

Letter to the Editor

September 24, 2018

Journal of the International Society for Respiratory Protection
Attn: Ziqing Zhuang, Ph.D.
105 EagleBrook Court
Venetia, PA 15367
JISRPeditor@ISRP.com

To the Editor:

We are writing in regard to the recent articles by Wu, et al. (2017, 2018). The articles present details of studies performed to validate respirator fit testing methods using a new optical particle counter. The authors state that these studies were conducted in accordance with the requirements of the validation protocol in ANSI Z88.10-2010, Annex A2. Unfortunately, it appears that the authors have misunderstood the requirements of this validation method and as a result we believe they have reached an incorrect conclusion in at least one case.

The first flaw in their procedure was the selection of a reference fit test method. Clause A2.1 of the standard states:

"While the ideal comparison procedure has yet to be proven, this annex provides a specific procedure for evaluating fit test methods against the current body of knowledge." ... "It involves a statistical comparison of a new fit test method against the generated aerosol method described in this standard."

The generated aerosol method referred to is described in Clause 7.1 of that document ("Generated Aerosol Quantitative Fit Test Procedure"). The PortaCount® method is defined as a separate method, as described in Clause 7.2 ("Particle-counting Instrument Quantitative Fit Test Procedure"). In general, generated aerosol methods use a photometer to measure the concentration of particles inside and outside of the respirator. Han, Willeke, and Colton (1997) describe several photometric-based fit test methods, although they are not explicitly referred to as "generated aerosol" methods, since that term was introduced by OSHA in 1998. This vagueness in the definition of the generated aerosol method in the ANSI document is a weakness of that standard.

Clause A2.1 of the standard does permit one exception to the selection of the reference test method:

"A proposed modification to an accepted QNFT procedure can be evaluated using the accepted protocol for that instrument as the reference standard."

If the MT-05U and the PortaCount® were comparable instruments then one might make the argument that substituting the MT-05U is a simple modification of an accepted QNFT procedure. However, as the authors note, there are multiple differences between the instruments in terms of the detection method, the lowest detectable particle size and the maximum measurable particle concentration. In view of these significant differences, one must conclude that the generated aerosol method is the only acceptable reference test method to be used for validation of a fit test method employing the MT-05U.

The second, more subtle, but equally important flaw in the procedure by Wu, et al. is in the treatment of outliers in their analysis of the data. Specifically, the authors used a statistical test to identify and eliminate outliers from their data sets. In doing so, they excluded pairs of data which affect the results. This is evident

most significantly in Figure 5C of the 2017 publication. As shown in that figure, there were at least five instances in which the reference fit test yielded a fit factor less than 500 (the Reference Fit Factor for full facepiece respirators) while the new method yielded fit factors greater than 500. The authors flagged these data as statistical outliers and eliminated them from the data set before calculating the contingency table. In the published figure the outliers are plotted with red Xs. As a result of excluding the outliers, the authors used values of $A=0$ and $C=67$ rather than $A=5$ and $C=69$ in calculating the Test Sensitivity. When the outliers are included in the calculation, the Test Sensitivity is 93.2%, which fails the ANSI performance criteria.

Our disagreement is not with the specific statistical method that was applied to identify outliers but rather with the automatic, uncritical elimination of data points that were deemed to be outliers. While the statistical test revealed that the data pairs were outliers, they do not reveal why. The authors infer from the statistical test that the outliers “originated from instrument errors or sampling bias”; however they offer no evidence to support those diagnoses. Outliers are a signal that something unusual has happened, e.g., equipment malfunction, data recording error, failure to follow the experimental protocol, etc. Statistical textbooks (e.g., Rawlings, 1988) counsel investigating the cause of outliers before eliminating them. It is not apparent from the article that the authors made such an investigation before eliminating the outliers. The charts in Figure 3 of the 2017 publication compare simultaneous readings from two different PortaCounts®. In each case, the regressions show excellent agreement between the two instruments. So, if an error occurred with one of the instruments during fit testing, then it was probably with the MT-05U. When a person is conducting a fit test in the field with that instrument, how will they know when an instrument error occurred and the fit test should be deemed invalid?

While the ANSI protocol does not explicitly address the handling of outliers, it does address the issue indirectly. As noted by the authors, the ANSI protocol establishes the requirement: “Any reference fit factor below 10% of the required fit factor accepted by the new fit test method shall disqualify the method.” Presumably a statistical test would identify the point ($FF_{Ref} < 0.1 \times RFF$, $FF_{New} > RFF$) as an outlier; however, the ANSI method does not permit exclusion of that outlier. Rather, if such a point exists in the data set then the new method is rejected. If that point can't be excluded from the data set, then less obvious outliers such as those identified in Figure 5C, should not be excluded either.

In general, the simultaneous sampling method employed in these studies is preferable to the sequential sampling method specified in the ANSI protocol. It offers the key advantage that the results obtained from the two instruments are from the same donning of the respirator at the same time. However, the fact that the two instruments were measuring the same aerosol sample during a fit test makes it difficult to see how an outlier can be attributed to sampling bias. In addition, the authors reported that the instruments passed the daily check procedures, so unless some error or irregularity was noted by the test operator at the time the test was conducted (e.g., if a sample hose dislodged from the respirator during the test or if the hose was kinked), there's no reason to believe that the results of that test are not valid. That is, there's no reason to believe that the results are not indicative of the normal instrument performance.

By comparison, Richardson, et al. (2013) conducted sequential fit tests. To ensure that the fit of the respirator did not change from one fit test to another, they adopted a screening technique that involved measuring the concentration of particles inside the respirator for 5 seconds during normal breathing before the first fit test, between the first and second fit test, and after the second fit test. If this “baseline fit” changed by more than a factor of 100 from the start of a given test to the end, then it was assumed that the fit of the respirator had shifted over the course of the test, so the assumption of a consistent fit during the two tests was not valid. In that case the data was not used in the analysis. Insofar as the goal of the ANSI protocol is to compare two test methods *for the same respirator fit*, the screening method was a reasonable approach to validate that assumption.

It is worth noting that the Sibata MT-05U is a count-based instrument. The authors may find it useful to define an acceptable margin of error for a fit factor of $FF_{New}=500$ (e.g., $\pm 5\%$) and then work backwards to establish a minimum required chamber concentration $C_{out,min}$ to achieve that margin of error, as discussed by TSI (2018). This will permit a *post hoc* analysis of the data, eliminating values of FF_{New} for which $C_{out} < C_{out,min}$, which may, in turn, eliminate some of the data points that were identified as statistical outliers. This approach would be justifiable because it is directly traceable to the raw data of the measurement and the variability associated with low particle counts.

In summary, the studies do not conform to the ANSI validation protocol because they did not use the generated aerosol method as the reference method. The published studies do provide interesting comparisons between the MT-05U and the PortaCount® (with and without the N95 Companion). We commend the authors for providing a complete accounting of the data that they collected and the logic they used in their analysis. Reanalysis of the data taking into the consideration the variability associated with count statistics may provide a legitimate method to identify and eliminate outliers.

Sincerely,

Andrew Viner
Craig Colton, CIH
Ginny Agresti
3M Personal Safety Division

REFERENCES

- American National Standards Institute. (ANSI) (2010) Respirator Fit Testing Methods. American National Standards Institute /American Industrial Hygiene Association, Fairfax, VA. ANSI Z88.10-2010.
- Han D-H, Willeke K, and Colton CE. (1997) Quantitative Fit Testing Techniques and Regulations for Tight-Fitting Respirators: Current Methods Measuring Aerosol or Air Leakage, and New Developments. *Am. Ind. Hyg. J.*, 58(3): 219-228
- OSHA (1998) 29 CFR 1910.134, Appendix A.
- Rawlings, JO (1988) *Applied Regression Analysis: A Research Tool*. 1st Ed., Wadsworth, Inc.
- Richardson AW, Hofacre KC, Weed J, Holm R and Remiarz R. (2013) Evaluation of a Faster Fit Testing Method for Full-Facepiece Respirators Based on the TSI PortaCount®. *J. Int. Soc. Resp. Prot.* 30: 116–128.
- TSI (2018) An Analysis of Utilizing Condensation Nuclei Counter (CNC) vs. Optical Particle Counter (OPC), Application Note RFT-023. (http://www.tsi.com/uploadedFiles/_Site_Root/Products/Literature/Application_Notes/RFT-023_QRFT_CNC_vs_OPC_US-web.pdf)
- Wu B, Leppänen M, Yermakov M and Grinshpun SA. (2017) Evaluation of a new instrument for aerosol quantitative fit testing. *J. Int. Soc. Res. Prot.* 34(2): 111-127.
- Wu B, Leppänen M, Yermakov M, and Grinshpun SA. (2018) Utility of an Optical Particle Counting Instrument for Quantitative Respirator Fit Testing with N95 Filtering Facepieces. *J. Int. Soc. Resp. Prot.* 35(1):14-25.