

生物医学实验室测量 不确定度评定 案例与分析

吕 京 陈宝荣 王惠民 主编



科学出版社

生物医学实验室测量不确定度 评定案例与分析

吕 京 陈宝荣 王惠民 主编

科学出版社

北 京

内 容 简 介

本书主要分为三个部分:第一部分(第一、二章)为医学实验室测量不确定度评定的基本概念、基本理论;第二部分(第三章至第六章)介绍检验医学各专业领域评定测量不确定度的典型案例;第三部分(第七章)重点介绍目前检验医学领域测量不确定度评定中应注意的问题。为便于读者理解和掌握,本书第三章至第六章给出了测量不确定度评定案例及其分析。本书不包含微生物项目测量不确定度的评定内容。

针对医学实验室的测量特点,本书介绍了 bottom-up 和 top-down 两类测量不确定度评定方法。由于检验医学各专业领域采用的测量技术不同,测量不确定度的来源不同,第三章至第五章分别给出临床生物化学、临床血液学、临床免疫学项目采用 top-down 方法评定测量不确定度的案例;第六章介绍采用 bottom-up 方法评定基于分光光度技术和质谱技术建立的各种参考测量方法测量不确定度评定的案例。

本书的读者范围包括医学实验室、医学校准实验室、生命科学实验室、动物医学实验室等的工作人员,大专院校相关专业的本科生和研究生,相关机构的管理人员等。本书可用作高等教育的补充教材,或用作继续教育的培训教材。

图书在版编目(CIP)数据

生物医学实验室测量不确定度评定案例与分析 / 吕京, 陈宝荣, 王惠民主编. —北京: 科学出版社, 2015. 9

ISBN 978-7-03-045400-3

I. ①生… II. ①吕… ②陈… ③王… III. ①生物工程-医学工程-实验 IV. ①R318-33

中国版本图书馆 CIP 数据核字(2015)第 196045 号

责任编辑: 罗 静 刘 晶 / 责任校对: 郑金红

责任印制: 张 倩 / 封面设计: 北京铭轩堂广告设计有限公司

科学出版社出版

北京东黄城根北街 16 号

邮政编码: 100717

<http://www.sciencep.com>

三河市骏杰印刷有限公司印刷

科学出版社发行 各地新华书店经销

*

2015 年 9 月第 一 版 开本: 720 × 1000 1/16

2015 年 9 月第一次印刷 印张: 8 3/4

字数: 177 000

定价: 68.00 元

(如有印装质量问题, 我社负责调换)

《生物医学实验室测量不确定度评定案例与分析》

编委会名单

主 编	吕 京	陈宝荣	王惠民		
副主编	王建兵	孙慧颖	季伙燕	孙可其	黄宪章
编 委	陈宝荣	孙慧颖	邵 燕	吕 京	贾淑芬
	王惠民	季伙燕	黎万菊	蔡 华	刘冠财
	王俊卿	杨 涛	王建兵	庄俊华	史光华
	韩丽乔	林海标	张乔轩	姚尚武	左悦媛
	周玉贵	沈国荣			

自序

随着测量不确定度的应用被广泛接受，其作为检测结果的一部分和检测能力的表示，也逐渐引入生物和医学检验领域。

GUM 和 QUAM 是经典的测量不确定度评定指南文件，鉴于医学检验的特殊性，在具体应用时相关人员难免存在各种疑问。为此，各国医学检验学者结合测量不确定度评定的基本理论，针对医学检验的特点，已编写并颁布了多个测量不确定度评定标准或指南，有代表性的如澳大利亚 AACB 发布的 *Uncertainty of Measurement in Quantitative Medical Testing*(2002)、我国 CNAS 发布的技术报告《医学实验室测量不确定度评定与表达》(2012)、美国 CLSI 发布的 C51-P : *Expression of Measurement Uncertainty in Laboratory Medicine*(2012)、英国 LGC 颁布的 *Evaluating Measurement Uncertainty in Clinical Chemistry*(2012)，以及目前仍在制定中的 ISO 标准 *Medical Laboratories-Practical Guide for the Estimation of Measurement Uncertainty* 和我国的国标 GB/TXXXX《合格评定 生物样本测量不确定度的评定和表示》，以及卫生行业标准《酶学参考实验室参考方法测量不确定度评定指南》等。

编写本书的目的是帮助读者认识并在实际工作中应用测量不确定度，因此，本书的内容以介绍和分析案例为主。本书共分 7 章。第一、二章介绍测量不确定度的基本概念、不确定度基本理论；第三至五章分别介绍在临床生物化学、临床血液学、临床免疫学领域应用 top-down 方法评定测量不确定度的案例；第六章介绍以分光光度技术和质谱技术进行参考测量时应用 bottom-up 方法评定测量不确定度的案例。为便于读者理解，每一案例后均附有案例分析。另外，本书第七章还重点介绍了目前检验医学领域测量不确定度评定中应注意的问题。

编者

2015 年 3 月

前 言

国际检验医学溯源联合委员会 (JCTLM) 的研究数据表明, 目前医生 70% 的诊断信息来自于实验室, 因此, 如何保证用于诊断的实验室数据的质量是一个至关重要的问题。20 世纪, 继 ISO/IEC 17025《检测和校准实验室能力的通用要求》发布之后, 国际标准化组织 (ISO) 又发布了 ISO 15195《检验医学 参考测量实验室的要求》、ISO 15189《医学实验室 质量和能力的专用要求》等一系列规范和指导医学校准和检测实验室活动的国际标准, 我国已等同转化了上述国际标准为国家标准。

在上述标准中均涉及测量不确定度的评定问题。测量不确定度如此重要, 是因为完整的测量结果应包括表征测量结果分散性即不确定度的信息, 医学检验结果也不例外。对测量结果及不确定度的了解, 可帮助测量者了解、认识、控制影响测量的因素, 帮助使用者在诊断和治疗疾病时, 更恰当地解释测量结果的意义和控制风险。

关于医学领域测量不确定度的专著在我国尚属空白, 仅在专业杂志中有一些研究个例, 远远不能满足我国广大医学工作者的应用需求。在国家质检总局公益性行业科研专项经费项目“我国医学参考测量体系的建立与示范”(201210066)、国家标准制定任务等项目的支持下, 课题组系统研究了医学实验室测量不确定度的评定方法。为帮助相关人员了解医学测量结果的不确定度, 掌握如何科学地进行评定和应用测量不确定度, 本书作者基于项目研究成果、参考国际文件并结合多年的医学检测和校准实验室的实际工作经验及管理经验编写了本书。本书较为全面地介绍并分析了医学实验室不同专业领域评定测量不确定度的案例, 特点是重在结合实际逐步进行指引, 以期对医学检验工作者和其他生物医学领域的工作者提供一本实用的案头工具书。

在本书编写和出版过程中, 我们始终得到了“我国医学参考测量体系的建立与示范”(201210066) 项目组各位领导与专家的大力支持和关心。在本书编写过程中, 我们还得到了很多实验室人员的大力支持和悉心指点, 他们多数是一线工作人员, 具有丰富的实践经验, 为本书提供了大量相关素材, 并提出各种修改意见, 恕难一一列出致谢。

由于编者水平有限, 本书存在不足和错误之处在所难免, 还望广大读者批评指正并及时向作者反馈。

编 者
2015 年 3 月

目 录

自序

前言

第一章 概述	1
第一节 测量不确定度概念的提出	1
第二节 认识测量不确定度	2
第三节 检验医学领域测量不确定度的研究现状	5
第四节 检验医学领域测量不确定度的应用	7
第五节 关于生物样本测量不确定度评定的思考	9
第二章 医学实验室测量不确定度评定的基本理论	14
第一节 测量不确定度相关的基本概念与术语	14
第二节 医学实验室测量不确定度的来源	29
第三节 医学实验室测量不确定度的评定方法	32
第四节 关于检验医学领域应用 top-down 方法的讨论	69
第五节 医学实验室测量不确定度的报告	72
第三章 临床生物化学项目测量不确定度评定案例	74
第一节 基于参考系统研究数据评定 LD 测量不确定度的案例	74
第二节 基于 PT 数据评定 LD 测量不确定度的案例	79
第三节 案例分析	82
第四章 临床血液学项目测量不确定度评定案例	86
第一节 基于参考系统研究数据评定 RBC 测量不确定度的案例	86
第二节 基于 PT 数据评定 RBC 测量不确定度的案例	90
第三节 案例分析	94
第五章 临床免疫学项目测量不确定度评定案例	95
第一节 基于参考系统研究数据评定 PA 测量不确定度评定案例	95

第二节 基于 PT 数据评定 TSH 测量不确定度的案例	97
第三节 案例分析	101
第六章 医学参考测量项目测量不确定度评定案例	103
第一节 GGT 测量不确定度评定案例	103
第二节 Glucose 测量不确定度评定案例	111
第三节 甲状腺素 (T4) 测量不确定度评定案例	116
第七章 检验医学领域测量不确定度评定中应注意的常见问题	120
第一节 正确度评价中的问题	120
第二节 精密度评价中的问题	121
第三节 数据统计学问题	122
主要参考文献	126

第一章 概 述

第一节 测量不确定度概念的提出

“根据所用到的信息，表征赋予被测量量值分散性的非负参数”是我国计量检定规范 JJF1001—2011《通用计量术语及定义》中给出的“测量不确定度”的定义。该定义来源于 *International vocabulary of metrology—Basic and general concepts and associated terms* (VIM)。测量不确定度是反映测量结果可靠性的客观指标，也是量值溯源相关的重要计量学概念。

“测量不确定度”由原美国标准局 (NBS) 的数理统计专家埃森哈特 (Eisenhart) 在 1963 年研究“仪器校准系统的精密度和准确度估计”时首次提出。为使测量不确定度的概念、评定及表示方法达到国际一致，1980 年国际计量局在征求了 32 个国家及 5 个国际组织的意见后，推荐采用“测量不确定度”来表示测量结果的质量并形成建议书，1981 年第 70 届国际计量委员会讨论并通过该建议。1986 年，国际计量局 (BIPM)、国际标准化组织 (ISO)、国际法制计量组织 (OIML)、国际理论和应用物理联合会 (IUPAP)、国际理论和应用化学联合会 (IUPAC) 及国际临床化学委员会 (IFCC) 等 7 个国际组织成立专门的“测量不确定度”工作组，起草关于测量不确定度评定的指导性文件。1993 年，《测量不确定度表达指南》(*Guide to the Expression of Uncertainty in Measurement*, GUM) 以 7 个国际组织的名义联合发布。1995 年，欧洲分析化学中心 (A Focus for Analytical Chemistry in Europe, EURACHEM) 出版 *Quantifying Uncertainty in Analytical Measurement* (EURACHEM Guide)。1997 年，该组织与分析化学国际溯源性合作组织 (CITAC) 共同协商，并邀请了国际原子能机构 (IAEA)、欧洲认证机构 (EA) 和美国官方分析化学家协会 (AOAC) 的专家，经过 3 年的讨论、修改，在 2000 年出版《化学分析中不确定度的评估指南》(第二版) (*Guidance on Evaluating the Uncertainty in Chemical Analysis*)。该文件被我国 CNAS 等同采用，并于 2002 年 8 月以《化学分析中不确定度的评估指南》发布并实施。目前，GUM 在 2008 年发布了第二版，对 1995 年版进行了部分的修订；而 QUAM 在 2012 年发布了第三版。

20 世纪 80 年代初，我国计量学家开始参与国际有关测量不确定度评定与表达的讨论。1996 年，原国家技术监督局根据有关国际组织文件的要求，结合我国的实际情况，制定了“检定 / 校准证书”模板来统一全国范围测量不确定度的表达格式。1998 年，我国发布 JJF1001—1998《通用计量术语与定义》，前 6 章的内容等同采用第二版 VIM。1999 年，我国发布 JJF1059—1999《测量不确定度评定与表示》，等同采用 GUM。这两个文件奠定了“测量不确定度”在我国测量科学领域评定的重要基础地位。2002 年，中国实验室国家认可委员会 (现并入中

国合格评定国家认可中心, 简称 CNAS) 编制《化学分析中不确定度的评估指南》, 等同采用 EURACHEM/CITAC 发布的第二版 QUAM。

由于检验医学测量的高度复杂性、测量样本的特殊性(难以重复获取、取样量有限等)和检验报告的高时效性等现实问题, 以及应用测量不确定度的可操作性等问题, 检验医学领域一直未使用“测量不确定度”这一术语。21 世纪初, ISO 发布了一系列有关测量结果溯源性的导则文件, 对测量质量与量值溯源提出更明确的要求, 也促进了检验医学领域研究并尝试应用测量不确定度。在我国, 特别是 2005 年后, 在卫生部临床检验中心带领下, 北京协和医院、北京航天总医院等 6 家医疗机构和检验产品生产厂家实验室先后建立了医学参考测量实验室(即医学校准实验室), 他们在参加 IFCC 组织的 Ring-Trails 国际参考测量实验室比对时, 须报告参加比对项目的测量不确定度, 此时检验界在卫生部临床检验中心杨振华、陈文祥及计量界专家的带领下开始认识并学习评定测量不确定度。

2008 年, 针对医学参考测量实验室评定测量不确定度遇到的问题, 我国启动卫生行业标准《酶学参考测量实验室测量不确定度评估指南》的编写工作, 2011 年形成报批稿, 预计 2015 年该标准将正式发布。

2010 年, 为解决我国医学实验室寻求 ISO15189 认可遇到的测量不确定度的评定问题, CNAS 启动《医学实验室 测量不确定度评定与表达》技术报告的研制工作, 2012 年正式发布(CNAS-TR01, 2012)。2014 年, 国家标准《合格评定 生物样本测量不确定度的评定与表达》的制定工作正式启动, 预计 2016 年可发布。

第二节 认识测量不确定度

测量的目的是确定被测量的值。测量结果通常只是被测量值的近似值或估计值, 因此, 完整的测量结果需要同时附有结果的不确定度声明。影响测量结果的潜在因素很多, 例如:

- (1) 被测量的定义不完整;
- (2) 被测量定义的复现不完善;
- (3) 所测量样本的代表性不足;
- (4) 对测量受环境条件的影响认识不足或对环境条件的测量或控制不完善;
- (5) 人员读数误差;
- (6) 仪器分辨力或识别阈值的限制;
- (7) 测量标准和标准物质的量值不准确;
- (8) 从外部得到并在数据约简算法中使用的常数和其他参数的值不准确;
- (9) 测量方法和程序中的近似和假设;
- (10) 在看似相同条件下, 重复测量过程的变异性;
- (11) 尚未认知或未识别的因素。

上述因素不一定是相互独立的,如(1)~(9)中有些因素可能对(10)有贡献。未识别的影响因素不可能在测量结果的不确定度评定中予以考虑,但其实际对测量误差有贡献。

一、测量不确定度的评定方法

导致测量结果产生误差的贡献量的准确值虽然未知或不可知,但与引起误差的随机影响和系统影响有关的不确定度是可以评定的。测量不确定度可以用多种方法评定,包括:“自下而上”(bottom-up)的方法、“自上而下”(top-down)的方法和蒙特卡罗方法(MCM)等。从本质上看,bottom-up方法、top-down方法、MCM评定测量不确定度的广义原则是一致的。如果我们从不确定度的具体评定方式看,只有A类评定和B类评定两个途径。根据GUM,测量不确定度的A类评定,当测量系统/过程稳定时,可以利用代表测量系统的实验标准偏差 $s(x_k)$ 和较少测量次数甚至单次测量 n' ($1 \leq n' < n$) 评定被测量估计值由于重复性导致的A类标准不确定度,自由度仍为 $\nu=n-1$,数学公式为:

$$u_A(\bar{x}) = s(\bar{x}) = s(x_k) / \sqrt{n'}$$

对一个测量系统/过程,采用核查标准和控制图的方法使测量系统/过程处于统计控制状态,若第 j 次核查时测量次数 n_j (自由度为 ν_j),每次核查时的实验标准偏差为 s_j ,共核查 m 次,则统计控制下的测量过程的标准不确定度可以用合并样本标准偏差 s_p 表征:

$$s(x_k) = s_p = \sqrt{\frac{\sum_{j=1}^m \nu_j s_j^2}{\sum_{j=1}^m \nu_j}}$$

若每次核查的自由度相等,则:

$$s_p = \sqrt{\frac{\sum_{i=1}^m s_j^2}{m}}$$

以后任意次测量包括单次测量的标准不确定度为:

$$u_A(x) = s(\bar{x}) = s_p / \sqrt{n}$$

可见,top-down方法也是符合GUM原则的一种方法,因此,与之相对,经典意义上的GUM方法又被称为bottom-up方法。在采用标准化的测量方法进行常规检测时,用top-down方法评定测量不确定度更为实用和可行,特别是在检验医学领域。bottom-up方法是对本次测量结果影响量的评估,而top-down方法是利用既往的测量表现代表本次的测量表现,因此,利用bottom-up方法的风险可能更小。但也要意识到,即使应用bottom-up方法,也可能有来源于经验数据的输入。要正确认识和利用数据,数据的代表性和可靠性是关键。

二、对测量不确定度评定的要求

通常,对测量不确定度评定所需的严密程度取决于检测方法的要求、用户的要求、用来确定是否符合某规范所依据的误差限的宽窄。例如,体检时测量身高,基本不需要评定测量不确定度,若用户要求评定,一般也无需特别考虑温度、时间等对测量结果的影响。

对运行参考测量程序的实验室(即校准实验室)而言,应对其开展的全部参考测量项目(参数)评定测量不确定度,并在校准证书中报告参考测量结果的不确定度,这是提供计量学服务所需要的。对运行常规测量程序的检测实验室而言,其应有能力对每一项有数值要求的测量结果进行测量不确定度评定,当不确定度与检测结果的有效性或应用有关、用户有要求、不确定度影响到对规范限度的符合性或测量方法中有规定时,检测报告需提供测量结果的不确定度。

三、测量不确定度评定策略

参考测量活动的目的是提供计量学服务,应优先采用 bottom-up 方法评定参考测量结果的不确定度。对运行良好的测量程序,可用 top-down 方法或 MCM 验证 bottom-up 方法评定的测量不确定度。在非线性模型或输出量的概率密度函数(PDF)明显背离了正态分布或 t 分布的条件下,用 MCM 可能更适宜。

运行参考测量程序,需严格控制测量条件。对所有相关的材料、设备、过程、环境、人员能力等的要求应满足测量程序的规定。应在对影响测量结果的因素控制在可接受的条件下,实施测量并评定测量不确定度。测量不确定度评定范围应包含提供参考测量服务期间短期的不确定度分量和可以合理的归为来源于被测样本的不确定度分量。对无法获得的不确定度分量,例如运输产生的不确定度,赋值后样本保存产生的不确定度等,通常不包括在不确定度评定范围内。但是,由于生物源性样本的不稳定性特征,实验室若预计到这些不确定度分量将对用户产生有意义的影响,应根据 GB/T27025(ISO/IEC17025, IDT)中有关合同评审的要求通知客户。

参考测量实验室需对应参考测量范围评估在常规条件下能够提供给用户的校准和测量能力(calibration and measurement capability, CMC),其应是在常规条件下校准/参考测量可获得的对应该浓度的最小测量不确定度。一般情况下,CMC 应基于对稳定、均匀样本的测量,用包含概率约为 95%的扩展不确定度表示,单位与被测量一致或用百分比表示。为用户报告的不确定度一般不应优于实验室的 CMC。可以理解为,CMC 是实验室在统计学意义上的“最佳”测量表现,而为用户服务是一次测量表现,因此,从风险管理的角度出发,为用户报告的不确定度不优于实验室的 CMC 时,风险较低。包含因子 k 的值是根据 $y-U$ 到 $y +$

U 区间所要求的置信水平而选择的。当 y 和 $uc(y)$ 所表征的概率分布近似为正态分布, 且 $uc(y)$ 的有效自由度较大时, 可设 $k=2$, 所形成区间具有的置信水平约为 95%, 设 $k=3$, 所形成的区间具有的置信水平约为 99%。

常规测量程序与参考测量程序的应用目的不同。常规测量活动的目的是满足用户对测量数据的需求, 为其提供客观的检测结果。通常, 常规测量程序的计量学要求低于参考测量程序, 参考测量结果的不确定度可通过溯源链传递到常规测量结果。部分常规检测项目的自动化程度很高, 重复测量变异可能较小, 在这种条件下, 如果不考虑校准引入的不确定度分量, 只评定重复测量导致的不确定度, 则会出现量值传递过程中不确定度传播率的表现不完整。

对常规测量程序, 在初始引入实验室阶段, 宜考虑采用 bottom-up 方法评定测量结果的不确定度, 主要目的是可以了解影响测量的因素并改进。对运行良好的测量程序, 宜用 top-down 方法评定测量不确定度, 应考虑测量范围与不确定度的关系。需要时, 应对应测量范围, 评定在常规条件下每套检测系统获得的对应浓度的测量不确定度, 并合理表示。同样, 需严格控制测量条件, 对所有相关的材料、设备、过程、环境、人员能力等的要求应满足测量程序的规定。可利用的评定偏移的数据包括与参考实验室比对的数据、测量参考物质(包括自建的参考物质)获得的数据、参加正确性验证计划的数据。若利用参加一般的能力验证计划的数据, 宜先评估其适宜性。系统影响的已知修正量 b 应尽量用于修正报告的测量结果, 而不是通过放大结果的不确定度将这种影响考虑在内。

由于某些检测方法的性质, 决定了无法从计量学和统计学角度对测量不确定度进行有效而严格的评定, 这时至少应通过分析方法, 识别各主要的不确定度分量, 并做出合理的评定, 重要的是, 应保证测量结果的报告形式不会使客户造成对所给测量不确定度的误解。

应了解和识别测量前、测量、测量后过程中影响测量结果的所有有意义的因素, 需要时, 采取控制措施。对于临床常规检验, 全过程管理和质控特别重要, 是保证结果一致性的根本措施。

我们知道, 测量结果取决于被定义的测量条件, 而测量条件(人、机、料、法、环等)只是相对稳定。因此, 需要对实验室的 CMC 和不确定度定期评审、评定。不同实验室对测量条件控制的程度不同或用于评定测量不确定度的条件/数据不同, 只有在合理的条件下, 不确定度的大小可表征测量质量, 否则可能带来误导。合理评定、合理表达测量结果的不确定性同样重要。

须牢记, 认识和评定测量不确定度应基于质量改进和更好应用测量结果的主观需求, 了解风险、控制风险, 并追求卓越。

第三节 检验医学领域测量不确定度的研究现状

测量不确定度在计量行业的应用始于 20 世纪 90 年代, 但迄今为止检验医

学领域和生物学领域仍未广泛应用。

误差与测量不确定度虽然均为与测量结果相关联的参数,但二者的本质和特征有明显差异。误差是测量结果与真值的差值,由国际米制公约组织在 1875 年提出。由于误差实际不可得,因此在应用中若以误差作为与结果关联的信息,则在理论上不能自洽。在临床检验领域,依然广泛应用“总误差”。可以看出,“总误差”概念的提出也是试图解决误差实际不可得的问题,同时,解决临床一次测量即报告结果的质量控制问题,它似乎更像一个目标不确定度。

2003 年,ISO 发布 ISO15189 以后,国际认可合作组织(ILAC)将其作为国际间医学实验室互认基础的认可准则,这大大推进了各国的医学实验室认可和测量不确定度在医学检验领域的应用,包括美国病理学家协会,也推出了 ISO15189-CAP[®] 认可计划。可见,检验医学领域引入测量不确定度是发展趋势。测量不确定度对于检验行业特别是大多数常规医学实验室仍是一个比较新的概念,在评定中面临许多实际困难,尚需实用的指南文件指导实验室评定和应用不确定度,包括指导临床医护人员认识和在临床决策中合理应用不确定度。

为找出相对比较简便实用的方法,EURACHEM/CITAC 分别在 2002 年、2006 年和 2007 年发布了三个技术报告——《检测中的测量不确定度》、《对定量检测结果评定测量不确定度导则》和《对测量不确定度的再回顾——评估不确定度的替代办法》。这三个技术报告最大的特点是提出了“black-box”方法评估测量不确定度(也被称为“经验法”)。为区分测量不确定度的两种评定理念,人们形象地将 GUM/QUAM 的经典方法称为 bottom-up 方法,后者称为 top-down 方法。之后,GUM 又补充了蒙特卡罗法(Monte Carlo method, MCM),它是一种通用的数值方法,主要通过测量模型以概率分布传播作为基础,可实现测量不确定度评定中的部分计算,其与 GUM 方法的原则在广义上是一致的,它适用于有任意个输入量但只有单个输出量的模型。我国南通医学院的王惠民等尝试了 MCM 在检验医学领域的应用。

对于常规检验,目前国内外的研究表明,更适合应用 top-down 方法。澳大利亚生化学家协会(AACC)于 2002 年发布了《定量医学检测中的测量不确定度——实验室应用指南》。该指南曾在 2009 年 11 期《临床实验室》发布全文翻译稿。如果将一个完整的测量过程分为分析前、分析中和分析后三个阶段,此文件仅限于分析中测量不确定度的评定。澳大利亚国家病理认可顾问委员会(NAPPC)2007 年发布《评估测量不确定度的要求》,提出目标不确定度的概念,并提出应尽量减少分析前和分析后阶段对测量不确定度的影响。澳大利亚国家认可机构(NATA)于 2009 年发布《化学测量结果不确定度评定与报告导则》,明确提出单一实验室可以利用实验室的偏移与精密度数据来评定测量不确定度。美国临床和实验室标准协会在 2012 年发布了 CLSI-C51: *Expression of Measurement Uncertainty in Laboratory Medicine*, 也主要介绍了 top-down 方法。

CNAS 于 2012 年年底发布了《医学实验室 测量不确定度的评定与表达》技术报告 (CNAS-TRL-001, 2012), 重点阐述常规医学实验室如何利用 top-down 方法评定测量过程的不确定度。该文件主要参考了 2004 年挪威北欧测试合作组织 (NordTest) 发布的《TR537: 环境实验室计算测量不确定度手册》、2008 年比利时食品安全联邦局 (FASFC) 发布的《测量不确定度的导则》、澳大利亚临床生物化学协会 (AACB) 发布的《定量医学检测中的测量不确定度——实验室应用指南》及 ISO25680.8。

迄今为止, 尽管不少国家都在发布一些测量不确定度评定的标准 / 指南文件, 但国际各方仍未就医学实验室如何评定和应用测量不确定度达成共识。ISO 自 2004 年开始编写 ISO25680《医学实验室——测量不确定度的评定和表达》, 至 2012 年已九易其稿。目前 ISO 仍在组织各相关领域专家研讨、制定《医学实验室——测量不确定度的评定和表达》标准。

检验结果的测量不确定度评定从理论上讲应包含完整的测量过程, 但实际工作中很难实现。完整的检验过程可分为分析前、分析中、分析后阶段, 每个阶段又包含多个测量影响因素。分析前阶段包括患者准备、样本采集、运输、离心、样本测量前等候时间等; 分析中阶段包含校准、方法特异性、精密度、线性、抗干扰能力等。分析后阶段主要涉及数据处理。此外, 个体内与个体间生物学变异也是影响检验结果的一个重要因素, 但是与测量过程无关。上述多个因素对最终结果均有影响, 但完全用不确定度表征又不现实。检验医学领域迫切需要找到一种理论与实际均可行的测量不确定度的评定方法。医学实验室要从应用的目的评定不确定度, 比较可行的办法是通过控制测量过程使测量结果稳定在一个范围内, 而这个范围即是与测量不确定度密切相关的参数, 可用于评定不确定度, 也即可以通过测量程序的偏移性能参数和精密度参数评定测量不确定度。为了实际工作的可操作性, 通常只评定测量阶段的不确定度, 其他过程只严格按照标准化程序操作, 不再评定其对不确定度的贡献。当然, 这是一种妥协的办法, 如果有充分的数据可以利用或用户有要求时, 仍建议评定。

与常规医学实验室相比, 医学参考测量实验室的目的是提供计量学校准服务, 因此, 每次服务均应按照定义的被测量, 用 bottom-up 的方法评定测量条件等各输入量对测量结果的影响程度, 须将不确定度作为结果的一部分提供给客户。

第四节 检验医学领域测量不确定度的应用

一、医学实验室

(一) 改进测量质量

目前, 我国医学实验室普遍建立了基于统计学分析的质量控制体系。显而易见, 不确定度越小表明测量的质量和能力越高。为了保证测量结果可以有效利

用,对测量程序的性能应有基本要求,其中,目标不确定度是客观控制指标。如果目标不确定度达不到预期要求,实验室应采取必要的技术措施改进和完善测量程序。而通过评定测量不确定度,特别是应用 bottom-up 方法评定,是识别影响测量结果因素及其程度的有效途径。识别、评估和控制这些因素可明显提高测量结果的质量,获得合理的测量不确定度。

(二) 选择测量程序与检验产品

测量不确定度是测量结果可靠性的指标。测量程序/系统、测量服务、计量器具、参考物质等声明的与不确定度相关的信息是其质量的反映,也是实验室评定测量不确定度的信息来源。因此,计量产品和服务声明的相关不确定度的信息非常重要,直接影响实验室检测结果不确定度的大小和评定。了解其不确定度是否满足实验室质量控制的预期要求,是医学实验室选择检验产品和服务的重要依据。

(三) 加强与临床联系

及时向临床提供测量不确定度的信息,可帮助临床人员科学地应用检验结果,有助于实验室人员与临床人员的相互配合与密切合作。

二、临床医师

(一) 临床诊断

医师诊断疾病时一般先将报告结果与生物参考值或临床决定限进行比较,而仅用报告结果不考虑不确定度是有风险的。

例:成年男性全血血红蛋白含铁量 [Fe] 的参考区间为 7.5~9.5mmol/L。3 位患者 A、B、C 的测量量值分别是:7.0mmol/L、8.2mmol/L、9.5mmol/L。已知此 3 个结果的标准不确定度都是 0.2mmol/L,取置信概率为 95%,扩展不确定度是 0.4mmol/L,上述测量量值可表达为:A: $(7.0 \pm 0.4)\text{mmol/L}$,即 6.6~7.4mmol/L; B: $(8.2 \pm 0.4)\text{mmol/L}$,即 7.8~8.6mmol/L; C: $(9.5 \pm 0.4)\text{mmol/L}$,即 9.1~9.9mmol/L。

考虑测量不确定度,可明确判断 A 患者血红蛋白含铁量低, B 患者血红蛋白含铁量正常。虽然患者 C 的血红蛋白含铁量测量结果在参考区间内,但无法确定是否正常,因为 C 的被测量量值加上扩展不确定度,已有 50% 的概率高于参考区间上限,因此不能明确患者 C 的血红蛋白含铁量是否正常,需进一步观察。如果 A 患者结果的扩展不确定度为 0.6mmol/L,则不能确定异常。结果的测量不确定度小,应用效率就高。

但也应注意,不确定度是根据可获得数据评定的,由于被测物定义的不完善、测量方法的限制等,以及样本因素、干扰因素等的影响,并不意味着给出不确定度的结果一定是一个可以直接用于决策的结果。同一个参数,由于检测科学发展的限制,不同厂商提供的方法可能测的是不同的“标志物”,此时,不同方法

之间的结果难以比较。临床医生应客观认识实验室结果和不确定度,需结合临床及其他证据综合判断。

(二) 疗效观察

测量不确定度信息可用于判定治疗前、后测量量值是否有差异,以及测量量值与参考值的偏离有无显著性等。

当被测量量值与一个没有测量不确定度的固定临床决定限进行诊断性比较时,假如差值 $|\Delta| \geq U(k=2)$,说明差异具有显著意义(置信概率为 95%)。

当一个新被测量量值与一个以前获得的被测量量值比较时,如果前、后两个量值具有相同的合成标准不确定度,如差值 $|\Delta| \geq \sqrt{2} U(k=2)$ 时,说明差异具有显著意义(置信概率是 95%)。

例:一个患者血清中葡萄糖浓度为 7.5mmol/L,治疗后该患者再次测量时葡萄糖浓度降低为 7.1mmol/L。如果标准测量不确定度是 0.12mmol/L,前后差异 0.4mmol/L 有意义吗?

首先,二次结果差异的不确定度 $u_{\Delta}(\text{mmol/L})$ 为 $u_{\Delta} = \sqrt{(u_{7.5})^2 + (u_{7.1})^2} = \sqrt{0.12^2 + 0.12^2} = 0.17 \text{ mmol/L}$,扩展不确定度为 $U_{\Delta} = 0.34 \text{ mmol/L}(k=2)$ 。假设两次结果差值 Δ 有显著意义(按 95% 置信概率),则两个测量值之间的差异要大于差异的扩展不确定度, $\Delta = (7.5 - 7.1) = 0.4 \text{ mmol/L}$,大于 0.34,因此,该患者血清中葡萄糖从 7.5mmol/L 降低到 7.1mmol/L 按 95% 的置信概率计算有显著意义,证明治疗有效。

三、患者

患者多为非专业人士,应该知道,检验结果受多种因素的影响,其判断需要非常专业的知识,最好的做法是应听取医生或检验医师的解释和建议。

第五节 关于生物样本测量不确定度评定的思考

2014 年,国家标准《合格评定 生物样本测量不确定度的评定与表示》的制定工作正式启动,该标准定位于合格评定领域有关生物样本的测量不确定度的评定与表示。涉及生物样本测量的领域包括临床检验、公共卫生检验、动植物检验检疫、安全性评价、兴奋剂检测、司法鉴定、交通执法、生命科学研究等众多学科,这些学科直接关乎生命健康或社会公正,测量结果被用于临床决定、卫生执法、合格判定、药品登记注册、法庭证据等,测量不确定度信息十分重要。

测量不确定度表征的是测量结果的分散性特征,此处测量结果应理解为被测量之值的最佳估计,即全部不确定度分量均贡献给了分散性,包括那些由系统效应引起的分量,而分散性信息一般只能通过重复测量获得。

ISO 15189 自 2003 年发布以来,在临床检验领域引入测量不确定度已经成为