DISCUSSION DRAFT

Dose – Response Relationship Characterization in Current Drug Development: *Do we have a Problem?*

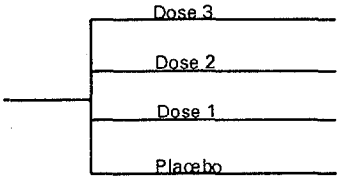
Part I: Inferences from animal and human data

Joga Gobburu, Ph.D. and Raymond Lipicky, M.D.
Division of Cardio-Renal Drug Products,
CDER, FDA

Introduction

Historically, dose (concentration)-response relationships have occupied a prominent role in antihypertensive drug development. Data from such trials (which some might classically classify as Phase 2) have been the primary basis of approval of all recently approved antihypertensive drugs.

Trials have, primarily, been double-blind, randomized, placebo-controlled, parallel–group in design, as depicted in the following diagram. Typically, such trials are of four to eight weeks in duration. The last office, cuff blood pressure measurements prior to randomization are taken as the "baseline", and the last (or approximately the last) blood pressures measured during the double-blind period are "the endpoint". Blood pressure for the randomized group is calculated (arithmetic mean or geometric mean) at "baseline" and again at "endpoint". The delta (baseline – endpoint) for each group are then calculated and the placebo group's delta is subtracted from each drug treated group (the delta of deltas) to represent the treatment effect. The measurements of blood pressure are those obtained just prior to the first dose of a day (medication is withheld until blood pressure has been measured, on the day of measurement). This measurement has historically, been called the trough blood pressure effect.



The purpose of such trials, methods of analysis, frameworks of reference that should surround the choice of methods of analysis, and inferences that one take from results, have not been the subject of much rigorous discussion. In order to provide a framework for beginning such discussion we have abstracted data from ten NDAs that were approved an antihypertensive indication for two classes of antihypertensive drugs. Namely angiotensin converting enzyme inhibitors (ACE inhibitors) and angiotensin II blocking agents (A2 blockers). No data was excluded, all data derived from all trials of the above were included (the analyses are therefore across trials).

Animal data should, we think, offer some basis for evolving a framework of reference; consequently, we have also abstracted the "best" description of dose (concentration)-response that

was submitted for effects on blood pressure. We arbitrarily limited such data to rat models of hypertension, primarily SHR hypertensive rats. We have not included other data from other animal experimentation.

All are clearly antihypertensive agents. The question is not, are these drugs antihypertensive? The question is, how adequately have the dose (concentration)-response relationships been identified?

Methods

Data

Dose, mean arterial blood pressure (MAP) in rats, and dose, diastolic blood pressure (DBP), and systolic blood pressure (SBP) in man were the data analyzed. All blood pressure data were group means (placebo subtracted in man, control subtracted in rats). All data from all trials were used, with no attention paid to how many data points came from what trial. This is the crudest of crude analyses.

The Problem Being Addressed

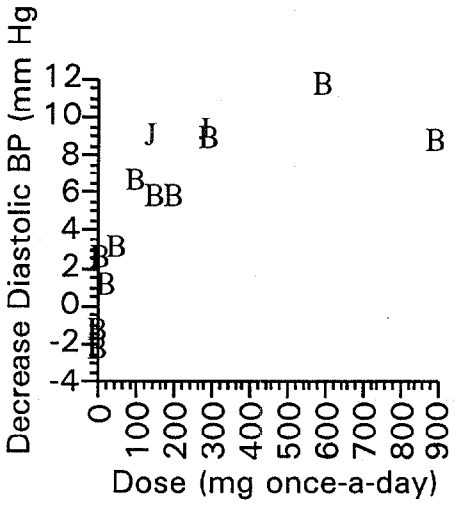
Generally, a table like the following is obtained where the dose for a group and the mean (placebo-corrected) response of that group is tabulated. Given sample sizes and variance (including the placebo group change from baseline), that is all statisticians need to decide if a given group had a response different from placebo, or whether one group differs from another group. Standard statistical methods (ANOVA, paired t, repeated measures, etc.) do this very well.

Dose Mg Once a Day	Drug Effect Diastolic Blood Pressure mm Hg Decrease
1.0	-2.2
5.0	-1.2
10.0	2.6
25.0	1.3
50.0	3.2
100.0	6.7
150.0	5.9
200.0	5.9
300.0	9.0
600.0	11.8
900.0	8.9

So, one can tell whether the drug was placebo or not, and one can also pretty well conclude that the drug had dose-related effects, the larger the dose the greater the effect.

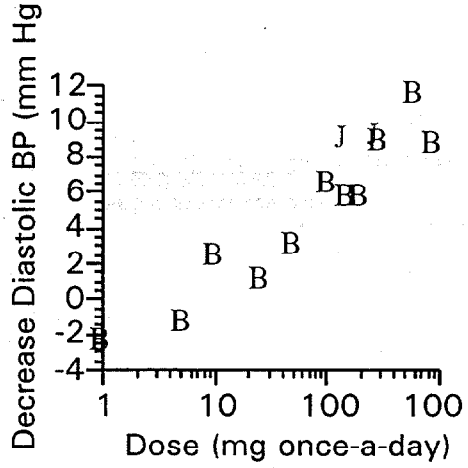
Many pages of figures follow this page. Should we be satisfied with the conclusions drawn by an inspection of the above table and standard statistical methods, the remaining pages would neither be necessary to read, nor to try to understand. This meeting is to discuss whether we should have stopped here.

For antihypertensive drugs, most persons have an expectation that a drug's effects should monotonically increase as a function of dose, and that there should be some maximum effect that cannot be exceeded, irrespective of how large the dose. That is, most persons have some model as basis for the expectation. That model is often used to make a prediction that is, in fact, outside of the limits of the data.



Often, one will not be satisfied with inspecting a table and concluding that it is not a table of random numbers and, in so doing, will look at a plot of the data. Two such plots are shown here. The plot immediately to the left plots the data from the above table and displays it using a linear scale for the X-axis. This plot certainly confirms the model described above. Namely, that effect on diastolic pressure increases monotonically and appears to reach a maximum effect. From this plot, one could reasonably expect that the next greater dose (say 1200 mg) would not produce a greater effect and "intuitively" conclude that the data support their notion of how that drug affects diastolic blood pressure as a function of dose.

The following plot of the same data, simply uses a log scale for the x-axis. This view of the data violates the notion that a maximum effect has been reached within the ranges of the doses explored, and that a greater dose (say 1200 mg) could reasonably be expected to produce a bigger effect on diastolic blood pressure. A very different conclusion.



Those familiar with modeling recognize that the tools needed to rule-in or rule-out classes of models are readily available. Therefore, the exercise shown in the rest of this document

represents an initial attempt to rule-in or rule-out classes of simplistic models from the data provided by recent, modern development programs.

Models

Dose – response data for all the drugs were fitted to Emax, log-linear, and linear models. The mathematical forms of the models are given below:

Emax model: $$E = E(0) + \frac{E\,max \cdot Dose}{Dose + ED_{50}} \quad (1)$$

Log-linear model: $$E = E(1) + Slope \cdot \log(Dose) \quad (2)$$

Linear model: $$E = E(0) + Slope \cdot Dose \quad (3)$$

Where, E is the effect (SBP, DBP, MAP), 'E(0)' is the effect when no drug is present (except for the log-linear model for which it is the effect when 1 unit dose is given), 'Slope' is the slope of the straight line describing the relationship between dose (or log(dose)) and effect, Emax is the maximal effect that can be elicited and ED_{50} is the dose required for half-maximal effect.
Fitting and simulation was performed using NONMEM (UCSF, ver 5.0, level 1.0).

The rules for deciding whether a model "fits" the data, or conversely, whether the data is sufficient to allow a model's appropriateness for attempting to summarize the data, are complex and could take days of discussion. We have chosen to limit considerations to several.

1) The residuals of the fit, that is how close do the predicted values calculated from the function come to the observed values. How much of the apparent relationship of the Y values to their respective X values is explained by the relationship (linear, log-linear, M_{ax}) being fitted. This is expressed numerically by the objective function and by the AIC. The closer the predictions come to the observed, the better. So, smaller numbers for these values indicate a better fit, thus one might conclude, on a mathematical basis that one model is better than another.
2) From the accuracy of predictions outside of the paired values used for the fit. For example, at zero dose there should be zero effect, so how well does the function used predict zero effect at zero dose, how well does it predict the maximum effect at a much larger dose than included in the fit, etc. Bias (%) is a formal means of estimating these notions.
3) For parameter estimates, say the ED_{50} for the E_{Max} model, how good is the estimate. For example if the estimate is 100 mg and the standard error of that estimate is 200 mg, one knows the "model" has very little predictive value.
4) Finally, visually inspecting the data points and the line generated by the fit.

Unfortunately, this results in many pages and many plots.

The objective function value calculated by NONMEM can be used for selection of an appropriate model, based on statistical reasoning. The Akaike information criteria (AIC) was determined for each of the model fittings using:

$$AIC = NOBS * LOG(WRSS) + 2 * NPARM \quad (4)$$

Where, NOBS is the number of observations, WRSS is the weighted residual sum of squares and NPARM is the number of model parameters. For both the objective function value and AIC, the lower the value the better.

The bias in the model predictions was calculated using the following formula:

$$\text{Bias}(\%) = \left[\frac{\text{predicted} - \text{observed}}{\text{observed}}\right] \cdot 100 \quad (5)$$

Results and Discussion

Shape of the dose – response data

Humans

The dose ranges studied are variable from between on order of magnitude and three orders of magnitude, as shown in Figures 1 and 2. Also shown in Figures1 and 2 are the Emax model predictions overlaid on the observed data. The shape of the dose – response curve appears different in Figures 3 and 4 than in Figures 1 and 2, where the dose axis is plotted on a logarithmic scale. For example, for drug F doses higher than 80 mg can be argued to have reached a plateau (apparently) from Figure 1. However, from Figure 2 it will be evident that the log(dose) – response is almost a perfect straight line. Figures 5, 6, 7 and 8 show the same observed and linear or log-linear model fits – response data. The estimates of the pharmacodynamic parameters are presented in Table 1.

Based on Figure 1 and 3, only drug C seems to barely demonstrate the maximal effect, probably due to the wider range of the doses studied (1 – 1000 mg). Visual inspection of the fittings suggests that the Emax model does not offer any great advantage over the linear or log-linear models. More quantitative evaluation comes to the same conclusion.

Rats

Figures 25 and 26 show the observed and Emax model fits overlaid on the observed data. All the above discussion for the human data is also pertinent to the rat data. Figures 27 and 28 show the observed and linear or log-linear model predicted dose – response data. The estimates of the pharmacodynamic parameters are presented in Table 1.

Smallest effect predicted by the models

Humans

The value E(0) in the case of the Emax and Linear models was employed as the smallest effect that the drug can produce. In the case of Log-linear model, E(0) was determined assuming a small dose of 0.0001 mg was given. The point #2 in the earlier section that deals with the extrapolation ability of log – linear models is pertinent here. Figure 9 shows the smallest effects as predicted by the 3 models, as histograms. Both Emax and Linear models predict effects closer to zero (but

about ± 5 mm Hg), while the Log – linear model predicts negative effects always. In some cases, the smallest effect is low as – 25 mm Hg (increasing the blood pressure).

Rats

Figure 29 shows the smallest effect that the drug can produce. The smallest effect predicted by the Emax and Linear models is closer to zero, but much higher (for many drugs as high as 25 mm Hg) than that predicted in humans. The Log – linear model predicted smallest effect is unrealistically low. This trend suggests that the slopes for the effects in the rats are much steeper than that in humans. Partly this could be (an artifact) due to the narrower dose range studied in rats.

Largest effect predicted by the models

Humans

Linear and Log – linear models have no upper bound, hence the dose required to produce 99% of the Emax was used to predict the effect using the Linear and Log – linear models and their parameters (Table 1). Figure 10 and 11 show the largest effect on SBP and DBP respectively, as predicted by each of the models. As expected, the Linear and Log – Linear models produce effects that are unrealistic, demonstrating their incapability to extrapolate. For drug B, all the models predict biologically implausible effects. The fact that prediction of the 99% effect of Emax itself could be erroneous.

Rats

The model predicted largest effect in rats follows similar trend as for the humans. All the earlier discussion is valid for rats too.

Relative goodness-of-fits

Objective function value and AIC

Humans/Rats

Figures 13 and 14 show the histograms of objective function values for each of the 3 models for effects in humans. Figure 31 shows the histograms of the objective function values for the 3 models in rats. Figure 15 and 16 show the AIC values for the 3 models for effects in humans and Figure 32 shows these values for rats. The smaller the objective function value the better the model. There is no clear superiority of Emax model over Linear or Log – linear models, suggesting that the dose range studied for most drugs is inadequate. Emax has one extra parameter than the other models. Table 1 provides the estimates of the pharmacodynamic parameters for the Emax, Log-linear, and Linear models. The Emax model has one parameter more than the Log-linear and Linear models. The difference in the objective function values between two given nested models is believed to follow a chi-squared distribution. Based on that distribution, a difference of about 10 in the objective function value is believed to reflect a $p<0.001$. However, it should be noted these 3

models are not 'nested'. Nevertheless, the objective function value can still be used to compare the goodness-of-fits as it is a function of the SSE. Most times Log-linear model was chosen based on the objective function value differences and/or AIC values. Although unexpected, for the rat data, Emax model almost always seems to offer superior fit to the data, based on the AIC values. This is different from the conclusion one would arrive at based on the objective function values, which suggest no difference among the 3 models (Figure 31). Such an observation calls for attention regarding the selection of doses and the method of assessment whether a 'true' maximal response is reached. Modeling dose – response data could be an alternative to a statistical test comparing the responses at various doses as 'groups'.

Bias

Humans/Rats

Figures 17 through 24 and Figures 33 through 36 depict the bias in the predictions of the effect by the 3 models. The bias was calculated according to equation 5 and expressed as %. For the data from humans, bias at the lower, middle, and upper portions of the dose – response curve do not allow discerning the superiority of any model over the other. However for the data from rats, there is seems to be slight hint of higher bias in the predictions by the linear and Log – linear models, particularly at the lower and upper ends of the curve.

General

Overall, it can be concluded from the above analyses:

1. Plotting dose – response data on a logarithmic dose axis gives a different appreciation of the curvature and seems to require selection of an appropriate model for data analysis.
2. Both animal and human data indicate that almost none of the drugs analyzed demonstrate maximal effect at the highest dose studied. Further, in most cases Log-linear and Linear dose – effect models provided adequate description of the data.
3. It is also to be noted that the data used for these analyses are average data per dose group. Ideally all the data collected in the experiment for all study subjects/animals should be used. Unfortunately, such data were not available. Further, the total daily dose is used as the independent variable as opposed to, say, concentration to reflect differences in the dosing regimens (e.g.: 100 mg/day vs. 50 mg bid). The models, ideally, should also reflect the current understanding about the in vivo characteristics of a given drug.
4. Nevertheless, the above analyses suggest that the clinical trials or experiments are not designed adequately in order to obtain the full range of the dose – response curves.

Table 1. Summary of pharmacodynamic parameters of various drugs given to humans estimated using Emax, Linear and Log-linear models.

			Emax Model				Linear Model			Log-Linear Model			
Drug	BP		E(0)	ED50	Emax	OBJ[c]	E(0)	Slope	OBJ.FN.	E(1)[d]	Slope	E(0)	OBJ
A	SBP	PE[a]	2.77	10.4	6.24	-13.55	4.22	0.116	-2.394	2.72	1.36	1.72	-13.472
		SE[b]	0	0.1	0		0	0		0	0.1		
A	DBP	PE	2.66	26.4	9.11	18.99	3.38	0.149	20.937	1.98	1.49	0.98	20.472
		SE	0.2	0.7	0.2		0	0		0.1	0.1		
B	DBP	PE	5.25	2340	353	3.742	5.25	0.149	3.678	4.13	1.54	3.13	8.132
		SE	0	50	49		0	0		0	0		
C	DBP	PE	0	103	11.8	23.4	2.95	0.0115	41.508	0	1.4	-1	31.149
		SE	5.00E+13	52	9		0.1	0		4.00E+13	11		
D	SBP	PE	1.06	22.1	4.92	21.9	3.39	0.0159	22.994	0.41	1.02	-0.59	22.065
		SE	0.1	0.1	0		0	0.1		0.1	0		
D	DBP	PE	0	8.93	9.41	33.3	7.13	0.0095	36.49	2.69	1.29	1.69	34.85
		SE	3.00E+13	62	51		0	0		0.2	0.1		
E	SBP	PE	3.51	2690	1040	50.7	3.55	0.346	50.747	0	3.6	-1	37.722
		SE	NE[e]	NE	NE		0	0		2.00E+13	77		
E	DBP	PE	0	14.6	11.6	36.06	2.22	0.202	48.783	0	2.14	-1	36.53
		SE	9.00E+10	84	7.4		0.1	0.1		1.00E+13	48		
F	DBP	PE	5.28	45.5	15.7	37.769	7.56	0.0891	42.006	1.91	2.99	0.91	39.35
		SE	0	0	0		0	0		0	0		
G	SBP	PE	1.18	88.8	47	17.7	1.29	0.468	17.622	1.04	1.93	0.04	18.748
		SE	5	50	44		0	0.1		0	0		
G	DBP	PE	0	1.59	4.81	-2.268	2.43	0.187	1.72	1.94	1.03	0.94	-1.162
		SE	5.00E+10	80	36		0	0		0.1	0		
H	SBP	PE	1.45	131	10.6	21.213	1.79	0.0484	21.395	0	1.1	-1	22.149
		SE	0.1	0.1	0.1		0.1	0.2		6.00E+10	13		
H	DBP	PE	2.74	245	12.3	9.013	2.88	0.0375	9.122	1.22	0.928	0.22	11.044
		SE	0	0.1	0.1		0	0		0	0		
I	SBP	PE	0	25	14.8	2.418	7.22	0.0364	9.042	0	2.49	-1	5.437
		SE	2.00E+10	102	41		0	0		1.00E+13	34		
I	DBP	PE	0	6.67	9.48	-1.97	7.68	0.0084	2.433	4.86	0.846	3.86	0.342
		SE	1.00E+13	50	38		0	0.1		0.1	0		
J	SBP	PE	0	22.8	9.22	4.525	4.35	0.0165	11.471	0	1.53	-1	6.511
		SE	2.00E+10	22	44		0.1	0.2		1.00E+10	28		
J	DBP	PE	0.4	67	6.84	-1.811	2.05	0.0147	3.105	0	0.936	-1	3.062
		SE	0.3	0.1	0		0	0		NE	NE		

[a] PE = point estimate; [b] SE = standard error (%); [c] OBJ=objective function value; [d] E(1)=effect when dose=1

[e] NE = not estimable

Note: Almost all analyses resulted in a warning "S matrix algorithmically singular"

Table 2. Summary of pharmacodynamic parameters of various drugs given to rats estimated using Emax, Linear and Log-linear models.

		Emax Model				Linear Model			Log-Linear Model			
Drug		**E(0)**	**ED50**	**Emax**	**OBJ**	**E(0)**	**Slope**	**OBJ**	**E1**	**Slope**	**E(0)**	**OBJ**
A	**PE**	10.4	0.0415	31.7	2.7	28.8	1.52	20.6	37.3	3.73	36.3	13.7
	SE	0.1	0.1	0		0.1	0.4		0.1	0		
B	**PE**	34.6	1.83	46.6	-19.0	41.9	3.26	12.1	54	8.04	53	7.5
	SE	NE	NE	NE		0	0		0	0		
C	**PE**	15.8	1110	1920	9.5	15.9	1.7	9.5	15.2	7.57	14.2	3.1
	SE	NE	NE	NE		0	0		0.1	0		
E	**PE**	17.1	2.49	54.9	11.5	33.7	1.34	28.0	37.4	9.08	36.4	9.9
	SE	0	0	0		0.1	0.2		0	0.1		
G	**PE**	0	3.56	70.9	-0.8	35	1	13.6	18.7	13.5	17.7	7.9
	SE	NE	NE	NE		0.2	0.2		0.1	0		
I	**PE**	11.3	120	626	3.6	11.4	5.08	3.4	18.8	5.56	17.8	9.0
	SE	0.2	50	49		0	0		0.1	0.2		
J	**PE**	0	5.02	115	14.4	15.7	6.27	17.6	13.1	28.1	12.1	11.7
	SE	1.00E+13	1.8	0.1		0	0		0.1	0		

[a] PE = point estimate; [b] SE = standard error (%); [c] OBJ=objective function value; [d] E(1)=effect when dose=1; [e] NE = not estimable

Note: Almost all analyses resulted in a warning "S matrix algorithmically singular"

Figure 1. Observed and Emax model fitted dose – DBP relationships (Linear x-axis) (Humans).

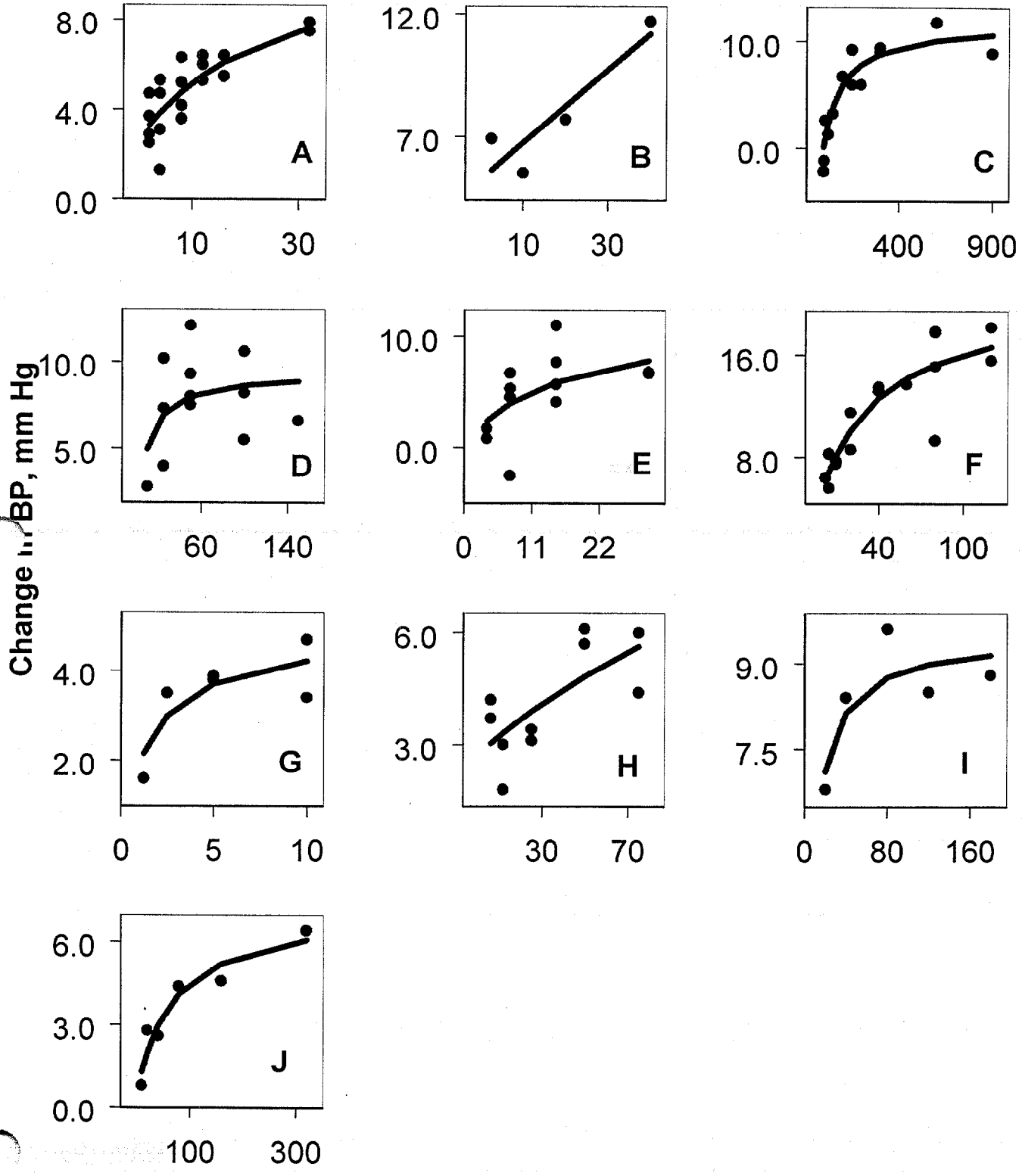


Figure 2. Observed and Emax model fitted dose – SBP relationships (Linear x-axis) (Humans).

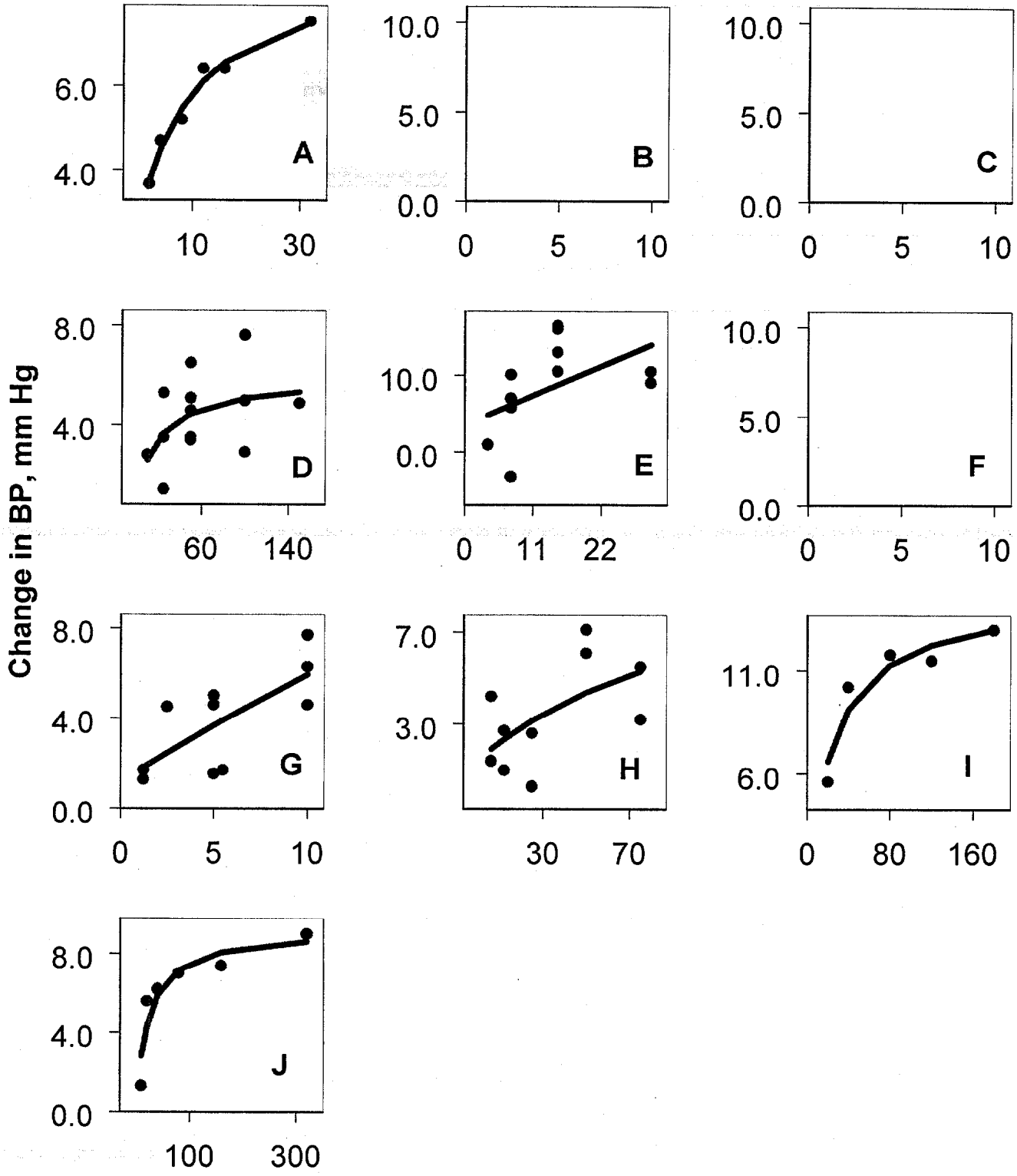


Figure 3. Observed and Emax model fitted dose – DBP relationships (Log x-axis) (Humans).

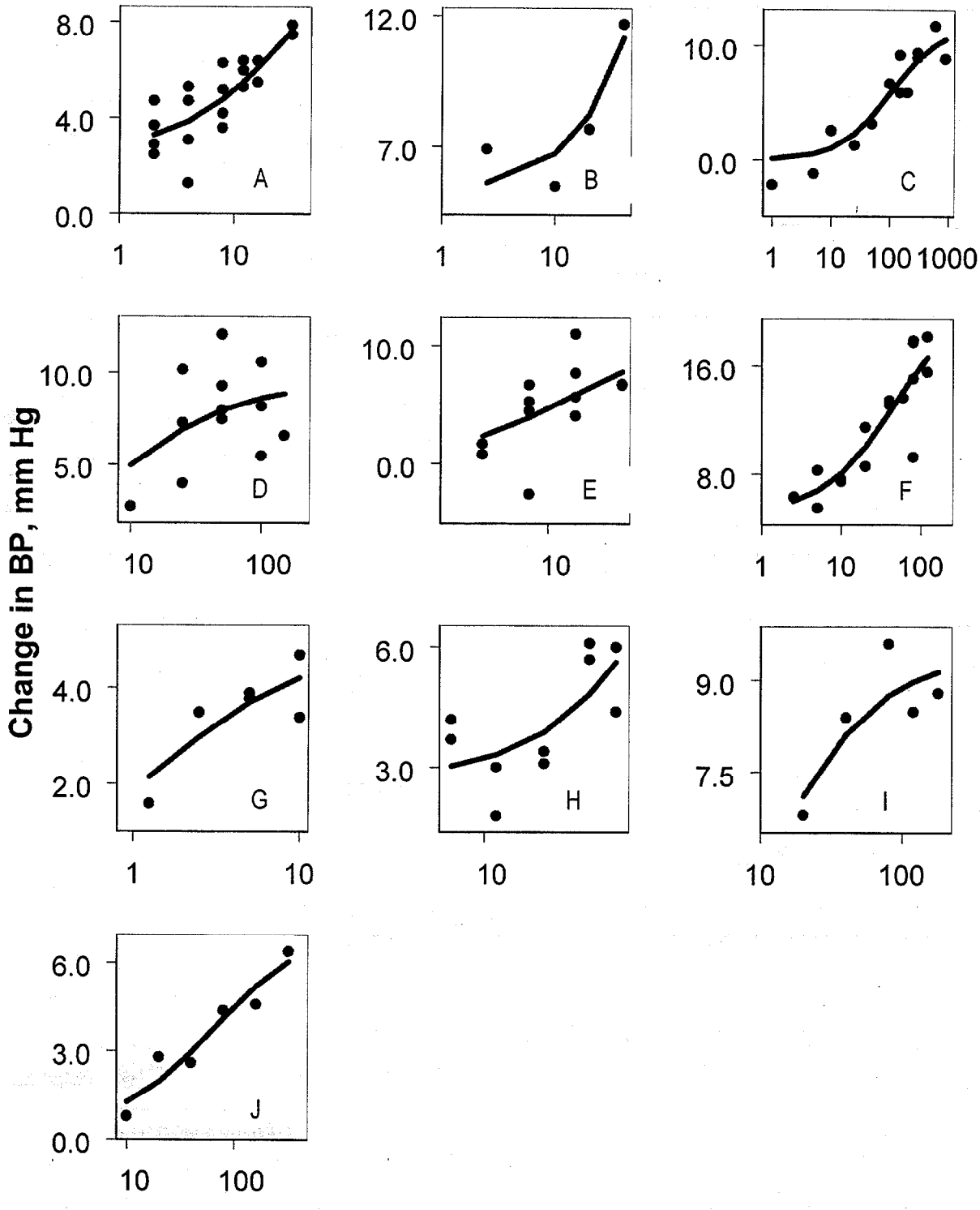


Figure 4. Observed and Emax model fitted dose – SBP relationships (Log x-axis) (Humans).

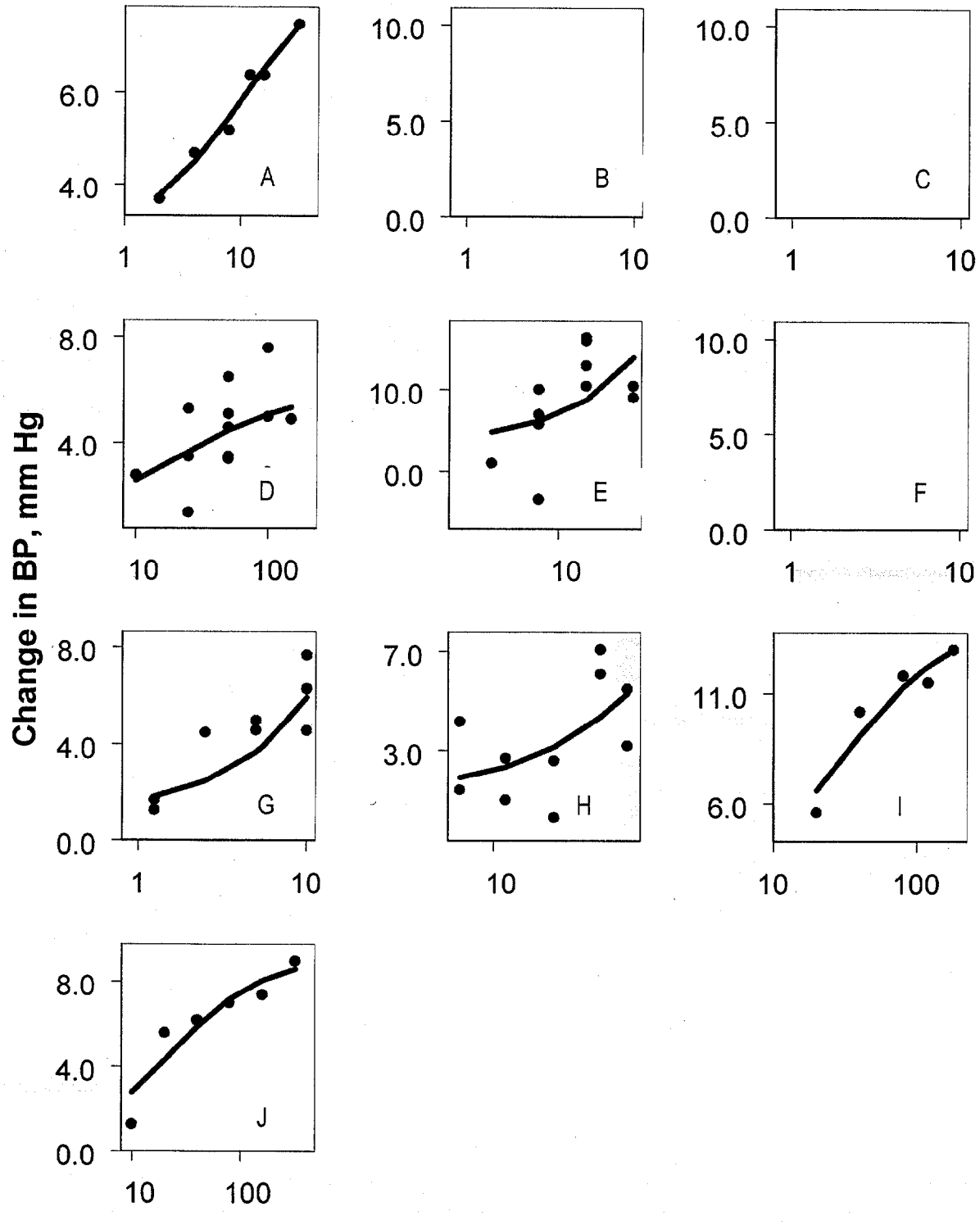


Figure 5. Observed and linear model fitted dose – DBP relationships (Humans).

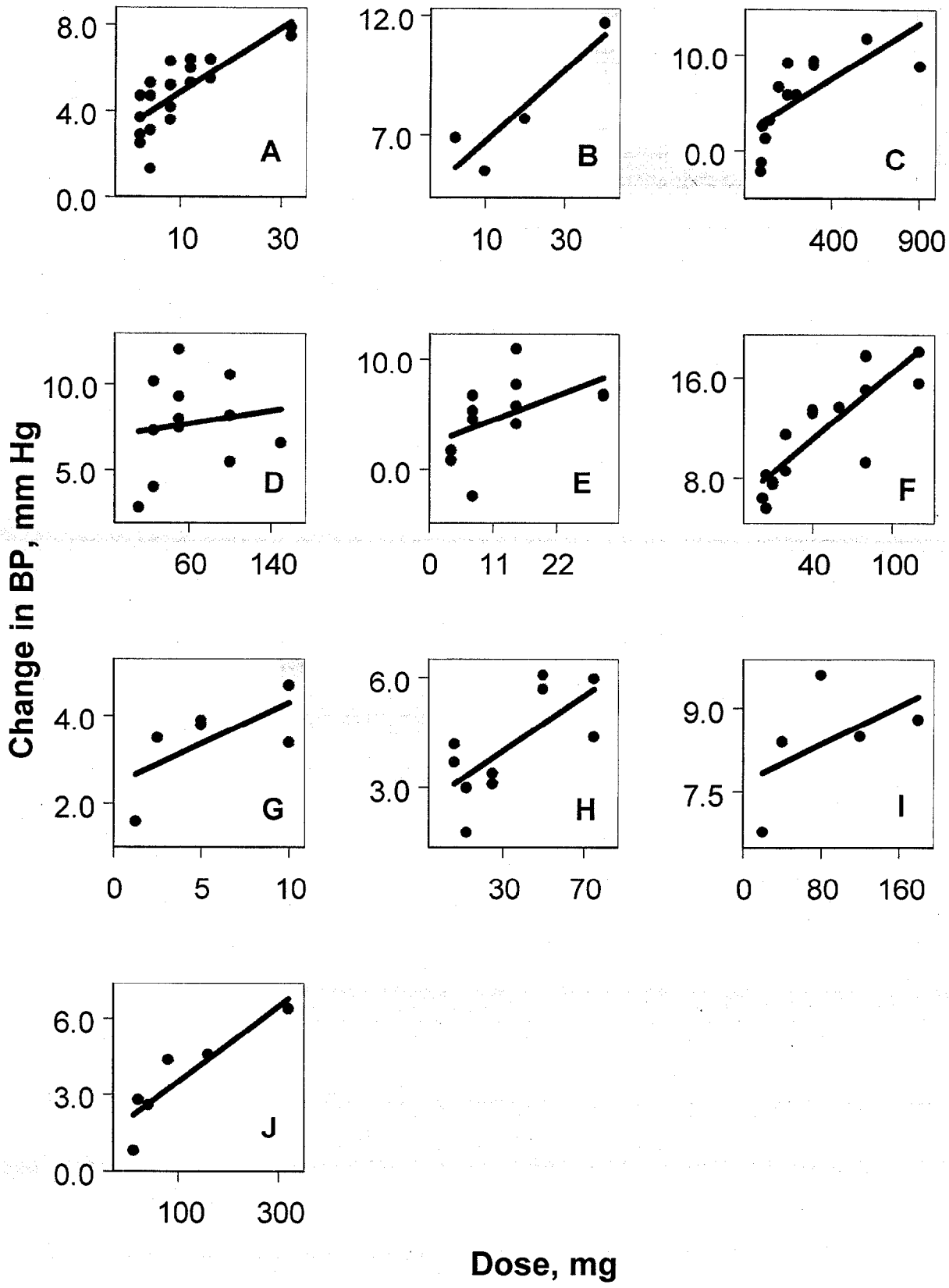


Figure 6. Observed and linear model fitted dose – SBP relationships (Humans).

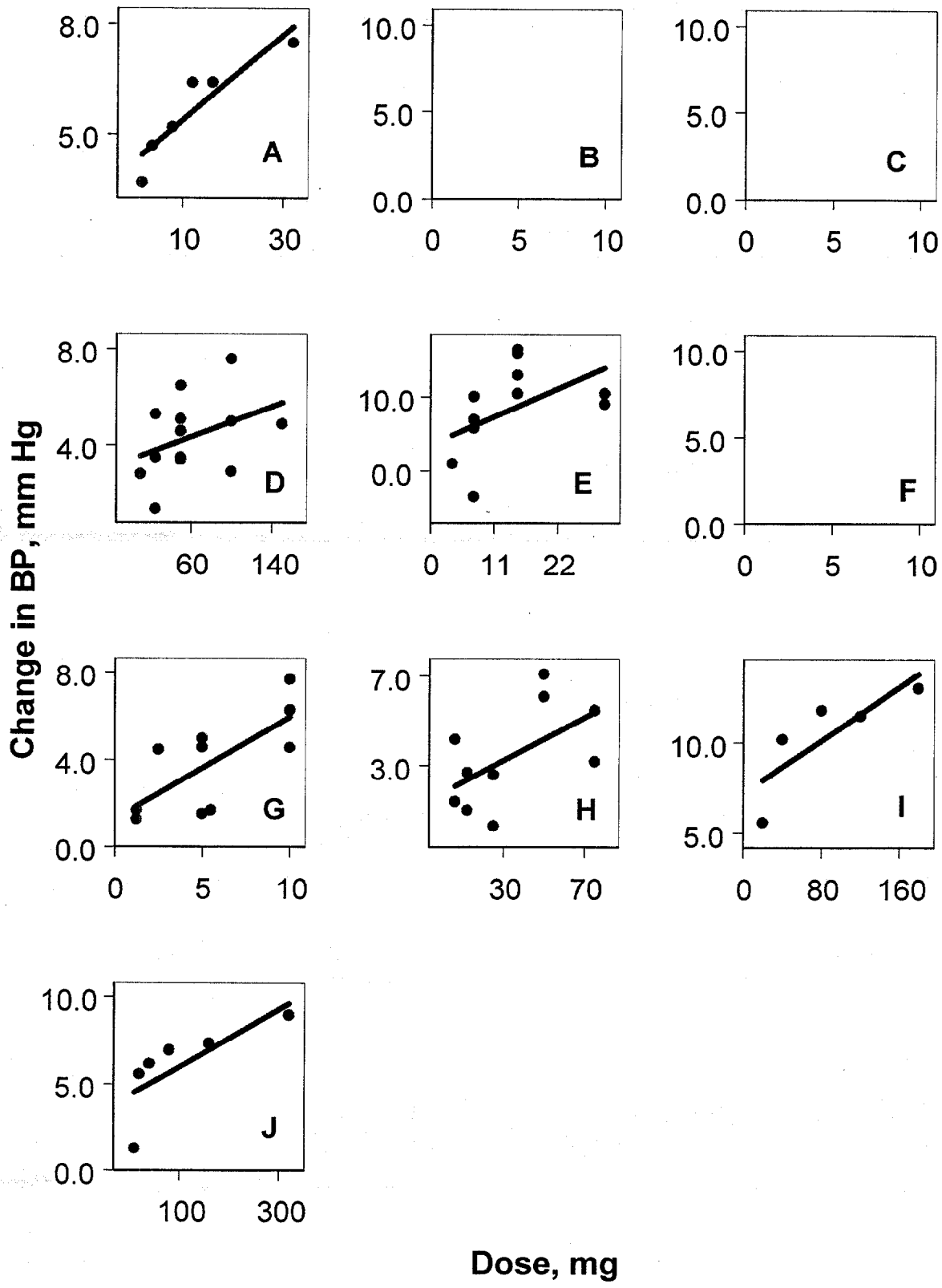


Figure 7. Observed and log-linear model predicted dose – DBP relationships (Humans).

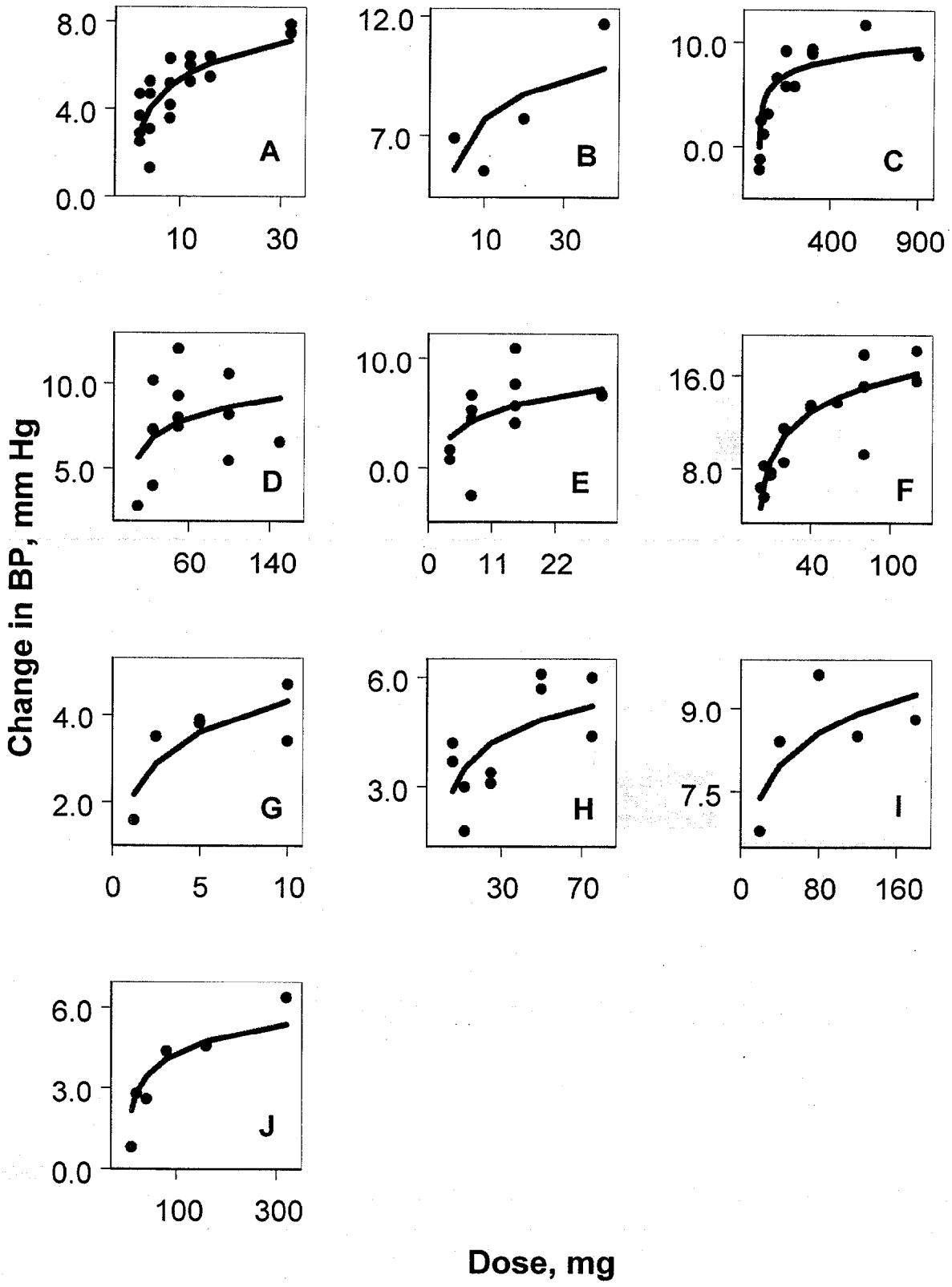


Figure 8. Observed and log-linear model predicted dose – SBP relationships (Humans).

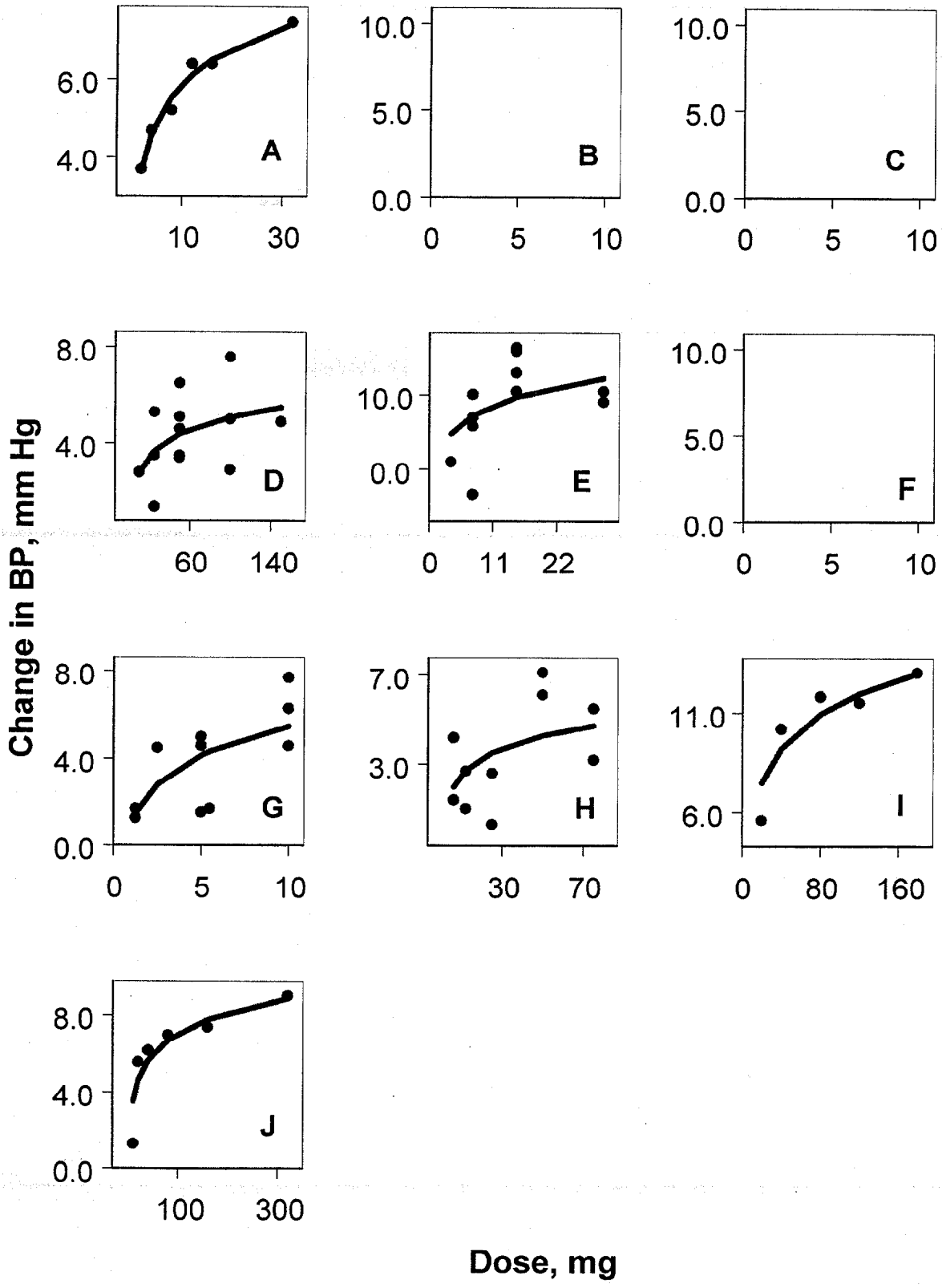
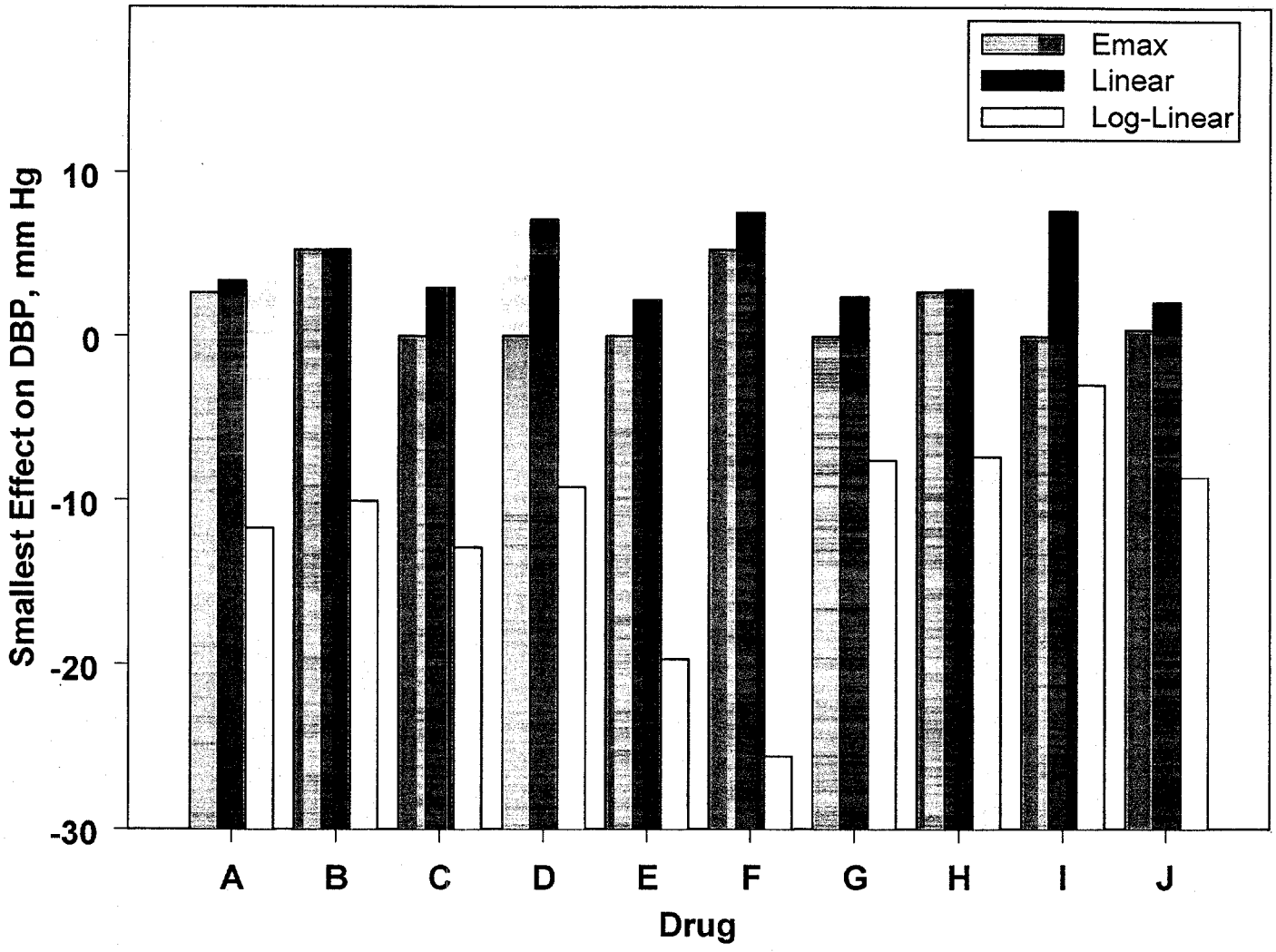
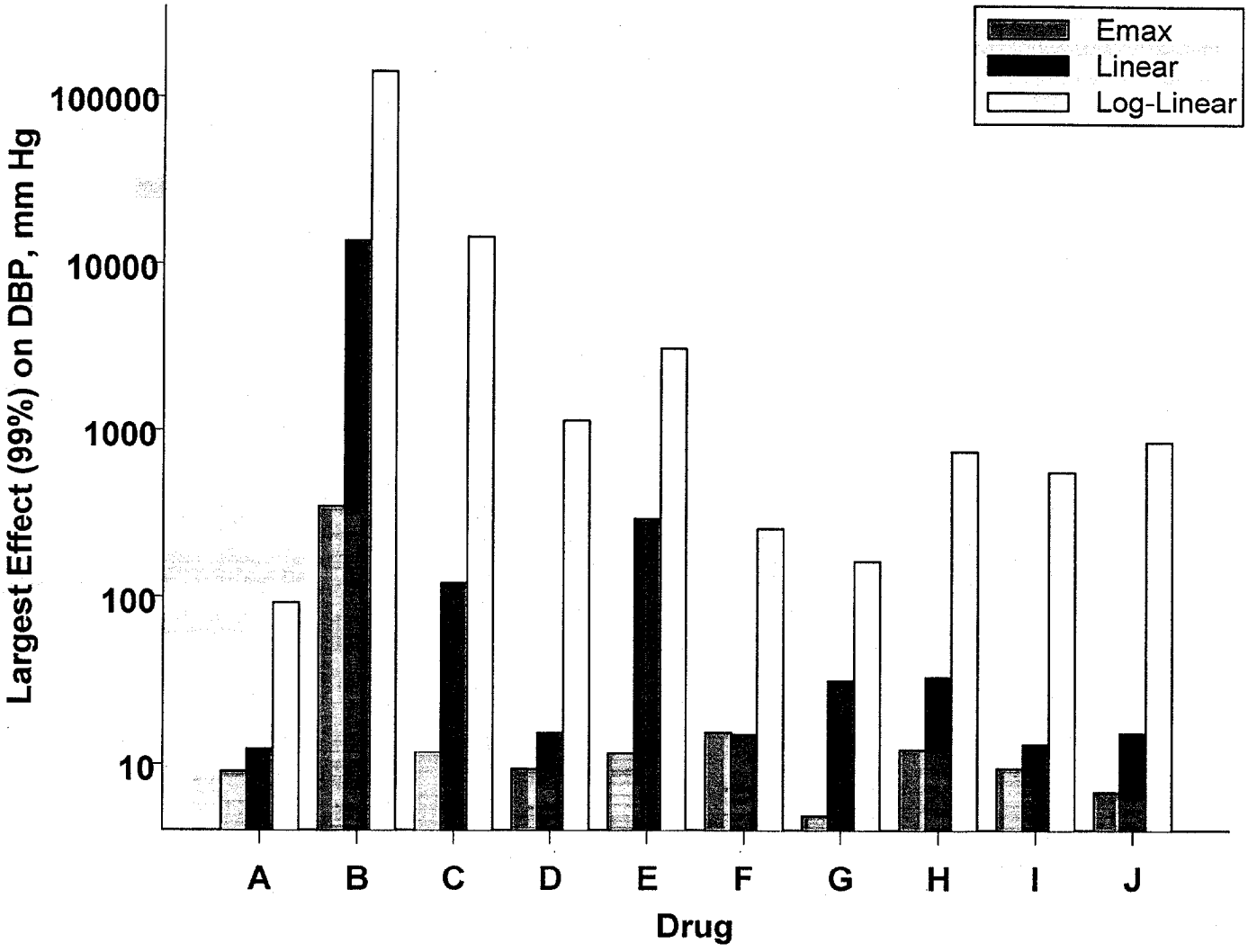


Figure 9. Model predicted smallest effect on DBP (Humans).



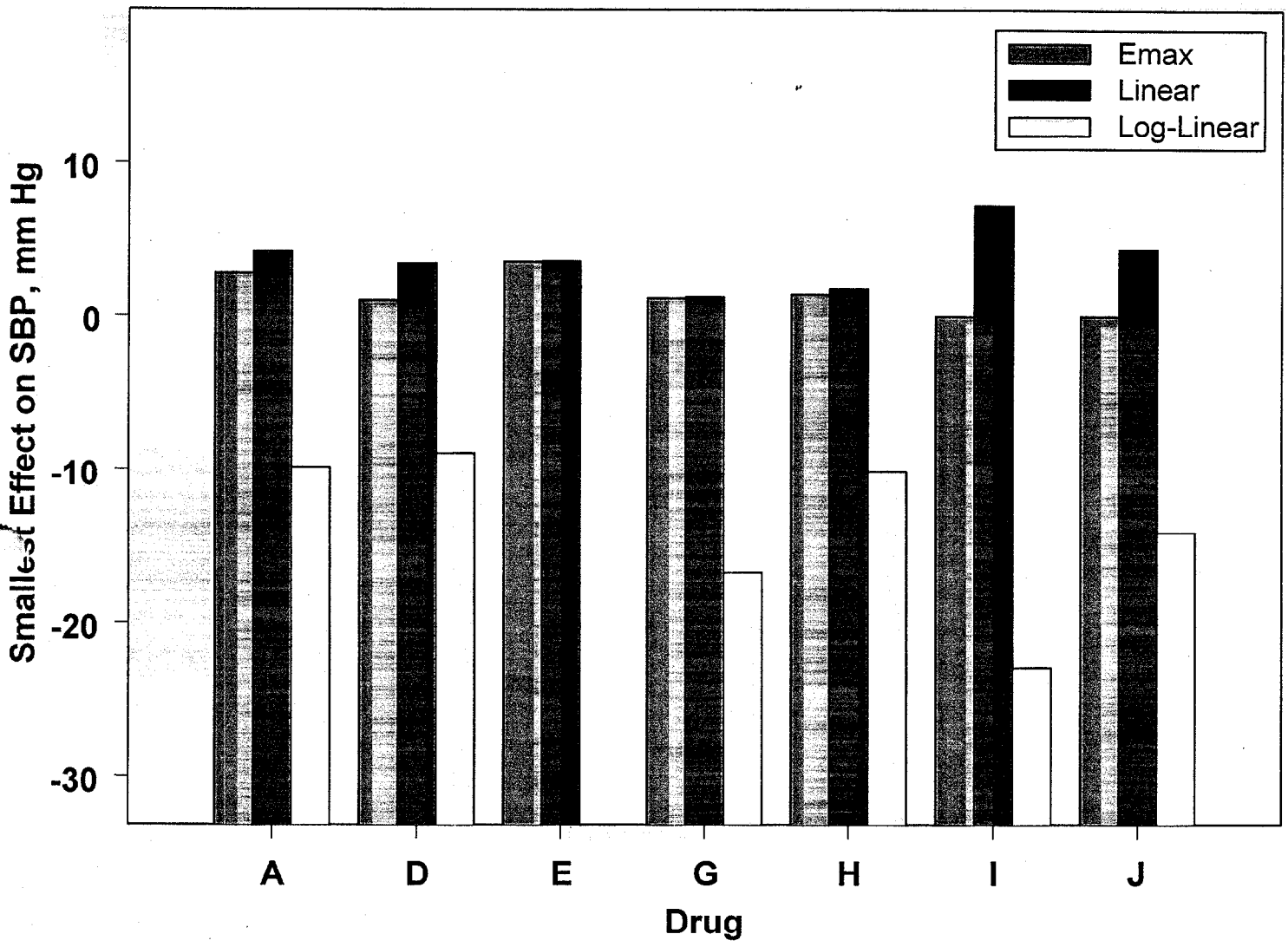
Inference: The smallest effect on diastolic BP predicted by the Emax and Linear models is close to zero (as it should be). The smallest effect predicted by the Log-linear model is atleast –10 mm Hg or lower, indicating a higher DBP (than baseline) at small doses.

Figure 10. Model predicted largest (99% of Emax) effect on DBP (log y-axis) (Humans).



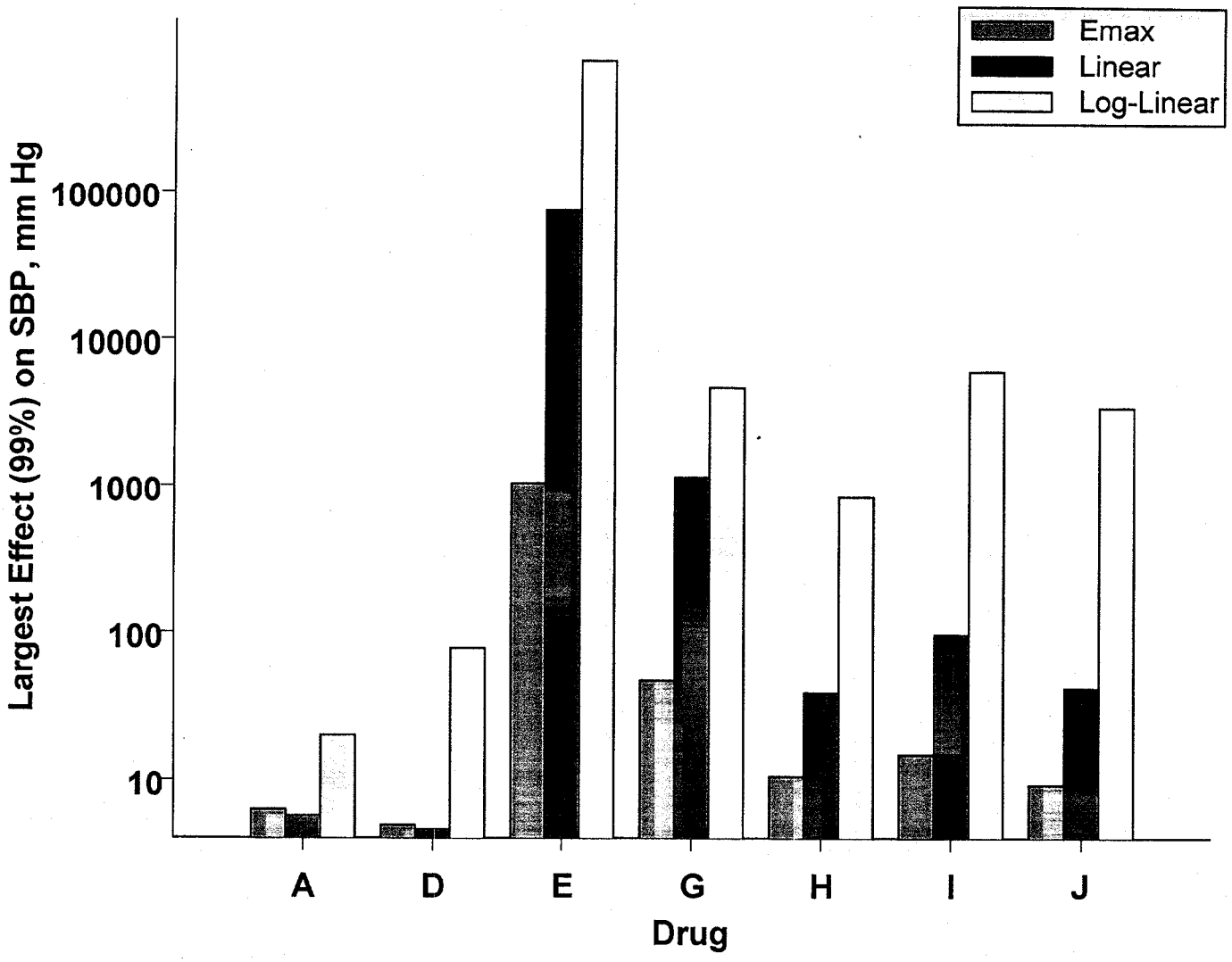
Inference: The largest effect on diastolic BP was predicted by 'believing' that the Emax model parameters are accurately estimated. The dose required to reach 99% of Emax was employed to predict the effect using Linear and Log-linear models. The Linear and Log-linear models, as expected, offer a very poor estimate of maximal effect. Of course, this observation is conditional upon belief that the Emax parameters of accurate (evidently they are not)!

Figure 11. Model predicted smallest effect on SBP (Humans).



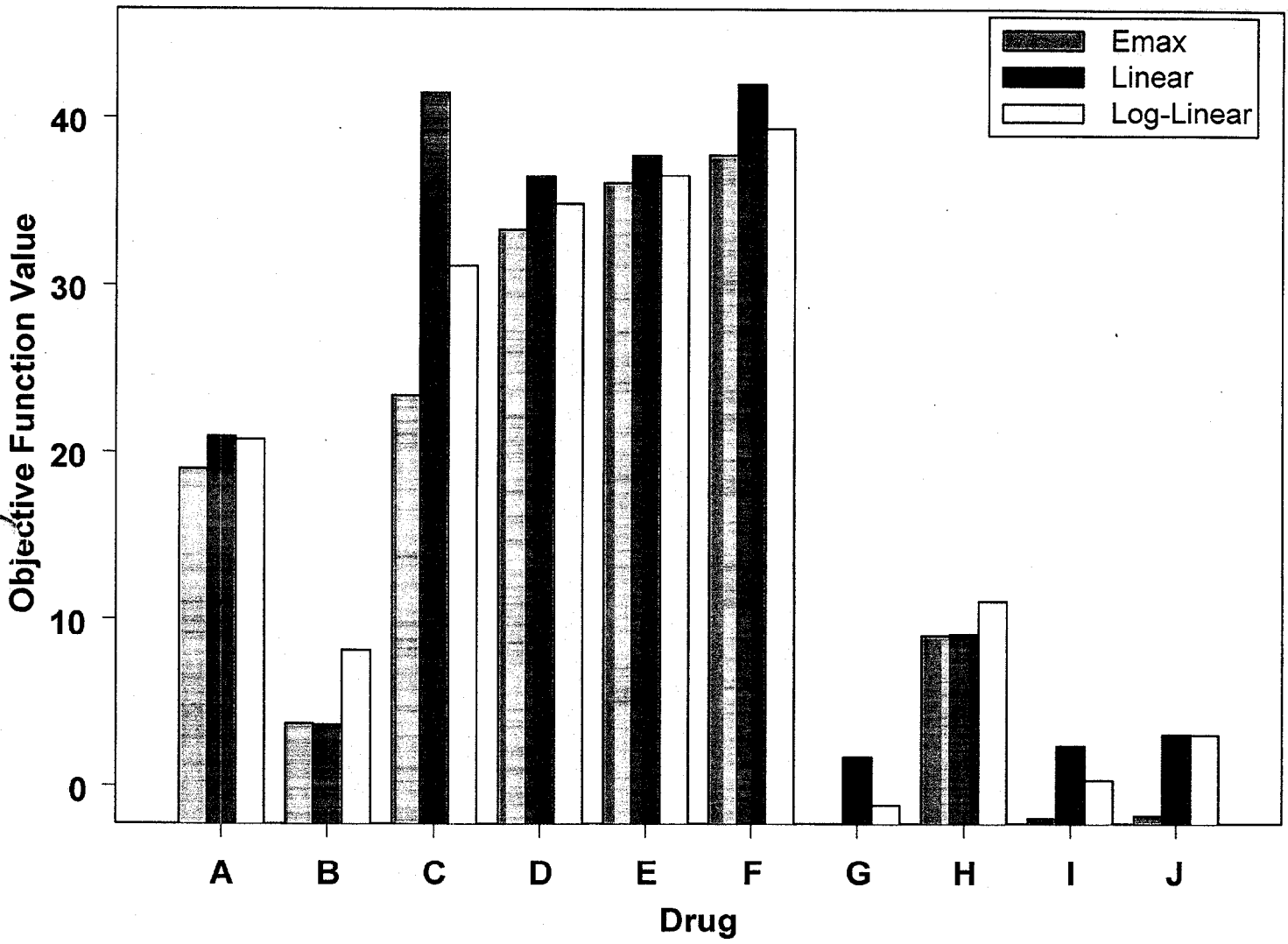
Inference: The smallest effect on systolic BP predicted by the Emax and Linear models is close to zero (as it should be). The smallest effect predicted by the Log-linear model is atleast –10 mm Hg or lower, indicating a higher SBP (than baseline) at small doses.

Figure 12. Model predicted largest (99% of Emax) effect on SBP (log y-axis) (Humans).



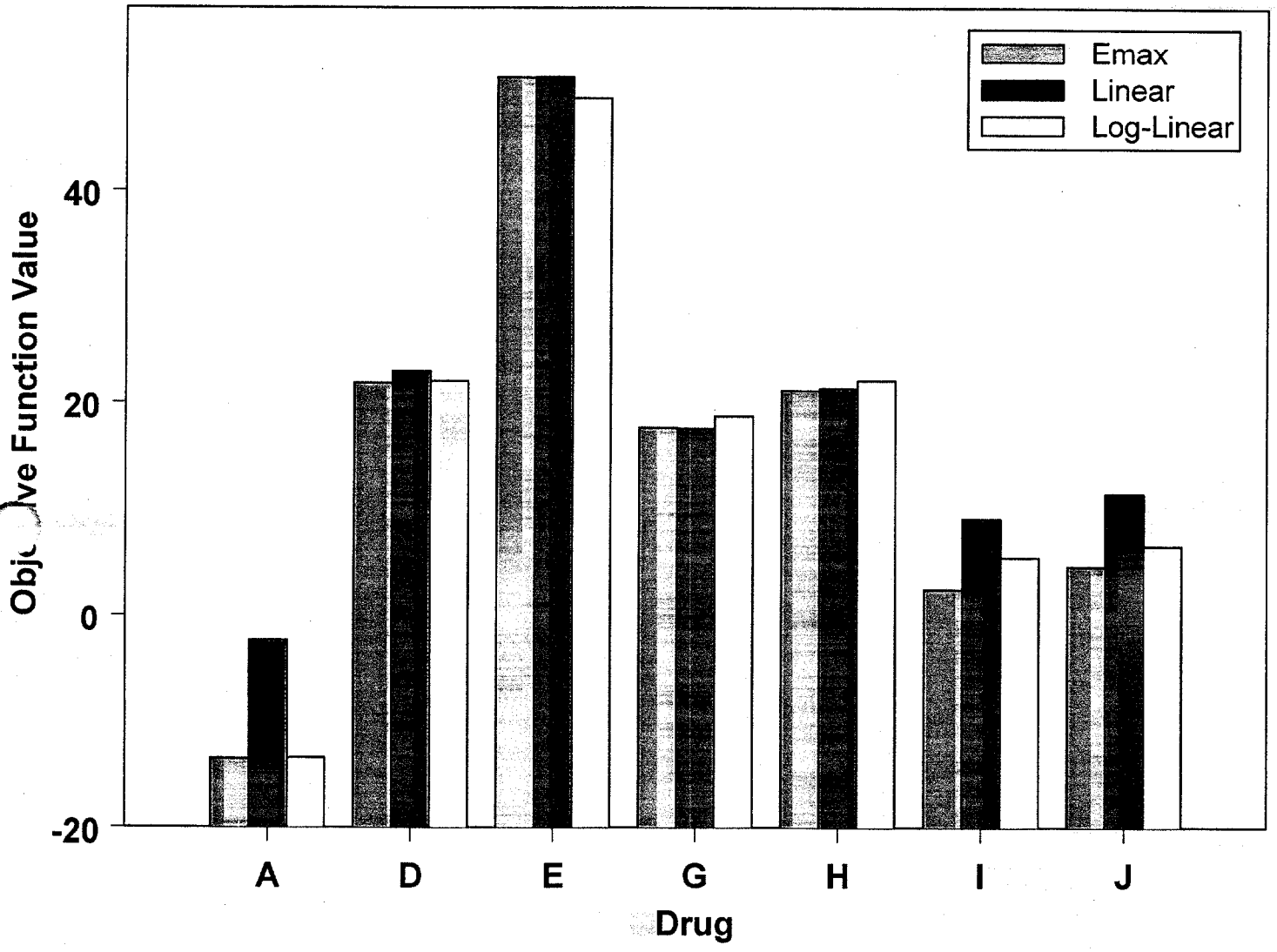
Inference: The largest effect on systolic BP was predicted by 'believing' that the Emax model parameters are accurately estimated. The dose required to reach 99% of Emax was employed to predict the effect using Linear and Log-linear models. The Linear and Log-linear models, as expected, offer a very poor estimate of maximal effect.

Figure 13. Objective function values (NONMEM) for the Emax, Linear, and Log-Linear models relating dose – DBP (Humans).



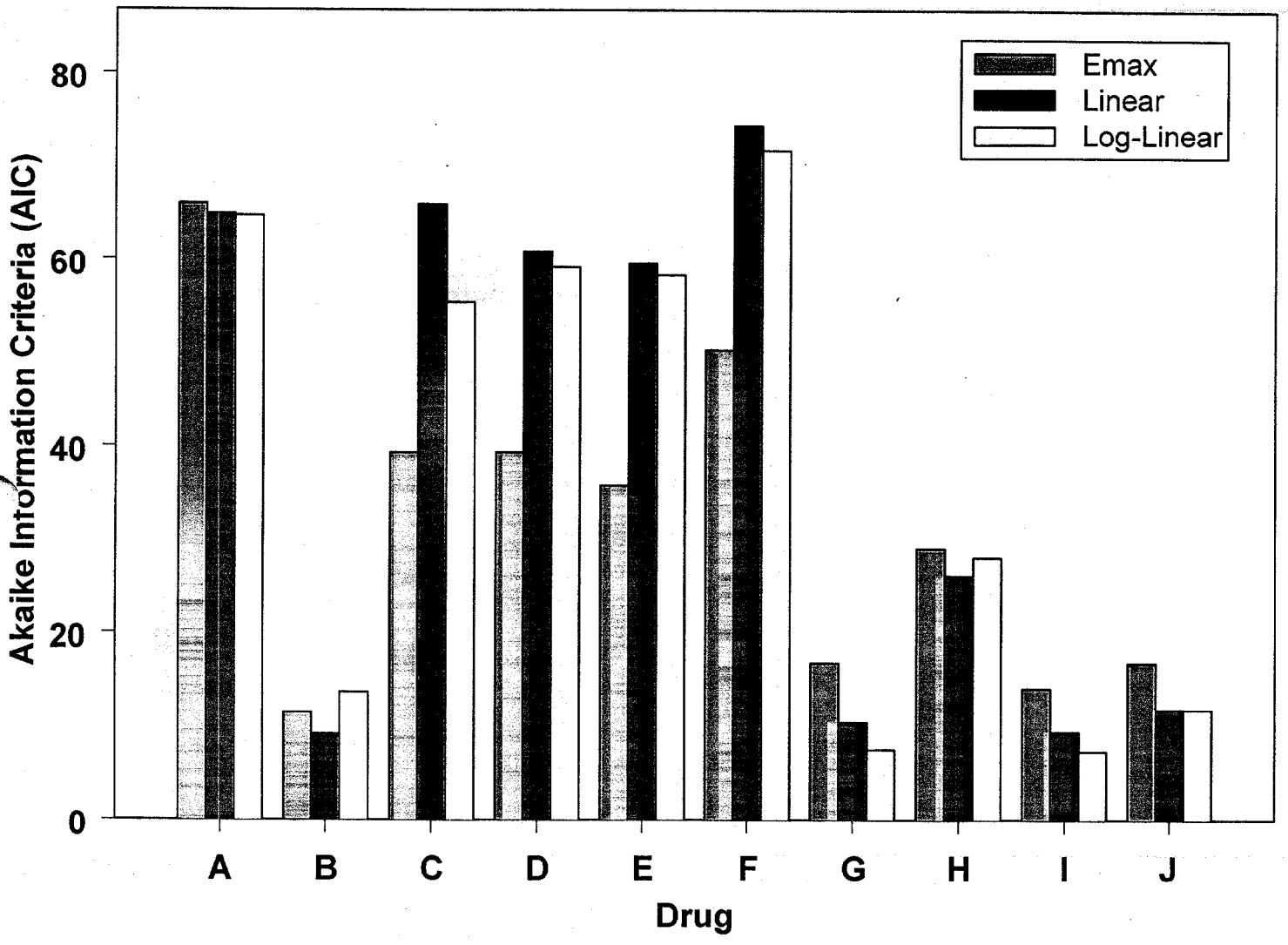
Inference: The objective function value (OFV, a routine measure of 'relative' goodness-of-fit) does not allow the selection of one model over the others. Emax model has 1 parameter more than the others, hence a difference of about 10 (between OFV of Emax and Linear models) would reflect a p=0.001. Should the dose – response be characterized adequately, a much clearer distinction among the 3 models (i.e., selecting Emax over others) could have been possible. By parsimony, a linear model is adequate to describe the dose – response data for the antihypertensives analyzed.

Figure 14. Objective function values (NONMEM) for the Emax, Linear, and Log-Linear models relating dose – SBP (Humans).



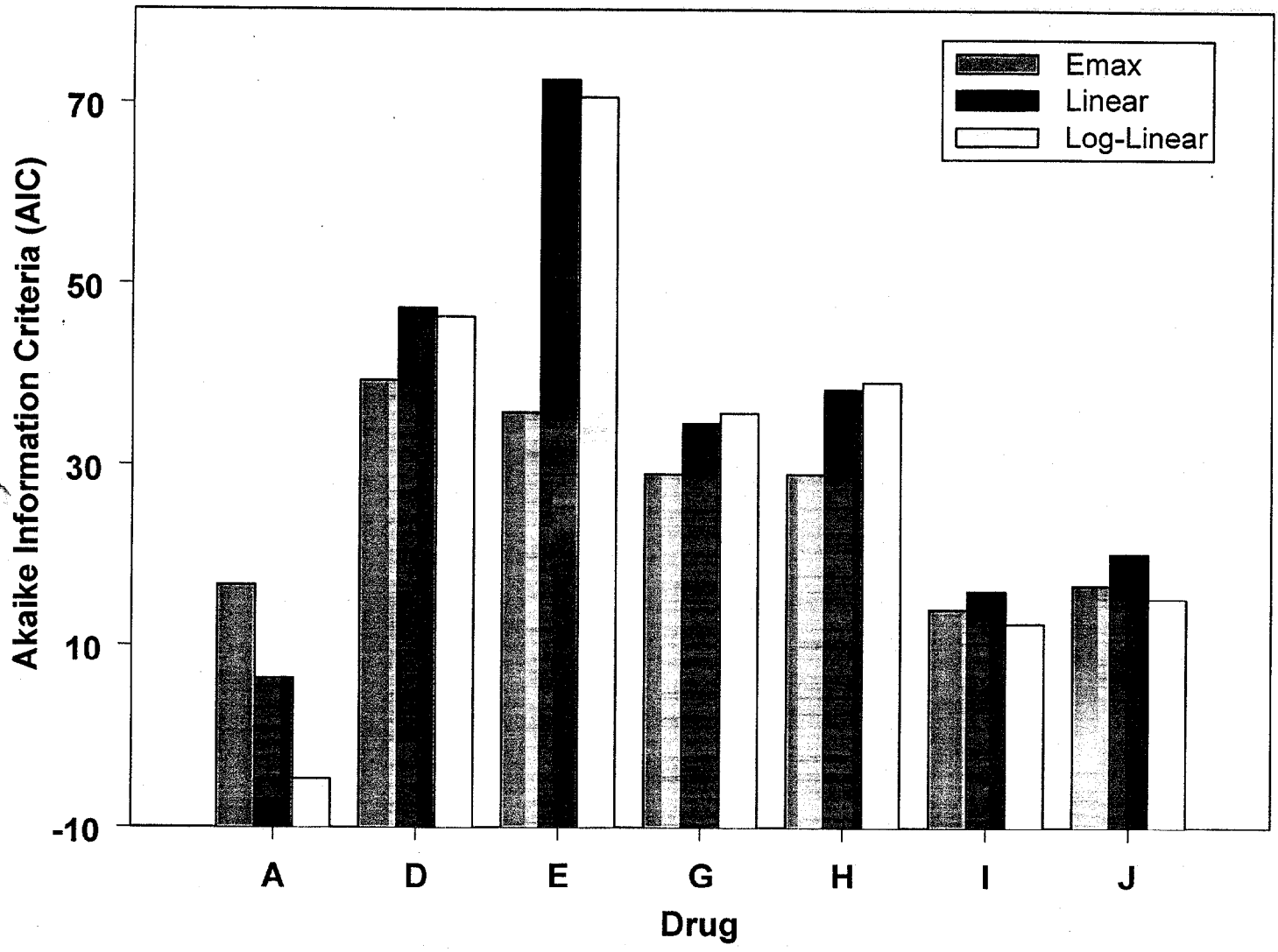
Inference: The objective function value (OFV, a routine measure of 'relative' goodness-of-fit) does not allow the selection of one model over the others. Emax model has 1 parameter more than the others, hence a difference of about 10 (between OFV of Emax and Linear models) would reflect a p=0.001. Should the dose – response be characterized adequately, a much clearer distinction among the 3 models (i.e., selecting Emax over others) could have been possible. By parsimony, a linear model is adequate to describe the dose – response data for the antihypertensives analyzed.

Figure 15. Akaike information criteria (AIC), adjusted sum of squares (for number of observations and number of model parameters), for each of the models (DBP) and drugs (Humans).



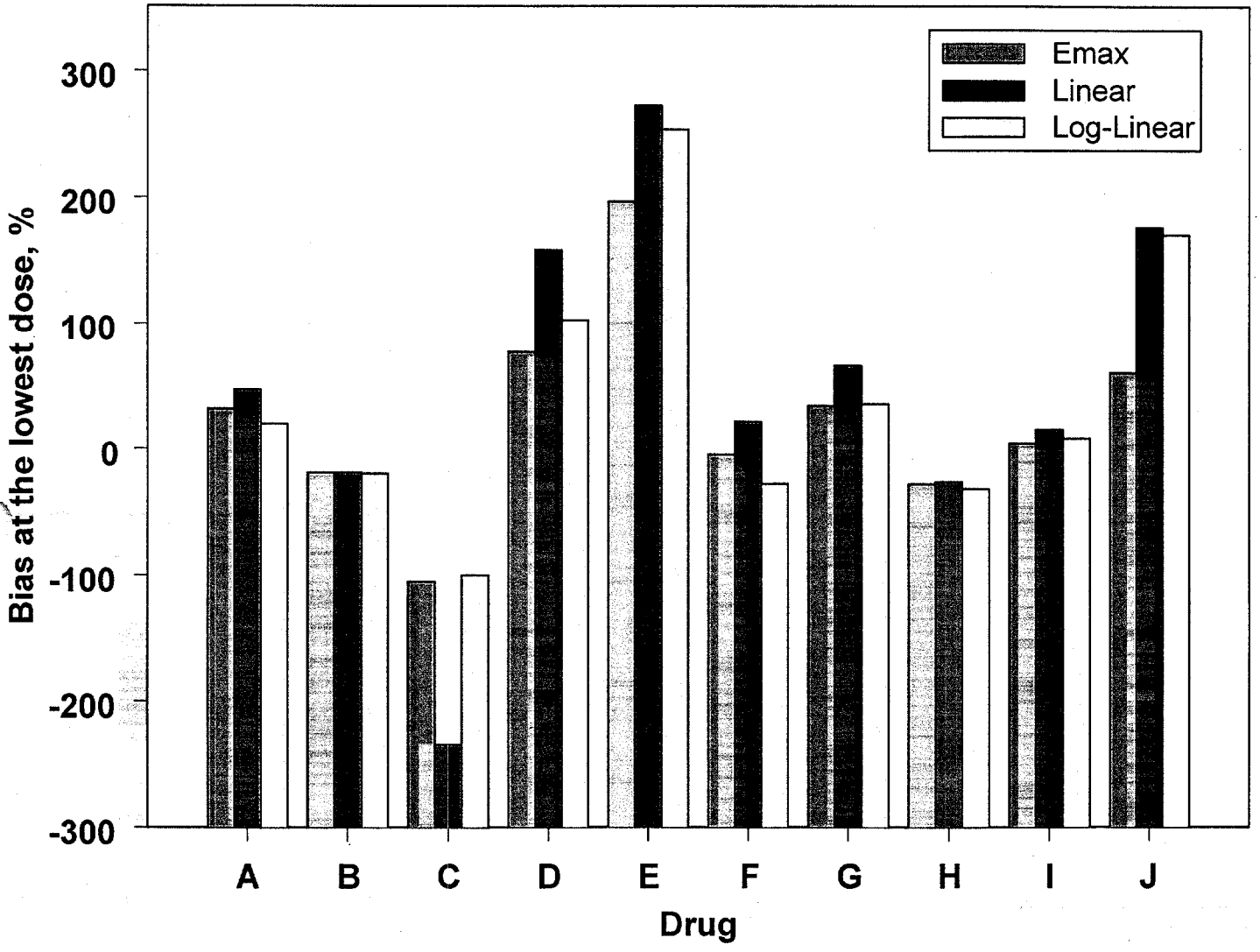
Inference: The AIC (a routine measure of 'relative' goodness-of-fit) does not allow the selection of one model over the others. The model which describes the data better than the others should have a lower AIC. Should the dose – response be characterized adequately, a much clearer distinction among the 3 models (i.e., selecting Emax over others) could have been possible. By parsimony, a linear model is adequate to describe the dose – response data for the antihypertensives analyzed. Caution: Emax, Linear and Log-Linear models are not nested!

Figure 16. Akaike information criteria (AIC), adjusted sum of squares (for number of observations and number of model parameters), for each of the models (SBP) and drugs (Humans).



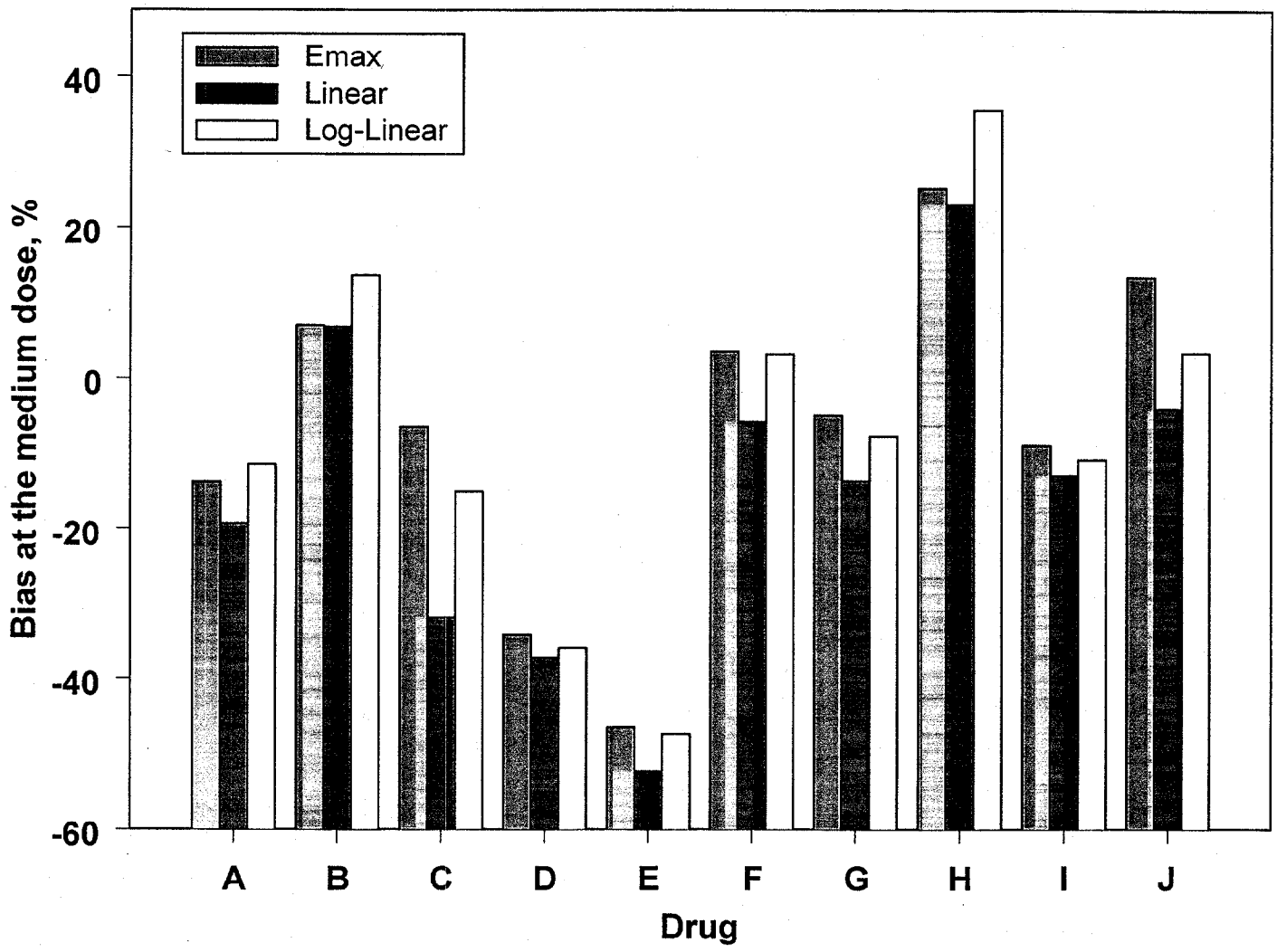
Inference: The AIC (a routine measure of 'relative' goodness-of-fit) does not allow the selection of one model over the others. The model which describes the data better than the others should have a lower AIC. Should the dose – response be characterized adequately, a much clearer distinction among the 3 models (i.e., selecting Emax over others) could have been possible. By parsimony, a linear model is adequate to describe the dose – response data for the antihypertensives analyzed. Caution: Emax, Linear and Log-Linear models are not nested!

Figure 17. Bias (%) in the model predictions at the lowest dose studied for each drug, for effect on DBP (Humans).



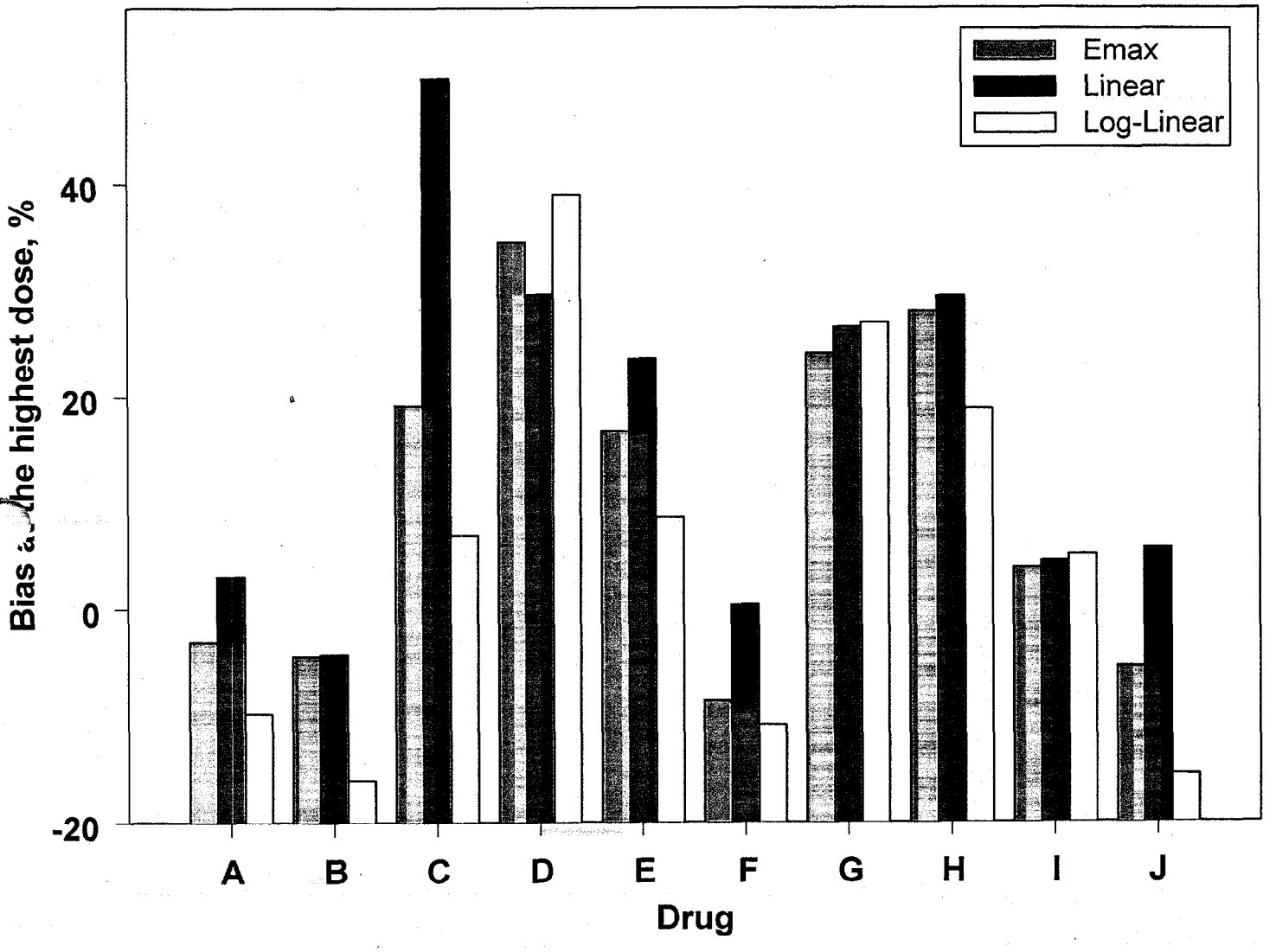
Inference: The 3 models support different signatures of dose – response data. Assuming the underlying 'true' model is Emax, the Linear model should have similar bias at the lower end of the dose – response curve, while the Log – linear model should be under-predicting the effect when compared to the Emax fittings. This expectation is not met. All 3 models behave similarly.

Figure 18. Bias (%) in the model predictions at the medium dose studied for each drug, for effect on DBP (Humans).



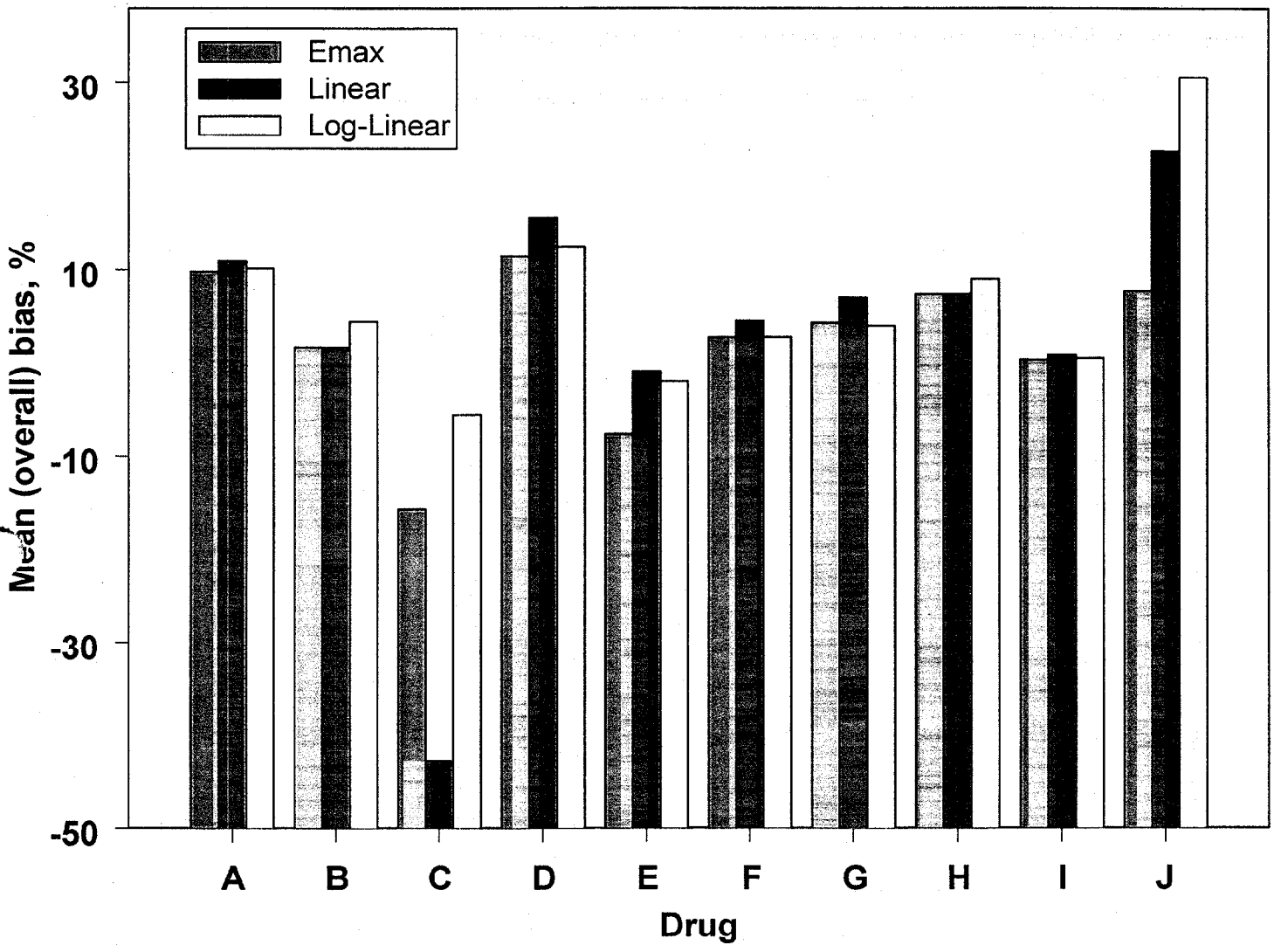
Inference: The 3 models support different signatures of dose – response data. Assuming the underlying 'true' model is Emax, all models should have similar bias. All 3 models behave similarly.

Figure 19. Bias (%) in the model predictions at the highest dose studied for each drug, for effect on DBP (Humans).



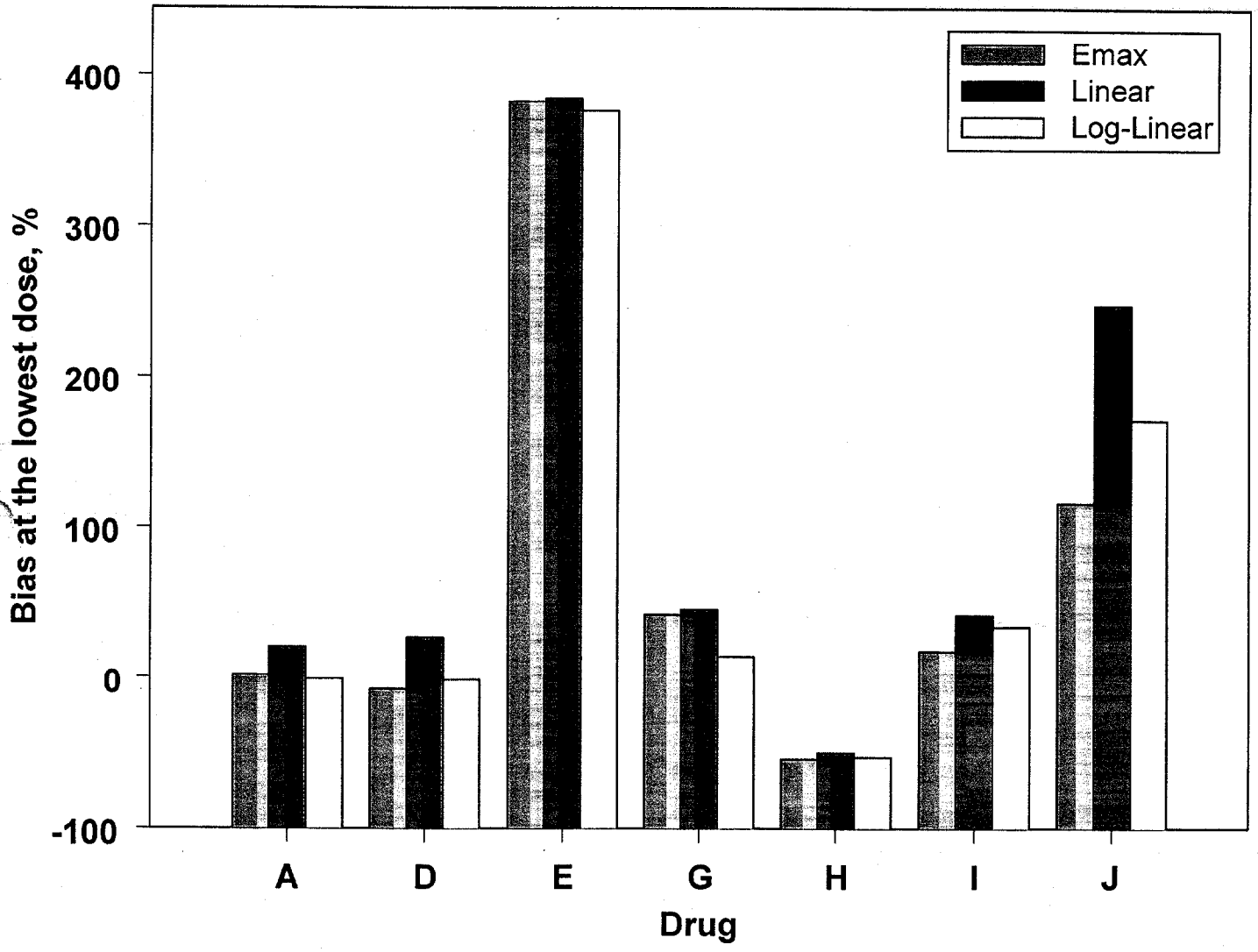
Inference: The 3 models support different signatures of dose – response data. Assuming the underlying 'true' model is Emax, the Linear and Log-linear models should be over-predicting the effect when compared to the Emax fittings. This expectation is not met. All 3 models behave similarly.

Figure 20. Overall bias (%) in the model predictions for each drug, for effect on DBP (Humans).



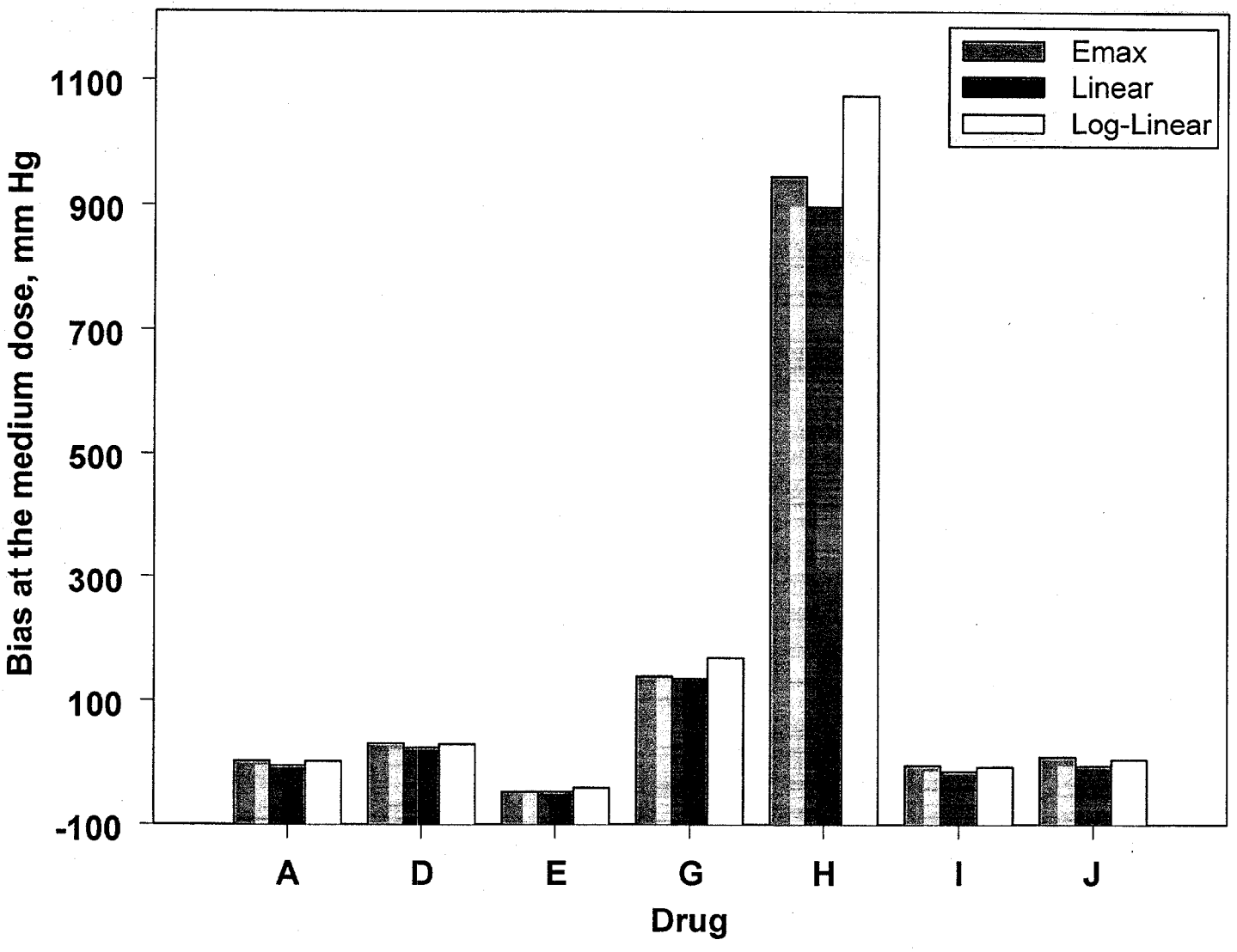
Inference:. Rejecting / accepting one model over the others is not feasible on the basis of overall bias, as shown above. By parsimony, a linear model adequately describes the dose – response data of the antihypertensives.

Figure 21. Bias (%) in the model predictions at the lowest dose studied for each drug, for effect on SBP (Humans).



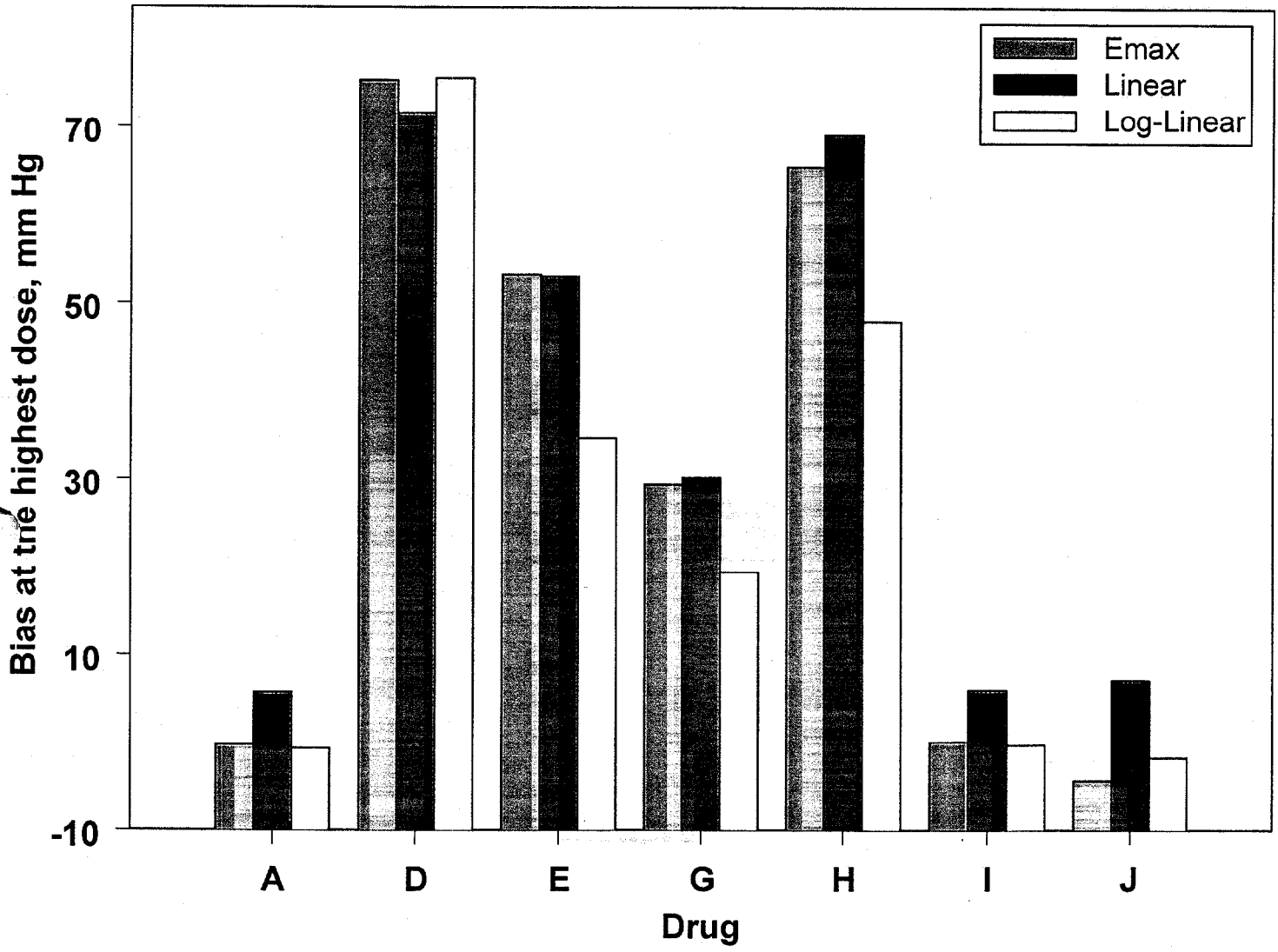
Inference: The 3 models support different signatures of dose – response data. Assuming the underlying 'true' model is Emax, the Linear model should have similar bias at the lower end of the dose – response curve, while the Log – linear model should be under-predicting the effect when compared to the Emax fittings. This expectation is not met. All 3 models behave similarly.

Figure 22. Bias (%) in the model predictions at the medium dose studied for each drug, for effect on SBP (Humans).



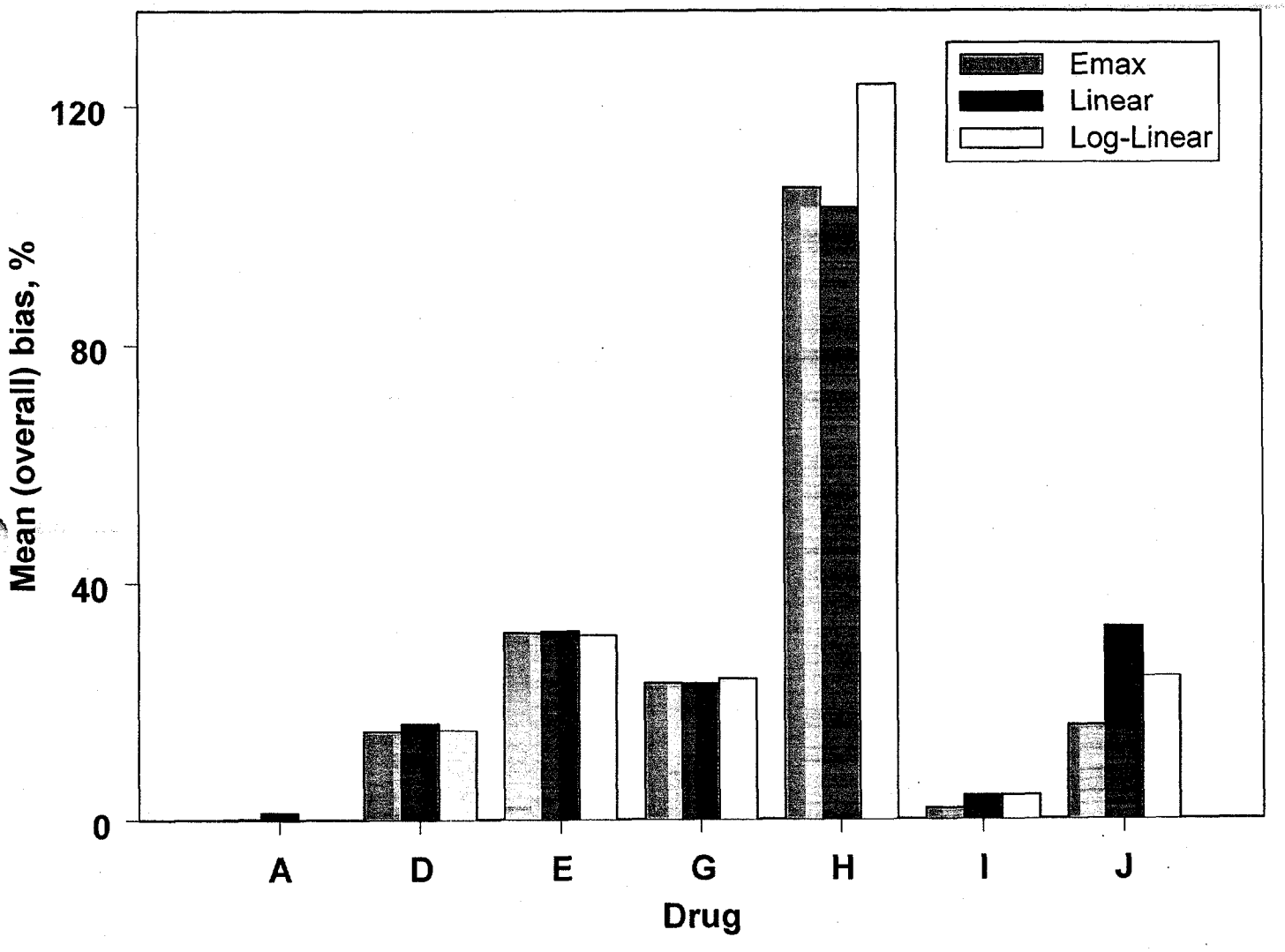
Inference: The 3 models support different signatures of dose – response data. Assuming the underlying 'true' model is Emax, all models should have similar bias. All 3 models behave similarly.

Figure 23. Bias (%) in the model predictions at the highest dose studied for each drug, for effect on SBP (Humans).



Inference: The 3 models support different signatures of dose – response data. Assuming the underlying 'true' model is Emax, the Linear and Log-linear models should be over-predicting the effect when compared to the Emax fittings. This expectation is not met. All 3 models behave similarly.

Figure 24. Overall bias (%) in the model predictions for each drug, for effect on DBP (Humans).



Inference:. Rejecting / accepting one model over the others is not feasible on the basis of overall bias, as shown above. By parsimony, a linear model adequately describes the dose – response data of the antihypertensives.

Figure 25. Observed and Emax model fitted dose – BP relationships (Linear x-axis) (Rats).

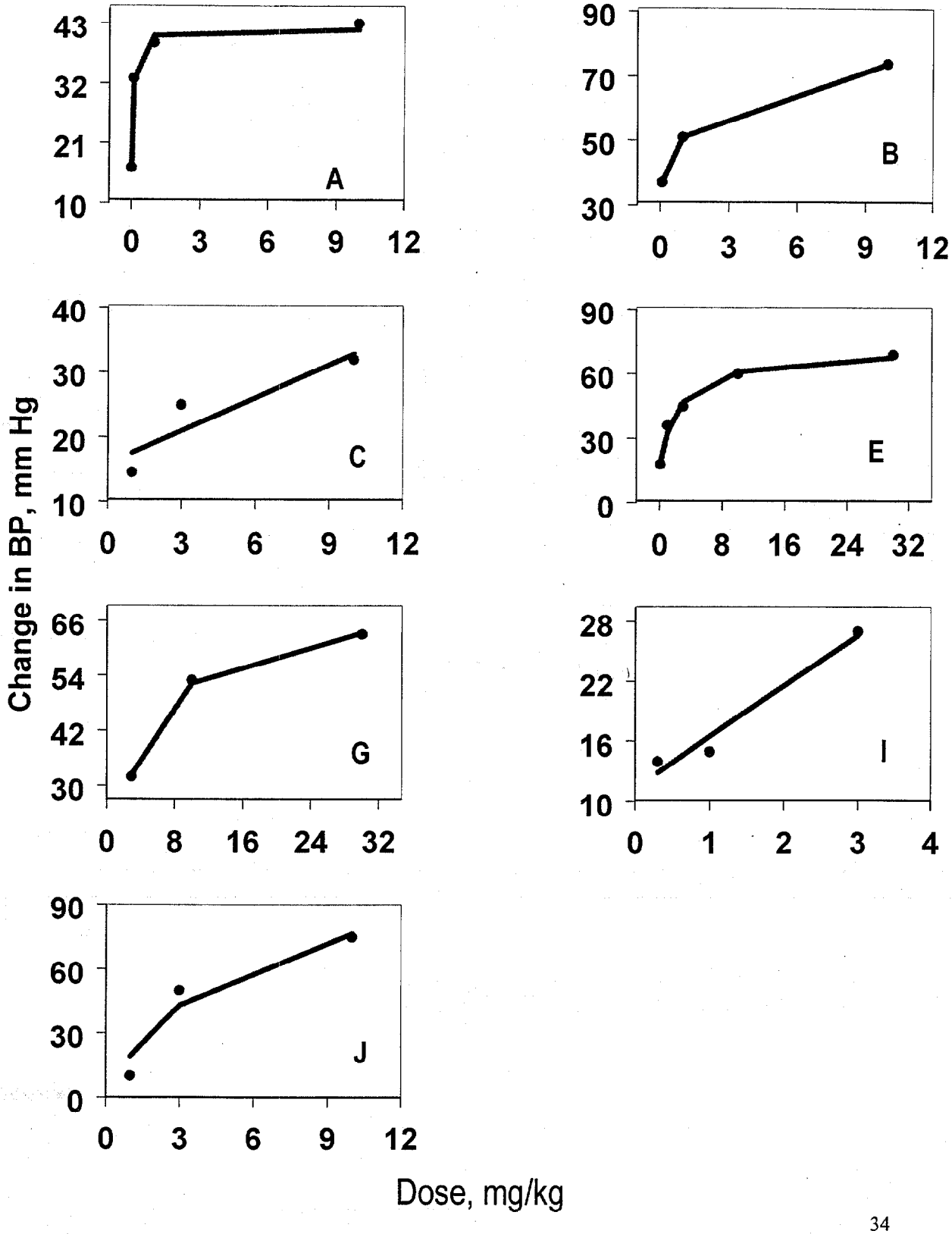


Figure 26. Observed and Emax model fitted dose – BP relationships (Log x-axis) (Rats).

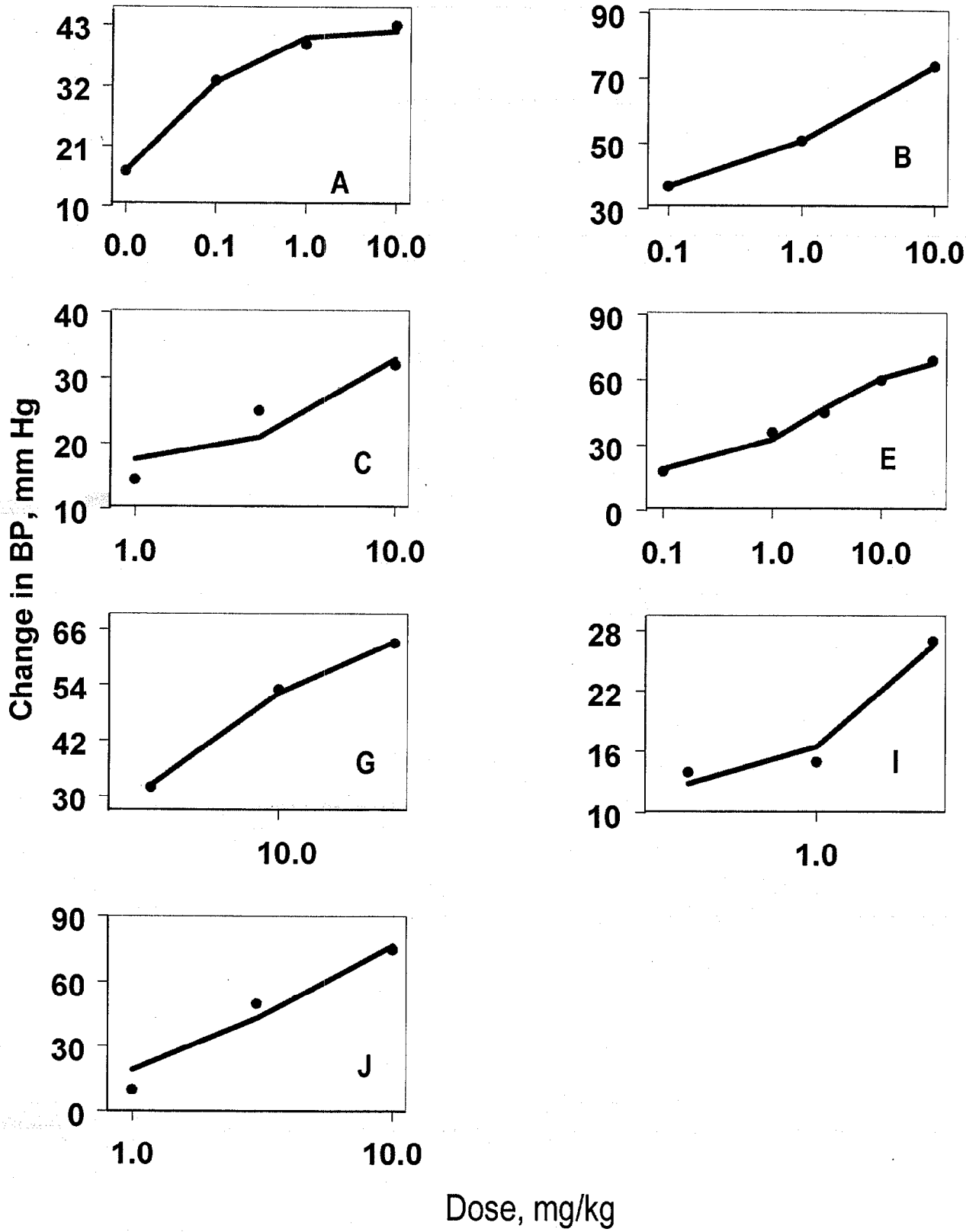


Figure 27. Observed and linear model fitted dose – BP relationships (Rats).

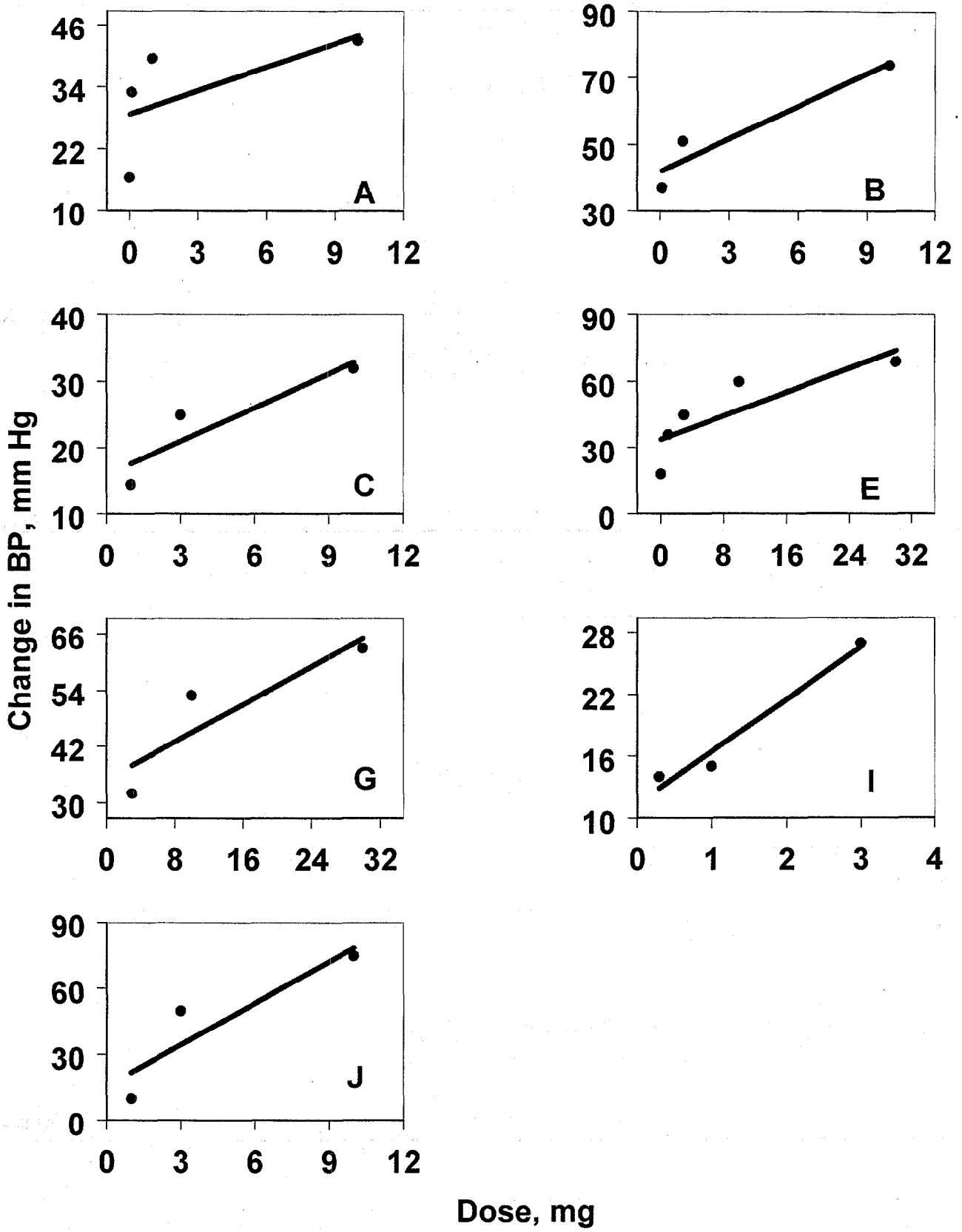


Figure 28. Observed and log-linear model predicted dose – BP relationships.

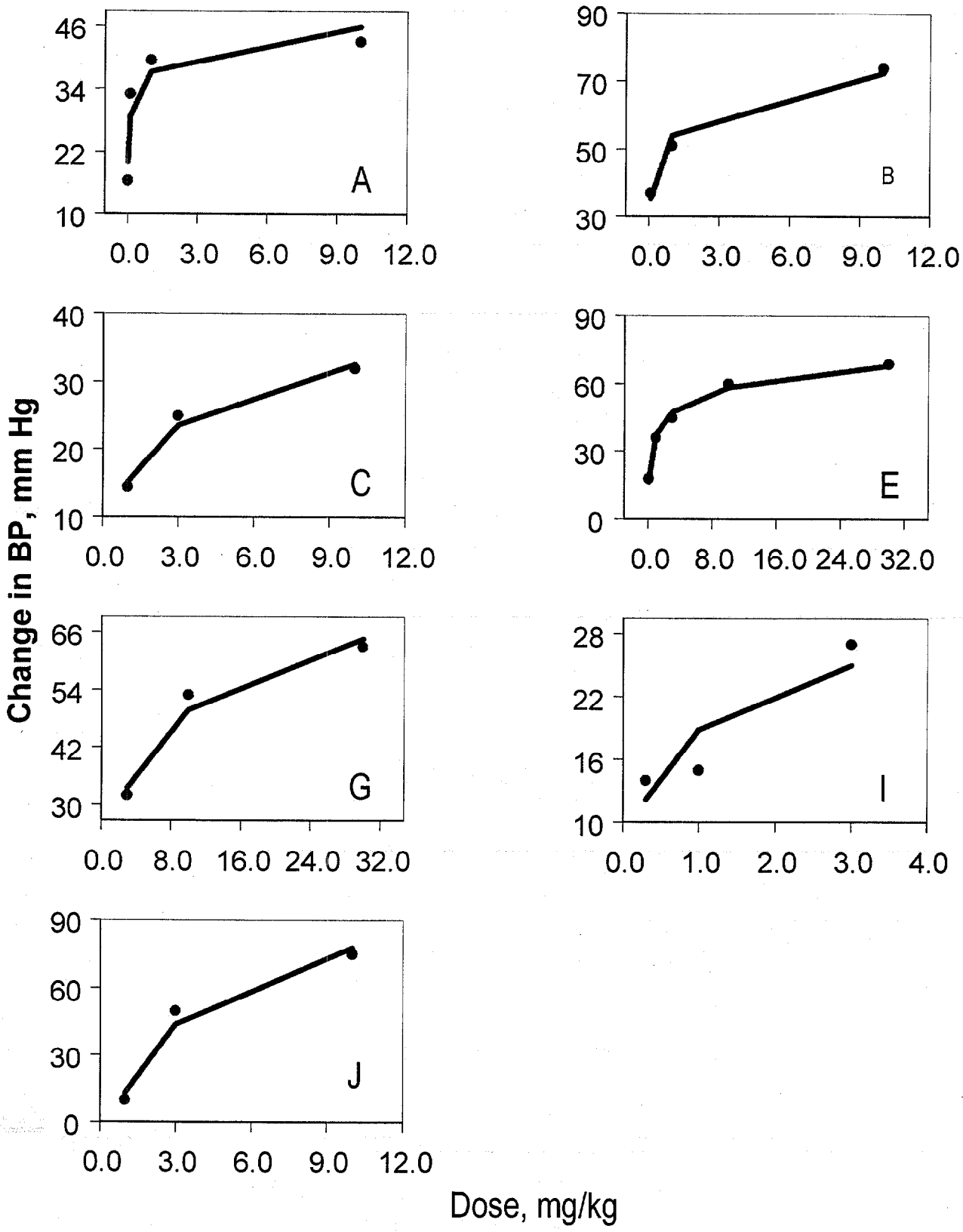
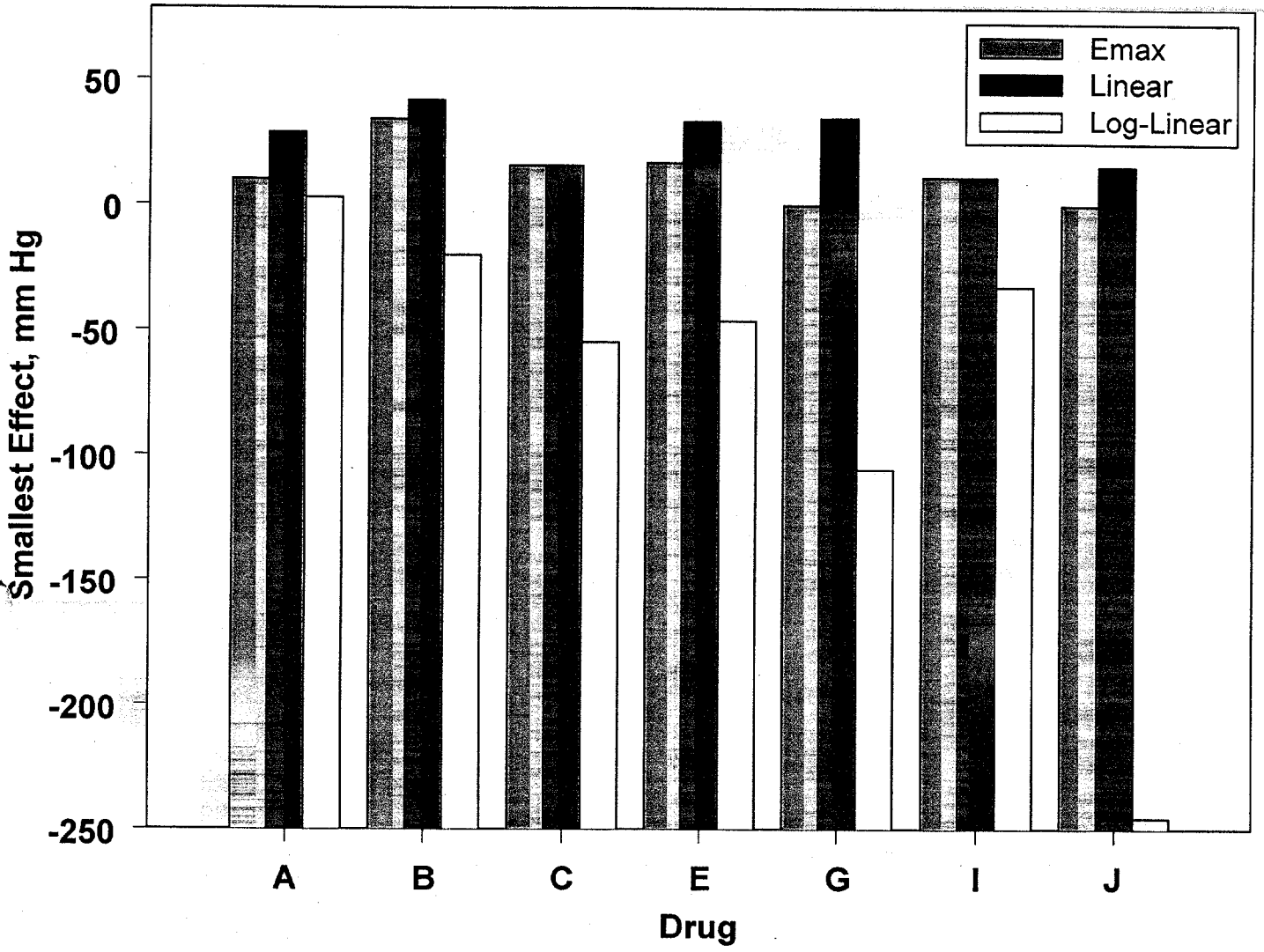
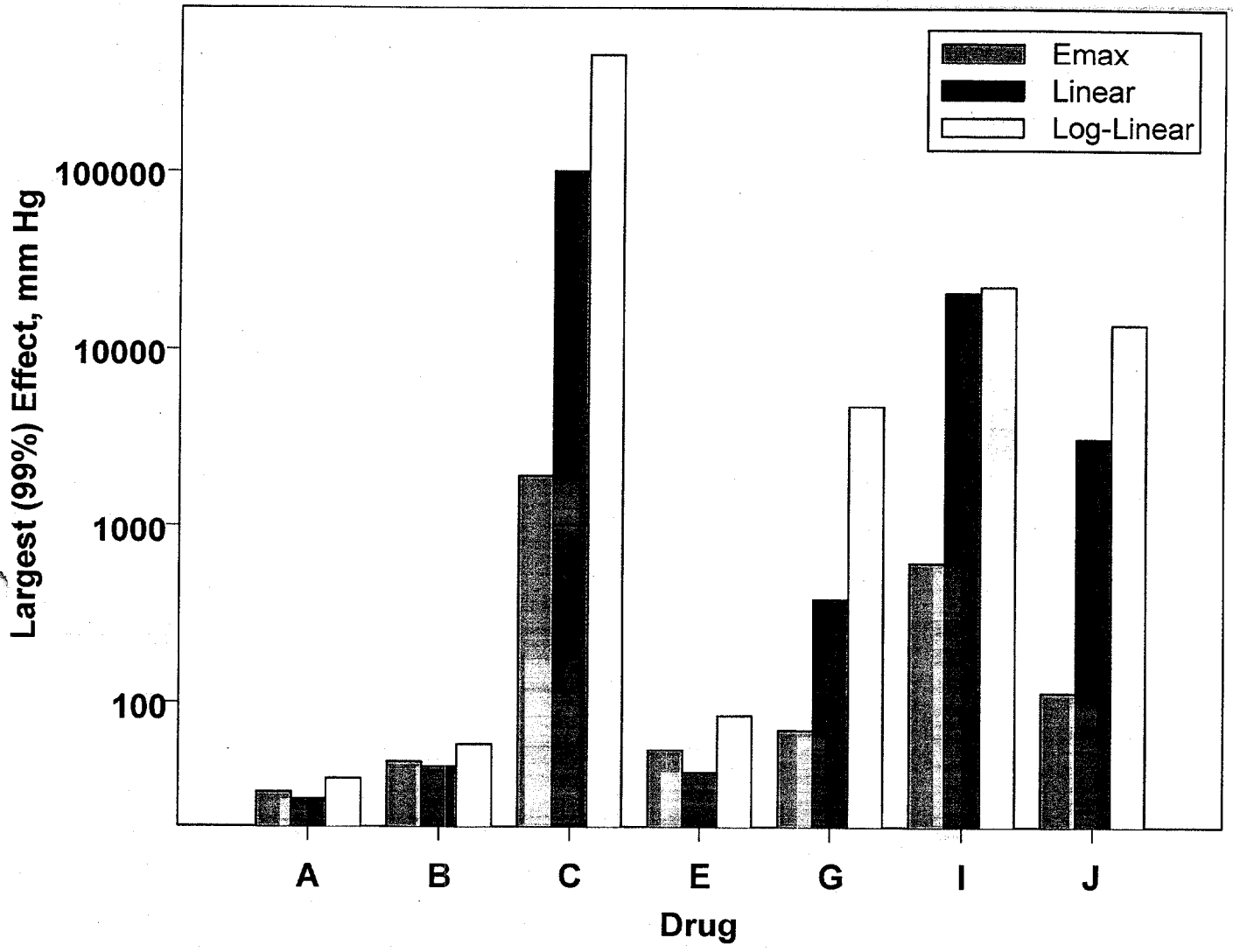


Figure 29. Model predicted smallest effect (Rats).



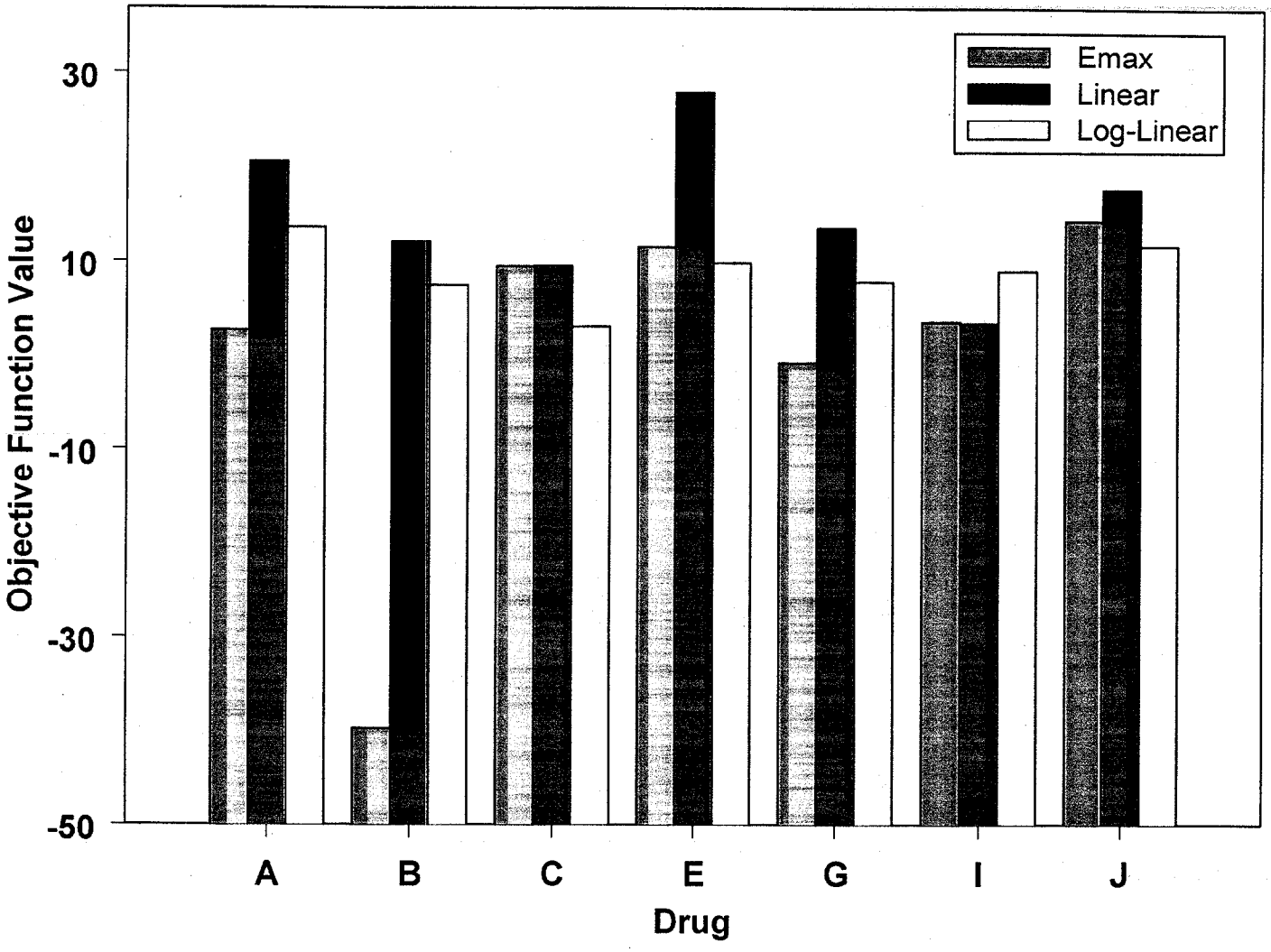
Inference: The smallest effect on BP predicted by the Emax and Linear models is close to zero (as it should be). The smallest effect predicted by the Log-linear model is atleast –20 mm Hg or lower, indicating a higher BP (than baseline) at small doses.

Figure 30. Model predicted largest (99%) effect (Rats).



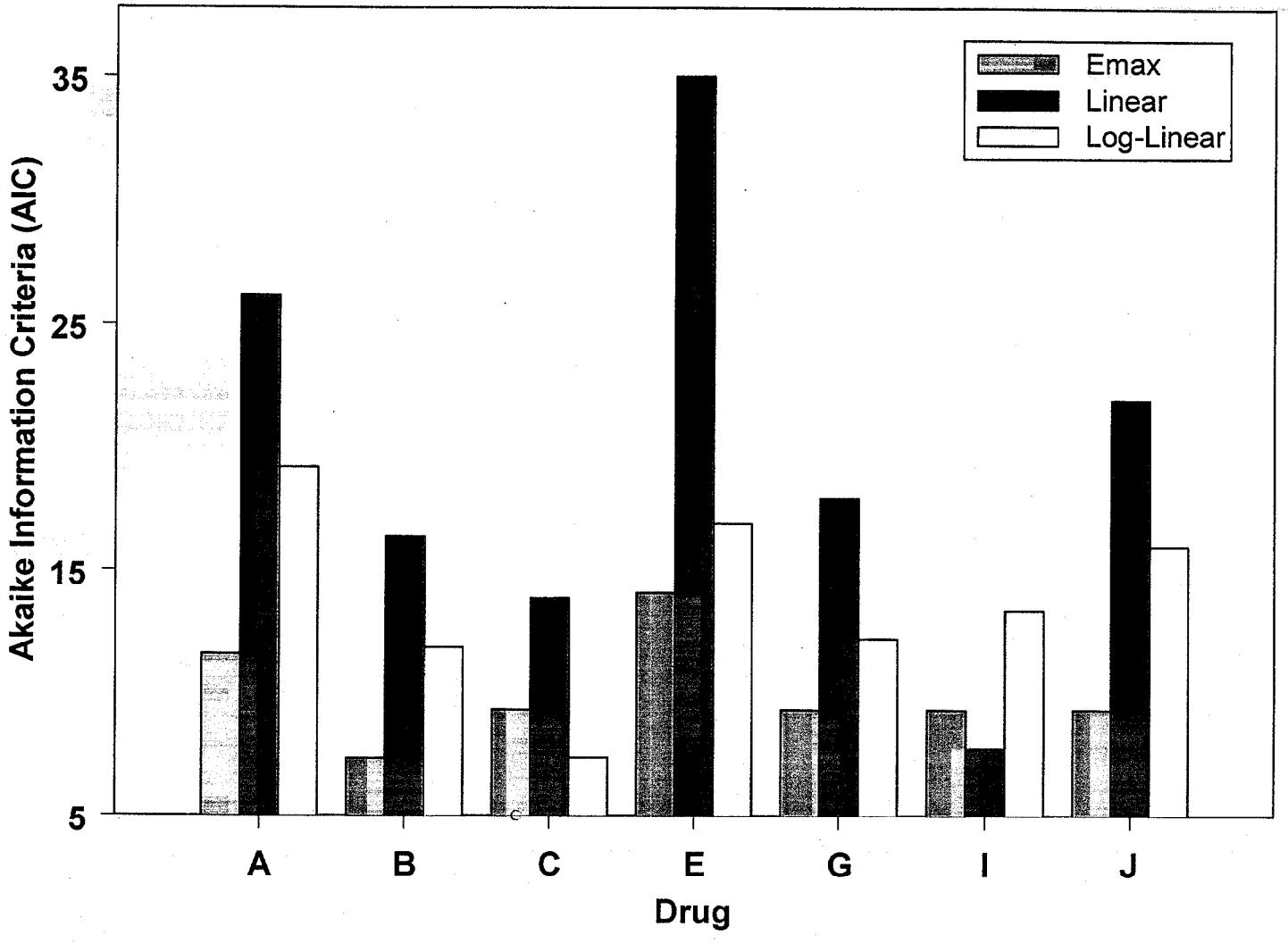
Inference: The largest effect on BP was predicted by 'believing' that the Emax model parameters are accurately estimated. The dose required to reach 99% of Emax was employed to predict the effect using Linear and Log-linear models. The Linear and Log-linear models, as expected, offer a very poor estimate of maximal effect. Of course, this observation is conditional upon belief that the Emax parameters of accurate (evidently they are not)!

Figure 31. Objective function values (NONMEM) for the Emax, Linear, and Log-Linear models relating dose – BP (Rats).



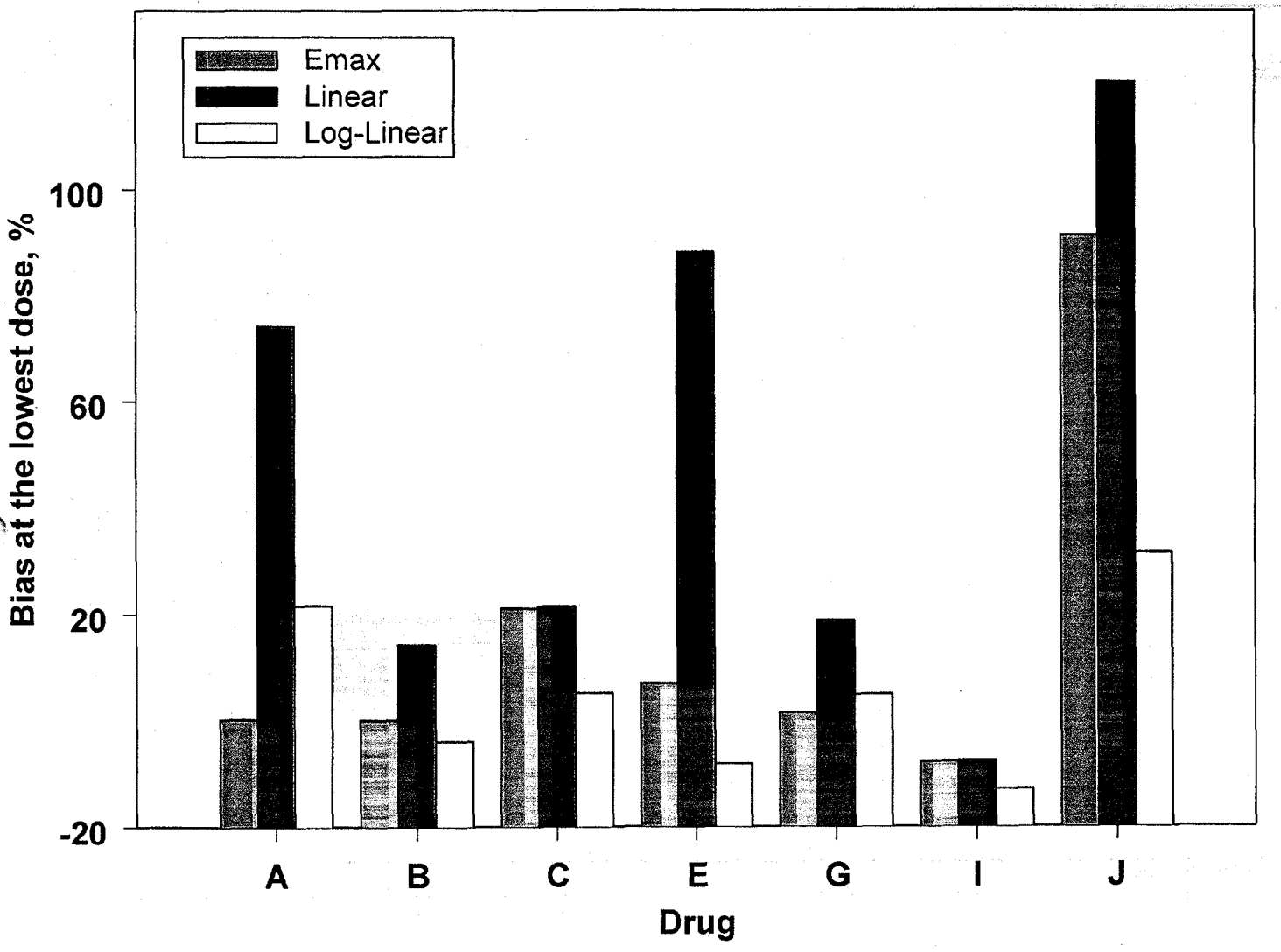
Inference: The objective function value (OFV, a routine measure of 'relative' goodness-of-fit) does not allow the selection of one model over the others. Emax model has 1 parameter more than the others, hence a difference of about 10 (between OFV of Emax and Linear models) would reflect a p=0.001 (surprisingly except for Drug B, see Figure 26). Should the dose – response be characterized adequately, a much clearer distinction among the 3 models (i.e., selecting Emax over others) could have been possible. By parsimony, a linear model is adequate to describe the dose – response data for the antihypertensives analyzed.

Figure 32. Akaike information criteria (AIC), adjusted sum of squares (for number of observations and number of model parameters), for each of the models and drugs (Rats).



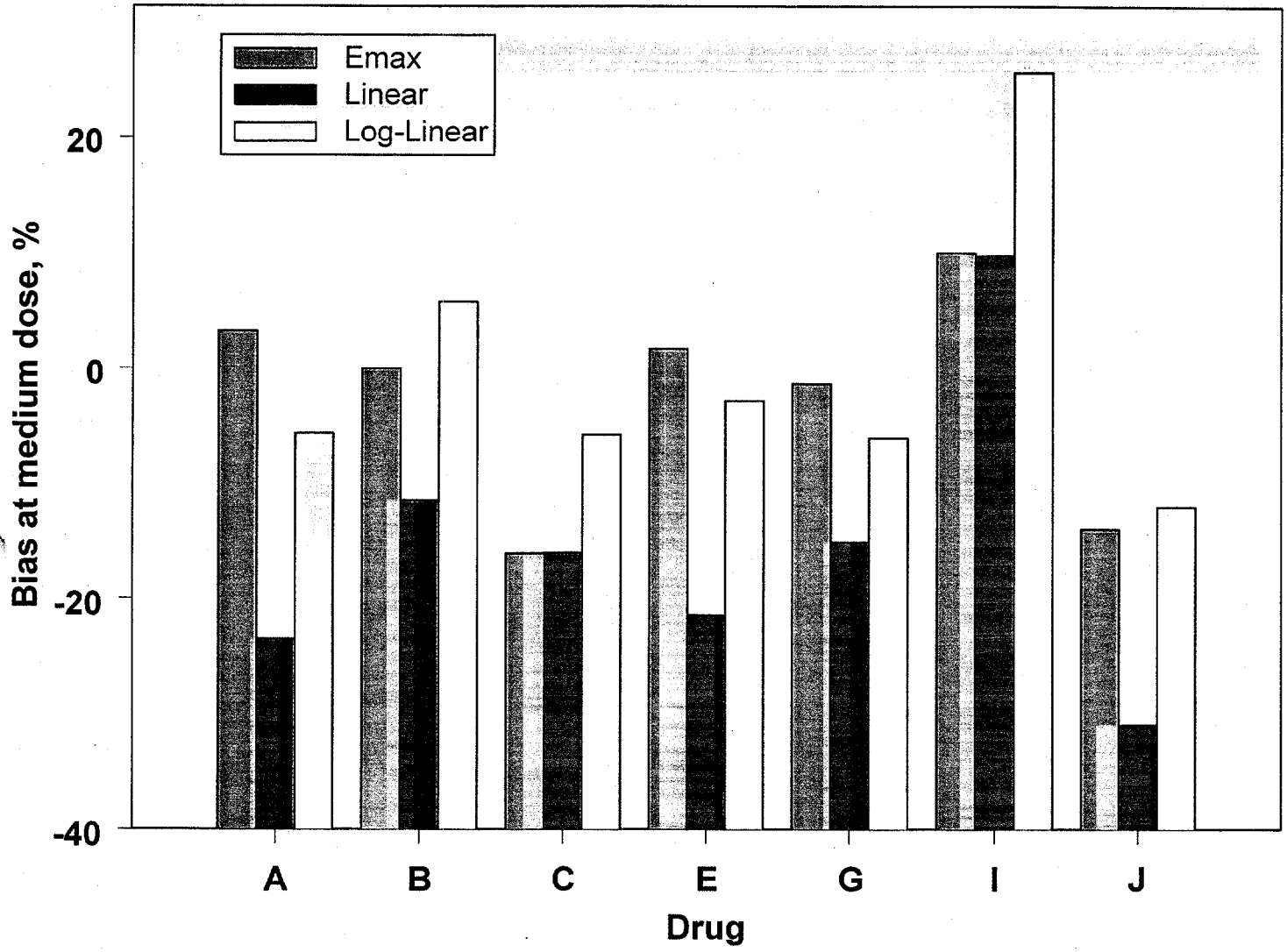
Inference: The AIC (a routine measure of 'relative' goodness-of-fit) does not allow the selection of one model over the others. The model which describes the data better than the others should have a lower AIC. The AIC for Emax model seems to be consistently lower than for the others. This observation is different than that for the humans. Probably this is due to fewer data in rats.

Figure 33. Bias in the predictions at the lowest dose for the Emax, Linear, and Log-Linear models relating dose – BP (Rats).



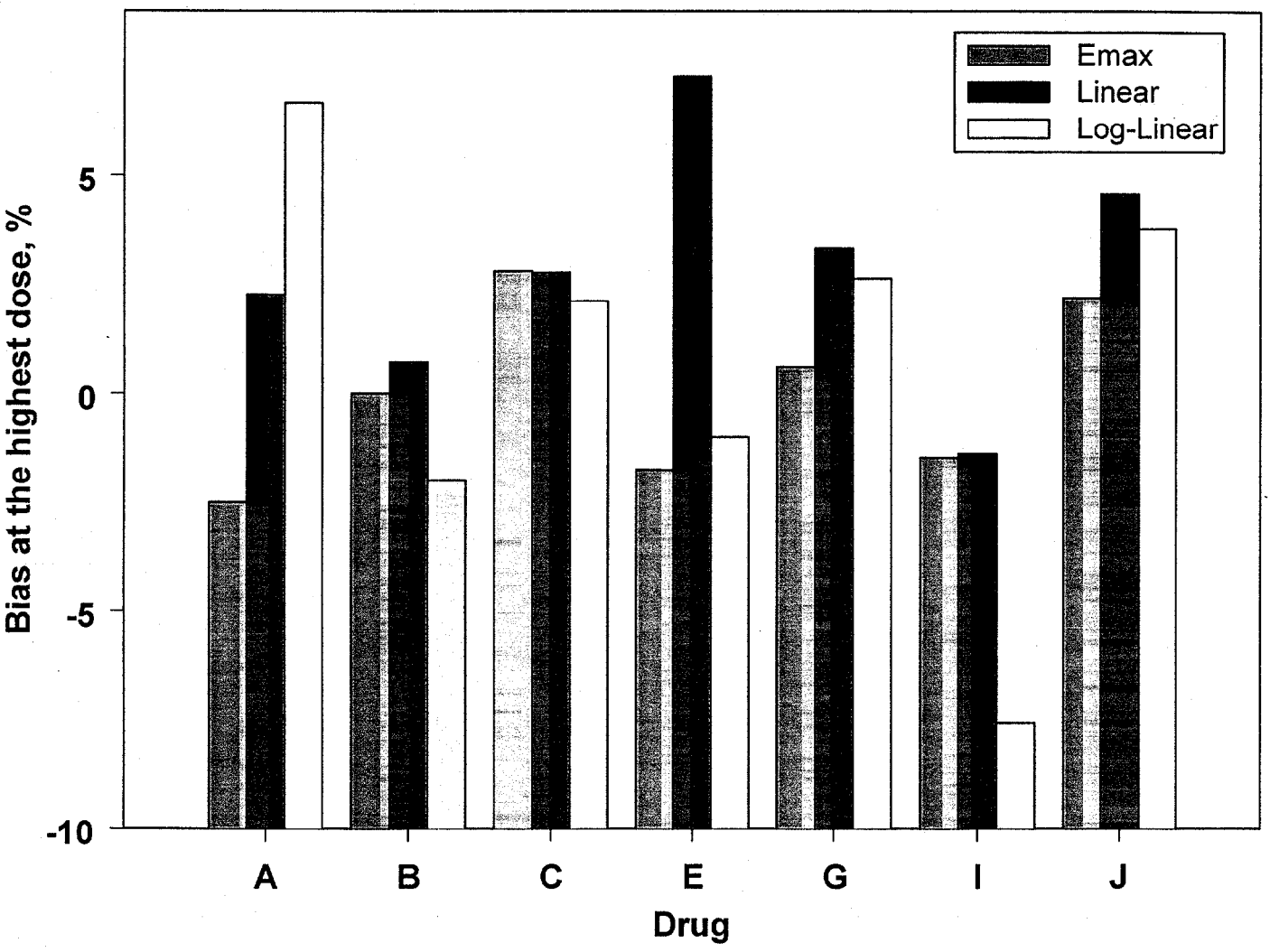
Inference: The 3 models support different signatures of dose – response data. Assuming the underlying 'true' model is Emax, the Linear model should have similar bias at the lower end of the dose – response curve, while the Log – linear model should be under-predicting the effect when compared to the Emax fittings. This expectation is not met. Linear model seems to be offering higher bias in the predictions while the Log – linear model offers similar bias as the Emax model. This could be a manifestation of not including 'truly' small doses and including doses which produce about 20% maximal effect or higher.

Figure 34. Bias in the predictions at an intermediate dose for the Emax, Linear, and Log-Linear models relating dose – BP (Rats).



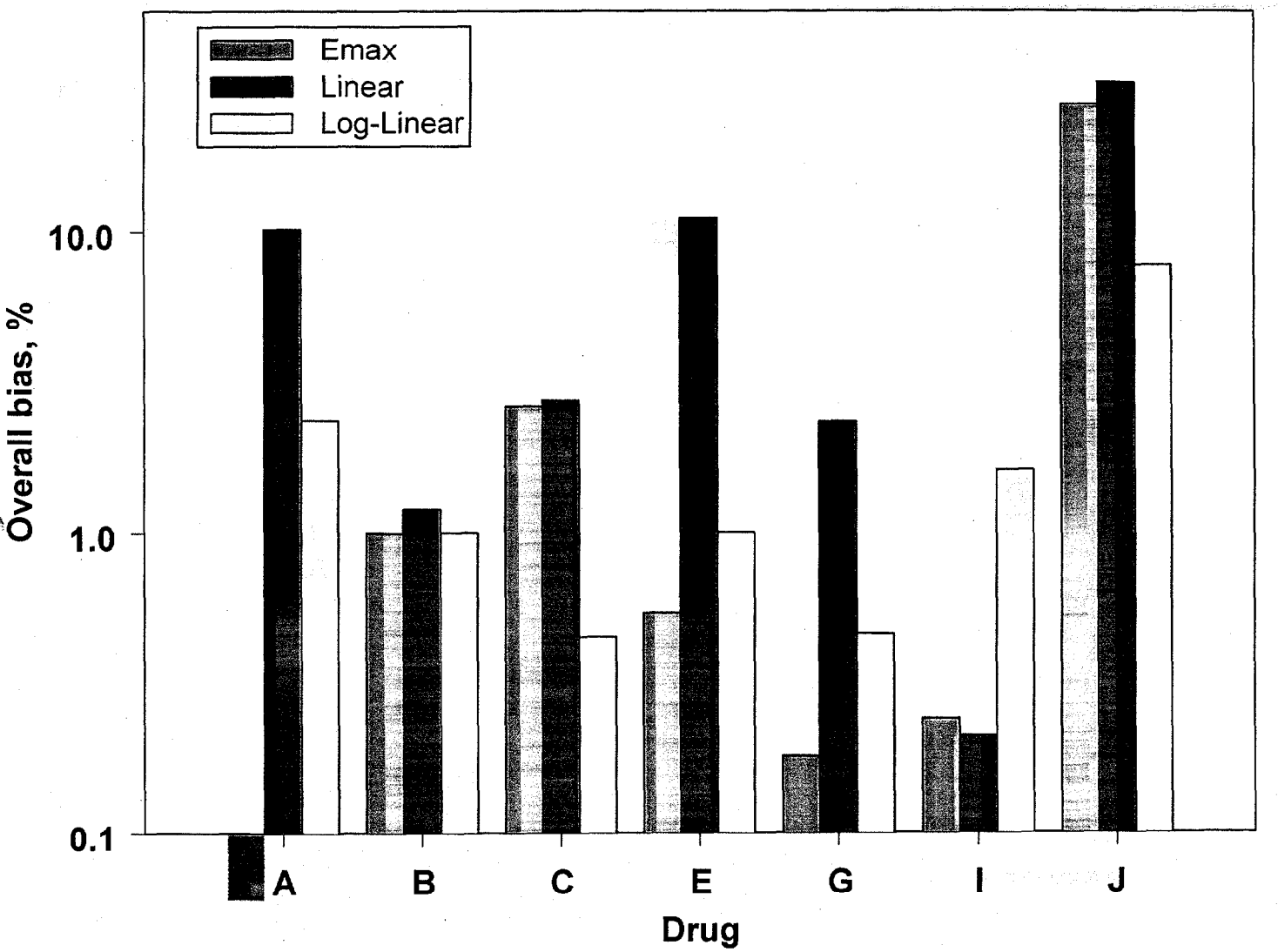
Inference: Assuming the underlying 'true' model is Emax, all models should have similar bias. Linear model, again seems to result in higher bias (almost always under predicting). This observation again corroborates the earlier hypothesis that the dose range does not start at a very small dose.

Figure 35. Bias in the predictions at the highest dose for the Emax, Linear, and Log-Linear models relating dose – BP (Rats).



Inference: Based on the inferences from Figures 33 and 34, if the dose – range does not indeed start at a small enough dose, the linear model should over-predict the responses at the highest dose when compared Emax model. However, the Linear model does so in only about 4 cases out of 7. The range of the bias is not too alarming, though.

Figure 36. Overall (mean) bias in the predictions for the Emax, Linear, and Log-Linear models relating dose – BP (Rats).



Inference: In about 3 cases, the Linear model results in a larger bias. Based on the overall bias, there is not clear indication for preferring one model to the others.

DRAFT

Dose – Response Relationship Characterization in Current Drug Development: Do we have a Problem?

Part II: Explorations via Computer Simulations

Joga Gobburu, Ph.D. and Raymond Lipicky, M.D.
Division of Cardio-Renal Drug Products,
CDER, FDA

Introduction

The inferences from the previous report entitled "Inferences from animal and human data" instigated further explorations using computer-based simulations. The main objective of the simulations was to explore the influence of dose range on the precision and accuracy of estimating the model parameters. The principal advantage in performing *in silico* Monte Carlo simulations, in this case, is that the 'true' values of the pharmacodynamic model parameters are known. Further several factors (for example: dose range, sample size, noise) affecting the power of the study to determine pharmacodynamic parameters (with appropriate accuracy and precision) can be controlled.

Methods

Simulation of Data

Data from a parallel, placebo-controlled, dosing ranging trial were generated using an Emax model, as shown in equation 1.

Emax model:
$$E = E(0) + \frac{E\max \cdot Dose}{Dose + ED_{50}} + \text{error} \qquad (1a)$$

Where, E(0) is the baseline response measured before administering the test drug, Emax is the maximal effect, ED_{50} is the dose required for half-maximal effect, and error is the measurement error in the pharmacodynamic variable. The true values of the model parameters are shown in Table 1. All noise in the simulations was confined to measurement error. No inter-individual error was assumed for the simulations.

Table 1. True values of pharmacodynamic model parameters used to simulate the data.

Parameter	Value
E(0), units	100
ED_{50}, mg	10
Emax, units	20
Error: Mean	0
Sd, units	10

Several simulation experiments with varying dose ranges (0 – 10000 mg) were performed each repeated 100 times using NONMEM (ver. 5, level 1.1).

Analysis of Simulated Data

The simulated dose – effect (placebo corrected) data were fitted to Emax, Log-linear, and Linear models. The mean of the responses in the placebo group was subtracted from the mean of the responses of each active treatment group. The mean data were fitted to the models. The mathematical forms of the Linear and Log-linear models are given below:

Log-linear model: $$E = E(1) + Slope \cdot \log(Dose) \quad (2)$$

Linear model: $$E = E(0) + Slope \cdot Dose \quad (3)$$

Where, E is the effect, 'E(0)' is the effect when no drug is present (except for the log-linear model for which it is the effect when 1 unit dose is given, E(1)), 'Slope' is the slope of the straight line describing the relationship between dose (or log(dose)) and effect.

Modeling was performed using NONMEM (UCSF, ver 5.0, level 1.0). The objective function value calculated by NONMEM can be used for selection of an appropriate model, based on statistical reasoning. The Akaike information criteria (AIC) was determined for each of the model fittings using:

$$AIC = EMAXOFV - LIN(or\ LOG)OFV + 4 \quad (4)$$

Where, EMAXOFV is the objective function value of the Emax model, LIN (LOG)OFV is the objective function value of Linear (Log-linear) model and the formula penalizes EMAXOFV for the 2 extra parameters of the Emax model. If the AIC value is greater than zero then Emax model is rejected. The power of the trial design to reject the true model used for the simulations was determined by counting the number of replications that rejected the Emax model. Further, the mean and standard deviation of the Emax model parameters (Emax, ED_{50}, E(0), and gamma) across the simulation replications were calculated. All statistical analyses were performed using SAS (ver. 6.12).

Results and Discussion

1. Figure 1 shows the dose – response curve without any measurement noise. Figure 2 shows the dose – response curve as in Figure 2, overlaid with simulated responses with measurement error. Similar dose – response curves with varying ranges of doses were simulated and fitted using Emax, Log-linear and Linear pharmacodynamic models.
2. Tables 1 and 2 show the results from the simulations. Increasing the dose range and sample size seems to decrease the probability of rejecting Emax model based on the AIC value. However, the precision of the estimate seems to be poor even with large sample sizes of 500 and 1000 and wide dose range. For example, in Table 1 simulation1 for a sample size of 1000 offers a high probability of selecting Emax model (about 90%). The estimates of the ED_{50} and Emax are quite reasonable, but the precision of the estimate is poor (about 100% CV). This

could be due to, at least, three factors: (1) inadequate dose range (maximum dose is only 10 fold higher than the ED_{50}), (2) inadequate number of doses (0.1, 10, and 100 mg), and (3) importantly only the mean data for all the active groups were fitted.

3. To investigate the influence of the factors 1 and 2, further simulations were conducted with wider dose ranges (up to 50,000 mg) and more number of doses (up to 9).

Figure 1. Dose – response curve with the true values of the parameters in Table 1 and with no error.

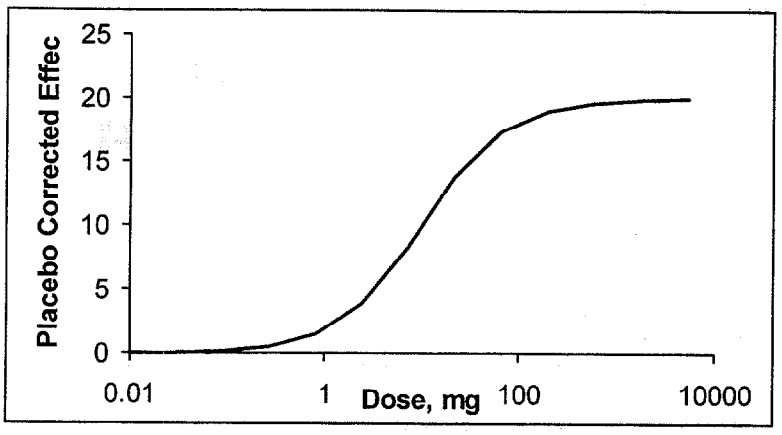


Figure 2. Dose – response curve with the true values of the parameters in Table 1 and with (squares) and without (solid line) measurement error.

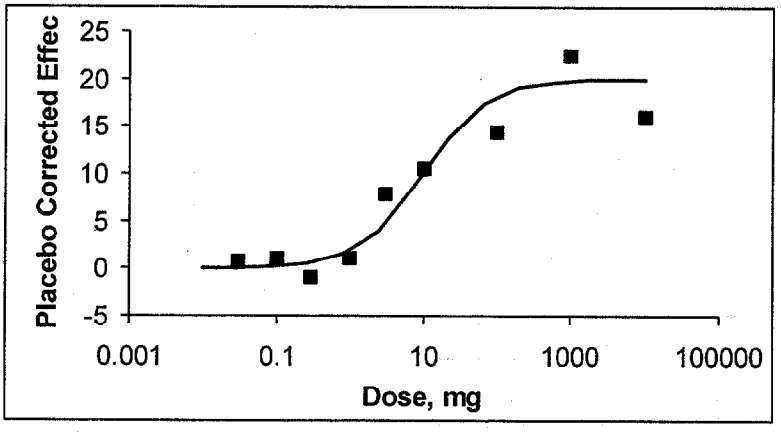


Table 1. Results of the simulations with varying range of doses (3 doses per design). The top part of the table shows the dose ranges of the simulation experiments 1 through 4. The estimated parameters from the simulated data are shown in the lower part of the table along with the power to reject Emax model over Linear and Log – linear models.

	The true ED50 = 10 mg, Emax = 20, Baseline=0, Gamma=1						
	Doses (mg)						
	0.1	**0.3**	**1**	**3**	**10**	**30**	**100**
Simulation1	x				x		x
Simulation2			x	x	x		
Simulation3					x	x	x
Simulation4				x	x	x	
Results of Simulation Experiments (No inter-individual variability)							
		median (sd)				**Power(%) (reject Emax)**	
	Sample Size	**ED50**	**Emax**	**Baseline**	**Gamma**	**Linear**	**Loglinear**
TRUE		**10**	**20**	**0**	**1**		
Simulation1	**20**	21 (3E6)	33 (362)	0.5 (3.5)	2.78 (9)	28	42
Simulation1	**48**	12 (723)	25 (320)	0.1 (2.35)	1.1 (7)	19	38
Simulation1	**100**	10 (4757)	20 (231)	0 (2)	2 (4)	15	37
Simulation1	**500**	10 (123)	19 (10)	0 (0.4)	1.2 (1.6)	0	25
Simulation1	**1000**	11 (17)	21 (5)	0 (0.04)	1 (1.5)	0	10
Simulation2	**20**	20 (5E17)	20 (2E17)	1.4 (3.5)	6.5 (19)	67	58
Simulation2	**48**	19 (6E18)	17 (6E17)	1 (2.4)	5.7 (10)	44	38
Simulation2	**100**	20 (3E21)	18 (2E20)	0.6 (1.8)	3.24 (14)	29	26
Simulation2	**500**	7 (2E15)	17 (8E14)	0.1 (1.2)	1.4 (2.3)	5	5
Simulation2	**1000**	23 (513)	30 (184)	0.1 (1)	1.1 (1.4)	1	1
Simulation3	**20**	203 (4487)	33 (1262)	5.3 (6.1)	3 (13)	42	40
Simulation3	**48**	60 (6978)	33 (1713)	6.5 (5.7)	4.2 (11.8)	27	26
Simulation3	**100**	46 (5E16)	25 (2E13)	4 (5.7)	1.5 (9.1)	22	21
Simulation3	**500**	15 (1551)	22 (422)	0.8 (6.3)	1.4 (8)	29	29
Simulation3	**1000**	14 (4E5)	22 (10E5)	0.7 (6.4)	1.3 (8)	30	31
Simulation4	**20**	49 (1E14)	35 (3E14)	3.4 (4.6)	4.7 (12)	33	28
Simulation4	**48**	17 (2E7)	23 (1E5)	3.1 (3.4)	2.6 (8)	12	12
Simulation4	**100**	10 (183)	19 (86)	0.9 (2.7)	1.7 (1.2)	3	2
Simulation4	**500**	9 (2E14)	19 (1E13)	0.2 (1.5)	1.2 (0.5)	1	0
Simulation4	**1000**	11 (74)	21 (13)	0.06 (0.7)	0.9 (0.5)	0	0

Table 2. Results of the simulations with varying range of doses (10 doses per design). The sample size was fixed at 20 subjects per dose group. The top part of the table shows the dose ranges of the simulation experiments 1 through 8. The estimated parameters from the simulated data are shown in the lower part of the table along with the power to reject Emax model over Linear and Log – linear models.

	The true ED50 = 10 mg, Emax = 20, Baseline=0, Gamma=1						
	Doses (mg)						
Simulation1	0.1,0.3,1,3						
Simulation2	0.1,0.3,1,3,10						
Simulation3	0.1,0.3,1,3,10,30						
Simulation4	0.1,0.3,1,3,10,30,100						
Simulation5	0.1,0.3,1,3,10,30,100,300						
Simulation6	0.1,0.3,1,3,10,30,100,300,1000						
Simulation7	0.1,0.3,1,3,10,100,1000,10000						
Simulation8	0.1,1,10,100,500,1000,5000,10000,50000						
Results of Simulation Experiments (No inter-individual variability)							
		median (sd)				**Power (%) (reject Emax)**	
	Sample Size	**ED50**	**Emax**	**Baseline**	**Gamma**	**Linear**	**Loglinear**
	(20/dose)						
TRUE		**10**	**20**	**0**	**1**		
Simulation1	**100**	13.38 (2E24)	48 (1E30)	0.01 (1.53)	5.64 (14)	84	75
Simulation2	**120**	3.717 (7E22)	11.6 (3E22)	0 (1.6)	3.5 (5.5)	64	47
Simulation3	**140**	9.9 (7E9)	17.9 (3E9)	0.4 (1.5)	1.7 (5.2)	38	32
Simulation4	**160**	10 (846)	18 (275)	0 (1.7)	1.5 (11)	7	39
Simulation5	**180**	10.6 (10.6)	18.8 (3.9)	0.14 (1.5)	1.3 (2)	1	30
Simulation6	**200**	11 (6.4)	18.4 (2.9)	0.16 (1.3)	1.4 (1.9)	0	27
Simulation7	**180**	10.6 (9.8)	18.9 (2.3)	0 (1.4)	1.2 (2)	0	15
Simulation8	**200**	10.6 (9.8)	18.9 (2.3)	0 (1.4)	1.2 (2)	0	15