database in 2017. Nucleic Acids Research.

555   2016;45(D1):D945-954.

556   28. Keiser MJ, Roth BL, Armbruster BN, Ernsberger P, Irwin JJ, Shoichet BK. Relating protein

557   pharmacology by ligand chemistry. Nature Biotechnology. 2007;25(2):197-206.

558   29. Guha R. Chemical informatics functionality in R. Journal of Statistical Software. 2007;18(5):1-6.

559   30. Zou H, Hastie T. Regularization and variable selection via the elastic net. Journal of the Royal

560   Statistical Society: Series B (Statistical Methodology). 2005;67(2):301-320.

561   31. Breiman L. Random forests. Machine Learning. 2001;45(1):5-32.

562   32. Ben-Hur A, Ong CS, Sonnenburg S, Schölkopf B, Rätsch G. Support vector machines and kernels

563   for computational biology. PLoS Computational Biology. 2008;4(10):e1000173.

564   33. Kuhn M. Caret package. Journal of Statistical Software. 2008;28(5):1-26.

565   34. He L, Kulesskiy E, Saarela J, Turunen L, Wennerberg K, Aittokallio T, Tang J. Methods for high-

566   throughput drug combination screening and synergy scoring. Methods in Molecular Biology.

567   2018;1711:351-398.

568   35. Berenbaum, MC. What is synergy? Pharmacological Review. 1989;41(2):93-141.

569   36. Bliss CI. The toxicity of poisons applied jointly. Annals of Applied Biology. 1939;26(3):585-615.

570   37. Loewe S. The problem of synergism and antagonism of combined drugs. Arzneimittelforschung.

571   1953;3(6):285-290.

572   38. Yadav B, Wennerberg K, Aittokallio T, Tang J. Searching for drug synergy in complex dose–

573   response landscapes using an interaction potency model. Computational and structural

574   biotechnology journal. 2015;13:504-513.

575

576 **Supporting Information**

577 Table S1. Drug target profiles including experimentally-validated primary and secondary targets, and

578 SEA-predicted secondary targets for the 38 compounds [SupTable1.xlsx]

579 Table S2. The chemical information including MACCS fingerprint profiles for the 38 compounds

580 [SupTable2.xlsx]

581 Table S3. The correlations of the CSS values obtained using the average of four viability replicates,

582 with the CSS values obtained from the replicate separately [SupTable3.xlsx].

583 Table S4. CSS drug combination sensitivity scores and S synergy scores for each drug combination

584 [SupTable4.xlsx]

585 Fig S1. The heatmap of the CSS1-CSS2 correlations sorted by drug combinations. Drug classes are

586 shown in different colors at the edges of the heatmap.

587 Fig S2. Replicability of CSS values over the replicates. The line plot of the minimal and maximal values

588 for the CSS replicates combined with CSS values over the standard deviation of the CSS replicates.

589 Fig S3. The coefficient of variation (CV) for each drug in the single drug screens. The correlations

590 between CSS1 and CSS2 for the drug combinations that involve a given drug were shown on top of

591 each bar.

592 Fig S4. The correlation between the variable importance of TOP1MT and the average CSS for

593 TOP1MT inhibitor for all the 39 cell lines. LNCAP is the only line which has a negative variable

594 importance for TOP1MT.
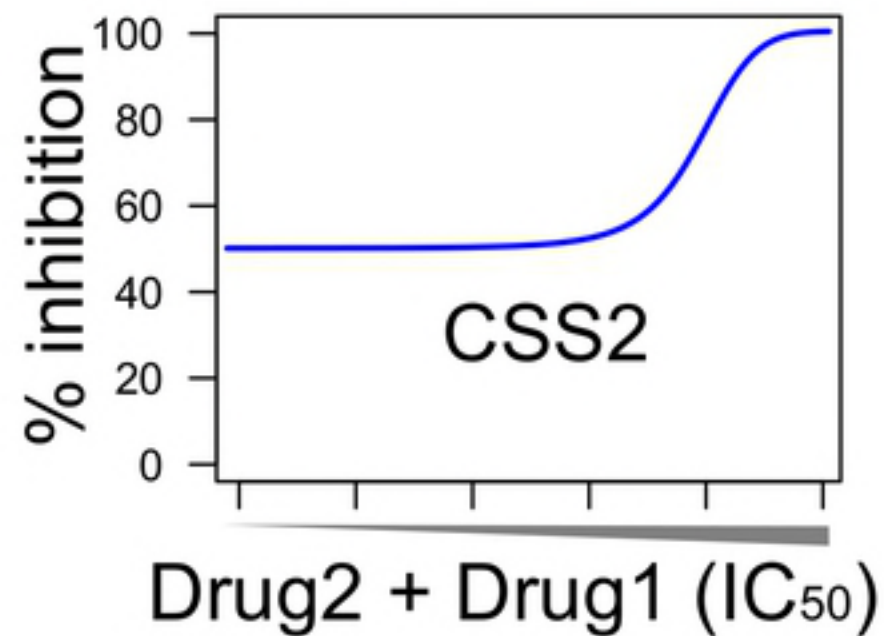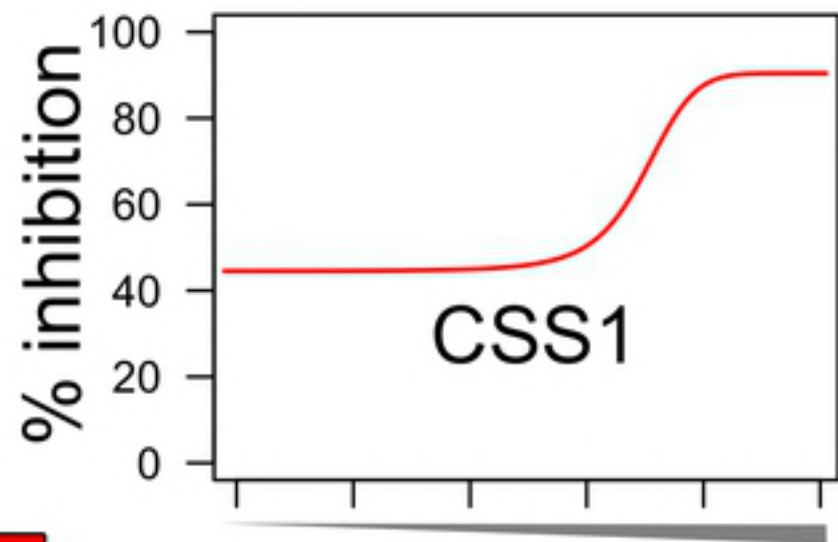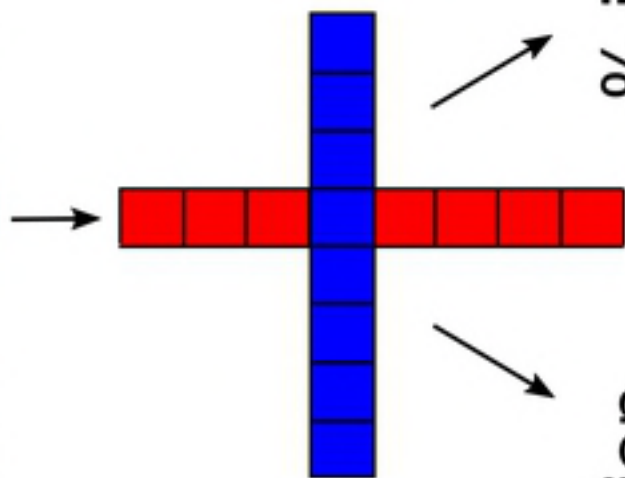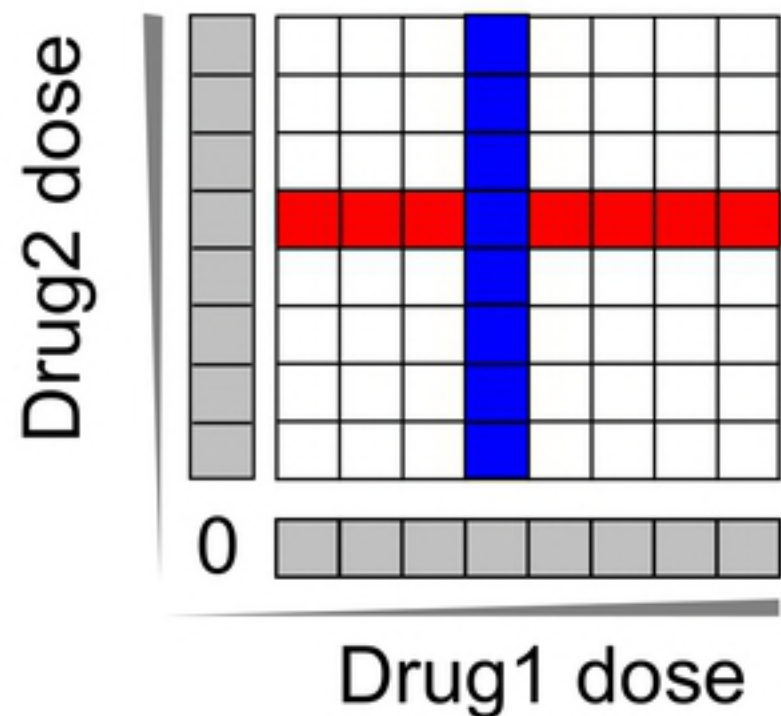
$$CSS = (CSS_1 + CSS_2)/2$$
$$Synergy = CSS - reference$$

Drug2 dose

0

Drug1 dose

% inhibition

CSS1

Drug1 + Drug2 (IC_{50})

% inhibition

CSS2

Drug2 + Drug1 (IC_{50})

Fig1

Fig2

Fig3

Fig4