Accred Qual Assur (2011) 16:311–322
DOI 10.1007/s00769-011-0767-0

PRACTITIONER'S REPORT

Mass balance method for purity assay of phthalic acid esters: development of primary reference materials as traceability sources in the Japan Calibration Service System

Keiichiro Ishikawa · Nobuyasu Hanari · Yoshitaka Shimizu · Toshihide Ihara · Akira Nomura · Masahiko Numata · Takashi Yarita · Kenji Kato · Koichi Chiba

Received: 18 January 2011 / Accepted: 23 February 2011 / Published online: 24 March 2011

Abstract Purity assay of high-purity materials (HPMs) of phthalic acid esters (PAEs) was carried out by means of a mass balance method. In this method, chromatographic methods such as gas chromatography-flame ionization detector (GC-FID) and/or high-performance liquid chromatography (HPLC) in combination with other methods such as Karl-Fischer (KF) titration and vacuum evaporation (VE) were applied. The sum of the impurities estimated by these methods allowed the estimation of the purity of the main component by difference. Seven PAEs with varying side chain structures and levels of impurities were analysed on a systematic way in which impurities were classified into several groups in terms of their abundance, availability of qualitative information and availability of authentic compounds, etc. The absolute quantity of each impurity was determined by GC-FID and/or HPLC based on the calibration made by the authentic compounds of impurities whenever available. The purities in mass fraction of these PAEs were certified at the National Metrology Institute of Japan (NMIJ), and the PAEs were registered as primary reference materials playing an essential role in linking the metrological traceability of the Japan Calibration Service System (JCSS) to the International System of Units (SI).

K. Ishikawa (✉) · N. Hanari · Y. Shimizu · T. Ihara · A. Nomura · M. Numata · T. Yarita · K. Kato · K. Chiba
National Metrology Institute of Japan (NMIJ), National Institute of Advanced Industrial Science and Technology (AIST), Tsukuba Central 3, Umezono, Tsukuba 305-8563, Japan
e-mail: ishikawa-keiichiro@aist.go.jp
URL: http://www.nmij.jp/

Keywords Certified reference material · Phthalic acid ester · Purity · Japan Calibration Service System · Mass balance method

Introduction

To avoid exposure to harmful substances, systematic assessment and regulation of hazardous chemicals as well as their monitoring are coming to more and more important. Phthalic acid esters (PAEs) constitute a class of synthetic compounds which has been widely used as additives in plastic products as well as pesticide carriers or insect repellents, etc. [1, 2]. They are suspected environmental hormones as well. As of 15th December 2010, European Chemicals Agency has listed four PAEs as a part of about 50 chemicals designated "Substance of Very High Concern (SVHC)" in the system for Registration, Evaluation, Authorization and Restriction of Chemicals (REACH) [3]. Some PAEs are also regulated in Japan by the Law of Standards and Criteria for Food and Food Additives, etc. [4]. Supply of reliable reference standards of PAEs to ensure their reliable monitoring is requested for enhancing both chemical safety and international trading [5].

Since 2001, the National Metrology Institute of Japan (NMIJ)/National Institute of Advanced Industrial Science and Technology (AIST) has been taking part in the development and maintenance of the high-purity materials (HPMs) with which primary reference solutions are prepared under the Japan Calibration Service System (JCSS) [6]. Since these solutions are ranked at the top of the traceability system in Japan, HPMs used for them play an essential role in establishing their traceability to the International System of Units (SI). Therefore, the HPMs have been developed as certified reference materials of

NMIJ (NMIJ CRMs) to have reliable purity values which are traceable to the SI whenever possible.

The freezing point depression method is one of the methods of choice for obtaining purity values traceable to the SI, since it is approved to have a potential to be operated as a primary method of measurement [7]. For most of the HPMs studied so far at NMIJ, this method has proved to be useful using a differential scanning calorimeter and/or an adiabatic calorimeter [8]. This method was, however, not easily applicable to the PAEs studied except diethyl phthalate (DEP) due to their non-crystalizing property. Thus, chromatographic methods such as gas chromatography-flame ionization detector (GC-FID) and/or high-performance liquid chromatography (HPLC) in combination with other methods such as Karl-Fischer (KF) titration for water analysis and vacuum evaporation (VE) for the analysis of non-volatiles are exploited for the quantification of organic as well as inorganic impurities. The sum of the impurities estimated by chromatography and other methods allows the estimation of the purity of the main component by difference [9, 10]. Here, we denote this method as a "mass balance method". In this report, the practical utility of this method was studied for obtaining reliable purity values of the candidate reference materials of PAEs.

Experimental

NMIJ has a quality management system based on the International Standards Organization (ISO) Guide 34 [11] and ISO/International Electrotechnical Commission (IEC) 17025 [12], which is accredited by the National Institute of Technology and Evaluation (NITE). The certification of the PAEs described in this paper has been carried out according to the technical requirements described in ISO Guide 35 [13]. Seven of the 8 PAEs defined in terms of the Japanese Industrial Standards (JIS) (JIS K 0450-30-10:2006, Testing method for PAEs in industrial water and wastewater) were studied: di-*n*-proyl phthalate (DPrP), di-*n*-butyl phthalate (DBP), di-*n*-pentyl phthalate (DPeP), di-*n*-hexyl phthalate (DHxP), di-2-ethylhexyl phthalate (DEHP), di-cyclohexyl phthalate (DCHP) and butyl benzyl phthalate (BBP). The purity assessment of DEP is not described because it had been independently assessed by the freezing point depression method and certified as an NMIJ CRM 4022-b prior to the development of other PAEs described in this report.

Candidate reference materials of PAEs

The candidate reference materials of the 7 PAEs were prepared and purified by a reference material supplier (Wako pure chemical industries, Ltd., Osaka, Japan) selected according to the criteria of NMIJ. DBP, DEHP and BBP were prepared by vacuum distillation of their raw materials synthesized using commercially available primary grade *n*-alcohols and phthalic anhydride. DPrP, DCHP, DPeP and DHxP were commercially available PAEs pre-selected by the quality evaluation by NMIJ. These materials were subjected to preliminary purity assessment by GC-FID and HPLC at NMIJ before use. The materials except DCHP were colourless oily liquid at room temperature, while DCHP was a white waxy substance. Each 1.5 mL of the material except DCHP was enclosed into 200–300 amber glass ampoules with argon gas, while each 1.5 g of DCHP was charged into 200 glass inserts which were encapsulated into polyethylene containers packed in laminated aluminium bags under argon gas atmosphere. The ampoules and vials were stored in a freezer at −20 °C.

Reagents and chemicals for impurity analysis

Phthalic acid (PAc), phthalic anhydride (PA), dimethyl phthalate (DMP) and dibenzyl phthalate (DBzP) were of reagent grade purchased from Wako Pure Chemical Industries, Ltd. Acetonitrile (ACN) was of HPLC grade purchased from Kanto Chemical Co., Ltd. (Tokyo, Japan). Acetone and toluene were of analytical grade purchased from Kanto Chemical Co., Ltd. Ultrapure water used for HPLC experiments was prepared with a Milli-Q Gradient system (Millipore Co., Billerica, MA, USA).

GC analysis for the quantification of impurities

A model 6890 GC (Agilent Technologies Inc., Santa Clara, CA, USA) equipped with a split/splitless injection port and a flame ionization detector (FID) was operated programming the column oven temperature from 40 °C (5 min hold time) to 300 °C (10 min hold time) at a rate of 10 °C/min. An HP-5 capillary column (30 m × 0.32 mm i.d., 0.25 μm film thickness; Agilent Technologies Inc.) was used for the analysis. The measurement samples of the candidate reference materials were prepared with ACN at a mass fraction of 10 g/kg. The authentic compounds for calibration solutions were prepared with ACN at mass fractions of 20, 40, 60, 80, 100 and 120 mg/kg, for example, depending on the mass fractions of the impurities to be determined. The injection volume and split ratio were set to 1 μL and 50 to 1, respectively.

Although the use of a second column of different polarity might be useful for detecting co-eluting substances, we considered the simultaneous use of GC and LC was enough for the purpose. This was confirmed, for example, by the agreement of the number of major impurities between GC and LC observations of a particular PAE. (Data not shown).

Gas chromatography/mass spectrometry (GC/MS) and direct introduction MS analysis for the identification of impurities

A model 5975B GC/MS (Agilent Technologies Inc.) equipped with a split/splitless injection port and a quadrupole mass detector was used. The capillary column and other measurement conditions were basically the same with those employed for the GC-FID analysis described above. Direct introduction field ionization mass spectrometry was conducted with an HX110/110A magnetic sector tandem mass spectrometer (JEOL Ltd., Tokyo, Japan) equipped with a field ionization source. Each 1 mg of the PAEs charged into a glass capillary filled with alumina powder was introduced to the ion source using a solid introduction probe. Mass spectra were acquired at an ion source temperature of 200 °C. The heater installed in the sample probe was kept off, while the sample was heated indirectly up to about 70 °C by the heat transferred from the ion source.

HPLC analysis for impurity determination

A model 1100 series HPLC (Agilent Technologies Inc.) equipped with a diode array detector was used; 10 mmol/L phosphate buffer (pH 6.0, solvent A) and ACN (solvent B) were used for the mobile phase with a linear gradient from 50% B (10-min hold time) to 100% B (10-min hold time) at a rate of 5%/min. An L-column (250 mm × 4.6 mm i.d., 5 μm octadecylsilica (ODS); Chemicals Evaluation and Research Institute, Tokyo, Japan) was used at a flow rate of 1 mL/min at an oven temperature of 40 °C. PAEs were prepared with ACN at a mass fraction of 1 g/kg. The authentic compounds for calibration solutions were dissolved with ACN at mass fractions 2, 4, 6, 8, 10 and 12 mg/kg, for example. The injection volume was set to 1 μL. Chromatographic data were acquired monitoring the wave lengths of 210, 254 and 275 nm, and data from 210 nm were used for the analyses.

Nuclear magnetic resonance (NMR) measurements for the identification of impurities

NMR measurements were performed on a Varian $^{\text{UNITY}}$INOVA 600A spectrometer (Varian Inc., Palo Alto, CA, USA) equipped with a Varian H{X} pulsed field gradient probe. The temperature was set at 303.15 K. The major impurity separated from DPrP by reversed phase HPLC under the same conditions with those employed for impurity analysis was evaporated and reconstituted with $CDCl_3$ at a mass fraction of 5 g/kg. Other conditions are as follows: spectral band width of 10 kHz, the number of transitions of 16, the integration time of 3.5 s, the delay time of 6.5 s and the pulse width of 3.25 μs at 45° pulse.

KF measurements for water analysis

The mass fractions of water in the PAEs were determined using a KF titrator (Model AQ-7, Hiranuma Sangyo Co., Ltd., Ibaraki, Japan) installed in a model GB2202 glove box (Dalton Co., Ltd., Tokyo, Japan). By flowing dry air, the dew point inside the box was kept below −30 °C. The solid sample of DCHP was introduced directly into the titrator using a polypropylene tubing, while other liquid PAEs were injected using a gas-tight syringe. The sample size was about 0.3–1 g per measurement. HYDRANAL-Aqualyte RS-A (Sigma–aldrich Inc., St. Louis, MO, USA) and Aqualyte CN (Kanto Chemical Co., Ltd.) were used as the anolyte and catholyte, respectively.

VE for the estimation of non-volatile impurities

The mass fractions of non-volatile organic/inorganic impurities in the PAEs were determined using a home-made vacuum evaporator; 0.1–0.2 g of the PAEs were loaded on glass boats weighed precisely using a micro balance (Model MC5, Sartorius Inc., Goettingen, Germany) and were subjected to vacuum evaporation (<500 Pa) at 100 °C for about 100 h. Along with the sample loaded boats, an empty boat was heated under the same condition for reference. After exhaustive evaporation of the PAEs, the difference of the mass of each boat before and after the assay was evaluated by the micro balance which gave an estimate on the absolute masses of non-volatile impurities.

Homogeneity assay

The homogeneity between ampoules was evaluated by analysing 8 ampoules selected randomly out of 200–300. Each of the ampoules was subjected to replicate analyses by KF and HPLC, and analysis of variance (ANOVA) over the data from all ampoules (the mass fractions of water from KF and the relative peak area percentages of the main component from HPLC) was performed independently. The standard deviations between ampoules (s_{ba}) and alternative standard deviations between ampoules in the case of insufficient repeatability of the measurement method (u_{ba}) [14, 15] were calculated as follows.

$$s_{ba} = \sqrt{\frac{MS_{among} - MS_{within}}{n}}$$

where n represents the number of measurements per ampoule. MS_{within} and MS_{among} represent the mean

squares within each group and among the groups, respectively.

$$u_{\mathrm{ba}} = \sqrt{\frac{MS_{\mathrm{within}}}{n}} \sqrt[4]{\frac{2}{v_{MS\mathrm{within}}}}$$

where $v_{MS\mathrm{within}}$ represents the degrees of freedom of MS_{within}. By comparing both values, the bigger one was accepted. Finally, the standard deviations between ampoules obtained from KF and HPLC were combined to give the uncertainty related to the inhomogeneity between ampoules. Here, the variations of water and organic impurities were assumed to be independent each other.

The use of GC instead of HPLC might have been better for the homogeneity assay for DPeP considering the presence of *i*-DPeP, which was one of the major impurities hardly separated by HPLC (see "Result of purity assay of PAEs"). The use of HPLC for homogeneity assay was, however, justified considering that DPeP was liquid at room temperature and, thus, it was very unlikely that this component, which was supposed to have physicochemical properties similar to the main component, was specifically distributed heterogeneously.

Stability assay

The stability of the PAEs was evaluated by analysing one to three ampoules every year for more than 3 years by using HPLC. The relative peak area ratios of the main component were processed by two-stage ANOVA leading to separate the uncertainties associated with between ampoule variation and between storage period variation. The standard deviations between different storage periods were taken as the uncertainties related to the instability of the PAEs.

Method for purity assay by the mass balance method

Classification of the impurities in PAEs

In general, experiments required for the certification of the reference materials include purity assay of the HPMs, homogeneity assay between ampoules and stability assay of the ampoules during long term storage [13]. Since the 7 PAEs studied were not amenable to the purity assay by the freezing point depression method, chromatographic methods such as GC-FID and HPLC in combination with other methods such as KF and VE were exploited to evaluate the mass fractions of all the impurities in each PAE. In planning the purity assay of the PAEs by this mass balance method, we placed a particular emphasis on the following points: (1) We should detect and evaluate as many impurities as possible without omission using multiple methods (2) quantification of each impurity should be performed preferably by more than two independent methods and (3) quantification of impurities must be based on the calibration against reference standards whose purities are assessed by sufficiently validated methods.

In this study, impurities were classified into 5 groups, A to E, as to their qualitative and quantitative nature in order to detect and evaluate various kinds of impurities mutually exclusive and collectively exhaustive. At the same time, the analyses for the purity assay were designed in which GC-FID, HPLC and other methods were exploited in parallel to check the results each other, and in complementary to secure the exhaustiveness of analysis. Parallel use of GC-FID and HPLC was essential for PAEs, because they contained residual phthalic acid and alcohols related to the starting materials for synthesis which could be detected only by HPLC and GC-FID, respectively. The overall scheme of the impurity classification is shown in Fig. 1.

Such impurities as water and non-volatile substances (non-volatile organics, metals, salts, etc.) are not amenable to GC-FID or HPLC analysis under the conditions employed in this study, although some of the non-volatile organics could possibly be detected by HPLC. They are classified into the "Group E" impurities in Fig. 1. For their quantification, KF and VE were used as described below. Residual solvents are also included into group E, although they are normally detectable by GC-FID by injecting the neat materials directly without solvent dilution and hence without huge solvent related peaks that might interfere with the small peaks of residual solvents. As to other organic impurities, some impurities, detectable by GC-FID and/or HPLC in principle, possibly present at levels lower than the detection limits of the measurement methods. These are

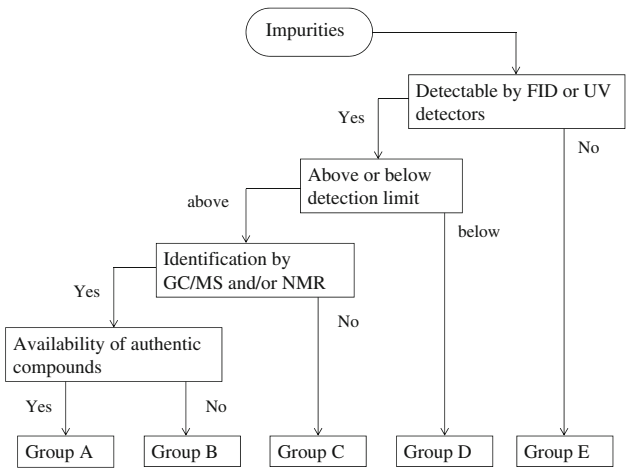

Fig. 1 Schematic diagram of the classification of impurities in PAEs. Impurities were classified into 5 groups A to E as to their qualitative and quantitative nature in order to detect and evaluate various kinds of impurities mutually exclusive and collectively exhaustive

named as "Group D" impurities. For the impurities above detection limit, the way of quantification depends on the success in the identification of their chemical structures by GC/MS and/or NMR. Unidentified impurities are classified as "Group C". Even if structural information of an impurity is available, it is not always possible to obtain an authentic compound(s) of the predicted structure(s). When the structural information of an impurity is available but its authentic compound is not, we assigned the impurity as "Group B". Finally, the impurities identified and quantified with their authentic compounds are categorized to "Group A" in Fig. 1.

Quantification of the impurities of each group

Figures 2, 3, 4, and 5 show the detailed schemes of the quantification of impurity groups A to E. As shown in Fig. 2, a group A impurity is quantified against the calibration curve drawn by the measurements of calibration solutions prepared with its authentic compound by means of both GC-FID and HPLC. The mean of both results gives the final result. The purities of the authentic compounds used are assessed by GC-FID (or by HPLC for compounds not amenable to GC-FID) based on their relative peak area ratios. In case of a group B impurity, an alternative reference compound of analogous structure is used for making calibration. In the making of calibration curves for both groups, a least-square fit of the data points represented by the observed peak areas versus the mass fractions of calibration solutions was applied. In the estimation of the uncertainties originated in the least-square regression, Eq. 1 was used,

$$S_{x0} = \frac{S_{y/x}}{a}\sqrt{\frac{1}{m}+\frac{1}{n}+\frac{(y_0-\bar{y})^2}{a^2\sum_{n=1}^{j}(x_i-\bar{x})^2}} \quad (1)$$

where S_{x0} denotes the estimated standard deviation of the mass fraction x_0 and x_0 denotes the mass fraction for the observed peak area of y_0 predicted by the calibration curve. The parameters m and n are the numbers of replicate measurements of the sample and calibration solutions (number of mass fraction levels (j, normally 6) times number of measurements at each level), respectively. The parameter a represents the slope of the regression line ($y = ax + b$). A parameter with an upper bar represents its averaged value. The statistical parameter $S_{y/x}$ is given by

$$S_{y/x} = \sqrt{\frac{\sum_{i=1}^{j}(y_i-\hat{y})^2}{n-2}}$$

where $\hat{y}$ represents the value of y at x_i predicted by the calibration curve ($\hat{y} = ax_i + b$).

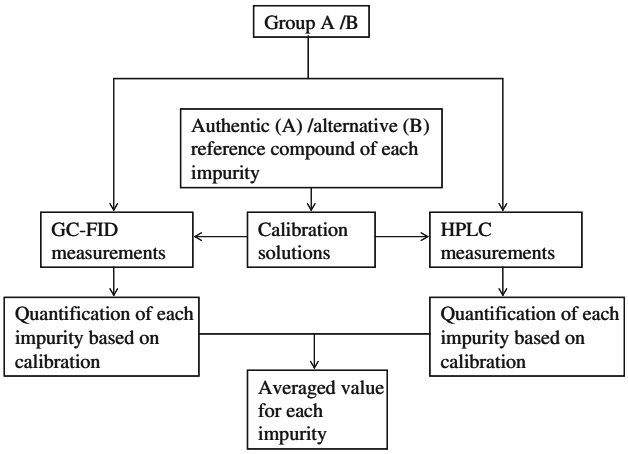

Fig. 2 Schematic diagram of the quantification of group A and B impurities

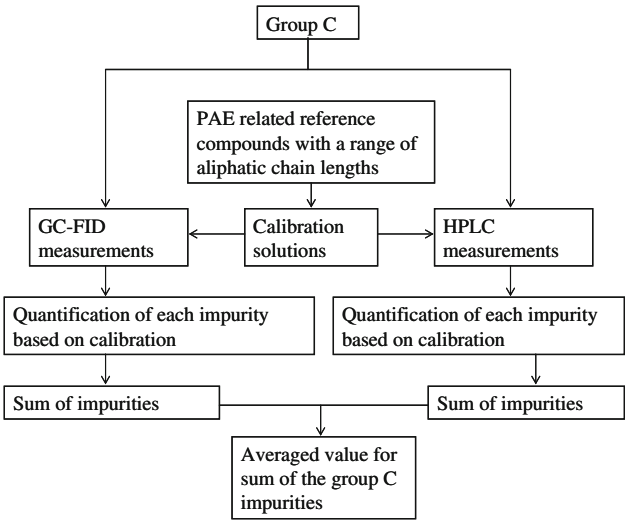

Fig. 3 Schematic diagram of the quantification of group C impurities

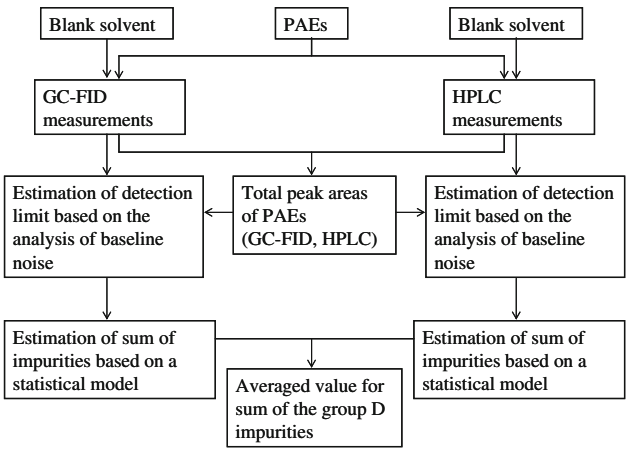

Fig. 4 Schematic diagram of the quantification of group D impurities

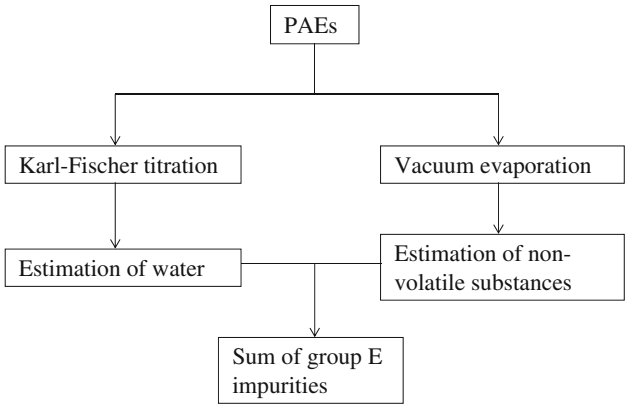

Fig. 5 Schematic diagram of the quantification of group E impurities

In addition to the uncertainty due to regression, repeatability of measurements of sample solutions estimated as the relative standard deviation of the peak areas (s_a) was considered as an additional source of uncertainties. In case of group B impurities, extra uncertainties associated with the evaluation of their relative response factors were included, because the calibration standards used were not exactly identical with the impurities under quantification. For example, this uncertainty was derived from the relative standard deviation (s_r) of the relative response factors of some alternative reference compounds as described in the case of DPrP in the next section.

$$u_c = \sqrt{S_{x0}^2 + s_a^2} \quad \text{(for group A)}$$

$$u_c = \sqrt{S_{x0}^2 + s_a^2 + s_r^2} \quad \text{(for group B)}$$

The quantification of the group C impurities is shown in Fig. 3. One of the difficulties in the qualitative analysis of PAEs is the similarity of mass spectral patterns with different aliphatic chain lengths: Typical mass spectra of PAEs exhibit characteristic fragment ions at *m/z* 149 with minor contribution from peaks related to molecular ions and other structure reflecting fragment ions. This makes it difficult to identify the side chain structures of trace-amount PAEs. Here, the majorities of unidentified impurities are assumed to be PAEs or their related compounds. Their response factors are approximated by PA and the PAEs with aliphatic chains composed of 2–8 carbon atoms each (corresponding to DEP to DEHP) referring to their retention times on the chromatograms. Except PA, they are candidate reference materials (DPrP, DBP and DEHP) or an NMIJ CRM (DEP) as denoted by "C_{std}" in Table 1. The gravimetric mixture of these reference standards was used for the calibration of the impurity peaks appearing around them on the GC-FID chromatograms as shown in Fig. 6. Here, a single-point calibration was used for simplicity. The sums of the mass fractions of the unidentified impurities estimated independently by GC-FID and HPLC are averaged to give the final result for the group C impurities (Fig. 3). Uncertainties associated with the results included measurement repeatability and uncertainty due to the unknown relative response factors of impurities. The latter was estimated by assuming a uniform distribution of response factors ranging from 0.5 to 2 relative to that of the reference PAE used. This assumption led to generate a relative standard uncertainty of about 25% around the centre value (1 ± 0.25).

Figure 4 shows the scheme of quantification of the group D impurities. The analysis of baseline noise on the GC-FID chromatograms proved that the distribution of signals can be approximated by a normal distribution with a standard deviation of about 4 μV. Thus, a minimum detectable peak height was set to the three times of this value, 12 μV. As a typical peak with a 12 μV height corresponded to a peak area of about 10 μV s, a peak area of 10 μV s was adopted as the detection limit (*DL*) expressed as peak areas of the GC-FID chromatograms in this study. Similarly, *DL* of HPLC measurement expressed as a minimum detectable peak area was estimated to be 0.07 mAU s, where "AU" denotes absorbance unit. The sum of undetectable impurities and uncertainty associated with *DL* were estimated as follows. Considering the peak widths of typical peaks (0.2 min for GC-FID) and the measurement time of 20 min, the maximum number of peaks observed during this measurement time was supposed to be 100. Thus, the most probable number of impurity peaks below detection limit was assumed to be the half of this value, 50. Assuming the intensities of undetectable impurities are distributed uniformly between zero and *DL*, sum of the peak areas of 50 peaks is expected approximately to have a normal distribution with a mean value of $50\frac{DL}{2} = 25\ DL$ with a standard deviation of $50\frac{DL}{2\sqrt{3}} = 14.4\ DL$, which is predicted by the central limit theorem [16, 17]. This theorem states conditions under which the mean of a sufficiently large number of independent random variables, each with finite mean and variance, will be approximately normally distributed [18]. By substituting *DL* with 10 μV s, the sum of the peak areas of the group D impurities was estimated to be (250 ± 144) μV s. For example, the sum of the peak areas of the main peak and detectable impurity peaks was 2074000 μV s for DPrP, the relative abundance of the group D impurities was estimated to be (0.12 ± 0.07) g/kg (see Table 3) assuming the same relative response factors of impurities with that of DPrP.

Figure 5 shows the scheme of quantification of the group E impurities. The mass fraction of water estimated by KF and non-volatile impurities estimated by VE were taken into account for the quantification of the group E impurities. For all the PAEs studied, non-volatile

Table 1 Purity in mass fraction of the NMIJ CRM of diethyl phthalate (DEP) and those of the 7 candidate reference materials of phthalic acid esters (PAEs)

CRM #	Name	Symbol	Purity (kg/kg)[a]	Remarks
4022-b	Diethyl phthalate[b]	DEP	0.9998 ± 0.0001	NMIJ CRM4022-b, C^{c}_{std}
4023-a	Di-*n*-butyl phthalate	DBP	0.9996 ± 0.0002	C_{std}, Type "H"[d]
4024-a	Di-2-ethylhexyl phthalate	DEHP	0.9994 ± 0.0002	C_{std}, Type "H"
4025-a	Di-*n*-propyl phthalate	DPrP	0.9841 ± 0.0012	C_{std}, Type "L"[e]
4026-a	Di-*n*-pentyl phthalate	DPeP	0.979 ± 0.005	Type "L"
4027-a	Di-*n*-hexyl phthalate	DHxP	0.979 ± 0.006	Type "L"
4028-a	Di-cyclohexyl phthalate	DCHP	0.9979 ± 0.0008	Type "H"
4029-a	Butyl benzyl phthalate	BBP	0.998 ± 0.0011	Type "H"

[a] Uncertainties expressed as expanded uncertainties determined using a coverage factor $k = 2$

[b] Purity determined by the freezing point depression method

[c] Used as reference standards in the quantification of group C impurities

[d] Purity higher than 0.99 kg/kg with many identified/unidentified impurities of low abundance

[e] Purity lower than 0.99 kg/kg with a few identified impurities of high abundance

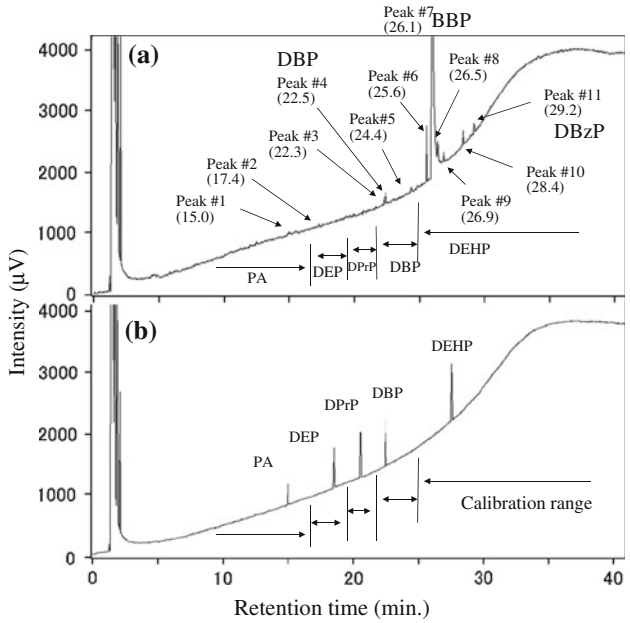

Fig. 6 Assignment of reference compounds for the quantification of group C impurities (a) and the chromatogram of butyl benzyl phthalate (b, BBP) observed by GC-FID analysis. Reference standards are phthalic anhydride (PA), diethyl phthalate (DEP), di-*n*-proyl phthalate (DPrP), di-*n*-butyl phthalate (DBP) and di-2-ethylhexyl phthalate (DEHP)

impurities were not detected. The amount of non-volatile impurities was estimated as a half of the detection limit of the micro balance, 26 μg, which was taken from the expanded uncertainty described in the JCSS calibration certificate of the balance. Uncertainty of the mass of non-volatile impurities was estimated assuming a uniform distribution between zero and this detection limit. When 0.2 g of PAEs were loaded, the mass fraction of non-volatile impurities was estimated to be (65 ± 38) μg/g ($26/(0.2 \times 2) = 65$ and $26/(0.2 \times 2 \times \sqrt{3}) = 37.5$, respectively). Since most of the PAEs studied are in the liquid phase at room temperature, the detection of residual solvents was performed by injecting their neat materials into GC without being interfered by the excess solvents as described above. The absence of residual solvents in DCHP (solid at room temperature) was confirmed by the GC-FID measurements of DCHP solutions prepared with acetone and then toluene as solvents. Here, residual solvents appeared at the retention time around acetone were observed by the measurement of toluene solution and vice versa. For all the PAEs studied, residual solvents were not detected. Considering the extremely low detection limit of GC-FID of neat PAEs ($<10^{-7}$ kg/kg), the contribution of the residual solvents below the detection limit was negligible and omitted in the evaluation of group E impurities (Fig. 5).

Result of purity assay of PAEs

Table 1 lists the purities in mass fraction of DEP and the 7 candidate reference materials of the PAEs studied in this paper. The uncertainties in the table are expressed as expanded uncertainties determined using a coverage factor $k = 2$, corresponding to an estimated confidence interval of approximately 95%. The purity of DEP already disseminated as an NMIJ CRM 4022-b had been determined by the freezing point depression method using an adiabatic calorimeter. DEP and three PAEs denoted by "C_{std}" in the table were used as the reference standards in the quantification of group C impurities as described above. There found roughly two types in the PAEs studied; DBP, DEHP, DCHP and BBP contained many unknown impurities of very low abundance leading to their purities around

Table 2 Estimation of the relative molar response factor of the group B impurity (di-*n*-propyl phthalate, MPrP) in DPrP in the GC-FID measurements based on those of dimethyl phthalate (DMP), diethyl phthalate (DEP) and di-*n*-propyl phthalate (DPrP)

	CN[a]	ECN[b] (2*C=O, 2*–O–)	RRF (calc)[c] (DEP = 1)	RRF (meas)[d] (DEP = 1)	$\frac{\text{RRF(meas)}}{\text{RRF(calc)}}$
DMP	8	8–1.2 = 6.8	0.773	0.782	1.012
DEP	10	10–1.2 = 8.8	1.000	1.000	1.000
MPrP	10	10–1.2 = 8.8	1.000	–	1.008 ± 0.007[e]
DPrP	12	12–1.2 = 10.8	1.227	1.244	1.014

[a] Carbon number

[b] Effective carbon number

[c] Relative response factor derived from calculation based on the effective carbon number

[d] Relative response factor derived from measurements

[e] Relative response factor estimated for MPrP with a standard uncertainty

0.998 kg/kg or higher (denoted by type "H" in the table); DPrP, DPeP and DHxP contained a few identified impurities of relatively high abundance degrading their purities down to around 0.98 kg/kg or lower (type "L"). Results of purity assay are described below in more detail for DPrP (type "L") and BBP (type "H") as typical examples.

In the GC-FID and HPLC chromatograms of DPrP, the most abundant impurity peaks exhibited about 2% of the total peak areas. Direct introduction field ionization mass spectrometry suggested that this major impurity had a relative molecular mass identical with that of DEP (222.2). The retention times observed in GC-FID and HPLC suggested, however, this impurity was not identical with DEP. For precise identification of this impurity, HPLC fractions were accumulated to obtain sufficient amount of the purified material of this impurity amenable to ^{1}H-NMR measurements. Figure 7 shows the assignments of major peaks observed for the impurity (a), DMP (b) and DPrP (c). As shown in Fig. 7, the aromatic protons α and β were common for all three PAEs, while protons γ were observed only for DMP and the impurity MPrP and protons γ', δ and ε only for DPrP and MPrP. The peak intensities of protons γ, γ', δ and ε for the impurity were about the halves of those observed for DMP and DPrP. This observation and the molecular weight information of the impurity strongly suggested that it had both methyl and *n*-propyl groups in the molecule with a deduced structure of methyl *n*-propyl phthalate (MPrP).

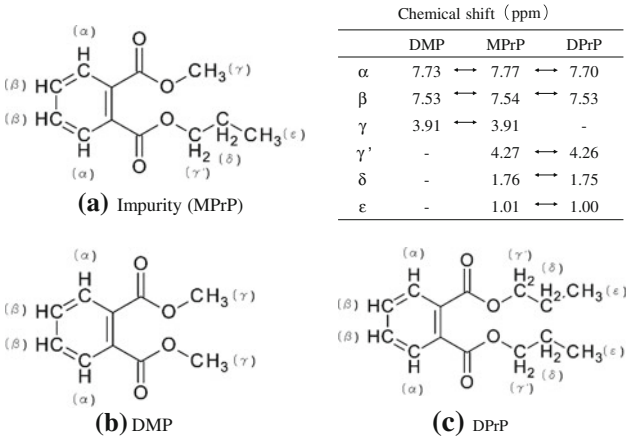

Fig. 7 Assignments of NMR peaks observed for the group B impurity of DPrP (**a**, methyl *n*-propylphthalate: MPrP), dimethyl phthalate (**b**, DMP) and di-*n*-propyl phthalate (**c**, DPrP)

Although the use of the authentic compound of MPrP might be feasible for the quantification of the impurity, this compound was not commercially available. Relative molar response factor (RRF) of MPrP was estimated based on those of DMP, DEP and DPrP as shown in Table 2. In GC-FID analysis, RRF of a compound is approximated by the effective carbon number (ECN). For example, ECN of DEP is calculated as 6.8 (=8 + 0.2×2 − 0.8×2) considering the presence of a couple of carbonyl groups (+0.2×2) and a couple of ether oxygens (−0.8×2). Similarly, ECNs of DMP, MPrP and DPrP were calculated to be 6.8, 8.8 and 10.8, respectively. Based on these values, RRFs of DMP and DPrP were calculated with reference to that of DEP as 1. The ratios of measured and calculated RRFs were approximately 1 for these three PAEs. The RRF of MPrP was estimated from the average of the three values (1.008 = (1.012 + 1.000 + 1.014)/3) with an uncertainty estimated by the standard deviation of these values (0.007). This RRF based on a unit mole is equal to an RRF based on a unit mass for MPrP, because the molecular weight of MPrP is identical with that of DEP. Thus, the impurity MPrP was quantified at the mass fraction using DEP with a correction factor of its RRF of 1.008 ± 0.007 relative to that of DEP. The RRF for HPLC analysis was estimated similarly based on the approximately equivalent mole sensitivities of the three PAEs. The RRF of MPrP was calculated to be 0.987 ± 0.025 for HPLC recorded at 275 nm.

Table 3 shows the summary of results from GC-FID and HPLC analyses of the "Type L" PAEs (DPrP, DPeP and

Table 3 Results of purity assay obtained for the phthalic acid esters (PAEs) with purities lower than a mass fraction 0.99 g/kg (di-*n*-proyl phthalate (DPrP), di-*n*-pentyl phthalate (DPeP), di-*n*-hexyl phthalate (DHxP): Type "L") Authentic compounds of phthalic acid (PAc), phthalic anhydride (PA) and 1-pentyl alcohol (PT) were used as reference standards. All values are given in g/kg

		A	B	C	D	E	Sum A-E
DPrP	GC-FID	–	15.3 ± 0.4	–	0.12 ± 0.07	–	
	HPLC	0.03 ± 0.001	15.6 ± 0.3	0.20 ± 0.14	0.09 ± 0.06	–	
	Mean	0.03 ± 0.001	15.45 ± 0.25	0.10 ± 0.07	0.11 ± 0.06	0.27 ± 0.003	15.96 ± 0.26
DPeP	GC-FID	2.99 ± 0.09	14.6 ± 2.2	2.1 ± 0.8	0.12 ± 0.08	–	
	HPLC	3.16 ± 0.17	–	3.7 ± 1.1	0.12 ± 0.08	–	
	Mean	3.08 ± 0.11	14.6 ± 2.2	2.9 ± 0.8	0.12 ± 0.06	0.37 ± 0.04	21.1 ± 2.3
DHxP	GC-FID	4.76 ± 0.09	7.8 ± 1.2	10.7 ± 4.0	0.12 ± 0.08	–	
	HPLC	4.12 ± 0.16	7.2 ± 1.1	7.7 ± 3.1	0.12 ± 0.08	–	
	Mean	4.44 ± 0.18	7.5 ± 1.2	9.2 ± 2.6	0.12 ± 0.06	0.15 ± 0.04	21.4 ± 2.8

A: PAc, PA in DPrP, DPeP and DHxP and PT in DPeP

E: Values for water and non-volatiles commonly used for both GC-FID and HPLC data

±: Figures after ± represent standard uncertainties

DHxP). The group A impurity of DPrP, phthalic acid (PAc) was quantified only by HPLC using the authentic compound. The mean value for PAc was taken from the result of HPLC, because the absence of the result for PAc in the GC-FID measurements was due to its non-volatility. In the group B column, the GC-FID and HPLC assays gave similar results on the mass fractions of MPrP. Cross-checking of the results from both chromatographic techniques serves for detecting the possible method-dependent biases and for ensuring the results of purity determination. There also found no peak for the group C impurity by GC-FID. The mean value for the group C impurity was estimated as the half of that determined by HPLC, because the identities of group C impurities were unknown and it was not clear that they were not observed because of their non-volatility. The group E impurities were from the results from KF and VE and were used for both GC-FID and HPLC in common. The sum of the group A to E impurities counted up to (15.96 ± 0.26) g/kg leading to the purity of DPrP to be (0.984 ± 0.00026) kg/kg. Major impurities in DPeP and DHxP also came from Group B. Relatively large uncertainty of the sum of the impurities in DPeP mainly owed to that of an impurity eluted just before DPeP. This impurity was supposed to be a structural isomer of DPeP based on the similar retention time observed by GC-FID and the identical mass spectral pattern observed by

Table 4 Summary of results from GC-FID and HPLC analyses obtained for the PAEs with purities higher than 0.99 kg/kg (DBP, DEHP, DCHP and BBP: Type "H"). All values are given in g/kg

		A	B	C	D	E	Sum A-E
DBP	GC-FID	–	–	0.048 ± 0.012	0.13 ± 0.08	–	
	HPLC	–	–	0.15 ± 0.04	0.10 ± 0.08	–	
	Mean	–	–	0.10 ± 0.03	0.12 ± 0.057	0.21 ± 0.04	0.42 ± 0.07
DEHP	GC-FID	0.10 ± 0.01	–	0.18 ± 0.04	0.11 ± 0.07	–	
	HPLC	–	–	0.21 ± 0.06	0.11 ± 0.08	–	
	Mean	0.10 ± 0.01		0.20 ± 0.04	0.11 ± 0.05	0.18 ± 0.04	0.59 ± 0.08
DCHP	GC-FID	–	–	1.80 ± 0.40	0.12 ± 0.08	–	
	HPLC	–	–	1.88 ± 0.53	0.09 ± 0.06	–	
	Mean	–	–	1.84 ± 0.33	0.11 ± 0.05	0.11 ± 0.05	2.05 ± 0.34
BBP	GC-FID	0.24 ± 0.02	–	1.2 ± 0.5	0.12 ± 0.08	–	
	HPLC	0.23 ± 0.02	–	1.8 ± 0.9	0.09 ± 0.06	–	
	Mean	0.24 ± 0.02	–	1.5 ± 0.5	0.11 ± 0.05	0.15 ± 0.02	2.0 ± 0.5

A: 2-ethylhexanol in DEHPand di-*n*-butyl phthalate (DBP) and dibenzyl phthalate (DBzP) in BBP

B: No peak found in this group

E: Values for water and non-volatiles commonly used for both GC-FIDand HPLC data

±: Figures after ± represent standard uncertainties

Table 5 Result of group C impurities in BBP determined with GC-FID (a) and HPLC (b)

(a) Peak #	RT (min)	Assignment	Peak area (μV s)	Standard	Response factor (μV s kg/g)	Mass fraction (g/kg)
1	15.0		60.7	PA	681	0.09
2	17.4		50.7	DEP	1202	0.04
3	22.3		160.3	↑	↑	0.10
4	22.5	DBP	188.7	DBP	1558	–
5	24.4		69.7	↓	↓	0.04
6	25.6		955.7	↑	↑	0.52
7	26.1	(BBP)	2064712			–
8	26.5		289.7	DEHP	1824	0.16
9	26.9		152.3	↓	↓	0.08
10	28.4		242.3			0.13
11	29.2	DBzP	183.7			–
Total						1.18
(b) Peak #	RT (min)	Assignment	Peak area (mAU s)	Standard	Response factor (mAU s kg/g)	Mass fraction (g/kg)
1	3.1		2.22	↑	↑	0.06
2	4.1		1.58			0.04
3	5.7		7.36	PA	35.66	0.21
4	6.2		0.88	↓	↓	0.02
5	7.2		2.58			0.07
6	10.4		8.31	↑	↑	0.25
7	13.8		5.17	DEP	33.37	0.15
8	14.5		0.58	↓	↓	0.02
9	17.7		2.22	DPrP	35.52	0.06
10	20.0	DBzP	4.95	↑	↑	–
11	20.4	(BBP)	18885			–
12	20.8	DBP	6.42	DBP	31.91	–
13	21.2		6.67	↓	↓	0.21
14	22.5		0.91			0.03
15	23.3		4.12			0.13
Total						1.26

Table 6 Summary of uncertainty estimation from the purity, homogeneity and stability assays of PAEs. All values are given in g/kg

CRM #	Symbol	Sources of uncertainties			Standard uncertainty u_c	Expanded uncertainty U ($k = 2$)
		u_p	u_H	u_s		
4023-a	DBP	0.07	0.05	0.004	0.09	0.17
4024-a	DEHP	0.08	0.04	0.018	0.09	0.18
4025-a	DPrP	0.3	0.08	0.5	0.6	1.2
4026-a	DPeP	2.3	0.06	0.32	2.3	4.6
4027-a	DHxP	2.8	0.11	0.4	2.8	5.7
4028-a	DCHP	0.34	0.19	0.024	0.39	0.78
4029-a	BBP	0.50	0.19	0.01	0.53	1.1

[a] The uncertainties derived from the purity (u_p), homogeneity (u_H) and stability (u_S) assay

GC/MS. This impurity tentatively denoted by *i*-DPeP was quantified using DPeP as the calibration standard with a large uncertainty due to the unknown RRF of the impurity. This resulted in a relatively large uncertainty for Group B impurity observed by GC-FID. On the other hand, this impurity was not observed by HPLC because of its limited resolution. Thus, the mean value for *i*-DPeP was taken from the result of GC-FID. The major impurity of Group A in DHxP, PA, was quantified both by GC-FID and HPLC using the authentic compound. The results from both methods, however, were not identical within their uncertainties, suggesting the presence of some unknown effects, for example, the presence of an impurity co-eluting with PA. In this case, the uncertainty (0.18 g/kg) for the mean value (4.44 g/kg) was estimated by taking the square root of the sum of squares of a half of the uncertainty obtained from GC-FID (0.09/2 g/kg = 0.045 g/kg) that from HPLC (0.16/2 g/kg = 0.08 g/kg) and a quarter of the difference of values between GC-FID and HPLC ((4.76 − 4.12)/4 g/kg = 0.16 g/kg).

$$\sqrt{0.045^2 + 0.08^2 + 0.16^2}\,\text{g/kg} = 0.184\,\text{g/kg}$$

Table 4 shows the summary of results from GC-FID and HPLC analyses of the type "H" PAEs (DBP, DEHP, DCHP and BBP). Most of the impurities were assigned to the groups C, D and E with a minor contribution from group A. Based on the sum of impurities, the purity of DBP was estimated to be (0.9996 ± 0.00007) kg/kg. In case of BBP, all but group B impurities were found at levels 0.1–1 g/kg. It was noted the impurities A1 (DBP) and A2 (DBzP) were symmetric diesters supposed to originate in the by-products of synthesis. This is in contrast with the presence of an asymmetric diester MPrP as the major impurity in DPrP. Impurities A1 and A2 were quantified using their authentic compounds both by GC-FID and by HPLC. Uncertainties were estimated combining those from regressions estimated by Eq. 1 and those from the repeatability of measurements. Table 5 shows the result of group C impurities in BBP determined with GC-FID (a) and HPLC (b). In the GC-FID and HPLC measurements, 11 and 15 peaks were observed, respectively. The peaks except those assigned to DBP, BBP and DBzP were calibrated against the response factors of PA and 3 PAEs (DPrP, DBP and DEHP) as reference standards. The peak #6 appeared just before the main peak in GC-FID could not be separated by HPLC, which was confirmed by the GC-FID measurement of the HPLC fraction taken at the main peak of BBP. The sum of group C impurity observed by HPLC was corrected for the amount of this impurity (0.52 g/kg) to give 1.8 g/kg (=(1.26 + 0.52) g/kg). Based on the sum of impurities, the purity of BBP was estimated to be (0.998 ± 0.00054) kg/kg as shown in Table 4.

Table 6 shows the summary of uncertainty estimation of the purity values of the PAEs studied. Sources of uncertainty included those from the purity assay (u_P), homogeneity assay (u_H) and stability assay (u_S). The combined standard uncertainty u_P was derived from the purity assay described above. The uncertainties u_H and u_S were the estimates derived from the homogeneity and stability assays described in the experimental section, respectively. As seen from the table, the uncertainties mostly came from those of u_P with minor contributions from those of u_H and u_S. These uncertainties were taken as those associated with the certified values.

Concluding remarks

Our goal as a CRM producer is to establish highly reliable certified values which are traceable to the SI units as far as possible. In this line, the freezing point depression method might be a method of choice so long as the method can be applied to the compound under investigation. There are, however, many classes of compounds like the PAEs studied in this paper which are not amenable to this method. A systematic way of a mass balance method proposed in this study is in principle more versatile and would be applicable to relatively wide range of compounds, although the presence of unexpected biases in the quantification of unidentified and/or undetectable impurities sets a limitation which should be cross-checked by other independent methods like quantitative NMR [19]. This method, utilized mostly for qualitative analysis in this study, is under development as a promising method for quantitative analysis at many laboratories including NMIJ [20] and expected to emerge as one of the versatile methods for purity assay complementary to the freezing point depression method and the mass balance method in the near future.

Acknowledgments We acknowledge the collaborators for the preparations of the candidate reference materials, and the NMIJ analysts who helped in the development of the CRMs described in this paper.

References

1. Mackintosh CE, Maldonado JA, Ikonomou MG, Gobas FAPC (2006) Environ Sci Technol 40:3481–3488
2. Liu Y, Guan Y, Yang Z, Cai Z, Mizuno T, Tsuno H, Zhu W, Zhang X (2006) Ecotoxicology 18:293–303
3. Internet: http://echa.europa.eu/chem_data/authorisation_process/candidate_list_table_en.asp
4. Internet: http://www.houko.com/00/01/S22/233.HTM
5. Kaarls R (2006) Accred Qual Assur 11:162–171
6. Internet: http://www.jcsslabo.or.jp/
7. Milton MJT, Quinn TJ (2001) Metrologia 38:289–296

8. Shimizu Y, Ohte Y, Bao X, Otsuka S, Kitamaki Y, Ishikawa K, Ihara T, Kato K (2008) Accred Qual Assur 13:389–396
9. Duewer D, Parris R, White E, May W, Elbaum H (2004) NIST Special Publication (SP 1012)
10. King B, Westwood S (2001) Fresenius J Anal Chem 370:194–199
11. International Organization for Standardization (2000) ISO guide 34: 2000 general requirements for the competence of reference material producers, 2nd edn. ISO, Geneva
12. International Organization for Standardization (1999) ISO/IEC 17025: 1999 General requirements for the competence of testing and calibration laboratories. ISO, Geneva
13. International Organization for Standardization (1989) ISO Guide 35: 1989 certification of reference materials-general and statistical principles, 2nd edn. ISO, Geneva
14. International Organization for Standardization (2006) ISO guide 35: 2006 reference materials-general and statistical principles for certification, 3rd edn. ISO, Geneva
15. Linsinger TP, Pauwels J, van der Veen AMH, Schimmel H, Lamberty A (2001) Accred Qual Assur 6:20–25
16. Fischer H (2011) History of the central limit theorem. Springer, New York
17. Rice J (1995) Mathematical statistics and data analysis, 2nd edn. Duxbury Press, Belmont
18. Internet: http://en.wikipedia.org/wiki/Central_limit_theorem
19. Wells RJ, Hook JM, Al-Deen TS, Hibbert DB (2002) J Agric Food Chem 50:3366–3374
20. Saito T, Ihara T, Koike M, Kinugasa S, Fujimine Y, Nose K, Hirai T (2009) Accred Qual Assur 14:79–86