Proposed International Guidelines for the Statistical Analysis of Key Comparisons

M. Cox, P. Harris

National Physical Laboratory, UK

maurice.cox@npl.co.uk

The technical supplement to the Mutual Recognition Arrangement (MRA) gives the following definitions:

1. The degree of equivalence of each national measurement standard is expressed quantitatively by two terms:

    a) Its deviation from the key comparison reference value.

    b) The uncertainty of this deviation at the 95 % level of confidence.

2. The degree of equivalence between pairs of national measurement standards is expressed quantitatively by two terms:

    a) The difference of their deviations from the key comparison reference value.

    b) The uncertainty of this difference at the 95% level of confidence.

BIPM Director's Advisory Group on Uncertainties is producing documents containing proposed guidelines for the statistical analysis of key comparison measurements. They constitute interpretations and implementations of the above definitions.

The first document is concerned with the analysis of key comparison measurements in the case where the following three conditions apply:

1. Each participating national institute provides a measurement of a travelling standard having good short-term stability and stability during transport, and the associated standard uncertainty.

2. Each institute's measurement is realised independently of the other institutes' measurements in the key comparison.

3. For each institute a Gaussian distribution (with mean equal to the institute's measurement and standard deviation equal to the provided standard uncertainty) can be assigned to the measurand of which the institute's measurement is an estimate.

A procedure provided within the document, based on the use of a weighted mean as the key comparison reference value, carries out a consistency check of the model and the data. If the check fails, another procedure based on the use of the median is recommended as a more robust solution in such circumstances. The latter procedure can also be used *a priori* if the third condition above is not applicable to the measurements from one or more institute.

Subsequent documents will address other cases including those where (a) the travelling standard does not necessarily possess the above stability properties and (b) the institutes' measurements contain mutual dependencies. A further document will treat the linking of key comparisons carried out under the auspices of the CIPM and those operated by regional metrology organizations.

This paper discusses the technical approaches proposed for the first document and possibilities for the subsequent documents.