Journal of Pharmaceutical & Biomedical Analysis
Vol. 7, No. 2, pp. 139–145, 1989
Printed in Great Britain

0731-7085/89 $3.00 + 0.00
© 1989 Pergamon Press plc

Bioassays in whole animals*

JACOBUS VAN NOORDWIJK

Taveernelaan 15, NL–3735 KA Bosch en Duin, The Netherlands

Abstract: The basis for whole animal bioassays as practised today was laid down by Paul Ehrlich in 1894. He introduced the concepts of a stable standard preparation and of the unit of activity as the activity of a defined mass of that standard preparation in the assay performed. Such assays have often provided a way of quantifying newly discovered active principles of biological origin, so that they could be applied in clinical medicine. Whole animal bioassays can be applied not only for the quantitative analysis of a biological product (analytical assays), but also for the comparison of different products intended for the same clinical indication (comparative or research assays). As such they have been the model for controlled clinical trial. For some products many different types of bioassay have been developed. They may produce heterogeneous results when more than one active principle is involved, and these are present in standard preparation and in the unknown preparation in different relative concentrations. In addition the precision of different assay methods for the same substance may vary markedly. An important source of variation in whole animal bioassays is the influence of some environmental factors on the individual animals during the assay. Thus the number of rats per cage markedly influences the variation of the response of rats to serum gonadotrophine. Careful studies are required to detect the discriminatory environmental factors in a particular whole animal bioassay. Keeping these discriminatory factors constant at their optimal level may increase the precision of the assay markedly and thus reduce the number of animals required to attain the precision specified, e.g. in pharmacopoeial tests.

Keywords: *Bioassay; whole animal; environmental temperature; influence of precision of bioassays; gonadotrophines.*

Introduction

Bioassays in whole animals belong to the oldest techniques for measuring quantities that could not be measured with simple physical or chemical methods. Finney [1] pointed out that reference to the basic principles of such an assay was made long ago:

> "And it came to pass at the end of forty days, that Noah opened the windows of the ark which he had made:
> And he sent forth a raven, which went to and fro, until the waters were dried up from off the earth.
> And he sent forth a dove from him, to see if the waters were abated from off the face of the ground;
> But the dove found no rest for the sole of her foot, and she returned unto him into the ark, for the waters were on the face of the whole earth: then he put forth his hand, and took her, and pulled her in unto him into the ark.

*Presented at the Symposium on "Biomolecules — Analytical Options", May 1988, Sollentuna, Sweden.

> And he stayed yet other seven days; and again he sent forth the dove out of the ark;
> And the dove came in to him in the evening: and lo, in her mouth was an olive leaf pluckt off; so Noah knew that the waters were abated from off the earth (Genesis 8:6–11)"

Finney recognizes this as an excellent account of an assay that, although only qualitative in character, has features of modern quantal response techniques:

> "The three essential constituents of an assay, stimulus and dose (depth of water), subject (the dove) and response (the plucking of an olive leaf), are well described. Knowledge of the response enabled Noah to estimate, or rather to place an upper limit on the size of the stimulus. His animal house could not provide the replication that would today be recommended, but in other respects his assay was admirable."

A biological assay of another type which has become famous was performed in 1747 by James Lind, a physician in the British Royal Navy. He used as his subjects patients with frank scurvy to test the antiscorbutic efficacy of cider, vitriol, vinegar, sea-water, oranges and lemons, garlic and mustard. The efficacy of the oranges and lemons resulted in the introduction of lemon juice in the diet of the British Navy 50 years later [2]. However, it was only in 1894 that Paul Ehrlich laid the basis for the type of whole animal bioassay that we use today [3]. Four years earlier, Behring had developed diphtheria antitoxin, but the attempts to measure the potency of preparations of antitoxin were unsuccessful. An appeal was made to Paul Ehrlich. Attempts to define a minimum lethal dose of diphtheria toxin for guinea pigs ran aground not only on the variability of the guinea pig, but also on the finding that diphtheria toxin may contain a component which is not toxic itself, but which may neutralise diphtheria antitoxin to a varying extent. Ehrlich solved this problem by showing that diphtheria antitoxin was stable and that a sample of this could be used as a standard against which other batches of antitoxin could be measured. In addition he made two other fundamental contributions to biological standardisation. He introduced the concept of the unit of activity as the activity of a well defined mass of the standard preparation under the conditions of the assay (which implies that an assay of an unknown sample must include one or more doses of the standard preparation). He also pointed out that a standard preparation must be fixed and stable. As liquid preparations of antitoxin tend to lose their potency, Ehrlich kept his standard in a dry form, at low temperature *in vacuo*.

Classification of bioassays

Whole animal bioassays may be classified according to the following three criteria.

Absolute versus relative assays

In absolute assays no standard preparation is employed; in relative assays the potency of the unknown sample is compared with that of the standard preparation in the same assay. Noah's dove assay is an absolute assay.

In such an assay the potency is expressed in a biological unit, such as a frog unit of digitalis leaf powder: the minimum amount per kg body weight required to kill a frog. Many of our classical units are in fact biological units, such as the foot and the horse power. More than 30 years after Ehrlich laid down the principle of the relative assay, Burn, Finney and Goodwin still had to point out that absolute assays are unreliable [4], because the standard animals which they imply do not exist. Nowadays an absolute assay is unacceptable.

Direct versus indirect assays

In a direct assay, repeated doses of the unknown sample are administered until a threshold effect is obtained [5]. The standard preparation is administered in the same way, and the potency of the sample under test follows directly from the ratio of the doses of standard and unknown required. Historical examples of a direct assay are the assay of digitalis on the cat and the rabbit head drop method for tubocurarine. However, practically all bioassays nowadays are indirect assays.

Indirect assays can be divided into three groups [5]: (a) assays in which each subject receives the whole series of doses of the unknown and of the standard preparation included in the assay scheme. The chicken blood pressure assay for oxytocin is an example; (b) assays in which each subject receives only one dose of the unknown sample or the standard, after which a graded response (e.g. body weight) is recorded; (c) assays in which each subject receives only one dose of the unknown or the standard, and in which an all-or-none response is recorded. The fraction of subjects responding to each dose is used to compute the potency. A well known example is the mouse convulsion assay for insulin.

The purpose of the assay: analysis or comparison

The purpose of an analytical assay is to estimate the concentration of active principle in the unknown sample, and this type of assay is the subject of our symposium. Ehrlich's assay of diphtheria antitoxin is a classical example of this type of assay. On the other hand the purpose of a comparative or research assay is to estimate the activity of a product by comparing it with that of one of the same nature, or intended for the same application. James Lind's assay was of this type.

Gaddum pointed out that the relative direct and indirect assays described above have been the prototypes of the controlled clinical assays developed in the past 50 years [6], and this applies in the first place to bioassays in whole animals. Comparative assays fulfil an important rôle in pharmaceutical research as pointers to important new products.

The merits of whole animal bioassays

A bioassay in whole animals often constitutes the first possibility for a quantitative estimate of a new active principle of biological origin.

In many cases the effect used in the bioassay on whole animals is more closely related to the activity which the clinician is interested in than the effect employed in an *in vitro* assay.

Moreover bioassays in whole animals often require only simple techniques to observe and record the response, such as weighing an animal or an organ.

Finally, some bioassays in whole animals enable the assayist to measure the responses of hormones with a diffuse site of action, such as insulin and calcitonin. On the other hand whole animal bioassays are often less precise per observation than *in vitro* bioassays; this implies that whole animal bioassays require more observations per assay than *in vitro* assays to obtain the same precision.

The problem of multiple choice in bioassays

Sometimes the assayist has a choice of bioassay methods for the same substance, e.g. the chicken blood pressure method, the rat uterus method and the guinea pig uterus

method for oxytocin. Can he or she expect to get homogeneous results with all these methods?

Theoretically standard and unknown samples in a true analytic assay may be regarded as dilutions of each other with respect to all active ingredients (hence the term "dilution assays"). This implies that strictly speaking an assay may be regarded as an analytic assay only when the unknown sample is assayed against a standard with an identical active principle. In this case it does not matter which biological system is used for the standard and the sample under test: the relative potency will be constant.

However, the pharmacokinetic profile of standard and sample under test may differ due to differences in their content of so-called "inactive ingredients", which in fact influence the rate of access to the site of action in whole animal bioassays. For this reason the potencies of digitalis preparations, assayed by intravenous injection in cats, differed from the potencies found on injection in the lymph sacs of frogs.

Discrepancies may also occur when more than one active principle is involved, the ratio of these in the unknown is different from that in the standard, and the biological indicators used in two types of bioassay have different relative sensitivities for these active principles [7]. This may occur in the case of hormones as discussed by Dr Bangham [8].

Of course it does not mean that different assay methods invariably yield heterogeneous results. In the collaborative study to establish a European Pharmacopoeia Reference Preparation for Insulin (a mixture of porcine and bovine insulin), use of the rabbit blood glucose method, the mouse blood glucose method and the mouse convulsion method resulted in homogeneous estimates [9].

Which method should the analyst choose, when different methods yield homogeneous estimates of potency? Not only the estimate of potency should be examined, but also its variance. In the parallel line assay, in which the response is a linear function of the logarithm of the dose, the variance of the log potency or var(M) is given approximately [1] by:

$$\mathrm{var}(M) = (s^2/b^2) \times [(1/N_s) + (1/N_u)]$$

in which s = standard deviation of the response; b = slope of the regression line; N_s = number of subjects receiving standard; N_u = number of subjects receiving unknown.

The ratio s/b, sometimes indicated by λ, is used as an index of precision. The smaller this index, the higher the precision of the assay method. This appears to be somewhat illogical, and therefore Gaddum [5] proposed to use not $\lambda = s/b$, but $L = b/s$.*

Table 1 shows some of the indices of precision which Bliss [10] and Gaddum [5] found in different types of assay for the vitamins A, C and D and for oxytocin, respectively. All methods for the vitamins shown are whole animal methods, and they differ widely in precision. In the case of oxytocin the whole animal chicken blood pressure method is much more precise than the two *in vitro* methods listed. An extremely interesting observation with respect to the variance of the response in biological experiments in whole animals was made by Chance in 1956 [11]. Chen *et al.* [12] had reported in 1943, that changing the environmental temperature influenced the mean effective dose of many drugs in mice. On reviewing their results, Chance noted that changing the environmental temperature had affected not only the mean effective dose, but also its variance. Next, Chance reviewed his own work from 1946, in which he had shown that

*Gaddum did not give his index a name; it might be called an index of the "signal to noise" ratio.

Table 1
The precision of different bioassays for one product

Product	Method	λ	L	Reference
Vitamin A	growth of rats	0.186	4.64	(10)
	vaginal cells	0.088	11.35	(10)
		0.063	15.87	(10)
Vitamin C	cure of scurvy	0.059	16.95	(10)
	dental structure	0.204	4.90	(10)
	odontoblast length	0.164	6.10	(10)
Vitamin D	epiphyseal line	0.22	4.64	(10)
	X-ray (Bourdillon)	0.175	5.71	(10)
	bone ash (21 days)	0.206	4.85	(10)
	bone ash (42 days)	0.095	10.53	(10)
Oxytocin	guinea pig uterus	0.2	5.1	(5)
	rat uterus	0.057	18	(5)
	blood pressure (chicken)	0.043	23	(5)

λ = index of precision, defined as s/b, in which s = standard deviation of the response; b = slope of the log dose-response regression line. L = index of precision, defined as b/s. All assays for vitamin C were performed on guinea pigs. The chicken blood pressure assay for oxytocin is an *in vivo* assay, the other two are *in vitro* assays. In all three all doses of standard and unknown are adminstered to each animal or organ preparation.

raising the temperature increases the toxicity of sympathomimetic amines in mice [13]. On reviewing his results, Chance discovered that in this case, not only the mean toxic dose, but also its variance was a function of the environmental temperature. Russell and Burch [14] in their discussion of these findings by Chance, quote Pasteur's opinion:

"Chance favours the prepared mind".

If making a discovery with its resultant euphoria is regarded as a favour, one may add that Chance's mind was apparently not prepared for this favour in 1946!

Methods to increase the precision of whole animal bioassays

So much attention is often given to the relationship between dose and effect in experiments on biological systems, that one tends to forget that the response depends not only on the dose, but also on the state of the animal.

The state of the animal in turn depends on the state of its somatic and autonomic nervous systems and its endocrine system. The state of its somatic nervous system in turn depends on the sensory input at that moment, the effects of previous experience (memory), and physical factors (including genetic factors, diet, environmental temperature, humidity, etc.). Russell and Burch [14] use the term dramatype to indicate the state of the animal resulting from its phenotype and what they call the "proximate environment"; the phenotype is the result of the genotype and developmental factors.

It is clear that such factors tend to increase the variance of the mean response of a group of animals, and that controlling this variance will tend to increase precision in whole animal experiments.

Three stages in the control of this variance can be distinguished [14]. In the first one, Fisher [15] showed that increasing the number of animals in a group reduces the variance. In the second stage, he developed efficient techniques for the analysis of

variance [15]. This made it possible to separate the contributions of diet, housing, etc. from the residual within group variance.

The third stage, which had barely begun when Russell and Burch wrote their book in 1959, consists of reducing the residual within group variance. One effective way of doing so is by improving the breeding of laboratory animals, but this is not enough.

Chance [16] studied the effect of the size of the cage, the frequency of changing the cage, and of the number of rats per cage on the variance of the response to a constant dose of serum gonadotrophine. He used the coefficient of variation of the response as a measure, i.e. the standard deviation as a percentage of the mean response. Housing in a small cage increased the coefficient of variation and so did changing the cages daily. The most striking effect was the increase in the coefficient of variation when either more or less than two rats were put together in one cage (see Fig. 1). To the author's knowledge this finding is not used at present. An interesting finding was that the number of disturbances of the rats by the experimenter had no appreciable effect. Hence Chance's conclusion that the size of the variance is directly related to the exact nature of the conditions. It is not enough to keep all environmental factors and the most important factors constant, but they should be kept constant at their optimal level.

Unfortunately few studies have been done on the reactivity of the experimental animal to environmental factors. One might argue that time has run out for such studies, as whole animal bioassays are being replaced more and more by *in vitro* assays or physico-chemical assays. However, whole animal bioassays are still carried out at present for the purpose of quality control; increasing their precision may lead to less laboratory animals being used without changing in essence the prescribed pharmacopoeial methods.

In addition, one should not forget that whole animal bioassays are used not only as analytical assays, but also as comparative assays in pharmaceutical research. Finding an unexpected influence of an environmental factor, or the unexpected absence of such an effect in comparative assays of a new product may give a new dimension to the mode of action of the new product. Now that Chance has prepared our mind, let us make a favourable use of his findings!

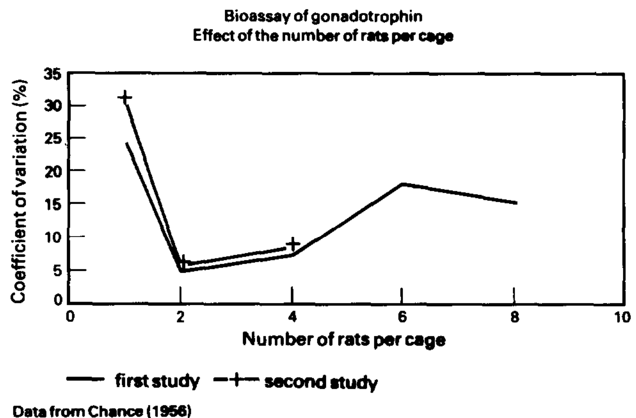

Figure 1
The effect of the number of rats per cage in the assay of serum gonadotrophine. Infantile female rats were used in this experiment, in which all rats received four injections of 50 International Units of serum gonadotrophine in four days. All the rats housed together were litter mates. On the fifth day the rats were killed and the weight of the ovaries was used as the response. Data from Chance [16].

References

[1] D. J. Finney, *Statistical Method in Biological Assay*, 3rd edn. Charles Griffin and Co., High Wycombe (1978).
[2] D. E. Danford and H. N. Munro, in *The Pharmacological Basis of Therapeutics* (A. G. Gilman, L. S. Goodman and A. Gilman, Eds), p. 1577. MacMillan, New York (1980).
[3] J. O. Irwin, *J. Hygiene* **48**, 215–238 (1950).
[4] J. H. Burn, D. J. Finney and L. G. Goodwin, *Biological Standardization*, 2nd edn. Oxford University Press, Oxford (1950).
[5] J. H. Gaddum, *Pharmacol. Rev.* **5**, 87–134 (1953).
[6] J. H. Gaddum, *Proc. Roy. Soc. Med.* **47**, 195–204 (1954).
[7] D. R. Bangham and P. M. Cotes, *Br. Med. Bull.* **30**, 12–17 (1974).
[8] D. R. Bangham, *J. Pharm. Biomed. Anal.* **7**, 169–172 (1989).
[9] D. R. Bangham, H. de Jonge and J. van Noordwijk, *J. Biol. Stand.* **6**, 301–314 (1978).
[10] C. I. Bliss, in *Vitamin Methods*, Vol. 2 (P. Gyorgy Ed.), pp. 45–275. Academic Press, New York (1951).
[11] M. R. A. Chance, *Coll. Papers Lab. Animals Bur.* **6**, 59–74 (1957).
[12] K. K. Chen, R. C. Anderson and C. A. Mills, *J. Pharmacol. exp. Ther.* **79**, 127–132 (1943).
[13] M. R. A. Chance, *J. Pharmacol. exp. Ther.* **87**, 214–222 (1946); see also *J. Pharmacol. exp. Ther.* **89**, 289–296 (1947).
[14] W. M. S. Russell and R. L. Burch, *The Principles of Humane Experimental Technique*. Methuen, London (1959).
[15] R. A. Fisher, *Statistical Methods for Research Workers*, 10th edn. Oliver and Boyd, Edinburgh (1948).
[16] M. R. A. Chance, *Nature* **177**, 228–229 (1956).

[*Received for review 4 May 1988; revised manuscript received 23 May 1988*]