

# Use and Implications of the Harmonic Mean Model on Mixtures for Basic Research and Drug Discovery

Radleigh G. Santos,<sup>†</sup> Marc A. Giulianotti,<sup>†</sup> Colette T. Dooley,<sup>†</sup> Clemencia Pinilla,<sup>‡</sup> Jon R. Appel,<sup>‡</sup> and Richard A. Houghten<sup>\*,†</sup>

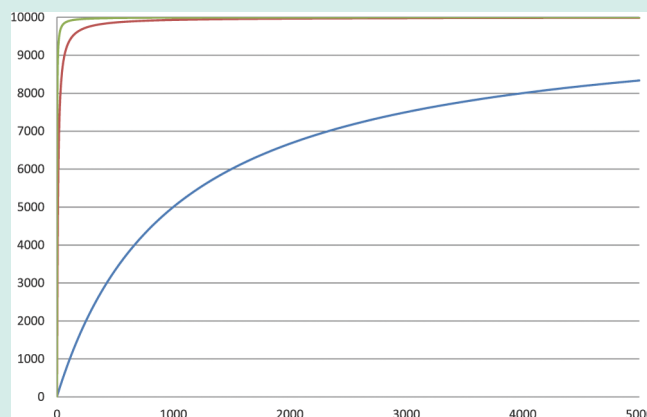
<sup>†</sup>Torrey Pines Institute for Molecular Studies, Port St. Lucie, Florida 34987, United States

<sup>‡</sup>Torrey Pines Institute for Molecular Studies, San Diego, California 92121, United States

**S** Supporting Information

**ABSTRACT:** The use of the harmonic mean model for predicting the activities of a given mixture and its constituents has not previously been explored in the context of combinatorial libraries and drug discovery. Herein, the analyses of historical data confirm the harmonic mean as an accurate predictor of mixture activity. The implications of these results are discussed.

**KEYWORDS:** harmonic mean model, mixture activity, drug discovery, combinatorial libraries



## INTRODUCTION

The use of mixtures is not a new concept in drug discovery; for example, approximately a 1,000 extracts derived from plants and used for the treatment of ailments are written about on tablets dating back to 2600 BC.<sup>1</sup> In modern drug discovery efforts mixtures are still being assessed to identify active compounds. These mixture samples include natural product extracts as well as systematically arranged mixtures of many compounds (such as mixture-based libraries<sup>2,3</sup>) and mixtures of few compounds such as cassette dosing and pooling.<sup>4–6</sup> In all these methods there is the need to distinguish the activity of one mixture sample from another. The activity of a mixture is of course driven by the individual components comprising the mixture. To this extent it is critical to understand how the individual components of a mixture contribute to the overall activity of the mixture sample. The predictive capabilities of averaging models on such mixtures are examined here using 36 case studies from 8 different publications.

## METHODS

The use of the harmonic mean as a method of modeling the activity of a mixture, given the activity of that mixture's constituents, is not new.<sup>7–11</sup> As first described by Finney,<sup>7</sup> the use of the harmonic mean as an averaging method is most mathematically suitable to model conditions based on the assumption of simple independent action. In fact, previous studies<sup>10,11</sup> have used the harmonic mean as a metric for determining the extent to

which simple independent action is present in a mixture. Its usefulness when applied to modeling the behavior of mixture-based combinatorial libraries associated specifically with drug discovery, however, is worthy of study. In particular, the effect of the mathematical properties of the harmonic mean on the efficacy of the use of mixture-based combinatorial libraries in drug discovery merits explicit exploration. This study therefore begins with a comparison of how the harmonic mean differs from other classical averaging methods, and these methods' relative accuracies when applied to biological data where the assumption of simple independent action is appropriate.

The classical methods of averaging compared herein are the arithmetic mean, the geometric mean, and the harmonic mean. The arithmetic mean is defined by the equation:

$$A = \sum_{i=1}^N f_i X_i$$

where  $f_i$  is the proportion of the  $i$ -th mixture constituent with dosing point  $X_i$ .  $N$  is the total number of mixture constituents; if constituents are present in equal numbers, then  $f_i = 1/N$  for all  $i$ . Similarly, the geometric mean is given by

$$G = 10^{\sum_{i=1}^N f_i \log(X_i)}$$

**Received:** February 11, 2011

**Published:** March 11, 2011

and the harmonic mean by

$$H = \frac{N}{\sum_{i=1}^N \frac{f_i}{X_i}}$$

Use of the arithmetic mean has been applied to simulated combinatorial libraries.<sup>12</sup> The geometric mean was calculated as an additional point of comparison, since it represents the use of the arithmetic mean on, for example,  $\log_{10}(\text{IC}_{50})$  values. In this study, because of the synthetic methods used to prepare the mixtures,<sup>13,14</sup> we assume that mixture constituents are present in equal proportions.

To demonstrate the relative effectiveness of the three above addition models, we use historical data which utilized the iterative process of deconvolution of mixture based libraries<sup>13–20</sup> reported by our laboratories and other groups. These data are ideal to compare the performances of each averaging method as a model for the activity of mixtures, because both the  $\text{IC}_{50}$  value of each of the mixtures and all of the constituent submixtures (or individual compounds) in most cases were determined and have been reported. In total, 36 different mixtures, each consisting of 4 to 19 constituents, were analyzed. The number of compounds in the constituents of the mixtures studied ranged from 1 to 6,859. In most cases, the measured  $\text{IC}_{50}$  value was reported and was used for this analysis. For those constituents whose  $\text{IC}_{50}$  is large and is only published as a lower bound, we use this lower bound as that constituent's  $\text{IC}_{50}$  if it is greater than the best  $\text{IC}_{50}$  of the previous iterative step. For example, in Houghten et al.<sup>13</sup> the mixture Ac-DVPAXX-NH<sub>2</sub> is reported as having an  $\text{IC}_{50}$  value “>1,400  $\mu\text{M}$ ,” and Ac-DVPXXX-NH<sub>2</sub> has a reported  $\text{IC}_{50}$  value of 41  $\mu\text{M}$ , which is less than 1,400  $\mu\text{M}$ , so the mixture Ac-DVPAXX-NH<sub>2</sub> was assigned a value of 1,400  $\mu\text{M}$ . In contrast, in Davis et al.<sup>19</sup> (egCB)(dG)XXT has a reported  $\text{IC}_{50}$  of “>10  $\mu\text{M}$ ” and (egCB)XXXT has a reported  $\text{IC}_{50}$  value of 40  $\mu\text{M}$ , which is greater than 10. Data such as this is discarded since no averaging method could return an accurate result; if the constituents were assigned values which made them all more active than the mixture itself, the average would necessarily have to be as well. For those constituents whose  $\text{IC}_{50}$  was too large to have any published data, we assign to them the  $\text{IC}_{50}$  of the least active measured compound. For example, in Dooley et al.<sup>14</sup> the  $\text{IC}_{50}$  value of Ac-rfgxxx-NH<sub>2</sub> is reported as “ND” since it was inactive at the highest dose tested. It is thus assigned a value of 69,000 nM, that of the least active measured constituent.

## RESULTS

For each mixture, the constituents of that mixture were added using each of the three addition models, and the results were compared to actual experimentally obtained mixture value. An example of these data, along with the outputs of each of the three addition models, is provided in Table 1. A summary of the results from all sources is presented in Table 2. The data in its entirety is in the Supporting Information, Table S1a–h. Data points which were altered to have assigned values (as described above) are shown in italics.

Because the data is taken across mixtures with varying complexities and activities, the numerical scale for each data point varies widely. If the error in a model-predicted value were simply the difference between the prediction and the experimental value, they would not be numerically comparable. Therefore, the error

**Table 1. Example of Data Used in This Study<sup>14</sup>**

mixture	$\text{IC}_{50}$ nM	
Ac-rfwink-NH <sub>2</sub>	110	
constituent compounds	$\text{IC}_{50}$ nM	
Ac-rfwink-NH <sub>2</sub>	18	
Ac-rfwir-NH <sub>2</sub>	27	
Ac-rfwina-NH <sub>2</sub>	37	
Ac-rfwins-NH <sub>2</sub>	130	
Ac-rfwinp-NH <sub>2</sub>	130	
Ac-rfwinn-NH <sub>2</sub>	130	
Ac-rfwinq-NH <sub>2</sub>	140	
Ac-rfwing-NH <sub>2</sub>	170	
Ac-rfwim-NH <sub>2</sub>	180	
Ac-rfwinh-NH <sub>2</sub>	200	
Ac-rfwint-NH <sub>2</sub>	230	
Ac-rfwihi-NH <sub>2</sub>	460	
Ac-rfwini-NH <sub>2</sub>	680	
Ac-rfwinf-NH <sub>2</sub>	770	
Ac-rfwiniw-NH <sub>2</sub>	790	
Ac-rfwine-NH <sub>2</sub>	960	
Ac-rfwind-NH <sub>2</sub>	1,100	
Ac-rfwinv-NH <sub>2</sub>	1,300	
Ac-rfwini-NH <sub>2</sub>	5,600	
mean	predicted value	scaled error
arithmetic	687	5.24
geometric	265	1.41
harmonic	106	0.04

was scaled by the experimental value, so that

$$\text{Scaled Error} = \frac{|(\text{experimental } \text{IC}_{50}) - (\text{model } \text{IC}_{50})|}{\text{experimental } \text{IC}_{50}}$$

Thus, the scaled error is as a fraction of the experimental value. The scaled error for each prediction is also included in Tables 1 and 2. The average scaled error for the arithmetic mean model is 144.59, the average scaled error for the geometric mean model is 59.02, and the average error for the harmonic mean model is 0.47. Thus the harmonic mean was the only addition model that consistently was capable of capturing to within an order of magnitude the  $\text{IC}_{50}$  value of the resultant mixture given the  $\text{IC}_{50}$  value of that mixture's constituents. The maximum harmonic mean scaled error was only 3.71, as compared to maximum scaled errors of 916.03 for the geometric mean and 2061.38 for the arithmetic mean. The harmonic mean has lower scaled errors than both other methods for all but five of the analyzed mixtures; in those five cases, the scaled error is below 0.60 for the harmonic mean, indicating the harmonic mean gave reasonably good approximations in these cases as well.

To compare the ability of each of the three averaging models to predict the experimental mixture  $\text{IC}_{50}$  value, a least-squares linear regression was performed. Least-squares linear regression allows each model to be evaluated as a whole, rather than looking at individual predictions. Because of the large difference in scale among the data points, the regression was performed on the logarithms of the  $\text{IC}_{50}$  values. The functional form of the fit

Table 2. Summary of the Results of the Prediction Models

reference	mixture value	arithmetic mean	scaled error	geometric mean	scaled error	harmonic mean	scaled error
Houghtenet. al. <sup>13</sup>	250	1,093	3.37	827	2.31	370	0.48
	41	1,322	31.26	1,016	23.79	75	0.83
	4.4	1,322.2	299.51	887.2	200.63	6.8	0.55
	0.38	11.19	28.44	4.90	11.89	0.44	0.17
Dooley et. al. <sup>14</sup>	14,000	22,642	0.62	12,414	0.11	5,618	0.60
	1,500	4,663	2.11	2,494	0.66	1,586	0.06
	480	2,049	3.27	1,122	1.34	729	0.52
	110	687	5.24	265	1.41	106	0.04
Dooley et. al. <sup>15</sup>	2,701	14,962	4.54	7,989	1.96	3,954	0.46
	907	1,323	0.46	1,077	0.19	828	0.09
	106	316	1.99	231	1.18	151	0.42
	24	30	0.26	19	0.20	13	0.45
Pinillaet. al. <sup>16</sup>	1,000	738,795	737.80	168,921	167.92	865	0.14
	400	4,267	9.67	3,118	6.79	1,882	3.71
	48	69	0.43	40	0.17	25	0.49
	6.6	9.1	0.38	7.7	0.17	6.9	0.04
Houghtenet. al. <sup>17</sup>	20,000	357,916	16.90	136,593	5.83	13,557	0.32
	860	1,089,794	1,266.20	560,506	650.75	1,705	0.98
	90	33,730	373.78	4,149	45.10	91	0.01
	6	15	1.45	12	1.04	10	0.72
Appelet. al. <sup>18</sup>	200,000	1,253,690	5.27	829,323	3.15	83,762	0.58
	7,329	840,278	113.65	208,754	27.48	5,813	0.21
	376	514	0.37	456	0.21	407	0.08
	257	299	0.16	255	0.01	230	0.11
	12,780	622,127	47.68	292,158	21.86	7,481	0.41
	408	841,453	2,061.38	374,148	916.03	702	0.72
	37	6,120	164.41	788	20.29	27	0.28
Davis et al. <sup>19</sup>	30	83	1.75	73	1.45	61	1.03
	20	41	1.04	36	0.82	30	0.50
Ecker et al. <sup>20</sup>	4	7.6	0.91	4.7	0.18	1.7	0.57
	0.5	7.5	14.08	3.5	6.00	0.6	0.15
	0.15	0.62	3.13	0.42	1.82	0.24	0.57
	0.08	0.28	2.53	0.18	1.22	0.11	0.37
	0.05	0.08	0.65	0.07	0.48	0.06	0.26
	0.03	0.04	0.35	0.04	0.27	0.04	0.17
	0.02	0.03	0.38	0.02	0.17	0.02	0.02

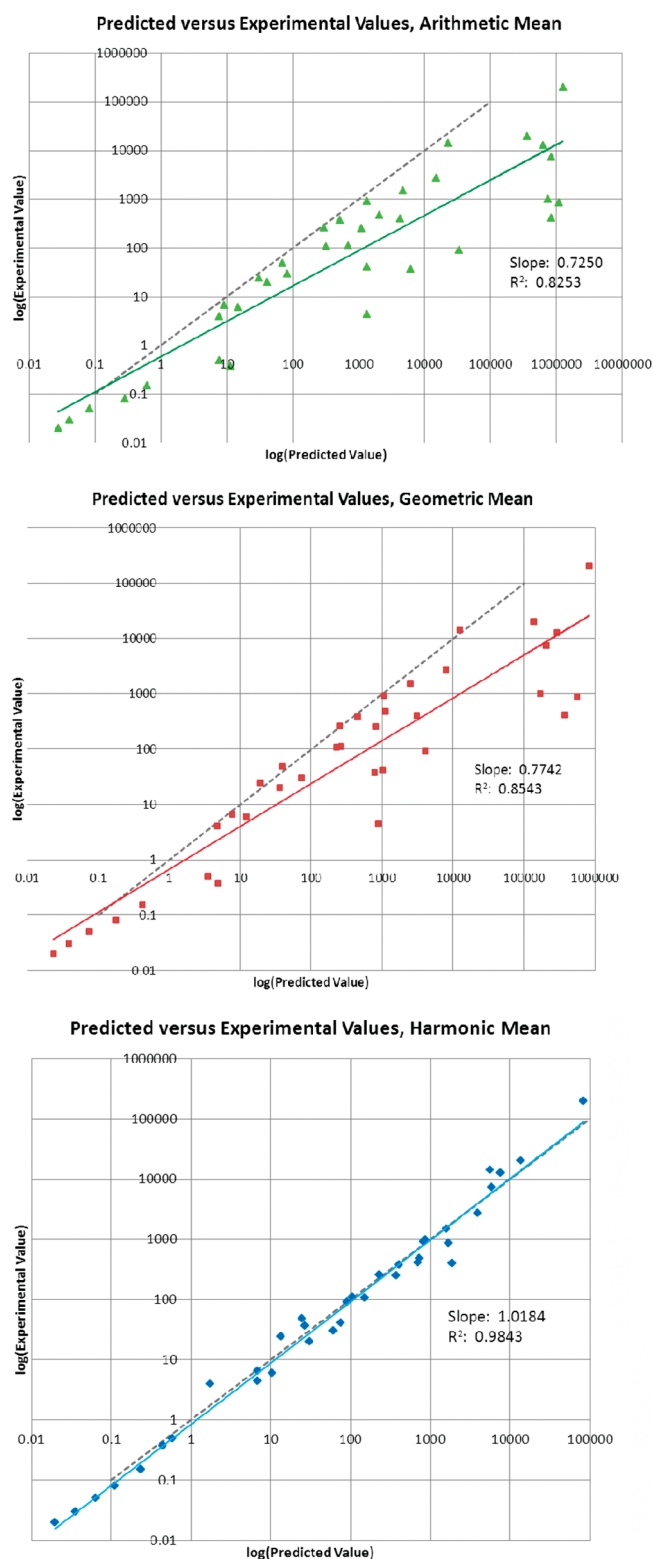
curves was therefore

$$\log(\text{experimental IC}_{50}) = a_1 \log(\text{model IC}_{50}) + a_0$$

Clearly, for a perfect model  $\log(\text{experimental IC}_{50}) = \log(\text{model IC}_{50})$ , and so a measure of how well the addition model is predicting the mixture  $\text{IC}_{50}$  is how close  $a_1$  is to one and  $a_0$  is to zero. Additionally, each fit also includes an  $R^2$  value. The  $R^2$  value is a measure of the goodness-of-fit, and represents the percentage of variance in the data that is explained by the model. For the arithmetic mean model,  $a_1 = 0.7250$ ,  $a_0 = -0.2291$ , and  $R^2 = 0.8253$ . For the geometric mean model,  $a_1 = 0.7742$ ,  $a_0 = -0.1685$ , and  $R^2 = 0.8543$ . For the harmonic mean model,  $a_1 = 1.0184$ ,  $a_0 = -0.0679$ , and  $R^2 = 0.9843$ . Plots of each of the addition model's predictions against the experimental values, along with the least-squares linear regression best fit line, are in Figure 1. The harmonic mean model both provides a better fit to the data given its superior  $R^2$  value, and also provides a slope and intercept very close to ideal.

## IMPLICATIONS OF A HARMONIC MEAN MODEL

The above analyses strongly suggest that the theory<sup>7,9,10</sup> which predicts the harmonic mean as the averaging methodology for mixtures under the assumption of simple independent action is valid in the data analyzed above. It should be noted that in the data surveyed above, the complexity of both the parent mixture and its constituents varies drastically from case to case. See Table 3: this variation does not affect the accuracy of the harmonic mean, which yielded significantly more accurate results than the other two standard averaging methodologies studied. In most cases the harmonic mean predicted the experimental results with high precision. This fact strongly suggests that the harmonic mean is an appropriate way of modeling the behavior of many of the mixture-based combinatorial libraries used in basic research and drug discovery.<sup>12–21</sup> The ability to predict the outcomes of combinatorial mixture-based experiments is useful for a variety of practical dosing and experimental design applications, such as deviation from the harmonic mean as a metric for determining



**Figure 1**

the existence of synergy or antagonism.<sup>10,11</sup> In the context of positional scanning, where each position is a different arrangement of the same constituents, the harmonic mean of the  $IC_{50}$  values at each position ought to equal one another. For example, Dooley and Houghten<sup>21</sup> presents positional scanning data for six positions, which has been reproduced

in Supporting Information, Table S2. Each position consists of 18 mixtures, each of which contains 1,889,568 individual compounds. The harmonic means of the  $IC_{50}$ 's of these mixtures at each position are 340, 271, 168, 248, 263, and 211. Thus it can be seen that the maximal scaled error in this set is only 1.02, and that most pairwise comparisons have significantly smaller scaled errors.

In addition to the above, the mathematical properties of the various averaging methods can help validate the usage of combinatorial mixture libraries in basic research and drug discovery and explain the impressive successes the process has already achieved.<sup>2–6,13–21</sup> The harmonic mean model differs from the other models primarily in its treatment of extreme numerical ranges within the data. Here the range of the data is defined to be the ratio of the highest constituent  $IC_{50}$  value to the lowest. A log–log plot of this value versus the scaled error for each of the three addition models is shown in Figure 2. This plot shows how the harmonic mean is able to outperform the other addition models: while at low ranges (with a ratio of most to least active less than 100) all three methods perform similarly, the scaled error of the arithmetic and geometric models rises steadily as the range increases, while the harmonic mean maintains a similar level of scaled error throughout. In particular, the harmonic mean is more influenced by active compounds having smaller  $IC_{50}$  values than the other addition methods; that the experimental data is well-predicted by the harmonic mean indicates that the experimental behavior of mixture based combinatorial libraries behaves similarly, with active compounds driving the activity rather than inactive compounds diluting it.

To further elucidate this point, we consider the hypothetical situation in which we define an active compound to have a fixed  $IC_{50}$  value of  $\alpha$  and an inactive compound to have a fixed  $IC_{50}$  value of  $\beta$ . Then the  $IC_{50}$  of a mixture containing  $N_{\alpha}$  active compounds and  $N_{\beta}$  inactive compounds is given by

$$A_{\alpha,\beta} = \frac{N_{\alpha}\alpha + N_{\beta}\beta}{N_{\alpha} + N_{\beta}}$$

for the arithmetic mean model,

$$G_{\alpha,\beta} = 10^{(N_{\alpha}\log(\alpha) + N_{\beta}\log(\beta))/(N_{\alpha} + N_{\beta})}$$

for the geometric mean model, and

$$H_{\alpha,\beta} = \frac{N_{\alpha} + N_{\beta}}{\frac{N_{\alpha}}{\alpha} + \frac{N_{\beta}}{\beta}}$$

for the harmonic mean model. From these equations, one can evaluate the ability to detect active compounds in a mixture governed by the harmonic mean model, and compare it to the abilities of the arithmetic and geometric mean models.

For example, if a single highly active compound with an  $IC_{50}$  of 10 nM ( $\alpha = 10$  and  $N_{\alpha} = 1$ ) were in a mixture with a number of inactive compounds with  $IC_{50}$ 's of 10,000 nM ( $\beta = 10,000$ ), the aforementioned equations would yield a predicted value for the  $IC_{50}$  of the resultant mixture. Figure 3 is a plot of the number of inactive compounds versus the resulting activity of the mixture, given each of the averaging models. As is clear from the graph, the  $IC_{50}$  of the resultant mixture increases much more slowly for the harmonic mean than for the other two methodologies; therefore, mixtures that are modeled by harmonic mean averaging are projected to be significantly more active than if the arithmetic or

Table 3. Mixture Complexities and Prediction Model Errors

reference	mixture constituents	compounds per constituent	arithmetic scaled error	geometric scaled error	harmonic scaled error
Houghten et al. <sup>13</sup>	18	5832	3.37	2.31	0.48
	18	324	31.26	23.79	0.83
	18	18	299.51	200.63	0.55
	18	1	28.44	11.89	0.17
Dooley et. al. <sup>14</sup>	19	6859	0.62	0.11	0.60
	19	361	2.11	0.66	0.06
	19	19	3.27	1.34	0.52
	19	1	5.24	1.41	0.04
Dooley et. al. <sup>15</sup>	19	6859	4.54	1.96	0.46
	19	361	0.46	0.19	0.09
	19	19	1.99	1.18	0.42
	19	1	0.26	0.20	0.45
Pinilla et al. <sup>16</sup>	19	361	737.80	167.92	0.14
	19	361	9.67	6.79	3.71
	19	19	0.43	0.17	0.49
	19	1	0.38	0.17	0.04
Houghten et al. <sup>17</sup>	19	6859	16.90	5.83	0.32
	19	361	1,266.20	650.75	0.98
	19	19	373.78	45.10	0.01
	19	1	1.45	1.04	0.72
Appelet. al. <sup>18</sup>	19	6859	5.27	3.15	0.58
	19	361	113.65	27.48	0.21
	19	19	0.37	0.21	0.08
	19	1	0.16	0.01	0.11
	19	361	47.68	21.86	0.41
	19	19	2,061.38	916.03	0.72
	19	1	164.41	20.29	0.28
Davis et. al. <sup>19</sup>	12	144	1.75	1.45	1.03
	12	12	1.04	0.82	0.50
Ecker et al. <sup>20</sup>	4	4096	0.91	0.18	0.57
	4	1024	14.08	6.00	0.15
	4	256	3.13	1.82	0.57
	4	64	2.53	1.22	0.37
	4	16	0.65	0.48	0.26
	4	4	0.35	0.27	0.17
	4	1	0.38	0.17	0.02

geometric mean were used. A mixture comprised solely of poorly active compounds (all having an  $IC_{50}$  of 10,000 nM) would have an  $IC_{50}$  of 10,000 nM in this scenario, so in order for the active compound to be detected the resulting mixture must have an  $IC_{50}$  value small enough to distinguish itself from the 10,000 nM mixture inclusive of experimental error. For example, if the experimental error is approximately 20%, then to ensure the mixture containing the active compound is more active than the 10,000 nM mixture when tested, it must have a true  $IC_{50}$  value of 6,600 nM or below. The arithmetic mean model would predict that only three 10,000 nM compounds would be needed to make the mixture with the 10 nM compound have an  $IC_{50}$  greater than 6,600 nM. Similarly, the geometric mean model would predict that only seventeen 10,000 nM compounds would be needed. In contrast, the harmonic mean model indicates that 1,939 compounds, each having an activity of 10,000 nM, would be needed. Since these numbers are independent of scale, these results may be restated: Under the arithmetic mean model, a mixture containing less than 25.0% active compounds would

not be detected. Under the geometric mean model, a mixture containing less than 5.5% active compounds would not be detected. But under the harmonic mean model, a mixture containing 0.052% of active compounds would still be detectable. It should be noted that this observation can be applied to mixtures in which the constituents comprising the mixture are not present in the same concentration as in the case of natural product extracts. In other words as long as 0.052% of the total composition of the mixture contains an “active component” that mixture will be distinguishable from a totally inactive mixture sample.

It is also possible to use the above arguments in a reversed fashion. If we continue with the above example in which an inactive compound has an  $IC_{50}$  of 10,000 nM, and we assume the validity of the harmonic mean, then given a mixture activity it is possible to mathematically derive a range of activity percentages associated with different activity levels of individual compounds. Such relationships are plotted in Figure 4 for differing mixture activity values. A mixture with an  $IC_{50}$  of 1000 nM, for example,

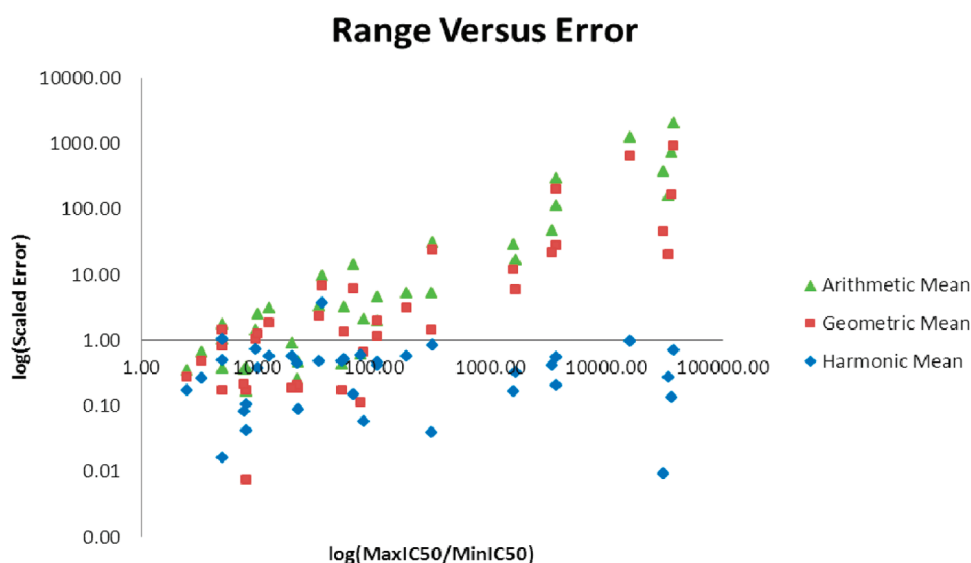


Figure 2

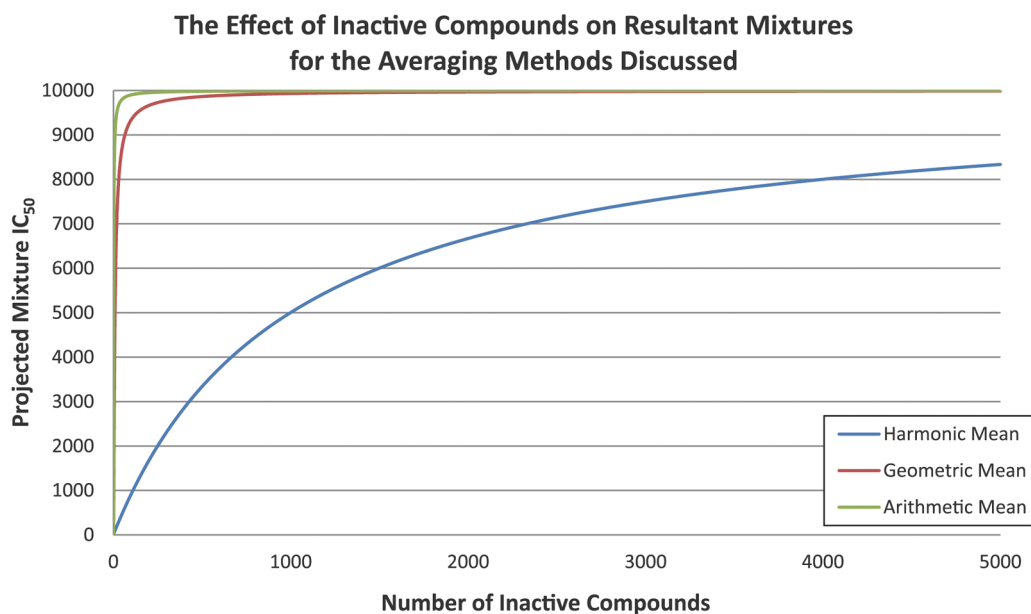


Figure 3

may contain 1% of individual compounds with  $IC_{50}$ 's of 10 nM, or 0.1% of individual compounds with  $IC_{50}$ 's of 1 nM. The maximal  $IC_{50}$  to guarantee detectability in this example is again 6600 nM, and so we can see such a mixture may contain 0.5% of individual compounds with  $IC_{50}$ 's of 100 nM, 0.05% of individual compounds with  $IC_{50}$ 's of 10 nM, or 0.005% of individual compounds with  $IC_{50}$ 's of 1 nM. Conversely, a mixture that is indistinguishable from inactive because its  $IC_{50}$  exceeds 6600 nM cannot even contain 0.5% of individual compounds with  $IC_{50}$  values of 100 nM, justifying the exclusion of such a mixture in further testing.

## DISCUSSION

The analyses presented here have significant implications on how mixtures can be used in drug discovery and, in part, explains the previous successes of research efforts where mixtures have

been used, such as natural product extracts and systematically arranged mixtures.<sup>2,3,13–23</sup> The active compounds isolated in various natural product studies<sup>24–27</sup> are often a very small percentage of the original material, and yet are still detectable. A clear distinction in the activity of mixtures containing highly active compounds and those that do not has been observed even in cases where the mixtures contained thousands of separate components. In the past, one of the reasons for this distinction has been posited as an abundance of similarly active compounds in a given mixture.<sup>28</sup> While this can be a valid statement in systematically arranged mixtures, it is surely not necessarily true in natural products, and to a large degree the reason for this distinction can rather be attributed to the harmonic meaning of the individual components. Indeed, the reactions occurring in living organisms mirror qualitatively this behavior, with only trace amounts of specific substances playing vital biological roles.<sup>29</sup>

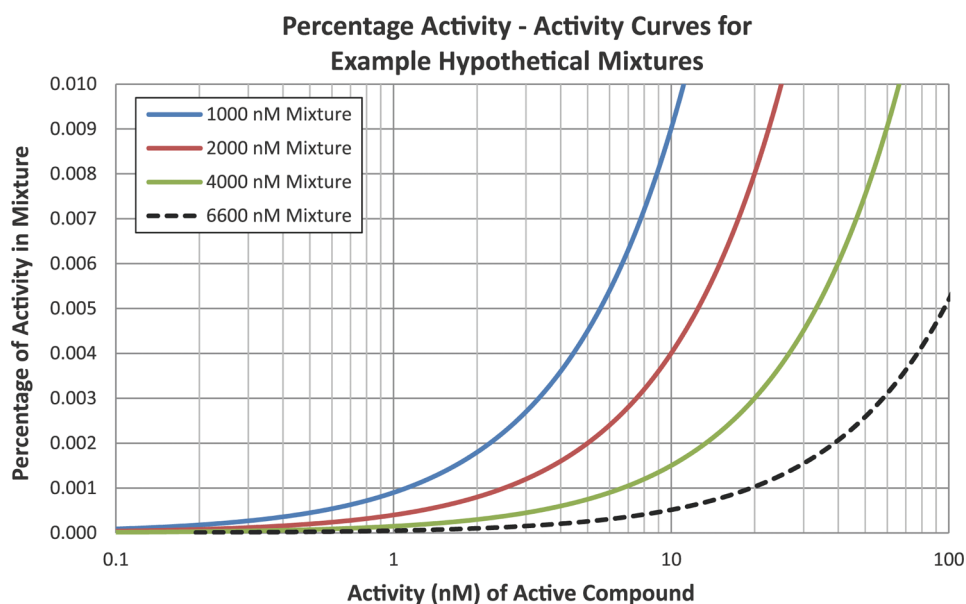


Figure 4

There are several additional applications of the harmonic mean currently being explored for future studies. As discussed above, in synthetic libraries arrayed in positional scanning format, the harmonic mean of the  $IC_{50}$  values at each position ought to equal one another; therefore, the integrity of the synthesis of a positional scanning library may be determined by comparing the harmonic means at each position and looking for significant deviation. A similar usage for the harmonic mean may be applied to pooling strategies. Algorithms could be developed for determining how to pool compounds in a given library in such a way as to identify large deviations from predicted harmonic means. Such deviations may indicate possible problems coming from pipetting errors or aggregation, and so these samples could be flagged for retesting. Finally, the harmonic mean could be used to determine if multiple compounds act in a simple competitive manner when tested in a given assay. If constituent compounds of known concentration are mixed, a substantial deviation of this mixture's activity from the expected harmonic mean could provide insights into possible synergy, antagonism, or multiple paths of action.

Although in this report the analysis was restricted to biological assays where a simple independent action is assumed, theory<sup>9,10</sup> suggests that, qualitatively, the harmonic mean's tendency to weight active compounds heavily is more widely applicable. The implications of other mathematical models, such as those for situations which are synergistic or anti-synergistic in nature, on the behavior of mixture-based combinatorial libraries associated with basic research and drug discovery are also currently being explored.

## ■ ASSOCIATED CONTENT

**S Supporting Information.** All data compiled in the study. This material is available free of charge via the Internet at <http://pubs.acs.org>.

## ■ AUTHOR INFORMATION

### Corresponding Author

\*E-mail: [houghten@tpims.org](mailto:houghten@tpims.org).

## ■ ACKNOWLEDGMENT

This work is supported in part by the State of Florida, Executive Office of the Governor's Office of Tourism, Trade, and Economic Development as well as NIDA R21 DA019620 and Multiple Sclerosis National Research Institute.

## ■ REFERENCES

- (1) Newman, D. J.; Cragg, G. M.; Snader, K. M. The influence of natural products upon drug discovery. *Nat. Prod. Rep.* **2000**, *17*, 215–234.
- (2) Nagasaki, H. The pharmacological properties of novel MCH1 receptor antagonist isolated from combinatorial libraries. *Eur. J. Pharmacol.* **2009**, *602*, 194–202.
- (3) Houghten, R. A.; Pinilla, C.; Giulianotti, M. A.; Appel, J. R.; Dooley, C. T.; Nefzi, A.; Ostresh, J. M.; Yu, Y.; Maggiora, G. M.; Medina-Franco, J. L.; Brunner, D.; Schneider, J. Strategies for the use of mixture-based synthetic combinatorial libraries: Scaffold ranking, direct testing, in vivo, and enhanced deconvolution by computational methods. *J. Comb. Chem.* **2008**, *10*, 3–19.
- (4) Kainkaryam, R. M.; Woolf, P. J. Pooling in high-throughput drug screening. *Curr. Opin. Drug Discovery Dev.* **2009**, *12*, 339–350.
- (5) Wolf, K. K.; Vora, S.; Webster, L. O.; Generaux, G. T.; Polli, J. W.; Brouwer, K. L. Use of cassette dosing in sandwich-cultured rat and human hepatocytes to identify drugs that inhibit bile acid transport. *Toxicol. in Vitro* **2010**, *24*, 297–309.
- (6) Smith, N. F.; Raynaud, F. I.; Workman, P. The application of cassette dosing for pharmacokinetic screening in small-molecule cancer drug discovery. *Mol. Cancer Ther.* **2007**, *6*, 428–440.
- (7) Finney, D. J. *Probit Analysis*, 3rd ed.; Cambridge University Press: Cambridge, 1971.
- (8) Smyth, H. F., Jr.; Weil, C. S.; Carpenter, C. P. An exploration of joint toxic action: twenty-seven industrial chemicals intubated in rats in all possible pairs. *Toxicol. Appl. Pharmacol.* **1969**, *14*, 340–347.
- (9) Hoel, D. G. Statistical Aspects of Chemical Mixtures. In *Methods for Assessing the Effects of Mixtures of Chemicals*; Vouk, V. B., Butler, G. C., Upton, A. C., Parke, D. V., Asher, S. C., Eds.; Wiley: New York, 1987; pp 369–377.
- (10) Talalay, P.; Chou, T. Quantitative analysis of dose-effect relationships: The combined effects of multiple drugs or enzyme inhibitors. In *Advances in Enzyme Regulation*; Webber, G., Ed.; Pergamon: New York, 1984; Vol. 22, pp 27–55.

(11) Feng, B. Y.; Shiochet, B. K. Synergy and antagonism of promiscuous inhibition in multiple-compound mixtures. *J. Med. Chem.* **2006**, *49*, 2151–2154.

(12) Frier, S. M.; Konings, D. A. M.; Wyatt, J. R.; Ecker, D. J. Deconvolution of combinatorial libraries for drug discovery: a model system. *J. Med. Chem.* **1995**, *38*, 344–352.

(13) Houghten, R. A.; Pinilla, C.; Blondelle, S. E.; Appel, J. R.; Dooley, C. T.; Cuervo, J. H. Generation and use of synthetic peptide combinatorial libraries for basic research and drug discovery. *Nature* **1991**, *354*, 84–86.

(14) Dooley, C. T.; Chung, N. N.; Wilkes, B. C.; Schiller, P. W.; Bidlack, J. M.; Pasternak, G. W.; Houghten, R. A. An all D-amino acid opioid peptide with central analgesic activity from a combinatorial library. *Science* **1994**, *266*, 2019–2022.

(15) Dooley, C. T.; Chung, N. N.; Schiller, P. W.; Houghten, R. A. Acetalins: Opioid receptor antagonists determined through the use of synthetic peptide combinatorial libraries. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 10811–10815.

(16) Pinilla, C.; Appel, J. R.; Houghten, R. A. Synthetic peptide combinatorial libraries (SPCLs): identification of the antigenic determinant of  $\beta$ -endorphin recognized by monoclonal antibody 3E7. *Gene* **1993**, *128*, 71–76.

(17) Houghten, R. A.; Appel, J. R.; Blondelle, S. E.; Cuervo, C. T.; Dooley, C. T.; Pinilla, C. The use of synthetic peptide combinatorial libraries for the identification of bioactive peptides. *Biotechniques* **1992**, *13*, 412–421.

(18) Appel, J. R.; Pinilla, C.; Houghten, R. A. Identification of related peptides recognized by a monoclonal antibody using a synthetic peptide combinatorial library. *Immunomethods* **1992**, *1*, 17–23.

(19) Davis, P. W.; Vickers, T. A.; Wilson-Lingardo, L.; Wyatt, J. R.; Guinasso, C. J.; Sanghvi, Y. S.; DeBaets, E. A.; Acevedo, O. L.; Cook, P. D.; Ecker, D. J. Drug leads from combinatorial phosphodiester libraries. *J. Med. Chem.* **1995**, *38*, 4363–4366.

(20) Ecker, D. J.; Vickers, T. A.; Hanecak, R.; Driver, V.; Anderson, K. Rational screening of oligonucleotide combinatorial libraries for drug discovery. *Nucleic Acids Res.* **1993**, *21*, 1853–1856.

(21) Dooley, C. T.; Houghten, R. A. The use of positional scanning synthetic peptide combinatorial libraries for the rapid determination of opioid receptor ligands. *Life Sci.* **1993**, *52*, 1509–1517.

(22) Li, J. W. H.; Vederas, J. C. Drug Discovery and Natural Products: End of an Era or an Endless Frontier? *Science* **2009**, *325*, 161–165.

(23) Butler, M. S. The role of natural product chemistry in drug discovery. *J. Nat. Prod.* **2004**, *67*, 2141–2153.

(24) Cantrell, C. L.; Richeimer, S. L.; Nicholas, G. M.; Schmidt, B. K.; Bailey, D. T. seco-Hinokiol, a new abietane diterpenoid from *Rosmarines officinalis*. *J. Nat. Prod.* **2005**, *68*, 98–100.

(25) Nicholas, G. M.; Molinski, T. F. Enantiodivergent biosynthesis of the dimeric sphingolipid oceanapiside from the marine sponge *Oceanapia phillipensis*. Determination of remote stereochemistry. *J. Am. Chem. Soc.* **2000**, *122*, 4011–4019.

(26) Marquez, B.; Verdier-Pinard, P.; Hamel, E.; Gerwick, W. H. Curacin D: an antimetabolic agent from the marine cyanobacterium *Lyngbya majuscula*. *Phytochemistry* **1998**, *49*, 2387–2389.

(27) Gerard, J.; Lloyd, R.; Barsby, T.; Haden, P.; Kelly, M. T.; Andersen, R. J. Massetolides A-H, antimicrobial cyclic depsipeptides produced by two pseudomonads isolated from marine habitats. *J. Nat. Prod.* **1997**, *60*, 223–229.

(28) Terrett, N. K.; Gardner, M.; Gordon, D. W.; Kobylecki, R. J.; Steele, J. Combinatorial synthesis — The design of compound libraries and their application to drug discovery. *Tetrahedron* **1995**, *51*, 8135–8173.

(29) Uppsala University Hospital (“Laborationslista”). Artnr 40284 Sj74a. Issued on April 22, 2008.

## NOTE ADDED AFTER ASAP PUBLICATION

There was an error in the equation describing harmonic mean in the version of this paper published March 11, 2011. The correct version published March 25, 2011.