arXiv:2303.06441v1 [stat.AP] 11 Mar 2023

利用血糖仪测量毛细血管血糖来进行口服葡萄糖耐量试验的贝叶斯统计分析

Nicol´as E. Kuschinski、J. Andr´es Christen、Adriana Monroy，CIMAT，瓜纳华托，瓜纳华托州，墨西哥

摘要

诊断2型糖尿病的一种常见检测是口服葡萄糖耐量试验（OGTT）。近期对口服葡萄糖耐量试验的研究进展，将其视为一个贝叶斯逆问题。这些数据分析方面的进步有望显著提升口服葡萄糖耐量试验的描述能力。口服葡萄糖耐量试验通常通过侵入性强、令人不适且价格较高的静脉血检测来完成。一个自然的问题是，改进的数据分析技术能否让侵入性较小、成本较低的血糖仪测量得以应用。本文将探讨这一问题。我们采用一个动态模型，为血糖仪毛细血管血糖测量建立误差模型，并将其结果与65名患者的静脉血糖检测结果进行比较；在超过90%的观测病例中，两者吻合。我们的结论表明，该模型（或与其非常相似的模型）或许能够以合理的准确度，利用毛细血管葡萄糖进行口服葡萄糖耐量试验。

1 引言

糖尿病是一种严重疾病，每年影响的人数众多且不断增加 [1]。根据世界卫生组织的说法，糖尿病是指胰腺无法产生足够的胰岛素，或身体无法有效利用胰岛素 [2]。

糖尿病分为两种类型：1型糖尿病与遗传性疾病有关，通常在生命早期确诊；2型糖尿病则会随着时间推移而发展，初期往往症状很少。及时诊断和治疗能够相对轻松地控制2型糖尿病，尽管它仍会带来困扰。然而，如果长期得不到治疗，2型糖尿病可能成为严重的健康风险，甚至致命 [3]。

诊断糖尿病时，常用的一项检查是口服葡萄糖耐量试验（OGTT）。检查时，先测量空腹患者的血糖，然后让其饮用含糖浆液（75 克葡萄糖）。接下来的两

小时内，会在多个时间点再次测量患者的血糖，以了解患者随时间处理血糖的情况 [3, 4, 5]。

基于口服葡萄糖耐量试验结果的传统诊断方法一直相当粗略，通常依据一些简单的描述性统计量得出结果，例如最大测量值、测量值平均值或最后一次测量值 [5]。这些分析方法存在一个显而易见的问题：它们都忽略了口服葡萄糖耐量试验测量值反映的是一个随时间变化的过程，且具有明确的相关结构。不过，近年来有人提出了数学模型，利用动力系统模拟患者在口服葡萄糖耐量试验过程中的血糖变化。血糖变化由一组常微分方程（ODEs）描述，再由统计模型描述测量值。口服葡萄糖耐量试验被构建为一个适用于贝叶斯推断的反问题。尽管使用数学模型研究口服葡萄糖耐量试验并非新进展 [4], ，但直到最近，这些模型才被认真用于推断。近期， [6]和 [7] 提出了模型，并用这些模型对部分口服葡萄糖耐量试验进行了推断，取得了令人鼓舞的结果。

口服葡萄糖耐量试验的一个问题是费用较高且操作繁琐。患者在整个试验期间必须在手臂上留置套管针，而结果则需经过实验室的缓慢流程处理。使用血糖仪可以廉价、轻松且快速地测量血糖，但人们一直认为血糖仪是不够可靠的方法，精度不足以用于口服葡萄糖耐量试验 [8]。然而，我们分析口服葡萄糖耐量试验结果的能力有所提升，这表明或许是时候重新探讨血糖仪能否用于口服葡萄糖耐量试验分析这一问题了。

本文首先介绍我们将采用的常微分方程模型，以及利用该模型进行推断的方法。接下来，我们提出一种针对血糖仪测量的误差分布模型，并尝试使现有模型适应这些新情况。我们对这一新模型进行了测试，结果令人鼓舞；随后构建的改进模型在 90% 以上的患者中都能给出令人满意的结果。

2 常微分方程模型

本文中的所有进展都与测量误差建模有关，而且在很大程度上应独立于所考虑的具体动态模型，前提是该模型能够合理地拟合数据。尽管如此，仍有必要使用某种动态模型来进行测试，否则便无法评估其对推断的影响。虽然我们预计本文中的结果适用于多种动态模型，但如果使用不同的动态模型，我们建议进行某种形式的验证。 [7], 中详细介绍了一种专为拟合口服葡萄糖耐量试验数据而设计的动态模型；另一种与之密切相关、且本文采用的模型则见于[6]。该模型已在许多真实患者病例中得到应用。下面我们将概述该模型的相关信息。该动态模型如下：

$$\frac{dG}{dt} = L - I + \frac{D}{\theta_2} \tag{1}$$
$$\frac{dI}{dt} = \theta_0(G - G_b)^+ - \frac{I}{a} \tag{2}$$
$$\frac{dL}{dt} = \theta_1(G_b - G)^+ - \frac{L}{b} \tag{3}$$
$$\frac{dD}{dt} = -\frac{D}{\theta_2} + \frac{2V}{c} \tag{4}$$
$$\frac{dV}{dt} = -\frac{2V}{c}. \tag{5}$$

该模型用 G表示血糖，其中 mg/dl。 I 表示患者血液胰岛素的作用。 L 表示当患者血糖过低时，患者身体作出的升糖反应（这通过多种机制实现，例如胰高血糖素）。 D 是消化系统中的葡萄糖，而 V是尚未摄入的葡萄糖。 a, b, c和G_b 的值是取自文献的已知常数，而 θ_0， θ_1， θ_2 和 G(0)则是推断参数。

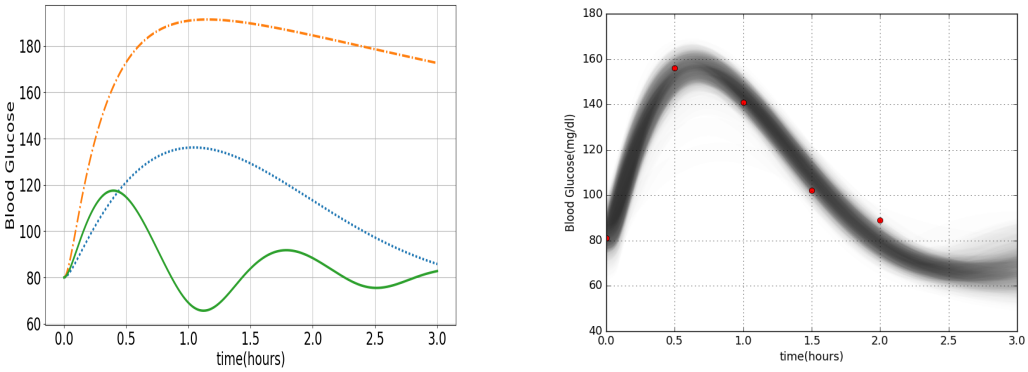
该模型遵循简单的三室结构和双重反馈逻辑，接下来我们将对此进行说明。患者摄入葡萄糖浓缩液后，葡萄糖进入消化系统（见（5）和（4））。进入消化系统后，葡萄糖以由 θ_2 决定的速率进入血液（见（4）和（1））。随着血糖升高，人体会产生胰岛素 I 来调节血糖，调节效率取决于 θ_0 （见（2）和（1））。当葡萄糖低于基础水平 G_b时，反应 L 会升高血糖，其效率取决于 θ_1（见（3）和（1））。

图2（a）展示了模型在推断参数取不同值时的表现（所有情况下G(0)都相同）。图中显示了一名健康患者（点线）、一名具有胰岛素抵抗的患者（虚线），以及一名胰岛素分泌量异常偏高的患者（实线）。健康患者的胰岛素水平升高后，会在2-3小时内恢复正常。可能患有糖尿病的患者在3小时后仍无法使血糖恢复至接近静息血糖水平，而胰岛素分泌量较高的患者在这段时间内会出现数次波动。按照以往的口服葡萄糖耐量试验分析模型，这类患者会被归为“正常”，尽管其葡萄糖代谢表现并不符合典型健康患者的预期。

3 静脉数据的模型推断

口服葡萄糖耐量试验中收集的数据，是在不同时间对 G（t）进行的测量。如 [6], 所述，这些数据采用以下模型:

$$y_i = G(t_i) + \sigma\epsilon_i;\ \epsilon_i \sim \mathcal{N}(0,1)$$



(a) (b)

图 1: （a）三位不同假设患者的行为模型 $G(t)$。点线表示健康患者，虚线表示存在胰岛素抵抗的患者，实线则表示一种振荡情形，即患者产生了异常大量的胰岛素。（b）对一位健康患者的口服葡萄糖耐量试验数据进行贝叶斯分析。红点表示血糖读数，“阴影”曲线表示 $G(t)$的后验分布（以时间 t 为中心的垂直切片代表时间 t时 $G(t)$后验分布的核密度估计）。这里展示了用 $G(t)$建模的口服葡萄糖耐量试验曲线估计不确定性，同时还显示了测试 2 小时后血糖的预测值。

其中 σ 设为 5。 $\mathcal{N}(0,\ 5)$ 的误差模型根据对设备标准误的临床意见确定。

例如，对于一位实际患者，在 t =0:00、0:30、1:00、1:30 和 2:00 小时，我们获得了葡萄糖测量值 $y = 81,\ 156,\ 141,\ 102,\ 89\ mg/dl$。我们的目的是利用这些数据推断参数 $\theta_0,\ \theta_1,\ \theta_2$ 和 $G(0)$。这是一个“逆”问题，也就是说，是一个关于回归模型的推断问题，其中回归量由常微分方程组驱动。

在 [6], 中， $\mathcal{N}(0,\ 5)$ 误差模型未经严格审查，主要是因为它能提供良好的结果而被接受。在本研究中，AM 能够获得一组 15 份重复血糖测量值。进行这些测量时，采集了一份静脉血样本，并对其葡萄糖含量测量了两次（这些并非完整的口服葡萄糖耐量试验结果，而只是 15 份血样）。这类数据很少，因为没有医学上的理由对同一份样本测量两次，而且这些数据仅为此目的而采集。这些样本之间的差异表明，测量值的方差实际上低于预期（我们的估计显示标准差为 3.09^{mg}_{dl} ），但其中包含两个离群值，距离零超过两个标准差。无论将差异视为绝对差值还是比例差值，都会发现这些离群值。现有数据量仍然相对较少，但有理由认为误差方差甚至低于先前预期，不过误差分布应当具有厚尾。

因此，我们对原始静脉血模型进行修改，将高斯误差替换为自由度为 4 的缩放 t 分布。因此，静脉血测量值的新误差分布为

$$y_i = G(t_i) + \sigma\epsilon_i;\ \epsilon_i \sim t(4)$$

其中 $\sigma = 2.2$，使误差的标准差为 $\sigma\epsilon_i = 3.09$。

3.1 先验

我们为每个参数指定如下先验分布：

$$\theta_0 \sim Gamma(2, 1)$$
$$\theta_1 \sim Gamma(2, 1)$$
$$\theta_2 \sim Gamma(10, 1/20)\ \mathbb{1}\{\theta_2 > 0.16\}$$
$$G(0) \sim \mathcal{N}(80, 100)\ \mathbb{1}\{G(0) \in [30, 400]\}.$$

这些先验旨在反映对真实患者参数的估计值。它们被设定得较为宽泛（高概率区域远超通常的估计范围），以避免数据即使在 $G(t)$ 的观测值很少，或患者的葡萄糖曲线极不寻常时，也会对结果产生过大影响。例外情况是 $G(0)$ 的初始先验。在这种情况下，已有患者被观察到空腹血糖水平极高。该先验为

仍然足够宽泛，因为任何初始测量值低于 30 或高于 $400mg/dL$ 的患者都不会接受检测，而是会被送往急诊护理（会先进行血糖仪血液检测，以排除并立即处理此类情况）。有关先验选择的更多详情，请参见 [6]。

为对模型进行推断，使用马尔可夫链蒙特卡罗模拟获取后验样本。使用 LSODA 软件包对常微分方程中的 $G(t)$ 进行数值计算。马尔可夫链蒙特卡罗模拟采用 t-walk [9]。图2 (b) 展示了对一名真实（健康）患者数据进行推断的结果。每条竖直的灰色切片都表示相应时间点 $G(t)$ 后验分布的核密度估计。可以看出，至少对于这名患者，模型与数据拟合得非常好。

4 毛细血管血糖数据

前文提到的常微分方程模型适用于标准血液检测，但这些检测具有侵入性，会给患者带来不便。检测速度较慢，费用中等，而且会占用实验室工作人员宝贵的时间。不过，测量血糖还有一种更实用的方法，称为血糖仪。血糖仪是一种容易获得的设备，可测量毛细血管中的血糖，而非静脉血糖。血糖仪最初是为糖尿病患者设计的，方便他们简单、快速且相对经济地自行测量血糖。血糖仪易于使用，几乎能即时给出结果，而且只需用采血针从患者指尖采集一滴血即可。通常可以在当地药店以较低价格买到血糖仪套装。然而，血糖仪并未普遍用于口服葡萄糖耐量试验的原因在于，人们通常认为其准确度不够。

虽然血糖仪的准确性确实明显低于静脉实验室检测，但口服葡萄糖耐量试验分析中常用的低效分析方法，可能才是这些检测无法使用的原因。用于分析的动态模型问世后，或许可以重新审视能否从血糖仪测得的口服葡萄糖耐量试验数据中获取有用信息。

为此，AM 收集了 65 名患者的数据，这些患者因预约进行常规口服葡萄糖耐量试验而前往墨西哥城的 Hospital General 就诊。除了常规的口服葡萄糖耐量试验测量外，这些患者还在采集常规检测用血样的同时，接受了血糖仪血液检测。血糖仪测量本质上是对同一患者进行的一次重复口服葡萄糖耐量试验，只不过使用血糖仪测量毛细血管血。为确保患者安全，AM 还会在给患者服用葡萄糖浓缩液前，照常使用血糖仪检测静脉血（这只需几秒钟，并能确保如果患者的葡萄糖水平过高，就能立即接受治疗）。这种做法很方便，因为它提供了一项测量，可用于区分血液类型（静脉血与毛细血管血）的影响和血糖仪的影响。

为适应质量较差的数据，我们调整了测量误差模型。有关血糖仪误差的文献表明，测量条件会极大地

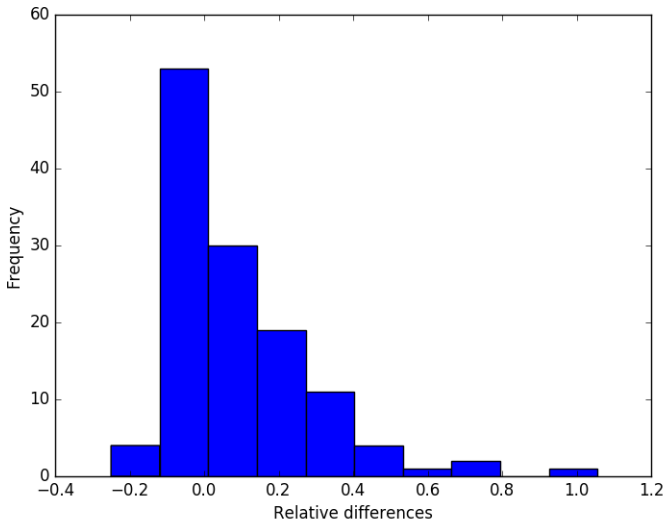


图2：AM 数据集中血糖仪毛细血管血糖测量值与实验室静脉血糖测量值之间相对差值的直方图。

影响血糖仪测量的可靠性。口服葡萄糖耐量试验的条件相当特殊：由专业人员在严格控制的环境中进行，但患者的身体正承受一种非常特定且多少有些不寻常的压力。

AM 数据集中血糖仪与静脉血糖测量值之间的比例差值以直方图形式呈现，如图2所示。

需要注意的一个重要细节是，除了数值分布范围很广之外，还存在明显可见的偏差。这种偏差在医学上是合理的，因为毛细血管中的血液会在进入静脉之前将葡萄糖输送到肌肉。与血糖仪检测误差相比，静脉检测误差可以忽略不计，因此，对血糖仪数据与静脉数据之间的差值建模，实质上就是对血糖仪误差建模。

选取了一种模型，将误差表示为两部分之和：一部分是由毛细血管中葡萄糖损失导致的伽马分布偏差，另一部分是正态分布的测量误差。由于这种偏差由患者体内的生物过程驱动，因此假设在同一次口服葡萄糖耐量试验中所有测量值的偏差相同，但不同患者之间有所差异。

根据数据获得偏差和误差参数的点估计，并建立毛细血管数据模型。该模型现在

$$y_i = G(t_i) + b + \epsilon_i;\ b \sim Gamma(0.4, 4.3)\ \epsilon_i \sim \mathcal{N}(0, 100).$$

图3展示了一位患者通过血糖仪测量得到的毛细血管口服葡萄糖耐量试验结

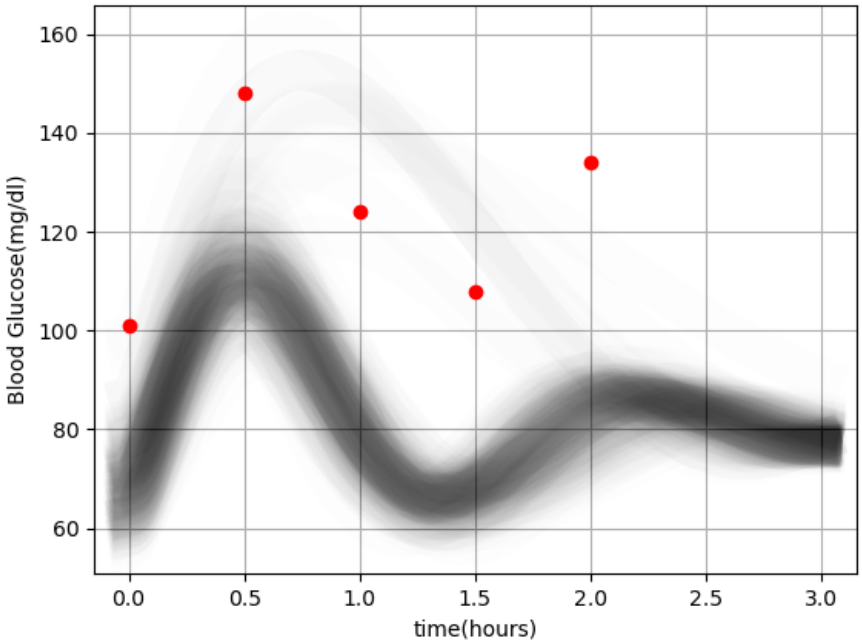


图3：基于毛细血管数据对患者进行推断。与静脉数据不同，血糖仪测量存在已知偏差，我们的误差模型将这一点纳入了考虑。因此，根据毛细血管数据估计的口服葡萄糖耐量试验曲线不一定能很好地拟合数据。仅凭这条曲线，很难一眼看出模型对数据的描述效果。一种判断方法是将其与同一患者的静脉口服葡萄糖耐量试验结果进行比较。

果。与图2相比，我们观察到，估计曲线的走势不 一定会贴近数据。这是因为已知数据存在偏差，而推断通过使用上述误差模型将其考虑在内。仅凭图表和数据，很难判断估计是否准确，因此必须将其与同一位患者的静脉口服葡萄糖耐量试验进行比较。

5 结果

该模型用于对65名患者的毛细血管血和静脉血数据进行推断。图4(a)和(b)展示了其中两名患者的结果。第一名患者的静脉数据与毛细血管数据差异很大。尽管如此，毛细血管血液误差模型仍能补偿

偏差严重的数据，并得出与静脉数据估计结果相似的后验估计。第二名患者的静脉数据与毛细血管数据相似。观察到的偏差极小，但模型足够灵活，能够判断在这种情况下无需补偿，因此得出的估计与基于静脉血获得的估计非常接近。

尽管我们的模型在拟合毛细血管血糖读数方面具有优势，但我们仍发现一些情况，其结果与静脉口服葡萄糖耐量试验并不完全吻合。可能出现两个主要问题。第一种情况是，同一患者的静脉估计值和毛细血管估计值均为 $G(t)$，但尽管毛细血管检测所得估计值与静脉检测的估计值重叠，使用毛细血管数据对 $G(t)$ 的估计却具有过大的方差，以至于很难据此判断患者的血糖调节状况。需要指出的是，这未必会构成重大问题。如果根据血糖仪数据对患者的 $G(t)$ 进行估计时方差过大，则可将毛细血管口服葡萄糖耐量试验判定为结果不确定，并建议患者接受常规静脉检测。补充材料中包含所有患者的后验估计图像。患者 1、3、18、19、36、39、54、55、56 和 58 均可见上述情况，占受检患者总数的 15.4%。图4(c) 展示了一个此类情况的示例（患者 55）。

第二个问题是，使用毛细血管血糖仪数据得出的估计值与静脉数据得出的估计值不一致。有时，仅通过目视检查数据就能发现这一点，因为毛细血管口服葡萄糖耐量试验和静脉口服葡萄糖耐量试验的数据表现出不同的特征，而调整误差模型（甚至很可能调整常微分方程组）都无法解决这个问题。在接受这项检测的65名患者中，只有3名出现了这种不可接受的误差。这些患者分别是患者4、患者31和患者51。图4(d)展示了这种情况的一个例子（患者31）。

6 结论

新的毛细血管模型在约5%的受测患者中产生了不可接受的结果。尽管这似乎是一个高得令人担忧的错误率，但必须将其与现有的基准错误率进行比较。虽然目前没有现成的关于经典口服葡萄糖耐量试验分析（不使用常微分方程模型）诊断错误率的数据，但有理由怀疑，错误发生率甚至可能高于使用毛细血管模型的检测（[5] 报告了某些版本的经典诊断标准的误诊率）。

使用毛细血管血和血糖仪具有显著的实际优势：静脉检测不仅费用高且令人不便，而且许多人可能根本无法进行这类检测，因为他们无法前往配备适当设备的医疗机构。相比之下，血糖仪在药店十分常见，适合家庭使用，价格也相对低廉。鉴于

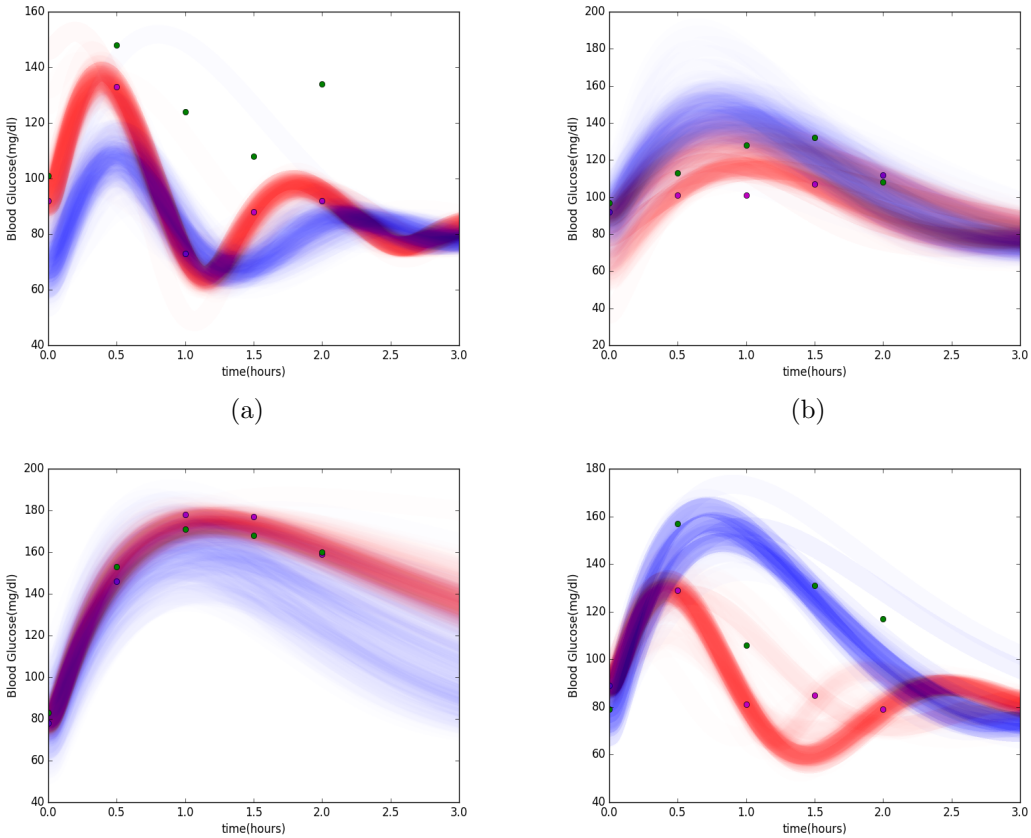


图4：四位真实患者的数据。使用静脉数据可得到洋红色圆点和红色曲线；使用毛细血管数据可得到绿色圆点和蓝色曲线。在（a）中，我们观察到两种血液来源产生的数据并不一致，但毛细血管血的误差模型能够对此进行调整，因此我们得到了相当相似的 $G(t)$ 估计值。这是图3中的同一位患者。在（b）中，我们观察到一位患者的毛细血管数据与静脉数据相似，偏差极小，而且模型足够灵活，也能适应这种情况，因此毛细血管口服葡萄糖耐量试验的结果与静脉口服葡萄糖耐量试验的结果也非常相似。在（c）中，我们看到的是基于毛细血管数据进行口服葡萄糖耐量试验时最常见的问题。对于这位患者，静脉口服葡萄糖耐量试验与毛细血管口服葡萄糖耐量试验的结果相当吻合，但毛细血管模型的不确定性过大，无法对患者状况得出明确结论。在（d）中，我们看到一种罕见情况：毛细血管口服葡萄糖耐量试验与静脉口服葡萄糖耐量试验的结果完全不一致。对于这位患者的数据，模型产生的误差可能无法接受。

只要有合适的葡萄糖浆、血糖仪和适当的分析工具，以便从毛细血管检测（例如本文建议的检测）推断结果，理论上就可以由护士在基础医疗机构以低廉的成本进行口服葡萄糖耐量试验。值得注意的是，可以轻松编写一款易于使用的软件来分析血糖仪读数，生成类似图3的图表，并得出所有必要结论。

值得注意的是，报告中的许多错误源于统计分析以外的环节；而且，至少有一个案例中，正是毛细血管口服葡萄糖耐量试验帮助发现了静脉口服葡萄糖耐量试验中的问题。具体而言，初次检查时发现了四个不可接受的案例，于是对这些案例的原始数据进行了复核。复核原始数据后，发现其中一个案例是由于填写静脉检测结果时出错所致。该患者最初接受的临床治疗有误；如果没有对毛细血管数据进行这番仔细审查，这一错误就不会被发现。找到正确数据后，我们重新进行了分析，发现静脉数据与毛细血管数据吻合良好。

如果实际条件允许进行规范的静脉口服葡萄糖耐量试验，显然应优先选择这种方式。然而，我们认为，在无法进行静脉口服葡萄糖耐量试验的情况下，可考虑将毛细血管口服葡萄糖耐量试验作为替代方案。

7 数据可用性

支持本研究结果的数据可向通讯作者索取。

参考文献

[1] Wild S、Roglic G、Green A、Sicree R、King H。糖尿病的全球患病率：2000年估算值及2030年预测值。http://care.diabetesjournals.org/content/27/5/1047.full；2004年。[2] 世界卫生组织。糖尿病。https://www.who.int/news-room/ fact-sheets/detail/diabetes；。访问日期：2023年2月14日。[3] 美国医学会 MM。美国医学会糖尿病生活指南：预防和治疗*2*型糖尿病——您和家人需要了解的基本信息。约翰·威立父子出版公司。第1版。2006年。[4] Jansson L、Lindskog L、Norden N、Carlstrom S、Schersten B。采用数学模型评估口服葡萄糖耐量试验的诊断价值。《计算机与生物医学研究》1980年；13：512–521。doi：10.1016/0010- 4809(80)90031-2

[5] Davidson MB、Schriger DL、Peters AL、Lorber B。重新审视糖尿病诊断的口服葡萄糖耐量试验标准。《普通内科医学杂志》2000年；15：551–555。doi：10.1046/j.1525-1497.2000.08024.x[6] Christen JA、Capistran M、Monroy A 等。用于口服葡萄糖耐量试验的糖尿病最小模型。https://arxiv.org/abs/1601.04753；2016年。[7] Flores-Arguedas H、Capistran MA。口服葡萄糖耐量试验（OGTT）期间葡萄糖动力学的贝叶斯分析。*MBE* 2021年；18(4)：4628–4647。doi：10.3934/mbe.2021235[8] Ginsberg BH。影响血糖监测的因素：测量误差的来源。《糖尿病科学与技术杂志》2009年；3：903–913。doi：10.1177/193229680900300438[9] Christen JA、Fox C 等。适用于连续分布的通用抽样算法（t-walk）。贝叶斯分析 2010年；5(2)：263–281。