Paper

EVALUATION OF EIGHT DECISION RULES FOR LOW-LEVEL RADIOACTIVITY COUNTING

Daniel J. Strom* and Jay A. MacLellan†

Abstract—**In low-level radioactivity measurements, it is often important to decide whether a measurement differs from background. A traditional formula for decision level (*DL*) is given in numerous sources, including the recent ANSI/HPS N13.30-1996, *Performance Criteria for Radiobioassay* and the *Multi-Agency Radiation Survey and Site Investigation Manual (MARSSIM)*. This formula, which we dub the N13.30 rule, does not adequately account for the discrete nature of the Poisson distribution for paired blank (equal count times for background and sample) measurements, especially at low numbers of counts. We calculate the actual false positive rates that occur using the N13.30 *DL* formula as a function of *a priori* false positive rate α and background Poisson mean $\mu = \rho t$, where ρ is the underlying Poisson rate and t is the counting time. False positive rates exceed α by significant amounts for $\alpha \leq 0.2$ and $\mu < 100$ counts, peaking at 25% at $\mu \cong 0.71$, nearly independent of α. Monte Carlo simulations verified calculations. Currie's derivation of the N13.30 *DL* was based on knowing a good estimate of the mean and standard deviation of background, a case that does not hold for paired blanks and low background rates. We propose one new decision rule (simply add 1 to the number of background counts), and we present six additional decision rules from various sources. We evaluate the actual false positive rate for all eight decision rules as a function of *a priori* false positive rate and background mean. All of the seven alternative rules perform better than the N13.30 rule. Each has advantages and drawbacks. Given these results, we believe that many regulations, national standards, guidance documents, and texts should be corrected or modified to use a better decision rule.**
Health Phys. 81(1):27–34; 2001

Key words: statistics; radioactivity; bioassay; dosimetry, internal

INTRODUCTION

SETTING the decision level for bioassay excreta analyses is an important function of the internal dosimetry program. If it is too high, potentially significant intakes will be missed. If it is too low, resources will be wasted on unnecessary resampling and reanalyses. The sampled population may also lose confidence in the program if they suspect the internal dosimetrist is "fishing" for the right answer.

This research was prompted by the realization that applying traditional decision rules to very low count rate data was giving unexpectedly high false detection rates. As radiation detection technology has improved, background count rates in alpha spectroscopy have dropped to levels not anticipated decades ago. One radiobioassay service contractor routinely reports 0 or 1 count in the ^{239}Pu α-window in 2,500 min, with an average near 0.7. This phenomenal capability leads to a need to examine decision rules for distinguishing activity from background when background is very low.

When counting particles, such as in alpha spectroscopy for measurement of ^{239}Pu, one typically subtracts an estimate of background counts from the counts of a sample. The resulting difference or net count value can then be compared to a statistic called decision level, *DL*. If the net count value is greater than the *DL*, then one makes the decision that there is activity present above background. One formula for *DL* is (HPS 1996)

$$DL_{\mathrm{N13.30}}(N_{\mathrm{b}}, \alpha) = k_{\alpha}\sqrt{2N_{\mathrm{b}}}, \tag{1}$$

where k_{α} is found from the cumulative Normal distribution:

$$1 - \alpha = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{k_{\alpha}} \mathrm{e}^{-x^2/2}dx, \tag{2}$$

and N_{b} is the observed number of background counts. In eqn (1), $2N_{\mathrm{b}}$ is assumed to be a good estimate of the variance of the net number of counts.

The probability of observing N counts when the underlying mean $\mu = \rho t$, where ρ is the underlying Poisson rate (e.g., counts per second) and t is the counting time (e.g., seconds), is given by the Poisson distribution,

$$\mathrm{Poi}(N|\mu) = \frac{e^{-\mu}\mu^N}{N!}. \tag{3}$$

* Risk Analysis and Health Protection, Pacific Northwest National Laboratory, Richland, WA 99352-0999; † Radiation and Health Technology, Pacific Northwest National Laboratory, Richland, WA 99352-0999.

For correspondence or reprints contact: D. J. Strom, Risk Analysis and Health Protection, Pacific Northwest National Laboratory, Richland, WA 99352-0999, or email at daniel.j.strom@pnl.gov.

(*Manuscript received* 10 *March* 2000; *revised manuscript received* 12 *December* 2000; *accepted* 24 *March* 2001)

0017-9078/01/0

Note that while N is an integer, μ is a non-negative real number.

It is important to verify that the DL employed is providing the desired results, and one method of doing this is suggested in the "Recommendations" section. The DL is given by

$$DL(N_n) = k_\alpha \sqrt{\sigma_g^2 + \sigma_b^2} = k_\alpha \sqrt{N_g + N_b}, \quad (4)$$

and when there is no analyte activity in the sample, by

$$DL_{N13.30}(N_b, \alpha) = k_\alpha \sqrt{2\sigma_b^2} = k_\alpha \sqrt{2N_b}, \quad (5)$$

as shown in eqn (1). We refer to eqn (5) as the "N13.30 DL" for counts.

An empirical approach to define the decision level has been used by some organizations. The empirical approach involves evaluating the actual net activity distribution for a large set of analytical blanks and selecting the count or count rate that corresponds to the selected false detection rate. We expect to examine this approach in future work as a potential solution to the difficulties identified here.

METHODS

Two methods were used to determine the actual false positive rates α' when eqn (1) is used for the paired blank counting problem.

Monte Carlo simulation

The first method was a Monte Carlo simulation. For each of 57 values of μ (0.01 through 100), a Poisson distribution was randomly sampled. This value was stored as the background observation. Then the same distribution, this time representing an unknown containing no analyte, was sampled again and stored as the unknown. A DL was computed using eqn (1), and the net rate (i.e., unknown—background) was compared to it. If the net result was greater than or equal to the DL, for that α, then the decision was "analyte activity was detected above background." All such decisions are false positives, since there is no net activity present. This procedure was repeated 10^6 times for each mean and for each of 18 values of α (0.5, 0.2, 0.1, 0.05, etc., down to 10^{-6}). The results were slightly noisy, but were in exact agreement with the analytical method described below.

Analytical solution

The cumulative Poisson distribution up through M is the sum of the Poisson distribution values:

$$\text{CumulPoi}(M, \mu) = \sum_{N=0}^{M} \text{Poi}(N, \mu). \quad (6)$$

The function Trunc(x) returns the integer part of non-negative real number x. The false positive rate for a Poisson mean μ and so-called Type I error probability α is given by summing over nonnegative integers N of the product of two probabilities: the probability of observing a background value of N counts given a Poisson mean of μ; and the probability of observing more than N plus the expected background counts in the sample count. The later probability is simply one minus the cumulative Poisson distribution up to $[N + DL(N, \alpha)]$. In symbolic terms, we have the actual false positive rate α' as

$$\alpha'(\mu, \alpha) = \sum_{N=0}^{\infty} \text{Poi}(N, \mu)(1 - \text{CumulPoi}\{[\text{Trunc}(N + DL(N, \alpha))], \mu\}), \quad (7)$$

a result we term "MacLellan's exact calculation."

RESULTS

Actual false positive rates from eqn (5), when counting blanks, are plotted in Fig. 1. The horizontal axis is the long-term number of mean background counts, μ_b, that one is trying to estimate when counting a reagent blank. Here, $\mu_b = \rho_b t_b$, where ρ_b is the background rate and t_b is the background count time. When subsequent blanks are counted, of course, any and all decisions that analyte activity has been detected are wrong decisions, i.e., "false positives." If the decision that analyte activity has been detected is based on the use of the ANSI N13.30 decision level, the actual false positive rates are shown on the vertical axis for various levels of the acceptable Type I error rate α. If the ANSI N13.30 formula were correct, each curve would be a horizontal line equal to the value of α, independent of background rate. Clearly, the N13.30 DL formula does not give good actual false positive rates for low numbers of observed background counts.

Fig. 1 shows that the actual false positive rate is essentially independent of α below 0.3 counts, and if $\alpha \leq 0.2$, this is true almost up to $\mu_b = 1$. Fig. 1 also shows that for values of α of 0.1 or less, the claimed false positive rate, that is, α, is not even achieved with a background value μ_b of 100! For very tiny values of α, the claimed false positive rate is not even close to α. For α equal 10^{-6}, the false positive rate at 100 background counts is 25.1×10^{-6}. For $\alpha \leq 0.1$, the maxima of the curves are about 0.25, and occur near $\mu_b = 0.7$ to 0.72 counts, depending only weakly on the value of α. In the interval $0.3 \leq \mu \leq 1.3$, the false positive rate is above 0.2 regardless of the value of α.

The false positive rate for very low background rates using the N13.30 DL is due almost entirely to the probability of observing zero background counts. Regardless of the α value applied, the square root of zero is zero, and the DL is zero. Therefore, any observed count is interpreted as "detected." For very low background rates, e.g., 0.01, one observes zero in about 99% of cases and one in the other 1% of cases. Similar rates pertain for a blank about which one is trying to make an inference. Thus for those 1 in 100 blanks for which one observes 1 count, the probability is 99% that the paired background measurement will have been 0, and that a false positive decision will be made. The false positive rate is then

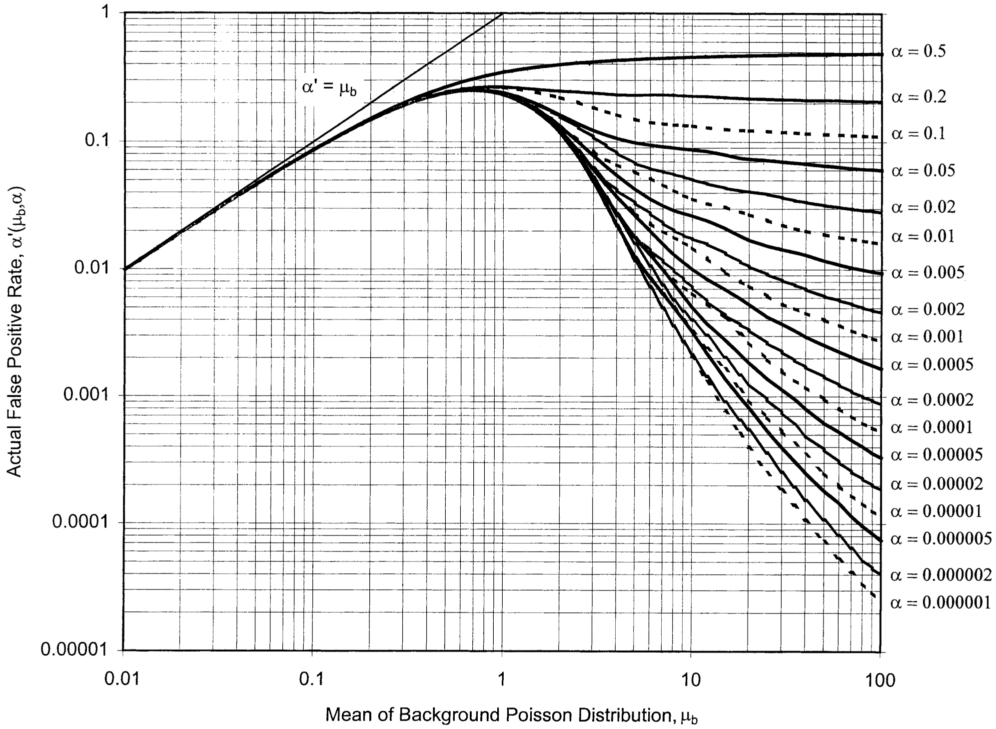

Fig. 1. Actual false positive rate α' for the N13.30 decision rule as a function of background count mean μ_b for 18 values of *a priori* false positive rate α. A perfect decision rule would result in horizontal lines with constant $\alpha' = \alpha$.

0.99×0.01, or approximately 0.01. For very low background rates the probability of observing a false positive with the N13.30 rule is approximately equal to the probability of observing one or more counts in the counting period.

APPLICATIONS OF THE TRADITIONAL FORMULA WITHOUT CONSIDERING ITS LIMITATIONS

Eqn (1) was popularized by, and is generally ascribed to, Currie (1968). It appeared earlier in a more general form for net count rate R_n (as opposed to net counts) with count times not necessarily equal, as Rule D_2 in Nicholson (1963),

$$DL_{N13.30}(R_n, \alpha) = k_\alpha \sqrt{\frac{N_b}{t_b}\left(\frac{1}{t_b} + \frac{1}{t_g}\right)}, \tag{8}$$

where t_b is the background count time and t_g is the gross count time. This equation has appeared in one form or another in countless publications including ANSI/HPS N13.30-1996 and MARSSIM (Brodsky 1986; Currie 1968; Currie 1984; Health Physics Society 1996; Hickey et al. 1993; Lochamy 1976; Strom and Stansbury 1992; U.S. NRC 1997). Given this history, we have arbitrarily chosen to dub eqn (8) the "N13.30" decision rule for count rate.

The problem arose from the assumption that one has a well-known estimate of the mean and standard deviation of the background. With low background rates one does not have a good estimate of either one. Currie's treatment of the "paired blank" (equal background and gross count times) case attempted to account for the increased uncertainty in the background when it was counted for only as long as the sample (as opposed to the well-known blank). But when the background rate is estimated from a measurement that is below the long-term mean, the use of Currie's *DL* (termed "critical level," L_C, in his 1968 paper) causes a large number of false positive decisions that are not offset by the fewer false positive decisions that result when the background rate is estimated from a measurement that is above the long-term mean. In particular, if the background is estimated from an observation of zero, one must decide that any gross count $N_g \geq 1$ results in a decision of "analyte activity has been detected above background."

Currie stated on page 22 of his NUREG document (1984) that an assumption underlying the N13.30 *DL* rule is that the estimated net signal is an independent random variable having a known distribution. Thus, knowing (or having a statistical estimate for) the standard deviation of the estimated net signal, one can calculate the *DL* given the distribution and alpha. He also stated on page 49 of the same document that if there are at least 5 counts in the background estimate, use of the Poisson variance as the estimate of the population variance is valid. Applying the rule to very low background count rates violates Currie's own assumptions for the *DL*.

Both Currie (1968, 1984) and HPS (1996) acknowledge limitations at low numbers of counts, but do not give any idea how good (or bad) the formula might be at low numbers of counts. Furthermore, a commonly-held notion that a Poisson distribution is well-approximated by a Normal distribution at $\mu \geq 30$ does not explain our results even at μ of 100 for small values of α. Even at higher background levels, the N13.30 *DL* gives a false positive rate $\alpha' > \alpha$ because the estimate of the mean (and therefore the estimate of the variance) of the distribution is biased low. The factor by which α' using the N13.30 *DL* overestimates α is particularly large for small α.

SEVEN OTHER DECISION RULES

Following our conclusion that the N13.30 decision rule does not give a good estimate of false positive results, seven other decision rules were investigated.

Most probable value of mean and variance

Rainwater and Wu (1947) showed that the most probable values of the mean and variance are not the observed value of the mean, but a value larger than the observed value. Although not intuitively obvious, an example was given for clarification. If zero is observed, the mean is not necessarily also zero; therefore, the average value of the mean that produces zero observations must be greater than zero, and the most probable value of the mean is larger than the observed value.

One formal way of addressing this is to use a uniform (uninformative) Bayesian prior probability distribution, which yields the result that the expectation value of background when N_b counts are observed is $N_b + 1$ (Friedlander and Kennedy 1955; Friedlander et al. 1963; Stevenson 1966; Little 1982). The variance is also $N_b + 1$. The Bayesian posterior probability density functions for observations of $N_b = 0, 1, 2, 3,$ and 4 are shown in Fig. 2. For each observation, the maximum likelihood value, i.e., the mode, is N_b, as predicted by classical statistics, but the expectation value, i.e., the mean, is $N_b + 1$.

The uniform prior Bayesian approach leads to a decision level for the net count rate of

$$DL_{N+1}(R_n, \alpha) = k_\alpha \sqrt{\frac{(N_b + 1)}{t_b}\left(\frac{1}{t_b} + \frac{1}{t_g}\right)}. \qquad (9)$$

If the observed value N_b is much greater than one, the distinction between N_b and $N_b + 1$ is not important. One approach we took was to modify the N13.30 rule by using $N_b + 1$ as the estimate of the variance instead of N_b.

Approach of Altshuler and Pasternak, and Turner

The decision level may also be defined in terms of the standard deviation of the net analyte activity. In his book *Atoms, Radiation, and Radiation Protection*, Turner (1995) describes a decision level similar to one originally proposed by Altshuler and Pasternack (1963). In this process, the decision is made on the basis of the

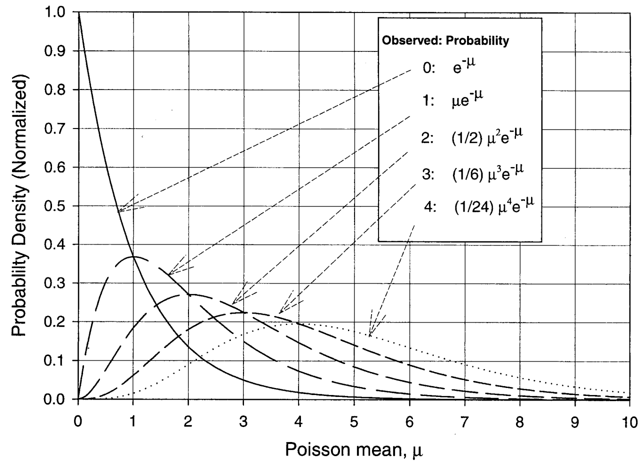

Fig. 2. Posterior probability densities for the background count mean $\mu_b = \rho_b t_b$ for observations of $N_b = 0, 1, 2, 3, 4,$ and 5 counts, derived using Bayes's theorem and a uniform (uninformative) prior. Formulas for each probability density are given in the legend. In each case, the most probable (maximum likelihood) values of μ_b are equal to N_b, but the expectation values (means of the posterior distributions) are equal to $N_b + 1$.

difference in the gross (R_g) and background (R_b) count *rates*, the net count rate R_n:

$$\frac{N_g}{t_g} - \frac{N_b}{t_b} = R_g - R_b = R_n = k_\alpha \sqrt{\sigma_{gr}^2 + \sigma_{br}^2}$$

$$= k_\alpha \sqrt{\frac{R_n + R_b}{t_g} + \frac{R_b}{t_b}}. \qquad (10)$$

Here, σ_{gr}^2 and σ_{br}^2 are the variances of the gross and background count rates, respectively. This is equivalent to Currie's detection level (minimum detectable count), when the decision level is set to zero. Solving the expression for R_n gives

$$DL_{A\&P/Turner}(R_n, \alpha) = \frac{k_\alpha^2}{2t_g} + \frac{k_\alpha}{2}\sqrt{\frac{k_\alpha^2}{t_g^2} + 4R_b\left(\frac{1}{t_b} + \frac{1}{t_g}\right)}. \qquad (11)$$

When $t_b = t_g$, the minimum significant count difference, Δ_1, is

$$\Delta_1 = k_\alpha \sqrt{2N_b}\left(\frac{k_\alpha}{\sqrt{8N_b}} + \sqrt{1 + \frac{k_\alpha^2}{8N_b}}\right). \qquad (12)$$

McCroan's rule

McCroan[‡] has developed a rule similar in form to that of Altshuler and Pasternak/Turner given in eqn (11), that can be derived from Nicholson's D_3 rule (eqn 22, below):

$$DL_{McCroan}(R_n, \alpha) = \frac{k_\alpha^2}{2t_b} + \frac{k_\alpha}{2}\sqrt{\frac{k_\alpha^2}{t_b^2} + 4R_b\left(\frac{1}{t_g} + \frac{1}{t_b}\right)}. \qquad (13)$$

[‡] Personal communication, Keith McCroan to Daniel J. Strom, 1 September 1999.

Binomial distribution

Nicholson (1963, 1966) and Sumerling and Darby (1981) describe equivalent processes for determining a decision level using the binomial distribution. As described in Sumerling and Darby (1981), if the background is not well known, the observation on the sample may be compared with the background observation, N_b counts in time t_b, to see if they are consistent with any single true count rate. They, and others, argue that using *both* the background and gross sample measurements to estimate the background increases the power of the test.

The probability of observing value N_g, when the mean of the quantity being measured is μ_g, is given by the Poisson distribution

$$\mathrm{Poi}(N_g|\mu_g) = \frac{e^{-\mu_g}\mu_g^{N_g}}{N_g!}. \tag{14}$$

The joint probability of making independent observations N_g and N_b when the respective means are μ_g and μ_b is given by

$$P(N_g, N_b|\mu_g, \mu_b) = \frac{e^{-\mu_g}\mu_g^{N_g}}{N_g!} \cdot \frac{e^{-\mu_b}\mu_b^{N_b}}{N_b!}. \tag{15}$$

Transforming to new variables N_g and $N_{total} = N_g + \mathrm{N_b}$,

$$P(N_g, N_{total}|\mu_{total}, Q) = \frac{e^{-\mu_{total}}\mu_{total}^{N_{total}}}{N_{total}!}\binom{N_{total}}{N_g}Q^{N_g}(1-Q)^{N_{total}-N_g}, \tag{16}$$

where $\mu_{total} = \mu_g + \mu_b$ and $Q = \mu_g/(\mu_g + \mu_b)$. The probability of observing N_g conditional on a particular value of N_{total} is given by

$$P(N_g|N_{total}, \mu_{total}, Q) = \frac{P(N_g, N_{total}|\mu_{total}, Q)}{P(N_{total}|\mu_{total}, Q)} = \binom{N_{total}}{N_g}Q^{N_g}(1-\mathrm{Q})^{\mathrm{N_{total}}-N_g} \tag{17}$$

Here, the binomial coefficient is denoted by

$$\binom{N_{total}}{N_g} = \frac{N_{total}!}{N_g!(N_{total}-N_g)!}.$$

This distribution is known as the binomial distribution with probability of success Q. If the sample is blank and N_g and N_b are both measurements of some unknown background with true count rate ρ_b, then $Q = Q_0$ where

$$Q_0 = \frac{\rho_b t_g}{\rho_b t_g + \rho_b t_b} = \frac{t_g}{t_g + t_b}. \tag{18}$$

For example, $Q_0 = ½$ when t_g equals t_b.

The inference about the presence of analyte activity in the sample is based on the conditional distribution of N_g given N_{total}. Hence, the null hypothesis that the sample is blank is rejected if a blank sample would have produced a gross count as large or larger than the observed 100 α% of the time or less, that is, if

$$\sum_{i=N_g}^{N_{total}}\binom{N_{total}}{N_g}Q_0^i(1-Q_0)^{N_{total}-i} \le \alpha. \tag{19}$$

To use this rule in practice, one may simply compute the function on the left hand side of eqn (15) to give the probability that the observed N_g was drawn from the same population as N_b for a given Q_0.

Nicholson uses eqn (19), which he terms his "D_e rule," as the basis of a "randomized decision rule." For the smallest value of N_g for which eqn (19) is true, Nicholson (1963, p. 25) modifies the result of this decision rule by means of "the academic device" of flipping a biased coin *after* the rule has been applied, to adjust the actual false positive rate α' to be exactly α in the long run. The results in this paper use Sumerling and Darby's form of the decision rule.

Nicholson also shows that, in the limit of large numbers of counts, eqn (19) is equivalent to his D_3 rule (eqn 22, below).

Stapleton's decision criterion

Stapleton[§] has proposed a criterion for estimating the standard normal deviate, $z_{Stapleton}$, of a set of observations N_b, t_b, N_g, t_g, that includes an arbitrary parameter d, $0 < d < 1$:

$$z_{Stapleton}(N_g, t_g, N_b, t_b, d) = 2\,\frac{\sqrt{\dfrac{N_g + d}{t_g}} - \sqrt{\dfrac{N_b + d}{t_b}}}{\sqrt{\dfrac{1}{t_g} + \dfrac{1}{t_b}}}. \tag{20}$$

If $z_{Stapleton} > k_\alpha$ then one concludes that analyte activity has been detected. Stapleton has recommended that the value of d be 0.4 for $\alpha = 0.05$.

Nicholson's D_1 rule

Nicholson (1963) gives two other decision rules for the net count rate. Nicholson's D_1 rule (1963) is k_α times the standard deviation of the net count rate:

$$DL_{\mathrm{Nicholson\ D_1}}(R_n, \alpha) = k_\alpha\sqrt{\frac{N_b}{t_b^2} + \frac{N_g}{t_g^2}}. \tag{21}$$

Nicholson states that D_1 uses the "obvious unbiased estimate of the variance with no restriction on" $\rho_n = \rho_g - \rho_b$.

Nicholson's D_3 rule

Nicholson's D_3 rule (1963) is k_α times the sum of the counts divided by the product of the count times:

$$DL_{\mathrm{Nicholson\ D_3}}(R_n, \alpha) = k_\alpha\sqrt{\frac{N_b + N_g}{t_b t_g}}. \tag{22}$$

[§] Personal communication regarding Stapleton, Keith McCroan to Daniel J. Strom, 1 September 1999.

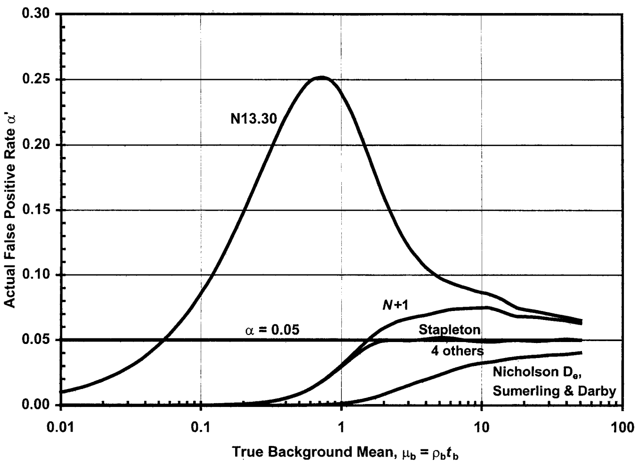

Fig. 3. Actual false positive rates (linear vertical scale) for eight decision rules for $\alpha = 0.05$ as a function of background count mean in the paired blank case. Decision rules Nicholson D_1, Turner/A&P, Nicholson D_3, McCroan, and Stapleton $d = 0.4$ all coincide for $\alpha = 0.05$.

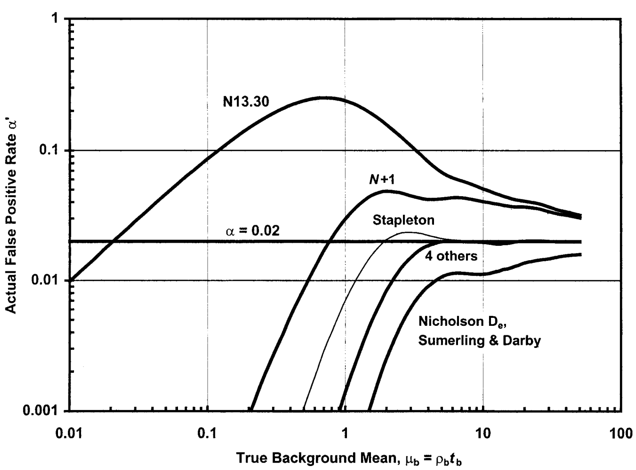

Fig. 4. Actual false positive rates (logarithmic vertical scale) for eight decision rules for $\alpha = 0.02$, as a function of background count mean in the paired blank case. Decision rules Nicholson D_1, Turner/A&P, Nicholson D_3, and McCroan ("4 others") all coincide for $\alpha = 0.02$.

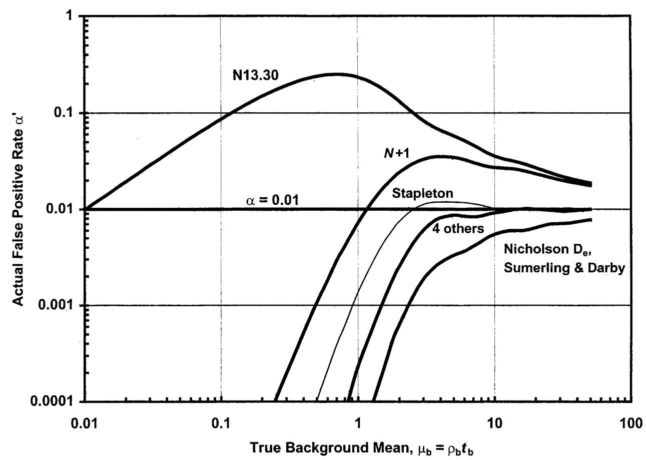

Fig. 5. Actual false positive rates (logarithmic vertical scale) for eight decision rules for $\alpha = 0.01$, as a function of background count mean in the paired blank case. Decision rules Nicholson D_1, Turner/A&P, Nicholson D_3, and McCroan ("4 others") all coincide for $\alpha = 0.01$.

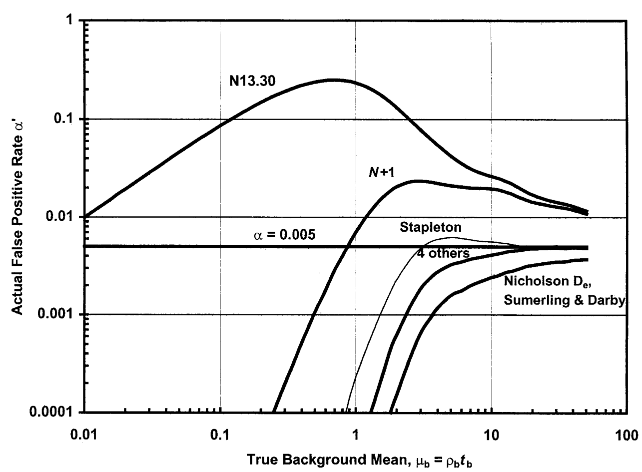

Fig. 6. Actual false positive rates (logarithmic vertical scale) for eight decision rules for $\alpha = 0.005$, as a function of background count mean in the paired blank case. Decision rules Nicholson D_1, Turner/A&P, Nicholson D_3, and McCroan ("4 others") all coincide for $\alpha = 0.005$.

Nicholson states that D_3 "optimally weights information about ρ_b in both t_b and t_g," but that its variance estimate is only unbiased if the underlying net rate (due to analyte activity in the sample), $\rho_n = 0$.

McCroan[‖] derived eqn (13) from Nicholson's D_3 rule.

COMPARISON OF EIGHT DECISION RULES

For the paired blank case ($t_b = t_g$), actual false positive rates α' were evaluated for each of the eight decision rules using the Monte Carlo simulation method described for the N13.30 decision rule. Six values of α (0.05, 0.02, 0.01, 0.005, 0.002, and 0.001), 57 values of μ_b (0.01 to 50) were used in 3,141,593 iterations each. The results are shown in Figs. 3–8. Note that while decision rules Nicholson D_1, Turner/A&P, Nicholson D_3, and McCroan coincide for the paired blank case, they may be distinct when background count time differs from gross count time.

We were not able to implement MacLellan's exact calculation, eqn (7), for exact (binomial) or Stapleton's tests or for Nicholson's D_1 and D_3 rules, because those tests use both N_b and N_g.

For the N13.30 rule using $N_b + 1$ as the estimator of the background mean and variance, the nominal α values

‖ Personal communication, Keith McCroan to Daniel J. Strom, 3 February 2000.

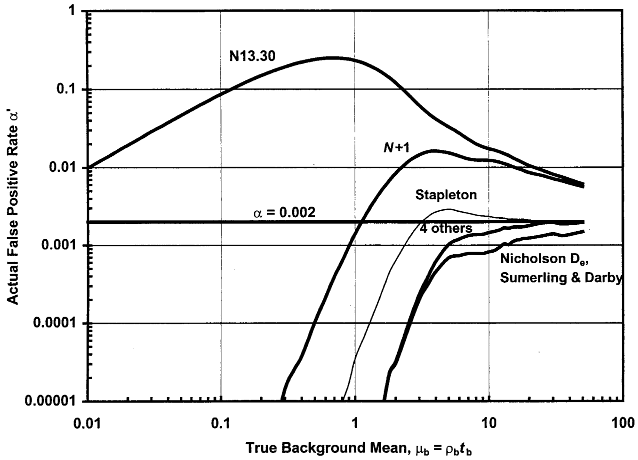

Fig. 7. Actual false positive rates (logarithmic vertical scale) for eight decision rules for $\alpha = 0.002$, as a function of background count mean in the paired blank case. Decision rules Nicholson D_1, Turner/A&P, Nicholson D_3, and McCroan ("4 others") all coincide for $\alpha = 0.002$.

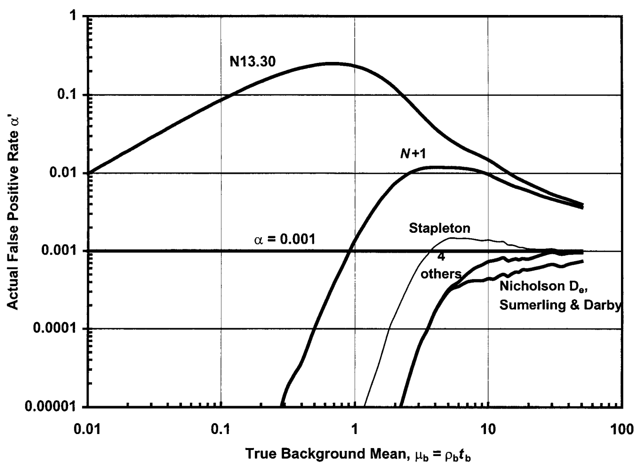

Fig. 8. Actual false positive rates (logarithmic vertical scale) for eight decision rules for $\alpha = 0.001$, as a function of background count mean in the paired blank case. Decision rules Nicholson D_1, Turner/A&P, Nicholson D_3, and McCroan all coincide for $\alpha = 0.001$.

consistently underestimate the observed false positive rates for all $\mu_b > 2$. At background means less than 1, the rule overestimates the false positives. This test is considered inadequate.

For the Nicholson/Sumerling and Darby binomial decision rule, the nominal alpha values overestimate the observed false positive rates for all background means, and grossly overestimate the false positives for $\mu_b < 10$. That is, using this decision rule results in a far smaller proportion of false positives than α. This test is considered inadequate for low background counting.

The Altshuler & Pasternak/Turner decision rule produces false positive rates that are relatively unbiased estimates of α down to $\mu_b = 4$ for the lowest α evaluated (0.001), and down to $\mu_b = 2$ with $\alpha = 0.05$. For $\mu_b < 2$, using the Turner decision rule results in a far smaller proportion of false positives than α. For equal count times, Nicholson's D_3 rule produces identical results, but not for unequal count times. These rules are almost as good as Stapleton's, especially for larger values of α and larger values of μ_b.

Nicholson's D_1 and D_3 rules give the same results as A&P/Turner and McCroan for the paired blank case. When the gross count time differs from the background count time, most rules give different results that are not presented here. No cases have been observed where the McCroan rule differs from Nicholson's D_3 rule, which leads to the need for further investigation. These rules are almost as good as Stapleton's, especially for larger values of α and larger values of μ_b.

RECOMMENDATIONS

None of the rules evaluated provides an unbiased estimate of the false positive rate at all background means μ_b. The N13.30 rule gives the poorest results. The computationally simple $N_b + 1$ rule gives much better, but not adequate results. The Nicholson's D_e/Sumerling and Darby rule based on the binomial distribution never overestimates the actual false positive rate, but it produces the most false negative results for small numbers of counts. The quartet of Nicholson's D_1 and D_3, A&P/Turner, and McCroan produce the fewer false negatives in some ranges of counts than the Nicholson's D_e/Sumerling and Darby rule, and Stapleton's criterion with $d = 0.4$ produces fewer still for the smaller values of α.

For all of the decision rules except the N13.30 and the $N + 1$ rules, α is much greater than the actual proportion of false positives at all but the very lowest background means. We are therefore presented with the conundrum in that our lowest background detectors may not be the most sensitive!

We found that Stapleton's rule is best overall for $\alpha \leq 0.05$ for low expected background counts. The quartet mentioned above did not perform as well at the lower numbers of counts, but were acceptable above 10 counts, and acceptable at lower numbers of counts for the larger values of α. The N13.30 rule does not perform acceptably for anything but the highest numbers of counts, and never works well for small values of α.

Based on these findings, we recommend that standards organizations such as the HPS/ANSI N13.30 committee reconvene to consider modifications to their recommended decision rules.

Furthermore, we recommend that laboratories monitor their method's actual false detection rate for blank samples to verify that the expected value is being met, no matter which decision rule is used.

Future work will investigate other existing decision rules, the effect of varying the background and sample count time ratios, Bayesian approaches, the use of preset counts rather than preset count time, and the effect on minimum detectable activity determinations.

Acknowledgments—The authors would like to express particular thanks to Keith McCroan, for many helpful suggestions and new ideas, as well as one anonymous reviewer who added significantly to our work. The authors would like to thank many other persons for encouragement and/or helpful discussions of this work, including Allen Brodsky, Lloyd A. Currie, Kenneth G. Inn, Mathias M. Lardy, Wesley L. Nicholson, James E. Turner, Charles A. Watson, and Dennis R. Weier. Pacific Northwest National Laboratory is operated for the U.S. Department of Energy (DOE) by Battelle under Contract DE-AC06-76RLO 1830. This work was partially supported by the DOE Office of Worker Protection and Hazards Management (EH-52).

REFERENCES

Altshuler, B.; Pasternack, B. Statistical measures of the lower limit of detection of a radioactivity counter. Health Phys. 9:293–298; 1963.

Brodsky, A. Accuracy and detection limits for bioassay measurements in radiation protection—Statistical considerations. Washington, DC: U.S. Nuclear Regulatory Commission; NUREG-1156; 1986.

Currie, L. A. Limits for qualitative detection and quantitative determination. Application to radiochemistry. Analytical Chem. 40:586–593; 1968.

Currie, L. A. Lower limit of detection: Definition and elaboration of a proposed position for radiological effluent and environmental measurements. Washington, DC: U.S. Nuclear Regulatory Commission; NUREG/CR-4007; 1984.

Friedlander, G.; Kennedy, J. W. Nuclear and radiochemistry. New York: John Wiley & Sons, Inc.; 1955.

Friedlander, G.; Kennedy, J. W.; Miller, J. M. Nuclear and radiochemistry. New York: John Wiley & Sons, Inc.; 1963.

Health Physics Society. Performance criteria for radiobioassay. An American National Standard. New York: American National Standards Institute; ANSI/HPS N13.30-1996; 1996.

Hickey, E. E.; Stoetzel, G. A.; Strom, D. J.; Cicotte, G. R.; Wiblin, C. M.; McGuire, S. A. Air sampling in the workplace. Final report. Washington, DC: U.S. Nuclear Regulatory Commission; NUREG-1400; 1993.

Little, R. J. A. The statistical analysis of low-level radioactivity in the presence of background counts. Health Phys. 43:693–703; 1982.

Lochamy, J. C. The Minimum-Detectable-Activity concept. In: Fivozinsky, S. P., ed. Measurements for safe use of radiation. Washington, DC: National Bureau of Standards; NBS Special Publication 456; 1976: 169–172.

Nicholson, W. L. Fixed time estimation of counting rates with background corrections. Richland, WA: Hanford Works; AEC R and D Report HW-76279; 1963.

Nicholson, W. L. Statistics of net-counting-rate estimation with dominant background corrections. Nucleonics 24:118–121; 1966.

Rainwater, L. J.; Wu, C. S. Applications of probability theory to nuclear particle detection. Nucleonics 1:60–69; 1947.

Stevenson, P. C. Processing of counting data. Livermore, CA: National Academy of Sciences—National Research Council; NAS-NS-3109; 1966.

Strom, D. J.; Stansbury, P. S. Minimum detectable activity when background is counted longer than the sample. Health Phys. 63:360–361; 1992.

Sumerling, T. J.; Darby, S. C. Statistical aspects of the interpretation of counting experiments designed to detect low levels of radioactivity. Chilton, UK: National Radiological Protection Board; NRPB-R113; 1981.

Turner, J. E. Atoms, radiation, and radiation protection. Second edition. New York: John Wiley & Sons; 1995.

U.S. NRC; U.S. EPA; U.S. DOE; U.S. DoD. Multi-Agency Radiation Survey and Site Investigation Manual (MARSSIM; http://www.epa.gov/radiation/marssim/). Washington, DC: U.S. Government Printing Office; 1997.

■ ■