**ISO /TS 20914:2019 – A critical commentary**

**Farrance, Frenkel, Badrick**

## Supplementary appendix

### Recalculation of two examples relating to 'pooled' variance from ISO/TS 20914

1. Section A.4 and Table A.5 use the following data: there are three measuring systems so that $m$ = 3 with $n_1 = 280, n_2 = 190,$ and $n_3 = 400$. We also have $N = n_1 + n_2 + n_3 = 870$, with respective means of $\bar{x}_1 = 5.15, \bar{x}_2 = 4.93,$ and $\bar{x}_3 = 5.28$. The respective standard deviations are $s_1 = 0.160, s_2 = 0.190,$ and $s_3 = 0.200$. Applying equation (10) and keeping extra decimal places for the evaluation as recommended, $s_{pooled} = 0.22948$ compared to 0.255 obtained in ISO/TS 20914 ("step 4"). Using equation (11), the pooled mean is $\bar{x}_{pooled} = 5.1617$ compared to 5.12 in ISO/TS 20914. This gives $u_{rel\,(pooled)} = 0.22948 / 5.1617 = 4.45\%$ and $U_{rel\,(pooled)} = 8.90\%$ for 95% confidence compared to ISO/TS 20914 values of 5.0% and 10.0% respectively.

2. Section A.8.1, Table A.12 (level 1 IQC only) uses the following data: there are two periods of measurement, February 2014 to March 2015 and April 2015 to March 2016. In this case $m$ = 2, $n_1 = 1390$ and $n_2 = 1216$, with $N = n_1 + n_2 = 2606$. The means for the two periods are $\bar{x}_1 = 28.32$ and $\bar{x}_2 = 27.14$, with standard deviations $s_1 = 0.82661$ and $s_2 = 0.84785$. This example demonstrates a situation where the within-lot variances are combined with the exclusion of any between-lot variability, on the assumption that the two within-lot variances do indeed represent the overall long-term imprecision. The two mean values are also combined even though they are from different samples; they are clearly statistically different even if analytically similar. The combination of means in this manner requires careful consideration regarding the intended interpretation, as the combination of significantly different mean values may not be statistically or analytically appropriate. In this situation however, it can be assumed that an approximate overall mean is to be used to define

the analytical range in which to apply the ‘new’ combined imprecision (the combined variances) and to derive a corresponding CV. Applying the first term in equation (10) for a weighted pooled variance gives $s_{intra\text{-}lot} = 0.83643$, compared to the unweighted pooled variance of 0.83730 (=0.837) in ISO/TS 20914. Using equation (11) gives the weighted pooled mean as $\bar{x}_{pooled} = 27.76939$ compared to 27.7 in ISO/TS 20914, with $u_{rel\,(pooled)} = 0.83643 / 27.76939 = 3.016\%$ and $U_{rel\,(pooled)} = 6.033\%$ or $\approx 6.0\%$ for a 95% confidence. In this example, where the *n*’s are large and similar, both the weighted and unweighted calculations provide the same result of 6.0% for a 95% confidence. However, not all laboratories (or even different tests within the same laboratory) will accrue such a large number of similar IQC observations. So as stated previously, there appears to be no good reason why the complete equations should not be used, as these will apply in all situations.