

临床检验总误差与测量不确定度*

张诗诗^{1,2}, 王 薇², 赵海建², 王治国^{1,2} (1. 北京协和医学院研究生院, 北京 100730;
2. 北京医院卫生部临床检验中心, 北京 100730)

摘要:在检验医学领域,探讨测量不确定度(measurement uncertainty, MU)和总误差(total error, TE)的区别是个值得关注的问题。多年来,国内外的检验医学专家们对 MU 进行研究并促进其在检验医学领域的发展,同时也使临床实验室开始关注 MU 的意义和作用。然而,由于 TE 在临床实验室中的习惯性使用以及 TE 与 MU 二者的概念有类似之处且容易导致混淆,因此许多实验室依然无法完全接受 MU。该文通过解释 TE 与 MU 的概念,并且分析 TE 模型与 MU 模型各自的优缺点和作用,帮助临床实验室进一步了解 TE 和 MU,并在实际工作中合理运用。

关键词:总误差;测量不确定度;允许总误差

中图分类号:R446 **文献标志码:**A **文章编号:**1671-7414(2016)05-153-04

doi:10.3969/j.issn.1671-7414.2016.05.046

Total Error and Measurement Uncertainty in Clinical Laboratory Medicine

ZHANG Shi-shi^{1,2}, WANG Wei², ZHAO Hai-jian², WANG Zhi-guo^{1,2}

(1. Graduate School of Peking Union Medical College, Beijing 100730, China;

2. National Center for Clinical Laboratories, Beijing Hospital, Beijing 100730, China)

Abstract: In the clinical laboratory medicine, the measurement uncertainty (MU) is a relatively new concept. Over the years, experts of clinical laboratory medicine from all over the world made a great number of further researches and promote the development of MU, which led clinical laboratories to pay more and more attention to the meanings and functions of MU at the same time. However, because of the habitual using of the total error (TE) in clinical laboratories and similarities between concepts of MU and TE which easily resulted in confusion, a lot of laboratories still cannot completely accept MU. By explaining concepts of TE and MU and analyzing the pros and cons of models of TE and MU as well as their functions, the objective of this paper is to help clinical laboratories make further comprehensions of TE and MU and understand how to properly use them in practice.

Keywords: total error; measurement uncertainty; allowable total error

在临床检验中,关于总误差(total error, TE)和测量不确定度(measurement uncertainty, MU)的争论存在已久。关于临床实验室是否应该评价 MU 以及如何评价 MU 一直很难得出结论。1993年,测量不确定度表示指南(guide to the expression of uncertainty in measurement, GUM)的发布得到了国际临床化学和检验医学联合会(IFCC)的大力支持。ISO 15189 也对 MU 提出了明确要求。2014 年,欧洲临床化学和检验医学联合会(EFLM)在意大利米兰举行了第 1 次战略性会议,主题为“在斯德哥尔摩会议 15 年之后规定分析质量规范”。该会议支持 1999 年斯德哥尔摩会议中质量规范设定策略的等级结构并对其进行简化和调整,同时还建立了 5 个任务完成小组分别从事与质量规范相关的主题工作,其中有 1 个小组专门负责探讨研究“总误差”理论^[1],需要澄清在何种情况下依然可以使用 TE 或者在何种情况下应该由

MU 替代 TE。显然,在临床实验室的质量管理中关注 MU 是必不可少的。

1 总误差和测量不确定度的定义 按照现代测量理论,对于测量过程中给出的测量结果,为了进行相互比较的需要以及评判测量能力或结果质量的需要,测量结果不仅需要给出测量值,还需要给出测量值的分散性或可靠程度,二者共同构成完整的测量结果^[2]。

测量值减去被测量的真值被定义为测量误差。真值是一个理想的概念,是永远无法确切知道的,所以实际上我们使用的是约定真值(有时称为“参考值”)。因此,测量误差更准确的定义应该是测量值减去参考量值^[3]。测量误差可以分为系统误差和随机误差:系统误差是指在重复测量时保持恒定不变或按可预见方式变化的测量误差分量^[3],可以被纠正或通过适当的校准进行消除;随机误差是指在重复测量时按不可预见方式变化的测量误差

* 基金项目:北京市自然科学基金资助项目(7143182),北京医院课题资助(BJ-2015-025)。

作者简介:张诗诗(1991—),女,在读硕士研究生,研究方向:实验室质量管理,E-mail:Athena0915@163.com。

通讯作者:王治国,硕士,研究员,主要从事临床检验质量控制方法研究和室内质量评价工作,E-mail:zgwang@ncccl.org.cn。

分量^[3],可以通过重复测量或改进测量方法而减少,但无法消除。在临床检验的实际工作中,每次检测的测量结果都会与被测量的真值有差异,也就是存在测量误差,由系统误差分量和随机误差分量组成。许多误差分量具有相加性,最终试验结果的总误差是各误差分量的总和并且大于任何单个的误差。确定的总误差是分析质量应该达到的要求,可用于判断实验方法的可接受性。针对临床实验室任何一个检测项目所用的检测方法,其测量结果的TE必须在可接受的界限内(即允许总误差, allowable total error, ATE),若超出界限则说明需要提高检测方法的性能以减小TE。

然而,测量误差仅是计算MU的基础,并不能用来表示测量值的分散性或可靠程度。在国际上,反映测量值分散性或可靠程度的公认指标是MU。MU是一个量值区间,它反映测量值不能肯定的程度(即可靠程度)并与其呈反比关系,是表示被测量量值的分散性的量度。国际通用计量学基本术语(VIM)第3版中将MU定义为:根据所用信息,表征被赋予被测量量值分散性的非负参数^[3]。

以误差理论为基础提出的MU虽然类似于临床检验中的TE,但二者并不等同。MU的合理性和科学性是强调全面考虑测量的影响因素和消除发现的明显误差,这也是MU被提出且逐渐被接受的原因^[2]。在临床检验中,TE是系统误差和随机误差两项之和;而MU在原则上只有一个组分类型,无论其来源类型如何均可用同一方式处理所有组分^[4]。GUM将测量的各种要素组合在一起,并以MU来表达量值不完整性的信息,而不去关注“真值”和“误差”不可知的本质^[4]。虽然TE和MU不等同,但是也不矛盾,正如MU的评定必须基于误差,而偏倚(系统误差)、不精密度(随机误差)等概念也不会因MU的提出而消失。

2 总误差和测量不确定度模型 传统的“TE模型”和新提出的“MU模型”,其观点略有不同,后者可通过误差合成的方法将随机误差与系统误差结合到MU这一概念中。

2.1 总误差模型 为了阐明测量结果中所包含的真实信息,C. F. Gauss于19世纪早期首次提出误差模型^[4]。在临床检验的实际工作中,产生误差是由于检测过程不够完善引起的,客观上这些误差并不能确切得到。

由Jim Westgard及其同事提出的TE/ATE模型在临床检验中已经运用了几十年,通常被称为TE模型^[5]。该模型曾经被寄希望成为一个完整的模式,但是依然存在一些缺点且越来越凸显出来。

在常规TE模型中,偏倚和不精密度(乘以 z 因子)呈线性增加,产生的总误差的值为: $TE = bias + 2 \times CV_A$ 。这个值代表着围绕真值(未知)附近的区间界限的估计,在该区间界限内测量结果以规定的概率(通常是95%的概率)出现。

一般,常规TE模式假定ATE来源于生物学变异,通常使用下列值计算ATE:

偏倚: $< 0.25(CV_I + CV_G)$; 不精密度 $CV_A: < 0.5CV_I$ (CV_I = 个体内变异; CV_G = 个体间变异; CV_A = 分析变异。)

虽然Westgard提倡使用这些方程式和常数,但值得注意的是,这些值是在相互排斥的条件下导出的。换句话说,就是术语“不精密度”只在零偏移(假设的条件)的情况下有效;相反,术语“偏倚”只在零不精密度(假设的条件)的情况下有效。使用这些方程式和常数已经被证明会导致ATE严重的过高估计,在临床检验中是不合适的。

若以TE模型推导最大误差限即ATE,在测量结果呈正态分布的区域,其显著特征是将超出某界限(例如, $z=1.65$ 时界限为5%,单侧)的测量结果的相应数量用于确定该界限的时候,满足界限的偏倚和不精密度(CV_A)的结合呈现为斜率是 z 的一条直线。因此,该模式也称为线性TE模式。在直线的一端, $bias = 0$ (假设值),则 $ATE = 1.65CV_A$; 在直线的另一端, $CV_A = 0$ (假设值),则 $ATE = bias$ 。进一步假设测量结果和真值之间的差(即TE)是已知的,若TE小于ATE,则满足所有临床预期的质量需求。

然而,这些假设均非全面的或适合目的的。虽然在许多能力验证(PT)/室间质量评价(EQA)计划中以ATE的形式表达“可接受界限”,但是通常并不知道ATE是否满足患者结果临床应用的预期质量。在多数情况下,这些“可接受界限”是基于当前技术水平且仅满足最低的监管要求,可能并不适用于诊断和监测。

根据线性TE模型,最大偏倚和 z 倍最大不精密度的总和等于常数ATE。这个特征在偏倚和不精密度的整个取值范围内是有效的。但是,该模型仅在所涉及变量为不精密度和偏倚时才有效,却不适合生物学变异和其他显而易见的附加变异原因。虽然在监测质控品的变异时,生物学变异是不相关的,但是在处理患者样品时,生物学变异则是有关的。若该模型涉及生物学变异和其他显而易见的附加变异,则需要通过合并来处理,这是该模型的一种严重缺点。

此外,在TE模型中:一方面,当已知偏倚的值并且可以将其用于计算TE时,为什么不纠正该偏

倚而继续维持它呢? 另一方面,在实际测量中,因为不精密度和偏倚不易分开,所以偏倚通常不是精确已知的。这是 TE 模型中的一个重要矛盾,但是可以通过下述的 MU 模型进行解决。

2.2 测量不确定度模型 在临床检验中,由于 TE 在客观上是不能确定的,因此用被测量值和 MU 来表示被测量的真值是有依据的。MU 模型起源于 20 世纪早期^[6]。该模型声称 MU 是测量结果的属性,其不关注被测量“真值”的不可知本质并且结合了影响测量结果的已知因素。估计合成 MU 必须阐明随机误差和系统误差的所有来源,并且通过适当的方法评估每个主要成分对 MU 的影响。

MU 模型受到支持的 2 个主要原因是:①在计量学的大多数领域中使用 MU 模型;②MU 模型鼓励对 MU 各主要成分的影响进行估计并且采取措施将这些成分的影响最小化。

MU 是一个量值区间,通过与一个覆盖因子 k (例如 1.96)相乘可以扩展为覆盖区间,即扩展不确定度,并且测量值的最佳估计以规定的概率落在该区间内。假设测量结果是正态分布,对于以覆盖因子 1.96 扩展的 MU,测量值的最佳估计以 95% 的概率(双侧)落在这个区间内。

GUM 中规定,所有的偏倚在可能的情况下应该进行纠正。针对上述提及的 TE 模型的重要矛盾,MU 模型可以纠正已知的偏倚,并且将纠正过程的不确定度与随机误差产生的不确定度成分结合起来^[4],见图 1。

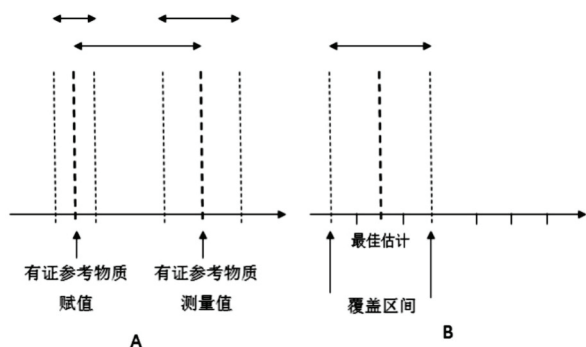


图1 将随机误差与系统误差相结合的 MU 模型^[4]

图 1 中,A 图第一排的 2 个箭头分别表示有证参考物质(CRM)的赋值不确定度和定量不确定度,第二排的箭头表示定量值与赋值的偏差。A 是 CRM 定量值的偏倚纠正过程,使得 B 图中的定量不确定度增加了,纠正后 CRM 的定量值与赋值一致。B 图中覆盖区间即扩展不确定度,最佳估计以一定的概率落在覆盖区间内。

MU 模型在许多方面均是合理适用的。在临床实验室内,MU 模型可以给出关于测量质量的信

息,用于比较临床实验室间的计量质量并且有助于解释测量结果,特别是在测量量值接近临床决定限的时候。在进行疾病诊断时,要注意测量量值并非真值,与临床决定限比较时应该考虑测量值的 MU。只要临床决定限不在测量结果的 MU 范围内,就可以给临床医生提供明确的信息。在临床实验室外,我们也依然会遇到 MU 模式,例如交通警察在对车辆进行速度测量时,只有以 MU 为基础进行纠正之后测量速度才可以与速度限进行比较。

然而,MU 模式也存在不足之处。关于临床实验室是否可以只评定测量过程的 MU、是否可以用“自上而下”的办法评定 MU、是否可以用临床实验室的实际工作数据来评定 MU 这 3 个主要技术问题,虽然国内检验医学专家们已经提出了一些合理的看法,但是还需要进一步的改进和完善^[2,7]。此外,在 IQC 为基础的质控程序下控制全部变异来源,MU 模型才能允许在测量患者样品时处理真正的总 MU^[8]。目前,国内临床实验室对 MU 的理解依旧不充分,这也导致许多实验室尚未进行 MU 的评定和报告。由于 MU 和 TE 有类似之处,在理解 MU 的定义时容易混淆,并且 MU 的计算也较为繁复,因此距离临床实验室完全接受并使用 MU 还有一个漫长的过程。

3 总误差和测量不确定度可否调和 目前已有一些文章对 TE 和 MU 之间的差异和相似之处进行了综述^[9~11]。TE 被定义为围绕参考(“真”)值的区域,在该区域内测量结果以规定的概率出现;MU 被定义为围绕测量结果的区域,在该区域内“真”值以规定的概率出现。两种模型的相似之处在于它们均表达试验结果的可靠性,但是却源自于不同的观点。

当 MU 模型消除偏倚并且尽可能简单地仅处理一些独立变异时,其与 TE 模型在数学上是相同的。然而,在临床实验室的常规操作中,变异来源是很复杂的,包括个体内和个体间的变异、试剂批号内和试剂批号间变异、测量系统内和测量系统间变异、方法内和方法间变异、实验室内和实验室间变异、校准变异以及检验前变异等。因此,MU 模型的使用可以对 TE 模型的缺陷进行补充。

Westgard 致力于调和 TE 和 MU^[12],他认为“TE 是准确度的度量,MU 是溯源性的度量”,并且他相信 TE 在医学实验室的质量管理中至关重要,MU 对于实现实验室间结果的可比性也必不可少。TE 模型已经在临床检验中适用了几十年,其需要考虑复杂的现实情况以满足 MU 模型的需要;而 MU 模型也需要发展成为现实可操作的方法。目前,MU 模型主要用于估计测量结果的可靠

性,但是在设定质量规范方面的应用尚未得到良好发展。在单一方法的室内质量控制(IQC)中,由于排除了在测量患者样本时会遇到的许多变异来源,因此两种模型变得完全一样,在TE模型上建立的质量规范设定均适用于2个模型。

4 总结 在溯源性的世界中,参考值应该是“无偏的”值,实验室测量的偏差将用于定义其不确定度^[13]。在这种情况下,努力提升TE框架用于规定临床测量的允许MU是非常值得赞赏的。

由Jim Westgard及其同事提出的TE/ATE模型几十年来很好地服务于临床检验,并且对临床检验的理论和实践的影响仍占据主导地位。然而,TE在VIM和GUM中是不承认的。除临床检验外,MU在计量学的其他领域已得到广泛应用,但其复杂的数学运算降低了其在临床实验室中的可操作性。然而,必须承认的是,计算测量结果的MU代表着一个重要的机会,有助于临床实验室在医疗诊断和治疗效果监测的方面提高质量。

若根据Westgard的观点将TE模型和MU模型分开或者用MU模型代替TE模型,这无法缓解临床检验中TE和MU的紧张关系。事实上,通过估计MU的分量以及应用计量学其他领域的MU计算方法,可以实质性地提高临床检验中已经广泛使用的TE模型。目前,以此为方向的研究成果也正在临床检验中实行^[10,11,14]。

MU的发展是TE概念的一种演变,是科学合理的一种进步,而非单纯的替代。因此,临床实验室应该充分认识TE和MU的优缺点,使二者相辅相成,而不是直接否认二者之一。

参考文献:

- [1] Panteghini M, Sandberg S. Defining analytical performance specifications 15 years after the Stockholm conference[J]. Clin Chem Lab Med 2015, 53(6): 829-832.
- [2] 中国合格评定国家认可委员会. 医学实验室-测量不确定度评定[S]. 北京: 中国标准出版社, 2015: 60-61.
- [3] ISO/IEC Guide 99-2007. International vocabulary of metrology-Basic and general concepts and associated terms(VIM)-third edition[S]. Switzerland: Geneva, ISO/IEC Guide 99-2007.
- [4] 王治国. 临床检验质量控制技术[M]. 3版. 北京: 人民卫生出版社, 2014.
- [5] Panteghini M, Sandberg S. Total error vs. measurement uncertainty: the match continues[J]. Clin Chem Lab Med, 2016, 54(2): 195-196.
- [6] Saty SM. From certainty to uncertainty: the story of science and ideas in the twentieth century[J]. Joseph Henry Press, 2002, 296(5571): 1244.
- [7] 杨振华. 是否可以对医学实验室评定不确定度争议做一小结[J]. 临床实验室, 2015, 9(11): 57-59.
- [8] O'Donnell GE, Hibbert DB. Treatment of bias in estimating measurement uncertainty[J]. Analyst, 2005, 130(5): 721-729.
- [9] Rozet E, Marini RD, Ziemons E, et al. Total error and uncertainty: friends or foes? [J]. Trac Trend Anal Chem, 2011, 30(5): 797-806.
- [10] Farrance I, Frenkel R. Uncertainty of measurement: a review of the rules for calculating uncertainty components through functional relationships [J]. Clin Biochem Rev, 2012, 33(2): 49-75.
- [11] Farrance I, Frenkel R. Uncertainty in measurement: a review of monte carlo simulation using microsoft excel for the calculation of uncertainties through functional relationships, including uncertainties in empirically derived constants[J]. Clin Biochem Rev, 2014, 35(1): 37-61.
- [12] Westgard JO. Useful measures and models for analytical quality management in medical laboratories [J]. Clin Chem Lab Med, 2016, 54(2): 223-233.
- [13] Panteghini M. Application of traceability concepts to analytical quality control may reconcile total error with uncertainty of measurement[J]. Clin Chem Lab Med, 2010, 48(1): 7-10.
- [14] White GH, Farrance I, AACB uncertainty of measurement working Group. Uncertainty of measurement in quantitative medical testing: a laboratory implementation guide[J]. Clin Biochem Rev, 2004, 25(4): S1-24.

收稿日期: 2016-03-21

修回日期: 2016-05-03

(上接 152 页)

- 暴露分析与防范[J]. 医学理论与实践, 2015, 28(2): 267-268.
- [6] Emmert EA, ASM Task Committee on Laboratory Biosafety. Biosafety Guidelines for handling microorganisms in the teaching laboratory: development and rationale[J]. J Microbiol Biol Educ, 2013, 14(1): 78-83.
- [7] 任碧琼, 朱浩稳, 邹国英, 等. 检验医学学生实习前生物安全强化培训效果评估[J]. 现代检验医学杂志, 2011, 26(5): 10-12.

Ren BQ, Zhu HW, Zou GY, et al. Evaluate the effect of laboratory medicine students bio-securaty intensive training before working in clinical laboratory[J]. J Mod Lab Med, 2011, 26(5): 10-12.

- [8] Merino-de la Hoz F, Dura-Ros MJ, Rodriguez-Martin E, et al. Knowledge and adherence to bio- safety measures and biological accidents by nursing students during their clinical practice[J]. Enferm Clin, 2010, 20(3): 179-185.

收稿日期: 2015-11-22

修回日期: 2016-07-30