Research Article

Drug Testing and Analysis

Received: 3 July 2009 Revised: 10 November 2009 Accepted: 10 November 2009 Published online in Wiley Interscience: 3 February 2010

(www.drugtestinganalysis.com) DOI 10.1002/dta.94

Estimating measurement uncertainty in quantitative methods not based on chromatography for doping control purposes

P. Van Eenoo,[a]* P. Van Renterghem,[a] C. H. Dimopoulou,[b] F.T. Delbeke[a] and C. G. Georgakopoulos[b]

Estimation of measurement uncertainty (MU) for quantitative results is a requirement of ISO/IEC17025. This concept is well established for chromatographic methods in doping control and forensic analysis. For non-chromatographic methods, however, very few practical methodologies have been published. In this paper, the applicability of a top-down model, established for estimating uncertainty in chromatography, was evaluated for two other methodologies with different sets of raw data as a starting point. The first case study involves the estimation of MU for the determination of haematological parameters. In this case, a large data set of quality control material and proficiency testing results was available to establish MU. The second case study involves the estimation of MU for the recently approved method for the determination of human growth hormone misuse. In this case the amount of data available to establish MU was limited to results from method validation and a basic set of analysis data. In both cases a methodology based upon long-term bias, long-term imprecision and – eventually – a correction for standard impurity is proposed. The proposed methodology can be regarded as a dynamic procedure, which allows re-evaluation of MU on a regular basis. Finally, a concept for the verification and evaluation of MU estimations using proficiency testing results is proposed. **Copyright © 2010 John Wiley & Sons, Ltd.**

Keywords: haematological passport; growth hormone; haemoglobin; decision limit

Introduction

Measurement uncertainty (MU) is a parameter associated with the result of a measurement that characterizes the dispersion of the values that could reasonably be attributed to the measure.[1–4] Hence, this parameter is an estimate that characterizes the range within which a measured value is asserted to lie within a certain probability. In doping control, MU needs to be determined for all quantitative tests[5] and needs to be taken into consideration in the evaluation process of compliance of samples with the anti-doping regulations.[6,7]

Therefore, since this parameter is an estimate, it should be noted that for MU there is – in contrast to other parameters analytical chemists deal with – no true value. As a consequence, several calculation methods can be used to derive this estimate. In general, two models are widely used: a bottom-up approach and a top-down approach.

The 'bottom-up'[8] or 'deconstructive' approach as proposed in the *Guide to the Expression of Uncertainty in Measurement*,[4] evaluates risks neglecting important uncertainty components resulting from matrix effects, sampling operations, or interferences.[9] It is widely used and suitable in metrology and physics. However, this approach is often considered to underestimate the complexity of analytical chemistry.[12]

An alternative approach for estimating measurement uncertainty is the 'top-down approach' where the total variability of the measurement results is obtained. The data used in this approach can be derived from several sources, including collaborative studies, method validation, and proficiency tests. In using this approach, care must be taken that all components, not only imprecision but also method bias, are accounted for in the uncertainty budget.[9]

Because MU is a prerequisite for all quantitative methods used in anti-doping laboratories by the World Anti-Doping Agency's (WADA) International Standard for Laboratories (ISL)[5] and ISO/IEC17025, several methodologies directly applicable to the anti-doping field have been published.[10,11] One of the described models is based on the use of three factors: long-term bias, long-term imprecision and an impurity factor (for the calibration material), as shown in Eqn (1).[10] However, these methodologies were set up for chromatographic methods. Lately, several other quantitative analytical techniques have been introduced to the field of anti-doping and MU methodologies need to be developed for these methods as well.

$$u = \sqrt{(u_{Precision})^2 + (u_{bias})^2 + (u_{impurity})^2} \quad (1)$$

The two most notable 'new' methodologies are the measurement of haematological parameters for the biological passport[17,19] and the detection method for the misuse of

* *Correspondence to: P. Van Eenoo, Ghent University (UGent), Department of Clinical Chemistry, Microbiology and Immunology, DoCoLab Technologiepark 30, B-9052 Zwijnaarde, Belgium. E-mail: Peter.VanEenoo@Ugent.be*

a *DoCoLab, Ghent University (UGent), Technologiepark 30, B-9052 Zwijnaarde, Belgium*

b *Doping Control Laboratory of Athens, Olympic Athletic Center of Athens 'Spyros Louis', Kifisias 37, 15123 Maroussi, Greece*

Copyright © 2010 John Wiley & Sons, Ltd.

recombinant human growth hormone, based on the ratio of different isoforms.[21]

Both methodologies are similar to those used in the field of clinical chemistry. Within this field, the concept of biological variation, which states that the contribution of the analytical determination is negligible compared to the intra-individual variation as long as the analytical variation is smaller than a reference value, is generally accepted.[13] Within this concept, it is accepted that an analytical goal might not be clinically or physiologically relevant. Also, according to this concept, the analytical coefficient of variation (CVa) needs to be significantly smaller than the intra-individual coefficient of variation (CVi), i.e., CVa $<$ k CVi. There are three levels of analytical goal for imprecision based on intra-individual biological variation (k $=$ 0.25: optimum – k $=$ 0.5: desirable and k $=$ 0.75: minimum).[13]

For many clinical parameters, reference values for CVi have therefore been established. However, the concept of biological variation is in contradiction to ISO/IEC17025 for doping control. Indeed, the International Standard for Laboratories from WADA and ISO/IEC17025 stipulate that MU needs to be established.[4,5] Moreover, it should be noted that even in clinical chemistry it is accepted that only when tests are clinically interpreted with reference values produced by the same analytical method, analytical bias should not introduce uncertainty additional to the imprecision of the method.[13] However, interpretation of results using clinical decision limits determined by other methods is in principle quite doubtful, and at least an estimate of bias needs to be included.

Finally, it is clear that the fundamental role of clinical chemistry and doping control laboratories is to routinely produce results that are fit for purpose; i.e., that they have appropriate analytical accuracy and imprecision for the purpose(s) to which they are applied. These purposes are clearly different between doping control laboratories and clinical chemistry laboratories and therefore suitable methodologies for the MU need to be established.

This paper discusses an approach to determine the MU for these types of measurements as well as an evaluation procedure for MU. It should, however, be noted that in cases where population-based thresholds (based upon a large number of analyses) have been established using identical tests, discussion exists whether or not in these cases MU isn't already incorporated into the thresholds.[23,24] Nevertheless, in these cases, establishing the MU of the analytical methodology can still be utilized to assess the correct implementation of a methodology in an individual laboratory.

Materials and Methods

Haematological profiling

A Sysmex XT 2000i (Sysmex, Chuo-ku, Japan) was used for the measurement of haematological parameters. e-Check (Sysmex) quality control (QC) stabilized blood samples (3 levels) were used for daily control of the instrument and the proficiency testing sample of the WHO recognized Centre Suisse de Controle de Qualité (CSCQ, Chêne-Bourg, Swizerland) were analyzed monthly (May 2006 to January 2009).

On a daily basis, the quality control samples (January 2006 to January 2009) from Sysmex were measured twice before use of the instrument and once more before shut-down. The measured QC-values need to be within the limits fixed by the supplier. Because the manufacturer used biological material as a starting material for the QCs, the assigned values differed from batch to batch. An example of the values for such QC-material is shown in Table 1. Prior to measurement of the blood samples, the samples were homogenized on a spiramix 5 (Denley Instruments Ltd. Billinghurst, UK) roller for a minimum of 15 min. The samples were consecutively inverted 8 times before measurement. The samples were analyzed in duplicate. Maximum differences between hemoglobin concentrations of both measurements were 0.1 g/dl and for the percentage of reticulocytes 0.15% (if %RET $<$1.0%) or 0.25% (if %RET $\geq$1.0%).

Table 1. Example of the reference values for the quality control samples (e-Check) at 3 different levels from Sysmex

Parameter	Level1	Level 2	Level 3
WBC ($10^3/\mu l$)	2.87	6.85	16.24
RBC ($10^6/\mu l$)	2.38	4.45	5.38
HGB (g/dl)	5.8	12.1	16.7
HCT (%)	17.9	36.1	48.8
MCV (fL)	75.2	85.2	90.7
MCH (pg)	24.4	27.2	31.0
MCHC (g/dl)	32.4	33.5	34.2
PLT ($10^3/\mu l$)	55	207	490

Measurement of the growth hormone isoforms

The QC samples (supplied with the kits) were analyzed using two different isoform differential immunoassays (kit 1 and kit 2) from CMZ (CMZ-Assay GmbH, Berlin, Germany)[23–26] and measured using a LB953 Universal Automatic Tube Luminometer (Berthold, Bad Wildbad, Germany) according to the manufacturer's and WADA's recommendations.[26]

Results and Discussion

Haematological profiling

The method for haematological profiling has been used for over three years in the laboratory and is currently used for the determination of blood parameters within the program of the biological passport of the International Cycling Union (UCI). The blood passport, as used currently by UCI, contains several haematological parameters (Table 2), and has evolved from the so-called health controls or haematocrit controls. In 1997, these controls started with the mere measurement of the hematocrit. The array of measured parameters continuously expanded and additionally, an OFF-hr score[14] based on the haemoglobin (Hgb) and percent reticulocytes (Ret%) parameters has been included in the evaluation procedure.

In contrast to the health controls, the blood passport uses individual reference ranges per athlete.[17]

Although the technology used for the determination of these parameters is the same as in clinical chemistry, in the field of anti-doping MU needs to be established. Because the measurement of the different parameters requires virtually no sample treatment, the application of a top-down approach seems most suitable. Therefore the model previously described for the determination of MU in quantitative chromatography methods (Eqn (1)) seems applicable. This model incorporates a long-term imprecision factor,

Copyright © 2010 John Wiley & Sons, Ltd.

Table 2. Long-term bias derived from proficiency tests (PT) and long-term imprecision, expressed as a CV, derived from quality control samples, intra-individual coefficients of variation (CVi) derived from the concept of biological variation[13] and measurement uncertainty for measured haematological parameters

Parameter	bias (%)	CV (%)	0,25 × CVi	0,5 × Cvi	0,75 Cvi	u (%)
WBC	1.8	2.9	2.6	5.2	7.8	3.41
RBC	0.77	0.98	0.8	1.6	2.4	1.25
HGB	0.52	1.05	0.7	1.4	2.1	1.17
HCT	1.2	1.21	0.7	1.4	2.1	1.70
MCV	0.86	0.7	0.325	0.65	0.975	1.11
MCH	0.88	1.1	0.4	0.8	1.2	1.41
MCHC	1.38	1.1	0.425	0.85	1.275	1.76
PLT	2.2	4.12	2.275	4.55	6.825	4.67
RET%	14.69	7.73				16.60

a long-term bias and an impurity factor for the reference standards used in the calibration of the instrument.

The controls used for the evaluation of the operation of the instrument are the e-Check quality control samples, supplied by the manufacturer (period January 2006 to January 2009). Although these QC samples comply with the IVD-directive 98/79/EC[15], the ranges representing intervals of suggested acceptable values for these QCs are not metrologically traceable and hence, these control materials do not comply with the standard EN ISO 17511:2000[16] for calibration purposes, since in biological materials small differences exist between the expected values from batch to batch. These samples remain nevertheless suitable for performance verification. Hence, a statistical analysis of the results from these quality control samples can be used for the assessment of long-term imprecision, but cannot be used for the evaluation of long-term bias because a 'true value' is not provided.

Since, the quality control material is made at three levels a coefficient of variation needs to be determined per level and an F-test can be used to test if there are significant differences ($\alpha = 0.05$) between each level. No significant differences were observed between the different levels and an overall coefficient of variation (expressed as a percentage) was calculated per parameter. These values are shown in Table 2. Although the concept of biological variation, as applied in the field of clinical chemistry, is not appropriate for the evaluation of MU in the field of anti-doping, the obtained CVs in the determination of long-term imprecision are suitable for comparison with the CVi from the concept of biological variation[13] in order to evaluate the validity of the method. As shown in Table 2, the obtained results are in agreement with the generally accepted criteria. Hence, the methodology used can be regarded as appropriate and valid. Thus far, no CVi has been noted in the literature for the percentage of reticulocytes; although for reticulocytes as absolute counts, CVi are available. Moreover, in contrast to other parameters like hematocrit and hemoglobin, for reticulocytes there seem to be differences between manufacturers with respect to measured values.[19] As generally accepted,[19] the reticulocyte measurements in this study show a relatively high CV, compared to the other parameters.

Because the results from the analysis of the quality control samples cannot be used to evaluate long-term bias due to the absence of traceability records, this variable needs to be derived from another source. One of the possibilities is the use of proficiency testing data. Indeed, if the number of participants is sufficiently high ($n > 30$)[20] and the technique used is the same, the consensus value of a PT-result can be regarded as an approximation of the true value and hence, the bias of an individual result can be determined. This bias can be expressed as a percentage and a long-term bias factor can be calculated. Moreover, this bias factor incorporates the impurity factor of the model in Eqn (1), because if there is an impurity in the calibration standard of a particular laboratory/instrument this will be reflected in the bias of the measurement of an individual laboratory, and hence the equation can be replaced by Eqn (2).

$$u = \sqrt{u_{precision}^2 + u_{bias}^2} \quad (2)$$

Based on Eqn (2), a standard measurement uncertainty (u) can be calculated per parameter as shown in Table 2.

It should, however, be noted that this uncertainty only comprises the analytical factors. Indeed, the sample collection and transport will also contribute to the overall uncertainty and stringent sample collection and transport protocols are necessary to minimize the impact of these variables on the overall u.

Growth hormone isoforms

The method for the detection of doping with human growth hormone (hGH) was developed to distinguish between the proportions of hGH isoforms found under normal physiological conditions and those found after recombinant human growth hormone (rGH) injection[21,22] Currently, this is the only methodology, approved by WADA, for the detection of misuse with hGH and uses two different kits especially dedicated for the anti-doping work. These kits each contain one 'recombinant' and one 'pituitary' assay. In the 'recombinant' (rGH) assay, the coated capture antibody preferentially binds to the 22-kDa hGH present in the samples, whereas the 'pituitary' growth hormone (pitGH) assay employs a capture antibody that recognizes a variety of other pituitary-derived hGH isoforms.

The respective assays are referred to as rec1, pit1, rec2 and pit2. The result of the test is expressed as the ratio of the concentration values rGH/pitGH for each particular kit. Each kit also contains two quality control samples which need to be analyzed with every batch. These QC-samples are lyophilized samples, which are reconstituted just before the analysis, ensuring their long-term stability

The uncertainty of the method for growth hormone may therefore be based on the long-term multiple measurements of these kit control samples QC1 and QC2. Based on repeated measurement of these samples ($n > 30$) over an extended period of time, long-term imprecision and bias can be combined to calculate the uncertainty, according to Eqn (2).

Because the threshold is based upon the ratio of rGH to pitGH, the uncertainty associated with the long-term imprecision, expressed as a coefficient of variation (%), of the ratio can be calculated by combining the long term imprecision of rGH and pitGH using Eqn (3).

$$CV_{ratio} = \sqrt{CV_{pit}^2 + CV_{rec}^2} \quad (3)$$

Besides the long-term imprecision, the quality control samples can also be used to determine the long-term bias. Indeed, the bias per individual result between the target value (as provided by the manufacturer) and the obtained value can be calculated

Copyright © 2010 John Wiley & Sons, Ltd.

Table 3. Results for long-term bias and imprecision and calculation of MU for the WADA approved method for the detection of growth hormone misuse

		Control 1			Control 2		
		rec hGH	pit hGH	ratio	rec hGH	pit hGH	ratio
Kit1	CV (%)	0,06	0,07	0,09	13,19	8,67	15,78
	bias (%)			12,72			−10,77
	u			12,7			19,1
Kit 2	CV (%)	9,05	8,89	12,69	6,57	5,60	8,63
	bias (%)			9,53			−8,80
	u			15,9			12,3

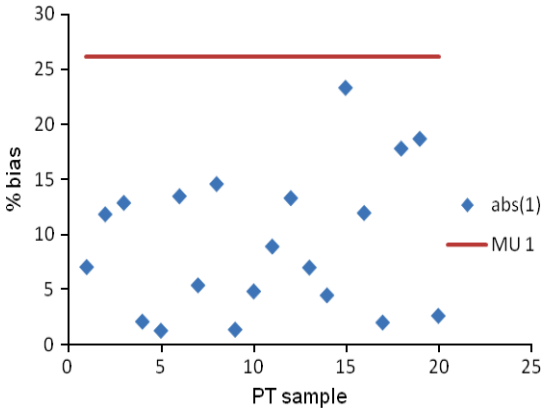

Figure 1. Measurement uncertainty evaluation using proficiency testing data for growth hormone isoform measurement (kit 1).

and expressed as a percentage. Based upon these results, the long-term bias can be calculated. Alternatively, if sufficient data exists, long-term bias might eventually also be determined via PT-results.

Because the threshold was set with the same methodology and technology and because the calibration standards accompanying the kits are calibrated from batch to batch to a traceable standard to give identical results, the impurity factor can be disregarded. Therefore the uncertainty can be calculated according to Eqn (2).

Hence, based upon this methodology, the uncertainty for the growth hormone method was determined. The results are shown in Table 3 for both control samples. As shown, two values are obtained per kit, dependent on the quality control sample used. However, for practical purposes a single measurement uncertainty value is more logical. Because the threshold values lie between quality control values, it is proposed that one may assume that the overall MU is the mean of the MU values from both control samples per kit, which can be calculated via Eqn (4).

$$u_{kit} = \frac{u_{QC1} + u_{QC2}}{2} \quad (4)$$

Although this assumption can be criticized, it should be noted that the concept of measurement uncertainty and the definition of the establishment of MU is based on estimations. Hence, assumptions can be made if the evaluation of MU shows that the estimation is valid. A procedure to evaluate MU for methods is enclosed in this paper.

Based upon these results, the standard uncertainties are calculated as 15.9% and 14.1% for kit 1 and kit 2, respectively.

Calculation of expanded uncertainty and compliance testing

In cases where threshold levels have been defined in the anti-doping regulations, doping control can be regarded as compliance testing, i.e., it is tested if the sample collected from an athlete complies with the regulations.[27] It should be taken into account that one-sided statistics are needed in compliance testing and that WADA's ISL requests a 95%-confidence level, therefore a value of $k = 1.64$ (assuming a Gaussian distribution[27]).

Consequently, values can be calculated for decision limits above which a sample is identified as non-compliant, i.e., the sample concentration or ratio exceeds the threshold by adding the expanded uncertainty to the threshold[27] as shown in Eqn (5).

$$\text{Decision Limit (DL)} = \text{threshold} + k \times u; \text{ whereas } k = 1.64 \text{ for 95\% probability} \quad (5)$$

Hence, based on the established thresholds, the decision limits can be calculated for both males and females using both kits. Any isoform ratio exceeding the decision limit can be regarded as a violation of the doping control regulations with a confidence level of 95% or higher.

In the framework of the blood passport haematological profiling, however, thresholds have not been set. Moreover, the analytical MU is only part of the overall MU and the biological variation within an individual. Besides analytical uncertainty, blood collection and transport (the pre-analytical phase) will also contribute to the overall MU of a measurement. Therefore strict protocols are applied for these steps. Hence, for the blood passport, all these steps as well as biological variation have to be considered when interpreting a haematological profile. It is clear, however, that – since either elevation or suppression of a haematological parameter can be indicative of doping misuse – for haematological profiling two-sided statistics need to be used when determining the MU. Therefore, the expanded MU per parameter at a 95% confidence level can be calculated by multiplying the unit MU (u) by a k-factor of 2[1].

Evaluation of uncertainty estimates

Because MU determinations are estimations, there is a clear need to evaluate the obtained estimations for correctness. Results of proficiency tests offer such a possibility. It can be assumed that the relative uncertainty in a small concentration range does not change significantly. Hence, if the concentrations of threshold substances in proficiency testing samples are similar to the threshold concentrations, this data can be used to evaluate the MU estimation.

Therefore the dispersion of the results of an individual laboratory should be in agreement with the true/consensus value of a proficiency test (PT), taking into account MU. Hence, the percent deviation of the individual result should be smaller than the expanded uncertainty expressed as a percentage. In Figures 1 and 2, the bias between the consensus value (mean of all participating laboratories after removal of outliers) and the value obtained by the individual laboratory are compared with the expanded measurement uncertainty (at a 95% confidence level) for 20 PT samples. Because the MU is expressed at a 95% confidence level, statistically it is expected that the bias for 95% of all samples (i.e., min. 19/20) should be lower than the MU.

The results in Figures 1 and 2 indicate that, for the 20 PT samples, the observed bias is lower than the expanded measurement

Copyright © 2010 John Wiley & Sons, Ltd.

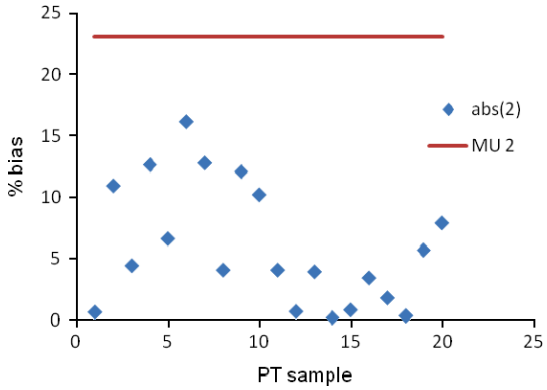

Figure 2. Measurement uncertainty evaluation using proficiency testing data for growth hormone isoform measurement (kit 2).

uncertainty. Nevertheless, there is, in general, a relatively equally spread within the uncertainty based tolerance band, especially in the case of kit 1 indicating that MU is neither overestimated nor underestimated. From Figure 2, it might be concluded that the uncertainty for kit 2 is slightly overestimated because no bias exceeds 20%, while the expanded MU is calculated at 23.1%.

It should be noted that the evaluation of the MU is based on a limited number of samples and that the evaluation process needs to be regarded as a continuous process. Not only the evaluation process, but also MU itself should be regarded as a dynamic rather than as a static concept.

As more data becomes available, this data needs to be included in the MU calculations and consequently, MU needs to be re-evaluated periodically.

Conclusions

In doping control laboratories, ISO17025 and WADA's International Standard for Laboratories require the determination of measurement uncertainty for quantitative methods. Several methodologies can be applied to estimate MU; this study provides evidence for the appropriateness of a top-down model incorporating long-term bias and long-term imprecision for this purpose. The model allows for a periodical re-evaluation of measurement uncertainty, which needs to be regarded as a dynamic analytical variable. Moreover, all MU estimations need to be evaluated for correctness and it has been shown that proficiency test results can be used for this purpose.

For threshold ratios or substances decision limits, above which an adverse analytical finding is established, can be calculated by the addition of the expanded MU to the threshold.

Acknowledgements

This work was funded by WADA. The purchase of the Sysmex XT-2000i instrument was funded by the Belgian National Lottery.

References

[1] L. S. R. Ellison, M. Rosslein, A. Williams (Eds). *Eurachem/CITAC Guide CG4. Quantifying Measurement Uncertainty in Analytical Measurement*, 2nd edn, **2000**.
[2] International Laboratory Accreditation Organisation. ILAC-G17. *Introducing the concept of uncertainty of measurement in testing in association with the application of the standard ISO/IEC 17025*.
[3] *International vocabulary of basic and general standard terms in metrology*. ISO: Geneva, Switzerland **1993**.
[4] ISO/IEC *Guide to the expression of uncertainty in measurement*, ISO: Geneva, **1993**.
[5] World Anti-Doping Agency. *The International Standard for Laboratories 2009, version 6.0*.
[6] A Van der Veen, *Accred. Qual. Assur.* **2003**, *8*, 334.
[7] P Van Eenoo, F. T Delbeke, *Accred. Qual. Assur.* **2003**, *8*, 477.
[8] Analytical Methods Committee, *Analyst* **1995**, *120*, 2303.
[9] S. L. R Ellison, V. J Barwick, *Analyst* **1998**, *123*, 1387.
[10] P van Eenoo, A Spaerkeer, F. T Delbeke, in *Recent Advances in Doping Analysis (15)*, (Eds: W Schänzer, H Geyer, A Gotzmann, U Mareck), Sportverlach Strauss: Cologne, **2007**, pp. 59–66.
[11] R Ventura, M Ventura-García, R Boqué, A Maroto, in *Recent Advances in Doping Analysis (11)* (Eds: W Schänzer, H Geyer, A Gotzmann, U Mareck), Sportverlach Strauss: Cologne, **2003**, pp. 131–140.
[12] W Horwitz, *J. AOAC Int.* **1998**, *81*, 785.
[13] G. H White, I Farrance. *Clin. Biochem. Rev.* **2004**, *25*, S1.
[14] C. Gore, R. Parisotto, M. Ashenden, J. Stray-Gundersen, K. Sharpe, W. Hopkins, K. Emslie, C. Howe, G. Trout, R. Kazlauskas, A. Hahn, *Haematologica* **2003**, *88*, 333.
[15] Directive 98/79/EC of the European Parliament and of the Council of 27 October 1998 on in vitro diagnostic medical devices. 1998-12-07 OJ No. L 331/1, http://eur-lex.europa.eu/LexUriserv/LexUriserv.do?uri=oj:L:1998:331:0001:0037:en:PDF accessed 25/01/2010.
[16] EN ISO 17511 : 2000, In vitro diagnostic medical devices. Measurement of quantities in biological samples. Metrological traceability of values assigned to calibrators and control materials, ISO, Geneva, **2000**.
[17] N Robinson, P. E Sottas, P Mangin, M Saugy, *Haematologica* **2007**, *92*, 1143.
[18] P Tarallo, J. C Humbert, P Mahassen, B Fournier, *J. Henny. Eur. J. Haematol.* **1994**, *53*, 11.
[19] N Robinson, L Schattenberg, M Zorzoli, P Mangin, M Saugy. *Int. J. Sports Med.* **2005**, *26*, 200.
[20] I. Kuselman, M. Belli, S. L. R. Ellison, A. Fajgelj, U. Sansone, W Wegscheider. *Accred. Qual. Assur.* **2007**, *12*, 563.
[21] M Bidlingmaier, Z Wu, C. J Strasburger. *Bailliere's Clin. Endocrinol. Metab.* **2000**, *14*, 99.
[22] Z Wu, M Bidlingmaier, R Dall, C. J Strasburger. *Lancet* **1999**, *353*, 895.
[23] M Bidlingmaier, J Suhr, A Ernst, Z Wu, A Keller, C. J Strasburger, A Bergmann. *Clin. Chem.* **2009**, *55*, 445.
[24] E. E Bassett, I Erotokritou-Mulligan, *Growth Horm. IGF Res.* **2009**, *19*, 361.
[25] O Barroso, I Mazzoni, O Rabin. *Asian J. Androl.* **2008**, *10*, 391.
[26] O Barroso, P Schamasch, O Rabin, *Growth Horm. IGF Res.* **2009**, *19*, 369.
[27] L. S. R Ellison, A Williams (eds). Eurachem/CITAC Guide CG4. *Use of uncertainty information in compliance assesment*, 1st edn, **2007**.
[28] C. G Fraser. *Biological variation: from principles to practice.* AACC Press: Washington DC, USA **2001**.

Copyright © 2010 John Wiley & Sons, Ltd.