The evaluation of degrees of equivalence in regional dosimetry comparisons

D. T. Burns and P. J. Allisy-Roberts, BIPM, March 2007

The following analysis assumes in the first instance a single linking laboratory, LINK, and a comparison of the quantity air kerma by measurement of the calibration coefficient N_K for a single transfer instrument. The generalization to multiple transfer instruments and linking laboratories is discussed in Sections 4 – 6. This is a simplified classical analysis that is relatively straightforward to implement and is considered to be appropriate for the level of uncertainty of the dosimetric input data. However, work is underway to enable an automatic analysis of the degrees of equivalence following the method in [1], which includes every measurement result in the comparison and uses the full measurement equations and uncertainties as input data to the uncertainties matrix so that correlations between each pair of results is taken into account.

1. Input data

Assume a travelling transfer instrument with long-term stability u_{stab} determined, for example, as the standard uncertainty of the measurements (not the standard uncertainty of the mean) for repeat calibrations at the pilot laboratory over the duration of the comparison (this laboratory should ideally be the linking laboratory).

For each participating laboratory i, we have the calibration coefficient $N_{K,i}$ with its combined standard uncertainty u_i (this must not include a component for the long-term stability of the transfer chamber). We have also the key comparison result K_{LINK} / K_{BIPM} and the combined standard uncertainty u_{BIPM} of the BIPM air-kerma determination.

Regarding the LINK uncertainties, the non-statistical components of the air-kerma determinations by the LINK laboratory both at their own laboratory and at the BIPM cancel. However, a component u_{LINK} will remain, comprising:

(i) the statistical uncertainty of the air-kerma determination at the LINK,
(ii) the statistical uncertainty of the LINK air-kerma determination at the BIPM,
(iii) the statistical uncertainty of the transfer instrument calibrations at the LINK, and
(iv) the non-statistical uncertainties for current measurements at the LINK (unless both the standard and the transfer instruments use the same current measurement system).

Note that if u_{stab} is derived from repeat measurements at the LINK, then component (iii) above will already be included in u_{stab}.

2. Degrees of equivalence with respect to the reference value

For the degree of equivalence of each laboratory i with respect to the key comparison reference value, evaluate the ratio

$$R_i = \frac{N_{K,i}}{N_{K,\text{LINK}}}\frac{K_{\text{LINK}}}{K_{\text{BIPM}}} \quad = \quad \frac{K_i}{I_i}\frac{I_{\text{LINK}}}{K_{\text{LINK}}} \quad \frac{K_{\text{LINK}}}{K_{\text{BIPM}}} \tag{1a}$$

and its variance

$$u_{R,i}^2 = \left(u_i^2 + u_{\text{BIPM}}^2 - \sum_n f_n^2 \left(u_{i,n}^2 + u_{\text{BIPM},n}^2 \right) \right) + u_{\text{stab}}^2 + u_{\text{LINK}}^2 \, . \tag{1b}$$

Here, the summation contains those components $f_n \, u_{i,n}$ and $f_n \, u_{\text{BIPM},n}$ that are correlated between laboratory i and the BIPM, with correlation factor f_n. The physical constants that enter in the air-kerma determinations are fully correlated ($f_n = 1$); certain correction factors, for example k_{wall}, might be considered partially correlated ($0 < f_n < 1$). When laboratory i is traceable to the BIPM, the summation contains all of the non-statistical components of u_{BIPM}, each with correlation factor $f_n = 1$.

The degree of equivalence for laboratory i is the difference $D_i = R_i - 1$ and its expanded uncertainty $U_i = 2\,u_{R,i}$.

3. Pair-wise degrees of equivalence

For each pair of laboratories i and j, evaluate the difference

$$R_i - R_j \tag{2a}$$

and its variance

$$u_{ij}^2 = \left(u_i^2 + u_j^2 - \sum_n f_n^2 \left(u_{i,n}^2 + u_{j,n}^2 \right) \right) + 2\,u_{\text{stab}}^2 . \tag{2b}$$

Here, the summation contains those components $f_n\,u_{i,n}$ and $f_n\,u_{j,n}$ that are correlated between i and j, including the physical constants. If both i and j are traceable to the BIPM (or to another laboratory, NMI), then the summation again contains all the non-statistical components of u_{BIPM} (or u_{NMI}).

The pair-wise degree of equivalence for laboratories i and j is the difference $D_{ij} = R_i - R_j$ and its expanded uncertainty $U_{ij} = 2\,u_{ij}$.

The above example is for an air-kerma comparison, for which the primary standards are all based on the same measurement principle. For a comparison of absorbed dose to water standards the equations above are the same, but the correlated components n and the correlation factors f_n are different, depending on whether each laboratory is traceable to a (or the same) graphite calorimeter, to a (or the same) water calorimeter, or to the BIPM ionometric standard. An example analysis, with choices for f_n relating to mass-energy absorption coefficients and to the heat defect in water calorimetry, can be found in [2].

4. Multiple transfer instruments

A ‘star’-type comparison, in which two or more transfer instruments are sent to the participating laboratories with regular return for re-calibration at the linking laboratory, is a robust arrangement and relatively straightforward to analyse. For p transfer instruments (which might include the same transfer instrument used with different current measurement systems), each laboratory i has p results $N_{K,i,p}$ and the linking laboratory has the p mean values $N_{K,\text{LINK},p}$, each with its own stability determination $u_{\text{stab},p}$. In this case relation (1a) gives rise to the p values $R_{i,p}$ and, if the uncertainties $u_{\text{stab},p}$ are reliable estimates, the best estimate for R_i is the weighted mean

$$R_i = \frac{\sum_p R_{i,p} \big/ u_{\text{stab},p}^2}{\sum_p 1 \big/ u_{\text{stab},p}^2} . \tag{3a}$$

The uncertainty $u_{R,i}$ is obtained using relation (1b) as before, where the stability estimate u_{stab} is now derived from the values $u_{\text{stab},p}$ using

$$\frac{1}{u_{\text{stab}}^2} = \sum_p \frac{1}{u_{\text{stab},p}^2} . \tag{3b}$$

Relation (3b) reflects the improvement in statistical uncertainty that is achieved by the use of multiple transfer instruments, although the principle reason for using multiple instruments is redundancy in the event of failure.

It is emphasized that the use of relations (3a) and (3b) requires reliable estimates for the statistical uncertainties $u_{\mathrm{stab},p}$. This should be the case if there have been a sufficient number of re-calibrations, say six or more, of each transfer instrument at the linking laboratory over the duration of the comparison. If this is not the case, then R_i should be evaluated as the unweighted mean of the $R_{i,p}$ and the uncertainty u_{stab} in relation (1b) taken as the mean of the $u_{\mathrm{stab},p}$ (not the r.m.s. value) divided by $\sqrt{p}$.

Evaluation of the pair-wise degrees of equivalence proceeds as in Section 3.

5. Multiple linking laboratories

Although a star-type comparison with a single linking laboratory is normally robust, an alternative that might be used in some cases is a 'round-robin' comparison with q linking laboratories (and a single transfer instrument). In this case relation (1a) gives rise to the q values $R_{i,q}$. In principle one can use relations equivalent to (3a) and (3b) to evaluate the weighted mean R_i and the combined uncertainty u_{LINK}. However, as noted in Section 4, the use of the weighted mean requires reliable estimates for the $u_{\mathrm{LINK},q}$. As these come from different laboratories and might involve non-statistical uncertainties, their self-consistency is not assured. Consequently, it might be more appropriate to evaluate R_i as the unweighted mean of the $R_{i,q}$ and to take u_{LINK} for use in relation (1b) as $u_{\mathrm{LINK,mean}}$ divided by $\sqrt{q}$, where $u_{\mathrm{LINK,mean}}$ is the mean of the $u_{\mathrm{LINK},q}$.

The value for u_{stab} to be used in relation (1b) should be derived from repeat measurements at one or more of the linking laboratories.

Evaluation of the pair-wise degrees of equivalence proceeds as in Section 3.

6. General case

The analyses of Sections 4 and 5 lead to the general case of p transfer instruments and q linking laboratories, yielding the pq values $R_{i,p,q}$ with transfer instrument uncertainties $u_{\mathrm{stab},p}$ and linking uncertainties $u_{\mathrm{LINK},q}$. In this case, the arguments presented in Section 4 can be used to derive the (weighted or unweighted) mean value $R_{i,q}$ and uncertainty u_{stab} for each linking laboratory q, and then the $R_{i,q}$ and $u_{\mathrm{LINK},q}$ can be combined as in Section 5 to give R_i and u_{LINK}.

The uncertainty $u_{R,i}$ for each R_i is evaluated using relation (1b) and the evaluation of the pair-wise degrees of equivalence proceeds as in Section 3.

References

[1] C. M. Sutton, Analysis and linking of international measurement comparisons, *Metrologia* **41** (2004) 272 – 277.

[2] P. J. Allisy-Roberts and D. T. Burns, Summary of the BIPM.RI(I)-K4 comparison for absorbed dose to water in ^{60}Co gamma radiation, *Metrologia* **42** (2005) Technical Supplement 06002.