

between respirator classes using the Co adjustment factor, he concluded that:

Observed 5th percentiles for WPFs, and their lower confidence intervals when adjusted for the Co effect, showed no clear evidence that any class was preferable to another. In particular, there was no indication that Class 5 (elastomerics) performed better than four disposable classes. (Ex. 5-1, p. 8)

The results of these analyses prompted a more detailed examination of the data. To control for study-related and agent-related factors that may contribute to variability, Brown performed an intra-study analysis on two different respirator classes used against the same workplace challenge agent under similar workplace conditions (Approach 2).

Approach 2. The second approach attempted to determine respirator performance after controlling for study-to-study and agent-to-agent sources of variability. Among the half-mask APRs, the chance of detecting performance differences appeared to be greatest for comparisons between elastomeric and filtering facepiece respirators. In implementing this approach, Brown assumed that controlling for study and agent sources of variability would result in WPF differences attributable, in large part, to variability in respirator performance.

Four of the studies compared the performance of elastomeric and filtering facepiece respirators against the same challenge agent in the same workplace. After reviewing these studies, a study by Meyers and Zhuang (Ex. 1-64-51) was selected for further analysis because it was recent, followed a protocol patterned after other published WPF study protocols, and was well documented. Brown's statistical analyses of this study (see Ex. 5-1, Appendix C) indicated large sources of variability within the study, making comparison of the two respirator classes difficult and tenuous. Based on plots of the data and the occurrence of several outliers, it appeared that even data on the same agent, obtained under similar workplace conditions, may not have come from the same parent distribution. In addition, the variability of WPFs within the study (regardless of adjustment for the Co effect) was large. Therefore, the results of this second approach led Brown to state that, at least in this analysis, "workplace studies may have too much intra-study variability for reasonably valid/accurate/reliable assessments and comparisons of respirator effectiveness." (Ex. 5-1, p. C-17)

Approach 3. Brown began the third statistical approach by dividing the data

into units smaller than respirator class, *i.e.*, units based on class of respirator, study, and workplace challenge agent (class-study-agent or CLSA). This procedure resulted in 42 CLSA combinations. After removing deficient data (*e.g.*, no data on Co), he narrowed the data set to 26 combinations. Again, he tested the data for each CLSA to determine if WPF increases with Co and, if so, whether the effect held for all respirator classes. Data analyses of the 26 CLSAs indicated that WPF increased with Co; Brown then derived a common estimate (across all CLSAs) of the Co effect. He subsequently estimated the means for the CLSAs within each class of respirator, both with and without adjustment for Co effect. Brown compared the means of these CLSAs within and between respirator classes. For each respirator class, he grouped the CLSAs that had no significant difference between their means into common subclasses, and plotted both the adjusted and non-adjusted means [*i.e.*, mean of $\ln(\text{PEN})$] of the subclasses, as well as their associated confidence intervals. The results of the comparisons showed that: the estimated means of CLSAs vary so much within a class that the mean of one CLSA is likely to be a poor predictor of the mean of another CLSA within the same class; and it was not visually apparent from the plots that one class of respirator performed better than another class. In general, the comparison indicated that study outcomes, even within the same class of respirator, are highly heterogeneous.

Final analysis. Since the three approaches discussed above could not distinguish between respirator effectiveness within or across classes, the data were viewed, as a whole, from the relationship of Ci and Co. Brown pooled the data for all 26 CLSAs and derived several functional relationships from the pooled data. This approach showed that the majority of the observed data pairs achieved a WPF of 10. (See Ex. 5-1 for more details.)

After performing the above analyses, Brown made a number of observations and conclusions. He noted that the range of WPF values within a CLSA was typically wide, and that the observations were highly variable. In addition, he believed that variability in WPF studies can affect the accuracy, validity, and reliability of study results, as well as the ability to compare study results. Brown noted several possible sources of variability in WPF studies, including: (1) Study characteristics related to study design, execution, sample analysis, and data management and reporting; (2) measurements of Ci at different outside concentrations (Co

effect), taken in conjunction with other poorly described factors (*e.g.*, particle size, temperature, humidity) that may affect the relationship of Ci and Co; (3) characteristics of the ambient agent itself (*e.g.*, possible effects of the agent occurring in a mixture with other agents); and (4) variations in data among studies related to using different study procedures (*e.g.*, repeated measurements on the same worker in some studies versus single measurements on each worker in other studies, random versus non-random selection of study participants). He also commented that the analyses assumed that the data were representative of workplace conditions; however, the data may not represent either current or future workplaces in which employees use respirators. Finally, Brown observed that studies with high Ci values, relative to Co, may have influenced his findings. He believed that these studies should be closely reviewed because some study weakness, unrelated to respirator performance, could be the reason for the high Ci values.

Brown also made some general observations about WPF studies. First, he believed that the role of WPF studies in assessing and comparing respirator effectiveness, and influencing APFs, should be reevaluated. He believed that a more refined instrument that is amenable to experimental design and control, such as chamber studies, is better suited for providing information during determination of assigned protection factors. Brown noted that the use of high concentrations of a challenge agent in chamber studies may minimize the uncertainty of extrapolating test results obtained at low outside concentrations to levels well above the observed range. Therefore, WPF studies would serve as a counterpart to chamber studies, *i.e.*, WPF studies would provide data on the respirator during actual use in the workplace, and identify workplace conditions in which a respirator may perform poorly. To improve comparability of results, he advocated using uniform procedures to: select the challenge agent; collect samples; record the data; and measure and interpret Ci and Co (Ex. 5-1, pp. 42-44).

Overall, the analyses led Brown to several conclusions. First, workplace studies have limitations for comparing respirator performance because of uncontrolled sources of variability. Support for this conclusion comes from the wide confidence intervals for the means of the CLSAs, and the wide range of those confidence intervals within the same respirator class. Second, Brown believed that the WPF has limits as a