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

KSA-SM-12

The importance of concurrent control carriers in laboratory tests of surface disinfectants

[*Key Words*: efficacy, untreated carriers or coupons, statistical guidelines, collaborative study, correlation, performance standard]

Background

Laboratory methods for testing a disinfectant against surface-associated microbes typically use easily manipulated, microbe-bearing carriers, e.g., glass disks (KSA-SM-02 Testing surface disinfectants: quantitative, semi-quantitative, quantal, and alternative methods, ver. 2011-10-25). For quantitative and semi-quantitative tests, some of the carriers are treated with the disinfectant and others serve as control (untreated or check) carriers. The assignment of carriers to treated or control groups should be done at random using an objective randomization technique. Disinfectant efficacy is quantified by comparing a measure of the number of viable microbes on the treated carriers to the measure on the control carriers (KSA-SM-02).

The use of control carriers eliminates some possible alternate explanations of the efficacy results. Colony forming unit (CFU) counts for control carriers demonstrate that microbes on the carriers are viable and culturable. Control coupon results allow the experimenter to compare treated carrier results against a control carrier baseline measurement. Good scientific practice for laboratory tests requires that influential variables are fixed or otherwise taken into account so that the conclusions drawn are valid and convincing. When the control carriers are observed *concurrently* with the treated carriers, the comparison effectively adjusts for variables associated with steps in the protocol that vary from day to day such as the preparation of the test microbe population, inoculation of the carriers, and the recovery of microbes from the carriers. The ability to adjust for those variables increases confidence that the lower number of viable microbes on the treated carrier is due to the disinfectant.

In order for a standardized disinfectant test method to be relevant to real-world applications, the inoculated carriers should carry a representative microbe challenge. Both the microbial class (phylum or species) and the amount of microbial contamination should be relevant (KSA-SM-03 Desirable attributes of a standardized method ver. 2010-06-10). The challenge used in laboratory tests should be within the range of microbial contamination measured on surfaces in environments where the disinfectant treatment will be applied in practice. However, because there is a lack of environmental sampling data suitable for determining that range, the specified microbial challenge on carriers usually is based on the judgment of experts or on negotiations among the various stakeholders. The amount of challenge typically is expressed in units of recoverable CFU per area, e.g., CFU per cm^2 of carrier surface area.

As an example of the recognition that concurrent controls are important to a disinfectant test method, consider the well-known Use Dilution Method (UDM). It initially was a quantal response test and no control data were collected. Now the UDM requires concurrent control carrier data that convert the UDM into a semi-quantitative test for which a log reduction measure of efficacy is calculable (AOAC International 2011; KSA-SM-08 The P/N formula for the log reduction when using a semi-quantitative disinfectant test of type SQ_1, ver. 23 June 2011). On each test day, for each of 6 control carriers, conventional dilution series and plates must be prepared for viable cell counting (Tomasino et al. 2012a). The viable cell density for each carrier, when expressed as CFU per carrier, is $\log_{10}$-transformed to arrive at the log density (LD) for the carrier. The mean of the six LD values is denoted by *TestLD*. It is required that the *TestLD* must be at least 6.0 (geometric mean density of 1.0×10^6) and not above 7.0 (geometric mean density of 1.0×10^7). This range of acceptability is achievable, provides a relevant challenge to the disinfectant, and mitigates potential variability in the efficacy results that may be caused by a wider range in carrier counts (Tomasino et al. 2012a).

It may be less expensive and more convenient to rely on *non-concurrent* control carrier counts; i.e., control carrier counts observed on a different day or under different laboratory conditions from the treated carrier counts. However, it is a practical truth that the use of concurrent controls in laboratory disinfectant tests is necessary in order to produce convincing results. An additional, but seldom discussed, bonus is that concurrent control carriers can increase the reproducibility of the test results. In fact, the extent of such improvement can be calculated using data from a collaborative (multi-laboratory) study.

The objectives of this article are to formulate the reproducibility calculation and to use collaborative study data for calculating the amount by which reproducibility is increased when control carriers are observed concurrently with treated carriers. The data indicate that concurrent control carriers increased the reproducibility for some, but not all, disinfectant test methods. A final Discussion section explains why data from some of the collaborative studies were not well-suited for calculating the benefit of concurrent control carriers.

Equation for quantifying reproducibility

Consider a quantitative or semi-quantitative disinfectant test method that requires both untreated control carriers and treated (disinfected) carriers. Following the notation of KSA-SM-10 (Assessing Resemblance, Repeatability, and Reproducibility for Quantitative Methods, ver. 23 June 2011), the random variable $\overline{T}$ denotes the mean of the LDs for the treated carriers and the random variable *TestLD* denotes the mean of LDs for the untreated carriers in a single test. The log reduction (LR) measure of efficacy for the test is defined by $\mathrm{LR} = TestLD - \overline{T}$. A test method is reproducible if tests of the same disinfectant treatment by different laboratories arrive at nearly the same result. Let SD denote "standard deviation." The reproducibility of a test outcome is quantified by the reproducibility SD of that outcome (KSA-SM-10). Because the SD is a measure of imprecision, a smaller reproducibility SD indicates increased reproducibility. Let S_{LR}, S_{TestLD}, and $S_{\overline{T}}$ denote the reproducibility SDs for LR, *TestLD*, and $\overline{T}$, respectively. This notation departs from the convention where a reproducibility SD has an *R* subscript to differentiate it from another SD such as the repeatability SD. Because this article is concerned only with reproducibility SDs, the *R* subscript is suppressed; instead, the subscript indicates the outcome to which the SD applies.

When *TestLD* and $\overline{T}$ were observed concurrently and have in common all laboratory variables for the same test day, one would expect that these two random variables are positively correlated; let *r* denote the correlation coefficient. Statistical theory shows that S_{LR} can be calculated by equation (1);

$$S_{LR} = \sqrt{S^2_{TestLD} + S^2_{\bar{T}} - 2r \cdot S_{TestLD} \cdot S_{\bar{T}}} \quad . \qquad (1)$$

The final term in equation (1) is negative and proportional to the correlation between *TestLD* and $\overline{T}$. Thus the reproducibility SD for LR is smaller when the test protocol creates a larger correlation. It is anticipated that the use of concurrent control carriers will generate a larger correlation than non-concurrent controls. Collaborative study data can be used to calculate the correlation coefficient and the SDs. [Note: The usual calculation formula for S_{LR} is equation (6) of KSA-SM-10; it is uses the LR values directly and does not require *r*, S_{TestLD}, or $S_{\bar{T}}$.]

Results

This section presents results for retrospective analyses of multi-laboratory collaborative studies for four different disinfectant test methods. The first two collaborative studies receive more intensive analysis based on equation (1). The calculations show that it is feasible to measure the extent to which the reproducibility SD for LR was improved by the use of concurrent instead of non-concurrent control carriers. In the plots displayed below, each plotted point is the result for a single test and shows the mean LD for treated carriers ($\overline{T}$) on the vertical axis and the associated mean LD for control carriers (*TestLD*) on the horizontal axis. The axes are scaled so that the length of one unit is nearly the same on both axes. The solid red line is the orthogonal least squares regression line (shows how *TestLD* and $\overline{T}$ vary together on the average), the dashed line is the least squares regression line (for predicting $\overline{T}$ based on the *TestLD*), *r* is the estimated correlation coefficient, and *p* is the upper, one-sided p-value for the test of no correlation. Plots are displayed only for those cases where there was a statistical association between $\overline{T}$ and *TestLD*. For cases where the relationship was not statistically significant ($p > 0.15$), the *r* and associated p-value are listed but the plot is not displayed.

Three Step Method (TSM) using spores (Tomasino et al. 2008)

Figure 1 displays some results from an 8-laboratory collaborative study of the Three Step Method (TSM) using spores of *B. subtilis*; each laboratory conducted 3 replicate tests of each of 3 efficacy levels (low, medium, high) of each of 3 sporicide formulations, 9 treatments in all. For the 5 treatments displayed in Figure 1, the correlations were large enough to suggest that the use of concurrent control carriers led to increased reproducibility of LR. For the other 4 treatments, however, there were no significant correlations ($r \leq 0.16$, $p \geq 0.23$).

Consider the data of Figure 1(d) for which $r = 0.53$. The estimates of the standard deviations were $S_{\bar{T}} = 1.0621$ and $S_{TestLD} = 0.2630$ in which case equation (1) yields $S_{LR} = \sqrt{0.0692 + 1.1282 - 2 \cdot 0.53 \cdot 0.2631 \cdot 1.0622} = 0.95$. If *TestLD* and $\overline{T}$ were non-concurrently observed in the same laboratory, equation (1) applies, but with a smaller (within-laboratory) correlation coefficient, $r = 0.02$, making $S_{LR} = 1.09$. Thus statistical theory applied to the data of Figure 1(d) indicates that the concurrent controls design was better because it produced a moderately smaller reproducibility SD for LR than if the controls were not observed concurrently. The improvement may be understated for reasons presented in the Discussion.

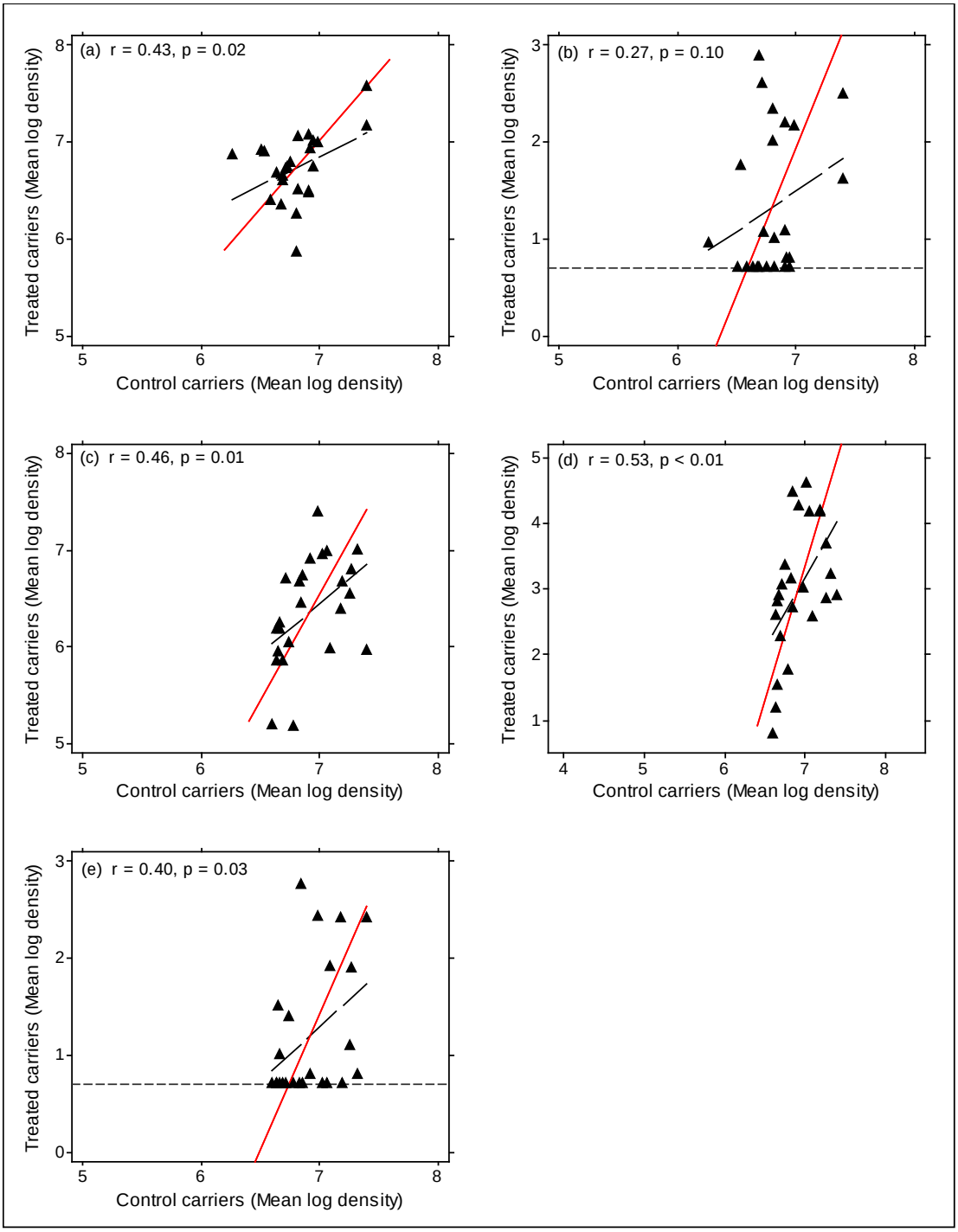

Figure 1. In this plot of data from the TSM collaborative study (Tomasino et al. 2008), the horizontal dashed line is the minimum possible treated carrier log density. The minimum occurred when all CFU counts were zero for a treated carrier in which case the log density was recorded as 0.7. Panels (a) and (b) are for sporicide 3 at low and high efficacy levels, respectively. Panels (c), (d), and (e) are for sporicide 1 at low, medium, and high efficacy levels, respectively.

Use Dilution Method (UDM) using vegetative bacteria (Tomasino et al. 2012b)

Figure 2 displays the *S. aureus* results from a 5-laboratory collaborative study of the semi-quantitative Use Dilution Method (UDM, AOAC Methods **955.15** and **964.02**) in which each of higher and lower efficacy levels of a disinfectant product were tested 3 times against each of the test microbes, *P. aeruginosa* and *S. aureus*. The correlations suggest that the use of concurrent UDM control carriers led to a more reproducible LR.

To measure the gain in reproducibility, consider the tests of the lower efficacy treatment shown in Figure 2(a) for which equation (1) yields $S_{LR} = 0.36$. If *TestLD* and $\overline{T}$ were non-concurrently observed in the same lab, the within-laboratory correlation drops to 0.13 and, by equation (1), $S_{LR} = 0.58$, which is 60% larger. The concurrent control carrier protocol substantially decreases the UDM reproducibility SD. On the other hand, no significant correlations were observed for the UDM *P. aeruginosa* tests of either efficacy level; the correlations were r = 0.24 and -0.09 for the lower and higher efficacy levels, respectively, neither of which was statistically significant (p > 0.6).

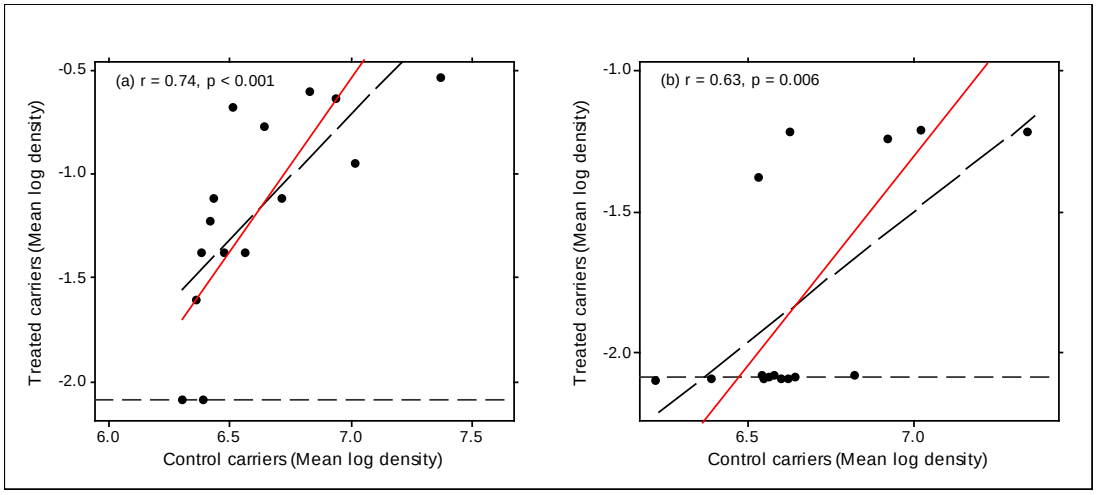

Figure 2. Plot of *S. aureus* data from the UDM collaborative study (Tomasino et al. 2012b), with tests of the lower efficacy treatment in panel (a) and tests of the higher efficacy treatment in panel (b). Points were jittered slightly to expose overlapping results. The line of short dashes shows the minimum possible treated carrier mean.

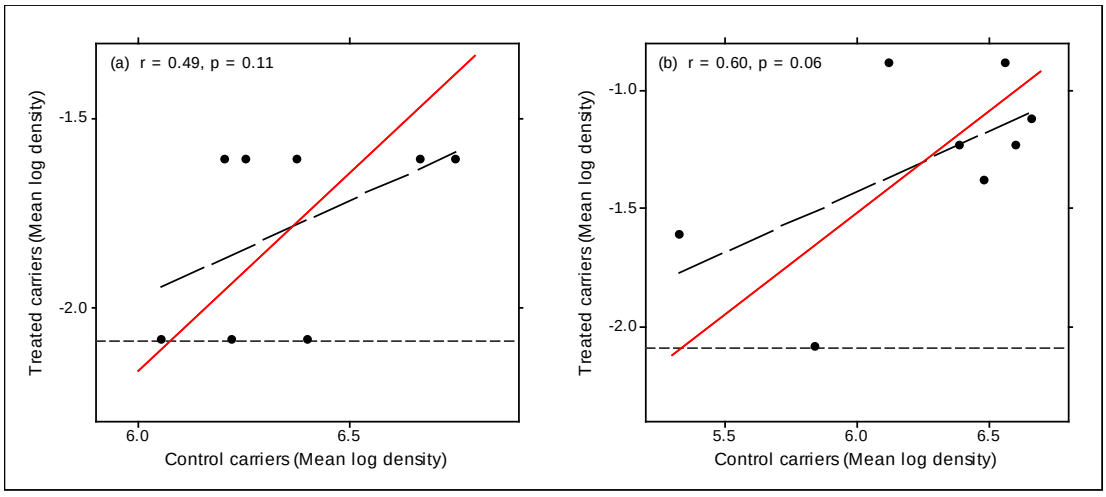

Figure 3. *P. aeruginosa* tests results for 2 disinfectant treatments; data from the collaborative study of the HSCT (Hamilton et al. 1995). Treatment 3 data in panel (a) and treatment 4 data in panel (b). The dotted line shows the minimum possible treated carrier mean log density for the 60 carrier HSCT.

Hard Surface Carrier Test using vegetative bacteria (Hamilton et al. 1995)

Figure 3 displays data from an 8-laboratory collaborative study of the Hard Surface Carrier Test (HSCT, AOAC Method 991.47, a semi-quantitative test) in which four disinfectant treatments were tested against each test microbe, *Pseudomonas aeruginosa* or *Staphylococcus aureus*. Two of the eight combinations of treatment and microbe are displayed, treatments 3 and

4 against *P. aeruginosa*. For testing the other two treatments against *P. aeruginosa*, the correlation was negligible, $r \leq 0.15$ ($p > 0.35$). For the *S. aureus* tests, the correlations were small, $r = 0.02$, 0.03, 0.11, and 0.08 for the four treatments ($p > 0.2$). The correlations suggest that the concurrent controls design led to a more reproducible LR. On the other hand, for 6 of the 8 combinations in this study, the correlations was negligible, indicating that LR for the concurrent carriers protocol was no more reproducible than if the control carriers observations had been non-concurrent with the treated carriers.

Quantitative Carrier Test using spores (Tomasino and Hamilton 2007)

Figure 4 displays data from a 3-laboratory collaborative study of the Quantitative Carrier Test (QCT, ASTM method E 2111 – 00) using spores of *Bacillus subtilis*. Each of 3 sporicide treatments was tested 3 times by each of the 3 laboratories, 9 tests in all. The correlations ranged from 0.55 to 0.65. These correlations suggest that the use of concurrent controls in the QCT led to increased reproducibility of LR.

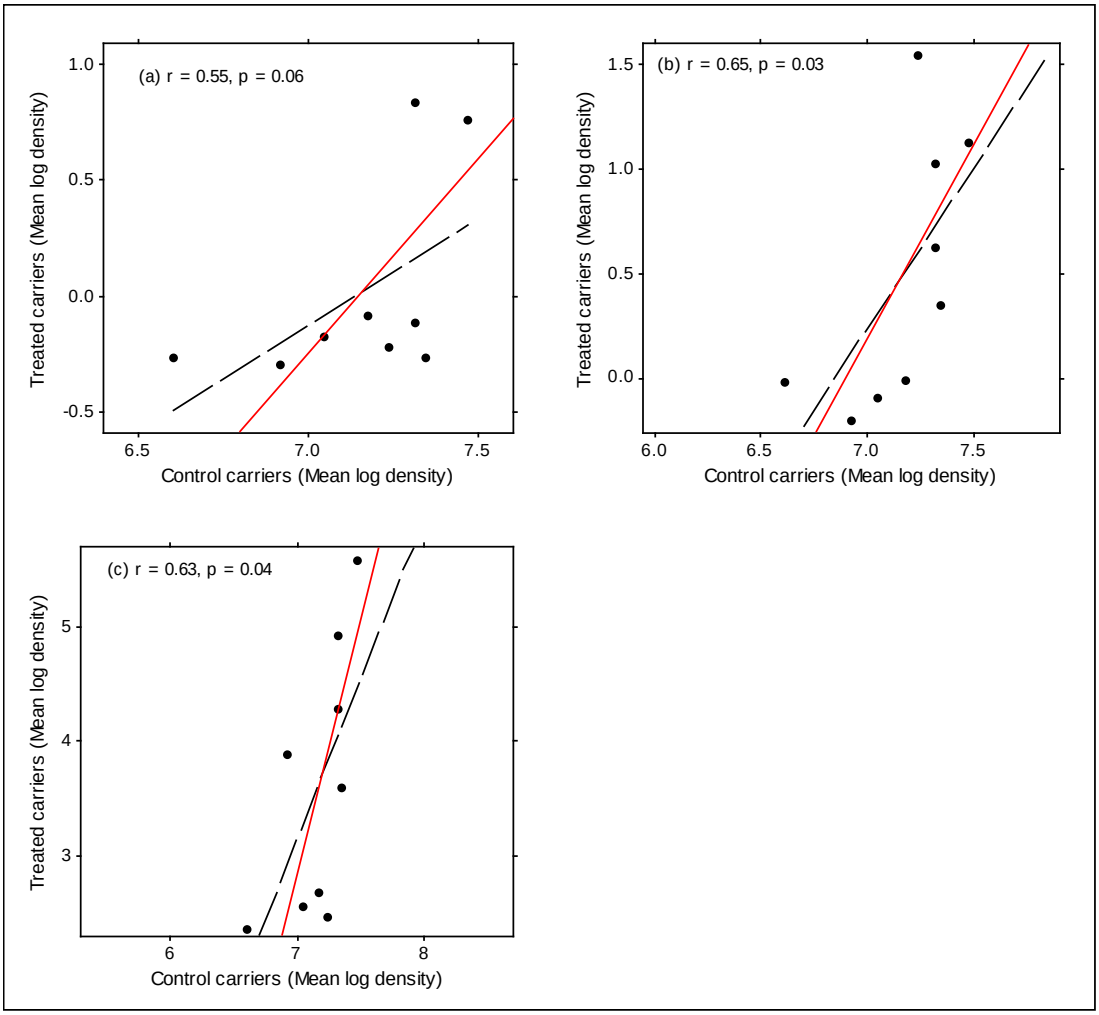

Figure 4. Data from a collaborative study of the Quantitative Carrier Test using spores (Tomasino and Hamilton 2007). Panels (a), (b), and (c) pertain to sporicides 1, 2, and 3, respectively.

Three Step Method (TSM) using spores (Tomasino and Hamilton 2007)

During the study underlying Figure 4, the same sporicides were tested in triplicate with the quantitative Three Step Method (TSM, AOAC Method 2008.05) using spores of *B. subtilis*. For the TSM, the correlations were small and statistically insignificant; specifically, r = -0.074 (p

= 0.57) for sporicide 1, r = -0.006 (p = 0.51) for sporicide 2, and r = -0.028 (p = 0.53) for sporicide 3. The mean log density for untreated carriers varied over a narrow range, min = 7.3 and max = 7.9 for the 9 tests. These negligibly small correlations suggest that the reproducibility of LR was not improved by the use of concurrent TSM control carriers (but see the Discussion).

Discussion

In principle, control carriers should receive the same manipulations and neutralization as the treated carriers, except that an inactive treatment such as sterile water is applied instead of the disinfectant. The motivation for this principle is that the manipulations may reduce the number of viable microbes per carrier and, if the control carriers do not receive the same manipulations as the treated carriers, then the log reduction would be biased upward. In practice, test methods do not always adhere to the "same manipulations" principle; e.g., control carriers may not receive a sham disinfectant or the neutralizer. When experience with the laboratory test method indicates that the effect of the manipulations is negligible, experimenters sometimes choose to avoid the extra cost associated with those manipulations.

Although the importance of controls is generally accepted, there may be special circumstances where the experimenter judges that the tasks of inoculating and measuring control coupons generate an unnecessary expense. Experience with the test protocol may indicate that trained microbiology laboratory technicians can prepare consistent carriers for which the viable microbe measurement is stable at the desired level. Factors such as microbial growth media, dilution water, laboratory equipment, and ambient laboratory climate conditions may vary so little from day to day that concurrent control carriers are judged to be unnecessary. The main disadvantage in conducting tests with no control carriers is that the results may have reduced credibility among people outside of the laboratory where the tests were conducted.

The results of the retrospective analysis of collaborative study data were not as conclusive as anticipated because some of the studies were imperfectly suited to the purposes of this investigation even though those studies was well-suited to their own objectives. It was not unusual for a collaborative study to use a different control carrier protocol than would be used in practice. Collaborative study protocols tended to remove some of the factors that are responsible for the correlation between *TestLD* and $\overline{T}$. The correlation is due to effects that cause the *TestLD* and $\overline{T}$ for a test to move together in departing from their overall means. For example, in some studies of sporicidal test methods, all carriers were prepared at one lab and sent to the participating laboratories (Tomasino and Hamilton 2007). In such a study the carriers were not affected by potentially important between-laboratories factors, including those associated with culture initiation, production of the spore suspension, carrier preparation, carrier inoculation, and carrier drying. Consequently, the *TestLD* and $\overline{T}$ did not depart as much from their means as they would if carriers are prepared separately by each laboratory, and the correlation is understated. In some collaborative studies of sporicide tests, each laboratory prepared in one batch all the carriers the laboratory would need during the study and stored them for subsequent use (Tomasino and Hamilton 2007); e.g., properly stored spore carriers could be used up to 3 months after inoculation (Tomasino and Hamilton 2006). The correlation is also understated for this design because within-laboratory day-to-day effects are suppressed, and the *TestLD* and $\overline{T}$ depart from their means less than when the carriers were prepared independently for each test day.

In some collaborative studies, the control carrier counts were not observed concurrently with the treated carrier counts. For example, in a study of a sporicide test, counts of viable spores

on a few of the untreated carriers from the stored set of spore-bearing carriers were observed prior to beginning the sequence of experiments. That one set of control counts was used for calculating the log reduction of each sporicide tested in the study. In some studies of sporicide tests, although the carriers were prepared in one batch and stored for use during the study, control carriers were counted on each test day alongside the treated carrier observations (Tomasino et al. 2008).

Conclusions

It is good standard practice to use concurrent control carriers. They are an important characteristic of a convincing disinfectant test method. In addition, the use of concurrent control carriers may increase, and will not decrease, the reproducibility of the test results. This investigation shows that it is feasible to measure the beneficial increase.

References

AOAC International (2011) Method **964.02** – Use dilution method for testing disinfectants against *Pseudomonas aeruginosa*. *Official Methods of Analysis 18th Ed*., AOAC International, Gaithersburg, MD,

Hamilton, M. A., DeVries, T. A., and Rubino, J. R. (1995) The Hard Surface Carrier Test as a Quantitative Test of Disinfection: A Collaborative Study. *JAOAC Int* 78:1102-9.

Tomasino, S. F. and Hamilton, M. A. (2006) Modification to the AOAC Sporicidal Activity of Disinfectants Test (Method 966.04): Collaborative Study. *JAOAC Int.* 89:1373-1379.

Tomasino, S. F. and Hamilton, M. A. (2007) Comparative evaluation of two quantitative test methods for determining the efficacy of liquid sporicides and sterilants on a hard surface: a pre-collaborative study. *JAOAC Int* 90:456-464.

Tomasino, S. F., Pines, R. M., Cottrill, M. P., and Hamilton, M. A. (2008) Determining the Efficacy of Liquid Sporicides against Spores of *Bacillus subtilis* on a Hard Surface Using the Quantitative Three Step Method: Method Validation Study. *JAOAC Int* 91:833-852.

Tomasino, S. F., Pines, R. M., and Hamilton, G. C. (2012a) Procedural Revision to the Use-dilution Method: Establishment of Maximum Log Density Value for Test Microbes on Inoculated Carriers. *JAOAC Int* (accepted).

Tomasino, S. F., Parker, A. E., Hamilton, M. A., and Hamilton, G. C. (2012b) Performance of the AOAC Use-dilution Method with Targeted Modifications: A Collaborative Study *JAOAC Int* (submitted).

Version date: 2012-05-17
Primary author: Martin A. Hamilton, mhamilton@biofilm.montana.edu