

Chapter 19

Assessment of the Significance of Low Count Rates

D. E. Case, D. J. Barnfield and P. R. Reeves

Imperial Chemical Industries Ltd., Pharmaceuticals Division, Macclesfield, Cheshire, England

When considering possible errors in scintillation counting, one is naturally concerned with those errors which are associated with the measurement of levels of radioactivity in samples of various kinds. We wish to focus attention on a topic which might therefore be considered out of place in this symposium; we are concerned to define how to reach the conclusion that a sample contains no significant level of radioactivity above the background level.

Frequently, metabolism and tissue residue studies of medicinal and veterinary drugs at some stage involve the measurement of very low count rates of soft β -emitters in tissues and biological fluids. In studies designed to evaluate tissue residues, low count rates are often the norm rather than the exception.

In any overall assessment of the significance of low count rates, three questions must always be posed:

- (i) Is there a significant level of radioactivity to be found in the sample as a result of the prior administration of the labelled compound?
- (ii) Is the radioactivity present due to the drug or its metabolites, or to incorporation of the radioactive atom(s) into normal endogenous metabolic pathways?
- (iii) Is the level of radioactivity present of any pharmacological or toxicological significance?

The second question, where relevant, can lead to very extensive and complex metabolic studies.^{1, 2} The third has to be answered by logical argument based on acceptable levels in edible tissues for example, or on the known toxicological and pharmacological properties of the drug and its metabolites. We wish for the moment to consider only the first question.

The necessity to evaluate very low count rates arises either by stumbling unexpectedly into the predicament or by setting out deliberately to assay samples anticipated to contain low or zero levels.

An example of the first situation arose with 'Vivalan'^a (viloxazine), an anti-depressant whose metabolism has been studied in animals and in man.^{3, 4} This compound was labelled in three different positions with ¹⁴C. One of these radiolabelled forms gave rise to blood levels in dogs with a terminal phase having an extraordinarily long half-life.

^a 'Vivalan' is a Trade Mark, the property of I.C.I. Ltd.

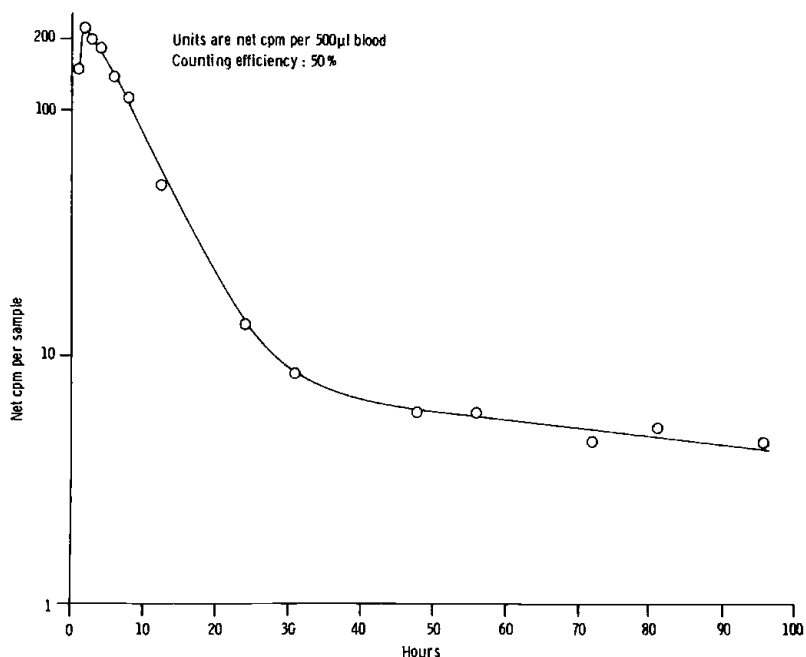


Fig. 1. ^{14}C blood level profile in an animal dosed orally with $40\ \mu\text{Ci}$.

This was shown not to occur when the ^{14}C -label was placed in a different position in the molecule and the effects are due to metabolic incorporation of the ^{14}C due to degradation of the molecule by a very minor metabolic pathway.² In this work we were unexpectedly having to measure levels of radioactivity in blood corresponding to 4–40 counts min^{-1} above background (Fig. 1), and particular attention was required in assessing the data.

Let us now consider the situation in which one deliberately sets out to measure very low or zero levels of radioactivity. Data will be presented which is derived from work with a veterinary product in cows. The objectives of the study were to monitor the elimination of the drug and its metabolites from the animal and to define the clearance from edible tissues and milk. The molecule involved could not be satisfactorily labelled with ^3H and was eventually prepared with ^{14}C at a high isotopic abundance at one position.

Preliminary studies indicated that any residues present would be of a very low order and that the count rates to be measured would be near to, or indistinguishable from, background. Our conclusion from all the data subsequently obtained was that this product did not in fact give rise to any detectable residual level of radioactive components in the edible tissues. How does one reach this conclusion?

We have been disconcerted to learn the different arbitrary methods used in some groups to define a level of significance above background. Some workers will only accept a count rate as being significant if it is twice the background count rate, or 100 counts min^{-1} above background, or at some other arbitrarily chosen level. This to us seems attractive only in the measure of convenience conferred and several of the problems we have faced in the past would have evaporated had we applied such arbitrary limits. Publications often contain the expressions 'not detectable', or 'not significant', without further explanation.

There are mathematical expressions describing the error function associated with a net sample count rate which are derived from consideration of the Poisson distribution of radioactive events. These are of value in the design of an experiment, in that such expressions predict the minimum error associated with a given set of counting parameters, but do not take into account other sources of error, introduced, for example, by the sample preparation procedures. Having performed the experiment and counted the samples, our concern becomes the *total* variance of the results, which is inevitably larger than that derived from Poisson statistics, and can only be estimated from the results observed.

To develop the problem as we see it and to present our own methods it is useful to use data from the residue study mentioned previously. Milk samples were counted in 'Instagel', at least six aliquots being counted in new vials for at least 100 min on an Intertechnique Model SL30 with a counting efficiency of 80%. Background data, generated by using predose samples from the same animal, gave comparatively high but reproducible values (Table 1).

Table 1. Mean count rates for predose milk.

Cow no.	Mean background (counts min ⁻¹) (n = 9)	S. D.
1	56.84	0.66
2	54.87	0.69
3	56.74	0.53
4	56.01	0.84
5	56.17	1.13
6	55.49	0.62
7	55.35	0.88
8	56.20	1.87
9	55.95	0.90
10	55.94	0.58

Samples contained 10 ml 'Instagel' + 10 ml milk.
Counting time: 100 min per sample.

The background data for one particular cow is shown in more detail in Table 2. The levels of radioactivity found in the milk from this animal declined rapidly after

Table 2. Cow No. 10: Background count rates (counts min⁻¹)

Sample	Count rate	Sample	Count rate
1	55.57	6	56.14
2	55.75	7	56.44
3	54.84	8	55.43
4	56.20	9	56.52
5	56.56		
Mean	= 55.94	S.D.	= 0.58

Samples contained 10 ml 'Instagel' + 10 ml predose milk.
Counting time: 100 min per sample.

dosing to levels which were clearly at or near the background value (Table 3). The mean values and standard deviations derived from this data have been calculated for clarity (Table 4). The problem remaining is to determine which samples contain levels of radioactivity which are significantly in excess of the background value. At what point in time can it be concluded that no detectable level of drug or metabolite is present?

The method employed in our laboratories to assess low count rates is based upon the assumption that the variance of the replicate blanks is the same as the variance of the assayed samples. This must be a very reasonable assumption at extremely low count rates and under conditions where rigorous attention is paid to sample preparation. This assumption cannot be made, of course, where very much higher count rates are involved.

Making the assumptions that the samples are independent and normally distributed, and that the sample standard deviation (S) estimates the population standard deviation

Table 3. ^{14}C found in milk samples (six replicate aliquots taken from each sample).

Sample	Collection period (h)	Net counts min^{-1} per aliquot	
1	0-4	256.4	266.0
		257.6	264.2
		261.0	267.3
2	4-8	182.3	195.0
		185.6	187.4
		192.9	188.4
3	8-24	30.3	29.1
		30.3	28.1
		29.8	30.6
4	24-36	5.2	6.9
		7.1	7.5
		7.6	7.7
5	36-48	0.2	1.2
		0.6	-1.7
		-0.3	0.8

Samples contained 10 ml 'Instagel' + 10 ml milk, and were counted for 100 min/ 10^4 counts.

Table 4. ^{14}C found in milk samples (summary).

Sample	Collection period (h)	Mean net counts min^{-1}	S.D.
1	0-4	262.1	4.5
2	4-8	188.6	4.7
3	8-24	29.5	1.0
4	24-36	7.0	0.9
5	36-48	0.1	1.0
6	48-60	-0.6	0.3
7	60-72	0.6	0.4

(o) with n degrees of freedom, then the variance of the difference between sample gross count and the background is equal to the sum of the variances. Thus,

Variance of the difference = variance of assayed samples + variance of blanks, i.e.

$$\frac{S_A^2}{N_A} + \frac{S_B^2}{N_B}$$

where N_A refers to the number of assayed samples and N_B refers to the number of blanks.

Hence,

$$\text{Standard error of difference} = \sqrt{\frac{S_A^2}{N_A} + \frac{S_B^2}{N_B}} \quad (1)$$

with associated degrees of freedom $(N_A - 1) + (N_B - 1)$.

Frequently there are insufficient replicates of the assayed samples to allow a reliable estimate of S_A and the use of Eqn. (1) above. However, blanks and assayed samples are prepared and processed in an identical manner and produce count rates of the same order when extremely low count rates are involved. Thus it is reasonable to assume that both have the same variance, whence, in Eqn. (1), S_A becomes equal to S_B .

Thus Eqn. (1) reduces to

$$\text{Standard error of difference} = S_B \sqrt{\frac{1}{N_A} + \frac{1}{N_B}} \quad (2)$$

where the number of degrees of freedom is $(N_B - 1)$.

A more accurate measure of error may be obtained by combining estimates of error from blanks and assayed samples when this is possible, i.e. by using the expression

$$\text{S.D.} = \sqrt{\frac{\Sigma(N_i - 1) S_i^2}{\Sigma(N_i - 1)}}$$

where N_i is the number of replicates in the i th group of observations.

Normally, however, Eqn. (2) is used and this requires a satisfactory estimation of the background variance. It is worth noting, of course, that the entire problem with which one is concerned hinges upon the accurate determination of the background count rate and its variance.

Using the milk data previously presented, the following data is used in the example which follows: number of background samples = 9, number of replicates of the assayed sample = 6, degrees of freedom = 8, mean background count rate = 55.94 counts min^{-1} , standard deviation for the latter = 0.58.

The calculation using Eqn. (2) above is thus

$$\begin{aligned} \text{Standard error of difference} &= S_B \sqrt{\frac{1}{N_A} + \frac{1}{N_B}} \\ &= 0.58 \sqrt{\frac{1}{6} + \frac{1}{9}} = 0.31 \text{ count min}^{-1} \end{aligned}$$

This has now to be multiplied by the Student's *t*-value for 8 degrees of freedom ($p = 0.01$, single sided) which is 2.90. Hence the confidence limit, i.e. the limit of detection at the confidence level specified, is 0.90 count min⁻¹ above background.

In applying this treatment to the experimental data previously presented (Table 4), we now understand that at 99% confidence limits a mean count rate 0.90 count min⁻¹ above the mean background value of 55.94 is significant. By inspection it can be concluded that net count rates in excess of 0.90 count min⁻¹ are significantly above background. From Table 4 it is clear that samples 1 to 4 contain significant levels of radioactivity and that the remainder do not.

The method of assessment described may of course be evaluated by adding known amounts of radioactivity to control blanks and determining experimentally whether the levels at which radioactivity can apparently be measured reliably agree with those predicted.

As an example, data are given in Table 5 for blood samples spiked with known amounts of the ¹⁴C-labelled compound. It will be seen that levels in excess of 2 disintegrations min⁻¹ can apparently be detected reliably. Analysis of the background data (Table 6) indicates a limit of detection corresponding to 1.84 disintegrations min⁻¹. This is consistent with the experimental results (Table 5) which become unreliable below this level. These data were generated using an oxygen flask combustion method based upon that of Kalberer and Rutschmann,⁵ the samples being counted for 100 min in a Packard 2211 Tri-Carb Scintillation Counter.

Table 5. Experimental determination of limits of detection, 0.5 ml whole blood spiked with ¹⁴C standard.

Nominal dpm added	Net dpm found
0.5	-0.3
1.0	-0.2
2.0	2.4
5.0	6.0
10.0	9.8

The procedure described for blood was repeated for several tissues, with the results presented in Table 7. By calculation of the limit of detection from blank data, and comparison with the 'dpm found' figure, one may determine, for each tissue, the standard for which a significant result was just obtained. The experimentally determined limit of detection must thus lie between this standard and the next smaller one (for

Table 6. Limit of detection predicted from blank data.

Number of samples = 6	Combustion aliquot factor = 1.5
Counting time = 100 min	Mean counting efficiency = 70%
Mean blank = 29.44 ± 0.72 counts min^{-1} (S.D. : $n = 19$)	
Student t -value ($p = 0.01$), with 18 degrees of freedom = 2.55	
Calculated limit of detection	
$= 2.55 \times 0.72 \sqrt{\frac{1}{19} + \frac{1}{6}} \text{ counts min}^{-1}$	
$= 0.86 \text{ net counts min}^{-1} \text{ per aliquot} = 1.29 \text{ net counts min}^{-1} \text{ per sample}$	
$= 1.84 \text{ dpm per sample}$	

Table 7. Experimental estimation of limits of detection.

Standard added (dpm)	dpm found in spiked sample ^a				
	Blood	Muscle	Liver	Kidney	Fat
0.5	-0.3	NA	NA	NA	NA
1	-0.2	NA	NA	NA	NA
2	2.4	5.9	1.9	2.6	2.3
5	6.0	7.4	5.0	5.1	5.3
10	9.8	11.5	8.6	11.3	11.2
25	NA	27.0	25.2	26.0	23.3
Calculated limit of detection from blanks (dpm)	1.8	3.2	2.0	2.8	2.9
Actual dpm in '1st sig. sample'	2	2	5	5	5
Actual dpm in 'last non-sig. sample'	1	(1)	2	2	2

^a All results are means of 6 replicates, each counted for 100 min.

NA: Not assayed.

which a significant result was not quite obtained). It can be seen that, with the exception of muscle, the calculated limit of detection lies in this range and that the practically and theoretically estimated limits are in good agreement.

The result for muscle exemplifies the importance of the questions mentioned earlier. For the 2 disintegrations min^{-1} standard, there is a significant result. We are more fortunate than usual in being able to state that it is obviously inaccurate, but the result remains significant nonetheless.

Knowing the correct answer, one is in a position to question the source of this inaccurate result. Is the error due to a faulty standard preparation, contamination or an inappropriate background determination?

Had it not been a standard, then subject to the latter two qualifications it would have been accepted at face value. The ultimate importance of this result would then

have been evaluated in terms of its biological significance. The answers to such questions will be obtainable not from statistical considerations but from commonsense and experience.

When concerned with problems of very low count rates, one may be exhorted to remove the problem by using a higher specific activity, a higher dose or a different labelled isotope. This, unfortunately, is not always possible.

In some circumstances, one may be in the fortunate position of being able to predict, from toxicological data, or guidelines laid down by regulatory authorities, the minimum tissue levels which may be acceptable. This enables one to calculate a corresponding level of radioactivity below which the corresponding levels are of no concern. If this level is sufficiently removed from background values, any statistical analysis may be dispensed with.

We would like to consider that scintillation counting is a technique amenable to statistical analysis and to base our calculations upon a rational treatment of the data. It is fascinating to find that very little has been published in this area and we would welcome discussion of the method presented and of any other treatment currently used in assessing the significance of very low count rates.

ACKNOWLEDGMENT

The authors wish to acknowledge the invaluable advice and continuing assistance of S. H. Ellis of the Statistics Group, I.C.I. Pharmaceuticals Division.

REFERENCES

1. J. C. Potter, J. E. Loeffler, R. D. Collins, R. Young and A. C. Page, *J. Agric. Fd. Chem.* **21**, 163 (1973) and references cited therein.
2. D. E. Case, *Xenobiotica* **5**, 133 (1975).
3. D. E. Case, H. Illston, P. R. Reeves, B. Shuker and P. Simons, *Xenobiotica* **5**, 83 (1975).
4. D. E. Case and P. R. Reeves, *Xenobiotica* **5**, 113 (1975).
5. F. Kalberer and J. Rutschmann, *Helv. Chim. Acta* **44**, 1956 (1961).

DISCUSSION

L. A. Currie: With regard to measurement of detection limits and statistical detection powers and errors of the first and second kind, my congratulations on a rational and objective approach to the problem of detection in counting. It would also be valuable to consider the question of detection power, i.e. given a certain level of radioactivity, what is the probability of detecting it according to your detection criterion? In this way one may encompass both errors of the first kind (probability of false detection) and of the second kind (probability of false non-detection) in identifying the detection limit.

P. Johnson: You began by showing a slide in which you briefly mentioned the possible significance of low count rates in tissues and body fluids, although your talk rightly concentrated on the main topic under consideration, which is the techniques involved in measuring and interpreting low count rates. You also emphasised that not everyone can attain the high specific activities or administer the high doses that they might wish to in order to obtain reasonable count rates, such as in a residue study, and that assessment of low count rates is then essential. I agree. However, there is also a danger, which I feel should be emphasised, when material of high specific activity or high doses can be

used. This danger is that, in setting a very low absolute level of detectable counts, data will be generated from which a legislative authority may give a significance to a few photons detected in a scintillation counter which is out of all proportion to the biological significance; in fact, the biological significance may be zero.

D. E. Case: In reply to your last comment, I must say that I agree entirely. In all studies of this kind the biological and toxicological significance of the results must be assessed rationally.

Often, the amounts detected may be totally insignificant in terms of safety evaluation. However, in some circumstances, we may be asked, having concluded that no detectable level is present, what level *could* be present which you are not quite able to detect by your methods? We feel, therefore, that some method of assessment is necessary; whatever the outcome, commonsense must be used in evaluating the real toxicological significance of the results.

B. R. Twite: The low count levels in the milk samples occurred 24 h or longer after dosing the radiolabelled drug. This means that your control (blank) samples were 24 h older than the counted samples. At the very low count levels measured, perhaps the difference in age between the samples could be responsible for the difference in counts observed. As it is possible to obtain a statistical difference in counts between samples which in fact does not exist, is it not more valid to quote residue levels as less than a certain value, the value being set to give a greater certainty that a true difference in counts does exist (e.g. twice background)?

D. E. Case: We concluded that the use of predose milk taken from each cow was the best possible blank; it is, of course, impossible to determine the absolute blank at a time point after administration of the labelled compound. It was also found that background data for different cows (Table 1) were remarkably consistent despite the fact that these samples were of slightly different ages when actually counted. I would object to your proposal to use a 'certain value' (e.g. twice background) as this falls into the arbitrary kind of practice which we find difficult to justify on any theoretical grounds.

J. E. Noakes: Milk samples are analysed for natural radiocarbon content at the University of Georgia by freeze drying and pallet spraying and combusting samples to CO_2 followed by conversion to C_6H_6 . Perhaps greater sensitivity could be gained in your present mode of counting raw milk samples if enough material is present to concentrate the samples in this way. If a combustion apparatus is available using sample conversion to CO_2 and collecting in an amine (2-methoxyethylamine), this could also be counted with good efficiency and perhaps with lower background, thereby increasing the sensitivity of your analytical method.