

Testing Surface Disinfectants

This series of knowledge sharing articles is a project of the Standardized Biofilm Methods Laboratory in the CBE

KSA-SM-06

Enumerating viable cells by pooling counts for several dilutions

[Key Words: acceptable range, colony forming units (cfu), Horvitz-Thompson estimate, serial dilutions, weighted average]

When estimating the number of bacteria in a liquid volume, it is common in microbiology to perform a dilution series, pipette onto agar plates for some of the dilutions, and then count the number of colony forming units (cfu). Usually counts are recorded for the first dilution which satisfies a rule defining the acceptable range of counts (Blodgett 2008). However, counts of bacteria from more than one dilution, even when the counts are below the acceptable range, can be combined to form the best possible estimate of the true number of bacteria in the original volume. The purposes of this article are to present the calculation formulas, examples of the calculations, justification for the formulas, and our recommendations.

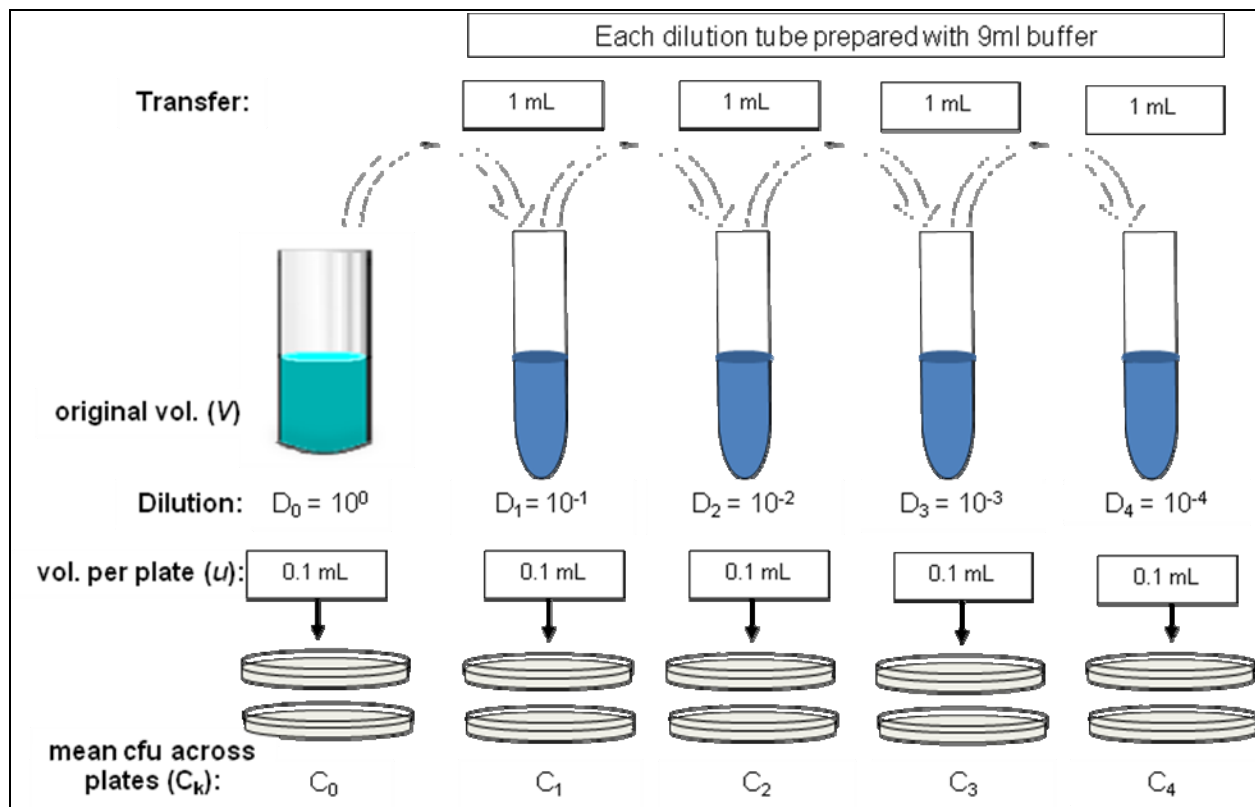


Figure 1. Decimal dilution series with $L=4$ dilutions and two spread plates per dilution.

Let V denote the original volume (mL) of the bacterial suspension. Let T denote the estimated density (cfu per original volume V). Let LD denote $\log_{10}(T)$, a relevant quantity because most statistical calculations for quantitative disinfectant tests are based on the \log_{10} -transformed density. We assume that the suspension is well-mixed and the bacteria are not aggregated. Let D_0, D_1, \dots, D_L denote L serial dilutions, where dilution D_k is the fraction of one mL of the original volume contained in one mL in the k^{th} tube in the dilution series, $k = 0, 1, 2, \dots, L$. For decimal dilutions, the value for D_k is 10^{-k} . Let u denote the volume (mL) of the aliquot transferred to each agar plate from the k^{th} dilution tube. Typically, $u = 0.1$ mL. Let C_k denote the mean cfu count across plates at the k^{th} dilution. Assume that the same number of plates is prepared for each dilution. The notation and nomenclature are illustrated in Figure 1 for a decimal dilution series.

How to estimate the density using a single dilution

If a plate contains so many colonies that a meaningful cfu count is impossible, the count is recorded as “too numerous to count” (TNTC), in which case, the mean cfu count for that dilution is also recorded as TNTC. Let m denote the dilution number of the first dilution for which the mean cfu per plate is not TNTC. Let T_m denote the estimated density of bacteria (cfu per original volume V) when based on that single dilution. Then T_m is calculated using equation (1).

$$T_m = V \times [C_m / (D_m \times u)]. \quad (1)$$

In (1), dividing the mean count C_m by $(D_m \times u)$ is the way to scale-up the count so it applies to one mL of the original volume. Then that “per mL” estimate is multiplied by V mL.

Table 1. Example data for a viable cell enumeration in which $V=40\text{mL}$ and $u=0.1\text{mL}$

k	$C_k^{(a)}$	D_k	T_k	LD_k
0	TNTC	10^0	NC ^(b)	NC
1	TNTC	10^{-1}	NC	NC
2	239.5	10^{-2}	9.58×10^6 ^(c)	6.99
3	25.0	10^{-3}	1.00×10^7	7.00
4	0.0	10^{-4}	0	NC
			T	LD
Pooled using weighted mean (equation 2)			9.53×10^6 ^(d)	6.98
Pooled using Horvitz-Thompson (equation 3)			9.53×10^6 ^(e)	6.98

^(a) Mean cfu count for two plates

^(b) NC = not calculable

^(c) Calculation in equation (1e)

^(d) Calculation in equation (2e)

^(e) Calculation in equation (3e)

Consider the data in Table 1 for which $L=4$, $m=2$, $V=40$, $C_2=239.5$, $D_2=10^{-2}$, and $u=0.1$. When these quantities are entered into equation (1), the result is the T_2 of equation (1e) (“e” is for “example”),

$$T_2 = 40 \times [239.5 / (10^{-2} \times 0.1)] = 40 \times 239.5 \times 10^3 = 9.58 \times 10^6. \quad (1e)$$

For the Table 1 data, the density estimates found by applying (1) to the observed counts at each of the 3rd and 4th dilutions ($C_3=25$ and $C_4=0$) are $T_3 = 1.00 \times 10^7$ and $T_4 = 0$. For each dilution, equation (1) produces an unbiased estimate of the density, but statistical theory shows that the estimate for the first countable dilution has the best precision by far. For the example of Table 1, the variance of T_3 is 10 times the variance of T_2 and the variance of T_4 is 10 times the variance of T_3 . Often the lab

specialist will record the counts only for the first dilution for which the mean cfu per plate is not TNTC.

[**Note:** Some laboratory specialists deviate from our definitions and equation (1) when they calculate the density estimate because they account for u when calculating the dilution, in which case u is removed from the scale-up formula. This alternative approach will produce the same numerical result as equation (1).]

How to estimate the density by pooling across multiple dilutions

When valid counts have been reported for multiple adjacent dilutions, it seems plausible that all the counts could be pooled to provide a more precise estimate than the estimate based only on the first countable dilution. For pooling, we recommend the “weighted mean” formula which pools the density estimates by taking the weighted average where the weight attached to a density estimate is proportional to the reciprocal of its variance; i.e., the more precise density estimate receives more weight. Suppose the n^{th} dilution was the last dilution in the series of dilutions for which valid counts were recorded. Then the weighted mean estimate of the density is shown in equation (2).

$$T = (T_m \times D_m + T_{m+1} \times D_{m+1} + \dots + T_n \times D_n) / (D_m + D_{m+1} + \dots + D_n). \quad (2)$$

Equation (2e) shows the equation (2) density estimate for the data in Table 1 where $m=2$ and $n = 4$.

$$\begin{aligned} T &= ((9.58 \times 10^6 \times 10^{-2}) + (1.00 \times 10^7 \times 10^{-3}) + (0 \times 10^{-4})) / (10^{-2} + 10^{-3} + 10^{-4}) \\ &= ((9.58 \times 10^4) + (1.00 \times 10^4) + (0 \times 10^4)) / (0.0111) = \mathbf{9.53 \times 10^6}. \end{aligned} \quad (2e)$$

There is a popular, alternative calculation formula that turns out to be mathematically equivalent to equation (2). It was created by altering equation (1) by replacing the dilution D by the sum of the countable dilutions and replacing the mean cfu count C by the corresponding sum of the C_k values. This alternative calculation formula is displayed in equation (3).

$$T = V \times [(C_m + C_{m+1} + \dots + C_n) / ((D_m + D_{m+1} + \dots + D_n) \times u)]. \quad (3)$$

It was first suggested by Farmiloe et al. (1954), who provide a more general version of the formula that allows a different number of plates for each dilution (Jarvis 2008). Equation (3e) applies equation (3) to the data in Table 1.

$$\begin{aligned} T &= 40 \times [(239.5 + 25.0 + 0.0) / ((10^{-2} + 10^{-3} + 10^{-4}) \times 0.1)] \\ &= 40 \times 264.5 / (1.11 \times 10^{-3}) = \mathbf{9.53 \times 10^6}. \end{aligned} \quad (3e)$$

Observe that the formulas (2) and (3) presume that $C_k = 0$ is a valid count (no colonies on any plates at dilution D_k) by including zero counts in the weighted average.

Equations (2) and (3) are recommended for pooling counts across dilutions (Hedges 2002, Maturin and Peeler 1998, Niemelä 1983, Niemi and Niemelä 2001, Parkhurst and Stern 1998). The choice between these two equations is only a matter of taste; the two approaches always produce the same numerical answer.

Do not pool by calculating the unweighted mean of densities

It is inadvisable to take an unweighted arithmetic mean of T_k values across multiple dilutions because, as the number of dilutions increase, the variance of the arithmetic mean grows exponentially larger than the variance of the estimator from a single dilution. For example, as the

number of dilutions increases from 2 to 3 to 4, the standard deviation of the arithmetic mean grows from 166% to 351% to 833% of the standard deviation of the estimator from the first countable dilution. The imprecision occurs because the relatively imprecise density estimates at the high dilutions can greatly influence the unweighted mean. For the data in Table 1, the arithmetic mean estimate of the density is $(T_2 + T_3 + T_4)/3 = (9.58 \times 10^6 + 1.00 \times 10^7 + 0)/3 = 6.53 \times 10^6$. The associated log density is 6.81, which is at least 0.17 below the other *LDs* in Table 1.

Statistical justification for the weighted mean

The weighted average of equation (2) is in a form suggested by least squares theory. According to that theory, equation (2) provides the general least squares estimate and therefore has the best precision possible among unbiased estimates that are linear combinations of the T_k values. In particular, the weighted average is guaranteed to be more precise than either the estimator from a single dilution or the unweighted arithmetic mean across all of the dilutions. Equation (3) is motivated by the Horvitz-Thompson formula, which is widely used in field ecology for efficient, unbiased estimation of a population density (Hanif and Brewer 1980; Horvitz and Thompson 1952; Witgen et al. 2006).

The weighted mean is only a slight improvement over the density for the first countable dilution

One might ask whether the increased precision of the weighted average estimator in equations (2) and (3) is worth the time, effort, and expense of plating, incubating and counting bacteria from the multiple dilutions. Or is estimation from counting only a single dilution “good enough?” It can be shown (assuming the counts from adjacent dilutions are uncorrelated) that the increase in precision when using the weighted average over multiple dilutions instead of using the count from a single dilution corresponds to a 10% decrease in variance (this corresponds to about a 5% decrease in the standard deviation). Moreover, this 10% savings only applies to a single experiment. Across multiple experiments, the savings applies only to the “within-experiment” component of variability, which experience shows is usually less than half of the total variability. Consequently, the pooled estimate is expected to decrease the total variance of bacterial density estimates by at most 5% (corresponding to a standard deviation decrease of about 2.5%). Thus, unless the data from multiple dilutions is already available, such as in a quantitative disinfectant test when the efficacy of the disinfectant is unknown, the time and cost of collecting the data necessary from the additional dilutions does not seem worthwhile. However, if the counts from multiple dilutions are available, then these counts should be incorporated into the weighted average formula (equation (2)), or equivalently, the Horvitz-Thompson formula (equation (3)) to produce the most precise estimate.

Comment on rules concerning the acceptable range for counts

It is conventional for a plating technique to be accompanied by a rule stating the range of acceptable plate counts (e.g., 30 – 300 is sometimes used for spread plates). There are practical reasons for the upper count. For example, when there are as many as 300 cfu, the colonies begin to overlap, thereby making accurate counting difficult. When there are many bacteria, they compete for nutrients and all viable cells may not make visible colonies. If there are more than 300 cfu, then the next dilution should have more than 30 cfu; therefore, one can save time and effort by counting that next dilution. The upper limit should be set by microbiologists, based on their knowledge and experience with the microbes under study.

For most laboratory disinfectant tests, a lower limit is not necessary. Consider the popular lower limit of 30. For disinfectant tests, we agree with Niemelä (1983): “The lower limit of 30 colonies must be taken only as a recommendation. If dilutions have been incorrectly selected, then colony

numbers below 30 may be the only ones available. [Besides, the low counts are] highly reliable because biological interactions become less as colony numbers decrease. It is only in the statistical sense that low colony numbers are somewhat unreliable because of their low precision. . . . It is foolish to disregard colony counts below 30 if they happen to be the only ones available.”

References

- Blodgett, R. J. (2008) Mathematical treatment of plates with colony counts outside the acceptable range. *Food Microbiology* 25:92-98.
- Farmiloe, F. J., Cornford, S. J., Coppock, J. B. M., and Ingram, M. (1954) The survival of *Bacillus subtilis* spores in the baking of bread. *J. Sci. Food Agric.* 5:292-304.
- Hanif, M. and Brewer, K. R. W. (1980) Sampling with unequal probabilities without replacement: a review. *International Statistical Review* 48:317-335.
- Hedges, A. J. (2002) Estimating the precision of serial dilutions and viable bacterial counts. *International Journal of Food Microbiology* 76:207-214.
- Horvitz, D. G. and Thompson, D. J. (1952) A generalization of sampling without replacement from a finite universe. *Journal of the American Statistical Association*, 47:663-685.
- Jarvis, B. (2008) *Statistical Aspects of the Microbiological Examination of Foods – Second Edition*. Academic Press: San Diego.
- Maturin, L. and Peeler, J. T. (1998) Chapter 3 – Aerobic Plate Count, Section: Conventional Plate Count Method, part d. in the *FDA Bacteriological Analytical Manual – Online* (January 2001). (<http://www.cfsan.fda.gov/~ebam/bam-3.html>)
- Niemelä, S. 1983. Statistical evaluation of results from quantitative microbiological examinations. NMKL Report no. 1, second edition. Nordic Committee on Food Analysis. Ord & Form AB, Uppsala. (formula 7 on p 9) (listed as available for purchase at <http://shop.nmkl.org/html/ProductBig.asp?Productno=rep1&Searchtype=Searchproduct&search=statistical&x=0&y=0>)
- Niemi, MR M. and Niemelä, S. I. (2001) Measurement uncertainty in microbiological cultivation methods. *Accreditation and Quality Assurance* 6:372-375.
- Parkhurst, D. F. and Stern, D. A. (1998) Determining average concentrations of *Cryptosporidium* and other pathogens in water. *Environmental Science and Technology* 32:3424-34-3429.
- Witgen, B. M., Grady, M. S., Nyengaard, J. R., and Gundersen, H. J. G. (2006) A new fractionator principle with varying sampling fractions: exemplified by estimation of synapse number using electron microscopy. *Journal of Microscopy* 222:251–255.

Version date: 22 Sept 2010

Authors: Martin A. Hamilton, mhamilton@biofilm.montana.edu
Albert E. Parker, parker@math.montana.edu