

UTILISER LES RESULTATS DES ESSAIS D'APTITUDE DERNIERE METHODE ALTERNATIVE POUR L'EVALUATION DES INCERTITUDES *Application à la biologie médicale*

Marc Priel*, Soraya Amarouche*, Michèle Désenfant*, Jacques De Graeve**

*Laboratoire national de métrologie et d'essais
1, rue Gaston Boissier 75024 Paris Cedex 15

**Laboratoire de Biochimie Groupe Hospitalier Rangueil – Larrey Toulouse
1 avenue du Pr Jean-Poulhès 31403 Toulouse Cedex 4

Résumé

Les essais d'aptitude (PT) connaissent un développement important car ils permettent d'évaluer et d'assurer la qualité des prestations d'essais et de mesures. Différentes organisations proposent des circuits de comparaisons inter laboratoires dans de nombreux secteurs tels que l'environnement, l'agroalimentaire et la santé. Jusqu'à présent l'exploitation des résultats permettait aux participants de se situer par rapport à leurs pairs en calculant une statistique de performance telle que le z-score.

Depuis la publication du Guide pour l'expression de l'incertitude de mesure (GUM) en 1995 qui fonde les principes de l'évaluation des incertitudes, de nombreux travaux ont été menés pour développer des méthodes alternatives, plus faciles à mettre en œuvre, notamment lorsque la modélisation du processus de mesures ou d'essais est impossible pour des raisons techniques ou économiques.

Outre la méthode décrite au chapitre 8 du GUM que l'on peut appeler "modélisation", l'exploitation des résultats de "validation de méthode" avait été présentée au Congrès de 2003.

En 2005 nous avons présenté l'approche « performances de la méthode » et en 2007 nous souhaitons présenter une dernière méthode qui se fonde sur l'exploitation des essais d'aptitude et des contrôles qualité interne.

Au cours de cette conférence, les auteurs feront part des réflexions et des travaux menés dans le domaine de la santé sur l'évaluation des incertitudes de mesures en biologie médicale. Une application au dosage de la glycémie dans le sang sera développée et les approches « GUM » et « essai d'aptitude » seront comparées.

Summary

There has been a strong increase in proficiency testing (PT) as a powerful tool to the quality assessment of testing and measurement activities. Many providers propose Inter-Laboratory Comparisons in various domains such as environment agri-food and health. In the classical former methods exploitation of results enabled the evaluation of each participant comparing their measurement results to their peers, calculating a performance statistic such as the 'z-score'.

Since the advent of Guide to the expression of Uncertainty in Measurement (GUM) in 1995 founding the principles of uncertainty evaluation numerous projects have been carried out to develop alternate practical methods, easier to implement notably when it is impossible to model the measurement process for technical or economical aspects.

Besides the method described in chapter 8 of the GUM so-called "modelling" exploitation of results to the "validation of method" was presented during the Metrology Congress of 2003.

In 2005 the approach "performance of method" was presented and in 2007 an other alternate method is presented in this paper based on the exploitation of proficiency testing and internal quality assessments.

Further considerations drawn from activities in the domain of health for the evaluation of measurement uncertainty will be presented in this talk. The example of the determination of total concentration of glucose in human serum will be described. GUM and Proficiency Testing approaches will be compared.

Typologie des méthodes d'évaluation des incertitudes de mesure

Le GUM présente les concepts nécessaires à l'évaluation des incertitudes (définition précise du mesurande, liste des facteurs d'influence...).

Il détaille aussi une méthode d'évaluation des incertitudes appelée approche "modélisation" qui est la méthode de référence en estimation d'incertitude. Des méthodes alternatives [1] de quantification de l'incertitude se sont développées, respectant néanmoins les concepts de base exposés dans le GUM.

Une typologie de ces méthodes est présentée en faisant la distinction entre les approches intra-laboratoire, le cas où le laboratoire est seul et utilise ses propres données pour évaluer l'incertitude de ses résultats et le cas de l'approche

inter-laboratoires qui se caractérise par un travail collaboratif entre plusieurs laboratoires.

L'approche intra laboratoire se subdivise ensuite en :

- Utilisation de la loi de propagation des incertitudes ou de la propagation des distributions (Simulation de Monte Carlo)
- Utilisation des données de validation de la méthode

L'approche inter laboratoires se subdivise ensuite en :

- Utilisation des données de performance de la méthode (NF ISO 5725[3] et ISO TS 21748[4])
- Utilisation des données d'essais d'aptitude (Guide ISO 43[5] et /NF ISO 13528[6])

La figure ci-après résume ces différentes approches

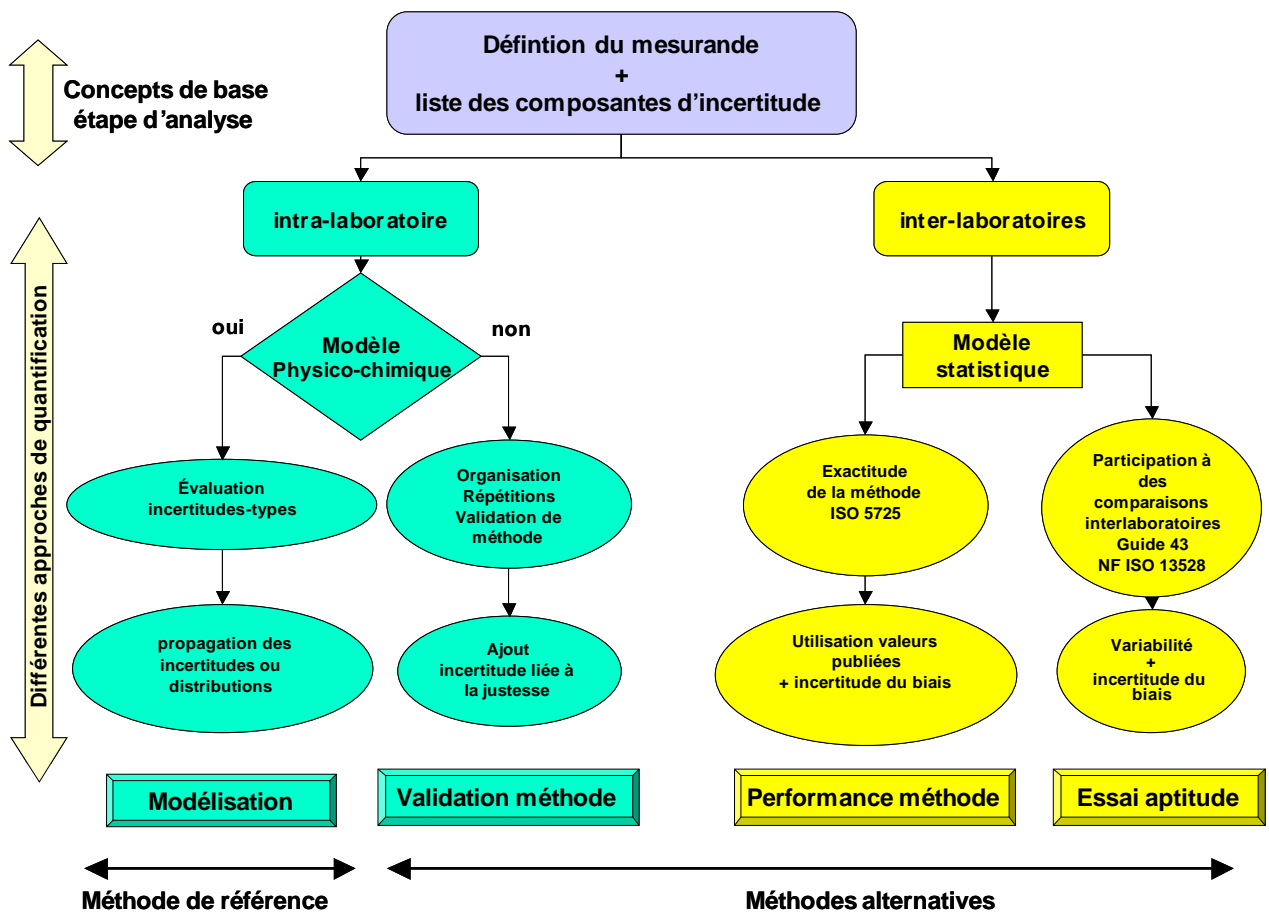


Figure 1 : Typologie des méthodes d'évaluation des incertitudes de mesure

Le GUM et les méthodes alternatives

Pour des raisons économiques ou techniques il est parfois impossible d'établir le modèle mathématique qui décrit le processus de mesure. Dans de tels cas des méthodes alternatives doivent être développées. En 2005 au Congrès de Lyon nous avons présenté l'approche « performance méthode » dans la conférence « *Tout se joue dans la définition du mesurande quelles que soient les méthodes de quantification de l'incertitude* » [2]. En 2003 au Congrès de Toulon, nous avons développé l'approche « validation de méthode » dans la conférence « *De la validation des méthodes d'analyse à l'évaluation de l'incertitude des résultats de mesure* » [3] et nous présentons dans cet article, la dernière méthode d'évaluation : l'utilisation des résultats d'essais d'aptitude.

Quelles justifications à l'utilisation de méthodes alternatives

La norme NF EN ISO/CEI 17025 Exigences générales concernant la compétence des laboratoires d'étalonnage et d'essais ouvre très explicitement des pistes pour le recours à des méthodes alternatives d'évaluation des incertitudes des résultats de mesure et d'essais :

[...] Dans certains cas, la nature de la méthode d'essai exclut un calcul rigoureux, métrologiquement et statistiquement valable, de l'incertitude de mesure. Dans de tels cas ; le laboratoire doit au moins tenter d'identifier toutes les composantes de l'incertitude et faire une estimation raisonnable, tout en assurant que la manière de rendre compte ne donne pas une impression erronée de l'incertitude. Une estimation raisonnable doit se baser sur une connaissance de la performance de la méthode et sur le domaine de mesure et faire appel à l'expérience acquise et aux données de validation antérieures [...]

L'ILAC (International Laboratory Accreditation Cooperation) qui est l'organisation qui regroupe les organismes d'accréditation au niveau international a publié un document ILAC G17/2002 Introducing the concept of uncertainty in testing in association with the application of the standard ISO/CEI 17025 qui fixe de manière assez précise les exigences en matière d'évaluation des incertitudes.

[...] The basis for the estimation of uncertainty of measurement is to use existing knowledge. Existing experimental data should be used (quality control chart, validation, round robin test, PT, handbooks etc.[...]

L'examen de ces deux textes montre très clairement que l'on peut s'appuyer sur des données existantes et les données citées sont : des données de validation de méthode, des données issues de contrôles qualité telle que les cartes de contrôle, des résultats d'inter comparaisons et des résultats d'essais d'aptitude. Ces deux textes justifient pleinement le développement de méthodes alternatives à la méthode de référence décrite dans le GUM.

Utilisation des résultats d'essais d'aptitude

Il existe de nombreux secteurs de la mesure et des essais où l'organisation d'essais d'aptitude (Proficiency Testing en anglais) constitue un des moyens privilégiés pour évaluer la compétence des laboratoires.

Actuellement, l'utilisation courante des essais d'aptitude consiste à calculer la différence entre le résultat d'un laboratoire et une valeur consensuelle, établie la plupart du temps, en prenant la moyenne arithmétique, pondérée ou robuste de l'ensemble des résultats puis à comparer cette différence à la dispersion des résultats.

Il semble évident que dans l'examen des écarts entre les valeurs obtenues par un laboratoire et la valeur assignée pour la comparaison, nous pouvons obtenir une information intéressante sur le biais ou la justesse du laboratoire. Cette information est d'autant plus pertinente que les comparaisons sont organisées régulièrement au cours d'une année.

On constate que les essais d'aptitude peuvent nous aider à appréhender l'erreur de justesse pour un laboratoire et pour un type de mesure, lorsqu'il n'existe pas de matériaux de référence. Si on a des moyens pour évaluer la « variabilité interne » sous la forme d'une évaluation de la reproductibilité interne alors en combinant Justesse et reproductibilité interne on pourra évaluer l'incertitude.

L'approche proposée en biologie médicale : le couplage de méthodes

Dans le domaine de la biologie médicale, les travaux menés par le CHU de Rangueil à Toulouse et le LNE ont conduit à proposer au Cofrac une méthode qui a servi de base à la rédaction d'un guide.

Le principe retenu se fonde sur l'exploitation de données existantes en utilisant les résultats de la validation de méthode (caractéristique méthode) et les essais interlaboratoire (essai d'aptitude).

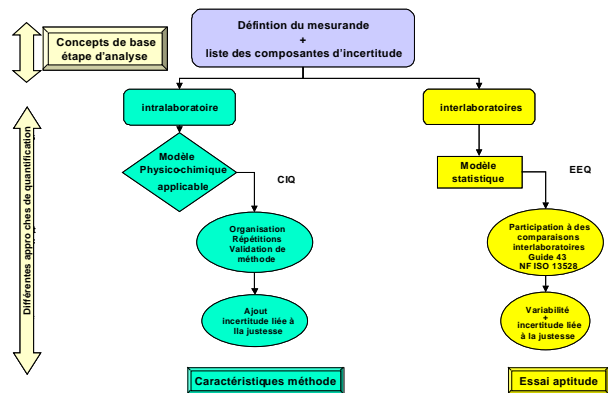


Figure 2 : couplage des méthodes d'évaluation des incertitudes de mesure

CIQ : contrôle interne de qualité

EEQ : évaluation externe de la qualité

Application à l'évaluation de l'incertitude d'une mesure de glycémie

Afin de valider cette approche d'évaluation des incertitudes par couplage des méthodes « essai d'aptitude » et « caractéristique méthode », le CHU

de Rangueil et le LNE ont traité un cas pratique celui du dosage du glucose dans le plasma humain.

La définition complète du mesurande ainsi que la liste la plus exhaustive possible des sources d'incertitude sont à la base d'une évaluation objective de l'incertitude du résultat de mesure. Cette étape commune à toutes les approches d'évaluation de l'incertitude nécessite la mise en commun de toutes les compétences (ingénieurs, techniciens, statisticien,...).

Dans l'exemple de cet article, il s'agit de la concentration en glucose C_g dans un spécimen de plasma humain en mmol/L. Le processus de mesure peut se résumer comme étant une méthode enzymatique UV (test hexokinase) après calibrage en deux points sur un Olympus AU 2700.

METHODE DE REFERENCE DU GUM (Chap8)

La définition du mesurande et l'analyse du processus sont présentées dans le diagramme des 5 M de la figure 3 :

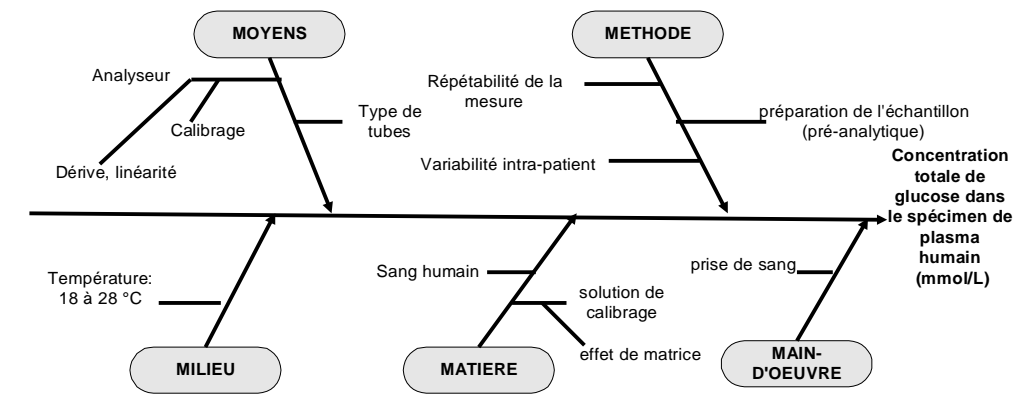


Figure 3 : diagramme des 5 M du processus de mesure de la concentration totale de glucose dans la prise de sang lors de l'analyse

Le processus de mesure après analyse des différents supports et de l'expérience dans le domaine peut être représenté par le modèle ci-dessous. Il est composé des principales sources d'incertitudes.

$$C_g = \left[\left(C_0 + \frac{A_s}{A_{cal}} (C_{cal} - C_0) \right) \times \left(\frac{V_1 + V_2}{V_2} \right) \times f_{matrice} \times f_{dérive} + k_{pre} \right]$$

Avec :

C_g : la concentration totale de glucose dans le spécimen de plasma humain [mmol/L],

et pour les grandeurs d'entrée :

C_0 : la concentration totale de glucose d'une solution (eau pure) établissant le point zéro de la courbe de calibrage [mmol/L],

A_s : l'absorbance corrigée du blanc du spécimen [AU : unité d'absorbance],

A_{cal} : l'absorbance corrigée des solutions de calibrage dans la cuve [AU: unité d'absorbance],

C_{cal} : la concentration totale de glucose lors du calibrage [mmol/L],

V_1 : volume du spécimen (μL),
 V_2 : volume des réactifs + eau (μL),
 $f_{matrice}$: facteur dû à la contribution des effets possibles de matrice (=1),
 $f_{dérive}$: facteur dû à la contribution d'une dérive permise dans la sensibilité d'instrument (=1),
 k_{pre} : correction due à la contribution de la phase pré-analytique [mmol/L] (=0).

En appliquant les différentes étapes de la méthode de référence du GUM après avoir défini le modèle mathématique du processus de mesure :

$$y = f(x_1, x_2, \dots, x_n)$$

• Evaluation des incertitudes types des grandeurs d'entrée et évaluation des covariances existantes :

$$u(X_1), u(X_2), \dots, u(X_n) \\ u(x_i; x_j)$$

• Propagation des incertitudes-types :

$$u^2(y) = \sum \left(\frac{\partial f}{\partial x_i} \right)^2 u^2(x_i) + 2 \sum \sum \frac{\partial f}{\partial x_i} \frac{\partial f}{\partial x_j} u(x_i; x_j) \bullet$$

Expression du résultat

$$y \pm U \quad (k = 2) \quad (\text{unité de mesure})$$

Le tableau 1 résume les différentes étapes avec une application à la concentration de 7,06 mmol/L.

Mesurande : la concentration totale de glucose dans un spécimen de plasma humain										
$C_g = \left(\left[C_0 + \frac{A_s}{A_{cal}} (C_{cal} - C_0) \right] \times \left(\frac{V_1 + V_2}{V_2} \right) \times f_{matrice} \times f_{dérive} + K_{pre} \right) \text{ mmol/L}$										
CODE X_i	Composantes	VALEUR X_i	UNITE	INFORMATION FOURNIE POUR L'EVALUATION DE $u(X_i)$		TYPE	LOI	$u(X_i)$	Ci	%
				valeur	source					
A_s	Absorbance corrigée blanc de la mesure	0.3	[AU]	0.006	Reproductibilité laboratoire évaluée CQI 2%	A		0.0060	23.5249	44.56%
A_{cal}	Absorbance de la sol étalonnage corrigée blanc	0.3445	[AU]	0.0046	Dispersion expérimentale sur 6 mois	A		0.0046	-20.486	19.86%
C_0	Concentration totale de glucose au pt 0 du calibrage (eau pure)	0	mmol/L	0	Par hypothèse pas d'incertitude sur l'eau pure	-	-	0.0000	0.13021	0.00%
C_{cal}	Concentration totale de glucose lors du calibrage	8.04	mmol/L	0.1437	Par hypothèse 3 fois l'incertitude du MR NIST 965	B	k=2	0.1437	0.87779	35.58%
V_1	Volume de l'échantillon (Automate)	1.6	μL	justesse : 0.03 Répétabilité : 0,014	Incertitude des pipettes incert.fabricant	B	k=2	0.0223	0.03501	0.00%
V_2	Volume 2 réactifs + eau (Automate)	200	μL	justesse : 1,6 Répétabilité : 0,3	Incertitude des pipettes incert.fabricant	B	k=2	0.9700	-0.0003	0.00%
$f_{matrice}$	Facteur dû à l'effet de matrice	1	-	0	pas d'information	-				
$f_{dérive}$	Facteur dû à la dérive (sensibilité de l'appareil)	1	-	0	info fabricant ou incluse dans la fidélité	-				
K_{pre}	Correction due à la phase pré-analytique	0	mmol/L	0	Plan d'expériences	-				
$C_g = 7.06 \text{ mmol/L}$								$u^2 = 4.47E-02$		
								$u = 2.11E-01$		
								$U = 0.423$	$(k=2)$	5.99%
								$U = 0.43$	$(k=2)$	6%

Tableau 1 : Calcul d'incertitude sur la concentration en glucose dans un spécimen de sang humain

$$C_g^1 = 7,06 \pm 0,43 \text{ mmol/L } (k=2)$$

L'incertitude ainsi évaluée sur ce résultat de mesure est de 6%, ce qui correspond a priori à un niveau attendu par les experts du domaine. Cette première quantification de l'incertitude va permettre d'avoir

une valeur de comparaison, plus largement de valider l'approche d'évaluation de l'incertitude utilisant les essais d'aptitude très largement pratiqués dans les laboratoires d'analyse en biologie médicale.

¹ Pour information cela équivaut a une concentration en g /l de $C_g = 1,271 \pm 0,078 \text{ g/l } (k=2)$

METHODE ALTERNATIVE : couplage "Caractéristiques méthode" + "Essai aptitude"

En biologie médicale, la terminologie intralaboratoire / interlaboratoires est remplacée par CIQ (Contrôle interne de la qualité) et EEQ (Evaluation externe de qualité).

La démarche d'évaluation de l'incertitude de mesure nécessite l'écriture d'un modèle quelle que soit l'approche retenue. Néanmoins dans cette nouvelle approche, le modèle reste très simplifié.

Le modèle s'écrit sous la forme suivante :

$$C = C_{lue} + K_{EEQ} \quad (1)$$

avec,

C : Concentration de l'analyte (résultat rendu),

C_{lue} : Concentration de l'analyte lue (résultat de l'analyseur),

K_{EEQ} : Terme de correction dû à la justesse

Nous considérons que dans le modèle exposé ci-dessus, il existe deux composantes d'incertitude, évaluées à partir du contrôle interne de qualité (CIQ) et du contrôle externe de qualité (EEQ). En appliquant la loi de propagation des incertitudes, on obtient :

$$u(C) = \sqrt{u^2(CIQ) + u^2(K_{EEQ})} \quad (2)$$

avec, $u^2(CIQ) = u^2(C_{lue})$, qui représente la variance de l'ensemble des résultats (C_{lue}) du contrôle interne de qualité (CIQ),

et avec, $u^2(K_{EEQ})$, la variance liée à la justesse

(biais), estimée à partir du contrôle externe de la qualité (EEQ).

Cette approche simple et réaliste fondée sur des données de contrôle interne et externe doit toutefois

pouvoir reposer sur des valeurs utilisées pour l'évaluation de la fidélité et de la justesse qui prennent en compte l'ensemble des facteurs d'influence susceptibles d'intervenir dans le processus de mesure.

Evaluation de la reproductibilité interne – Exploitation des CIQ

Lors de ces CIQ, les LABM quantifient, la variabilité du processus de mesure par un coefficient de variation de reproductibilité qui se calcule à partir de l'écart-type s de reproductibilité et de la moyenne m des résultats d'un CIQ :

$$CV\% = \frac{s}{m} \times 100 = \frac{s}{C_{lue}} \times 100$$

On a donc :

$$u^2(CIQ) = (CV_{CIQ})^2 \times C_{lue}^2$$

pour un résultat individuel.

Evaluation de la justesse et de son incertitude – Exploitation des EEQ

A chaque comparaison interlaboratoire i , le laboratoire obtient son écart $E_i = (x_{lab} - x_{ref})_i$

On en déduit lors de n comparaisons, la moyenne de l'écart et son écart-type,

$$\bar{E} = \frac{\sum (x_{lab} - x_{ref})_i}{n} \quad S_E = \sqrt{\frac{\sum (E_i - \bar{E})^2}{n-1}}$$

Deux procédures sont utilisées. Soit le laboratoire corrige ses résultats de l'erreur moyenne de justesse et intègre l'incertitude de cette correction, soit le laboratoire ne met pas en œuvre de correction et intègre totalement l'erreur de justesse dans son incertitude, ce qui a pour conséquence d'augmenter l'incertitude au final.

Données en mmol/L :		Calculs en mmol/L :	
		Correction	Pas de correction
Concentration :	6,00	$u(CIQ) = 0,12$	0,12
Contrôle interne de qualité (CIQ)		$K_{EEQ} = -\bar{E}$	$K_{EEQ} = 0$
$s =$	0,12	$u(K_{EEQ}) = S_E$	$u(K_{EEQ}) = \frac{\text{MAX } E_i }{\sqrt{3}}$
Contrôle externe de qualité (EEQ)		$u(K_{EEQ}) = 0,1648$	0,329089653
$\bar{E} =$	0,1242	$u(C) = \sqrt{u^2(CIQ) + u^2(K_{EEQ})}$	
$S_E =$	0,1648		
$\max E_i =$	0,57		
$u(c) =$		0,203860344	0,350285597
$U (k=2) =$		0,41	0,70
Résultats mmol/L (k=2):			
Absolu		5,88 ± 0,41	6 ± 0,70
Relatif		5,88 ± 7%	6 ± 12%

En incertitude élargie, les résultats sont de 7% lorsque l'on décide d'appliquer une correction et de 12% sans application de correction.

En appliquant une correction, il y a équivalence avec les 6% obtenus par la méthode de référence du GUM, tout en sachant que par l'approche essais d'aptitude l'évaluation des incertitudes est nécessairement un peu plus élevée puisqu'elle prend en compte a priori tous les facteurs d'influence.

Conclusion

L'incertitude élargie évaluée sur la mesure de la concentration en glucose dans un spécimen de plasma humain est de $\pm 6\%$ avec l'approche de référence (GUM) et de $\pm 7\%$ avec la nouvelle approche

La similitude des deux résultats apporte une première validation de l'approche essai d'aptitude couplée avec des données intra laboratoire.

Cette nouvelle approche ouvre des perspectives et permet :

- D'utiliser doublement les essais d'aptitude, dont les référentiels sont actuellement en cours de révisions.

- De diffuser plus largement le concept de qualité du résultat de mesure au travers d'un calcul d'incertitude.

- D'apporter un outil d'aide à la décision notamment dans le diagnostic au niveau médical.

La prochaine étape sera de diffuser cette approche à tous les utilisateurs d'essais d'aptitude, bien au-delà des laboratoires de biologie médicale.

Références

[1] Guide pour l'expression de l'incertitude de mesure (1995), International Organization of Standardization, NF ENV 13005

[2] Michèle Désenfant, Marc Priel, Road map for measurement uncertainty evaluation, Measurement 39 (2006) 841 – 848.

[3] NF ISO 5725 Exactitude (justesse et fidélité) des résultats et méthodes de mesure

[4] ISO/TS 21748 Lignes directrices relatives à l'utilisation d'estimations de la répétabilité, de la reproductibilité et de la justesse dans l'évaluation de l'incertitude de mesure

[5] Guide ISO 43 Essais d'aptitude des laboratoires par intercomparaison

[6] NF ISO 13528 Méthodes statistiques utilisées dans les essais d'aptitude par comparaison interlaboratoires

[7] Michèle Désenfant, Marc Priel, Tout se joue dans la définition du mesurande quelles que soient les méthodes de quantification de l'incertitude, Actes du Congrès International de Métrologie, Lyon 2005.

[8] Michèle Désenfant, Marc Priel, Cédric Rivier, De la validation des méthodes d'analyse à l'évaluation de l'incertitude des résultats de mesure, Actes du Congrès International de Métrologie, Toulon 2003.

[9] Cofrac LAB GTA 14 Novembre 2006, Guide d'évaluation des incertitudes de mesures des analyses de biologie médicale