

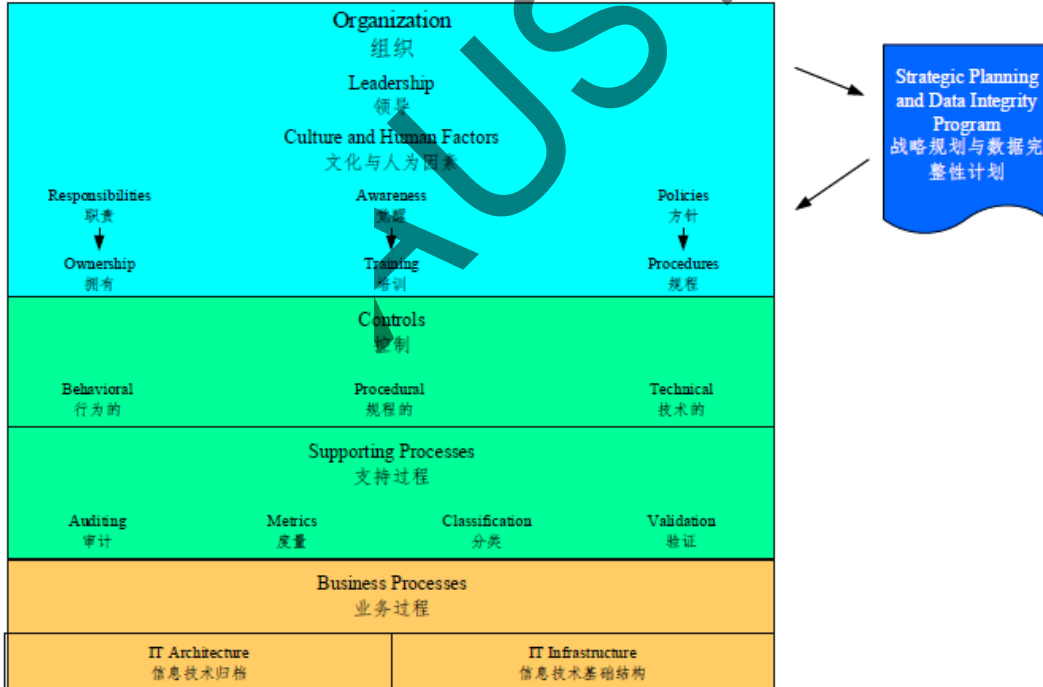
## 数据可靠性系列：数据准确真实

柯争先 贾晓艳

**A ( Accurate ) 准确的** 世界卫生组织 ( WHO ) 在其数据可靠性 ( DI ) 指南附件一中对 “A准确的” 定义是 “data are correct, truthful, complete, valid and reliable数据是正确的、真实的、完整的、有效的和可靠的” ，ISPE 《记录与数据可靠性指南》表达的是 “Free from error无错误、No editing performed without documented amendments没有在文件化修改之外进行编辑、Conforming to truth or standard符合事实或标准” 。

不管是纸质的，还是电子的，抑或影像资料，**数据的本质是“反映事实的真相”**。我们曾分享了“归属至人”、“清晰可溯”、“同步记录”、“原始一致”等数据可靠性的属性要求，但如果“准确真实”属性出了问题，其他的属性即使维持得再好也枉然，因为不准确真实的数据已经无法反映事实的真相，其本质意义也就不存在了。“准确”和“真实”其实是两个不同的概念，通俗来讲一个是说数据“准不准”，一个是说数据“真不真（没有经过操纵、篡改）”。

我们在前期的数据可靠性系列微信稿中分别从文化、质量体系、人员、ALCOA属性期望和要求、QC实践等不同角度，去分享解析数据综合治理的各方面要求及特殊考量；如下数据治理框架（来源：ISPE 《记录与数据可靠性指南》）可作为一个很好的参考。



“真实性”问题将不会在本篇中进行深入阐述，其涉及到如“作假”、“伪造”、“篡改”、“操纵”，已经属于道德、诚信和文化层面的深层次问题了，大家感兴趣的话可以回顾该DI系列的“文化篇”及“人员篇”，里面均有较为深刻的说明。本篇主要讨论的是“准确性”问题。

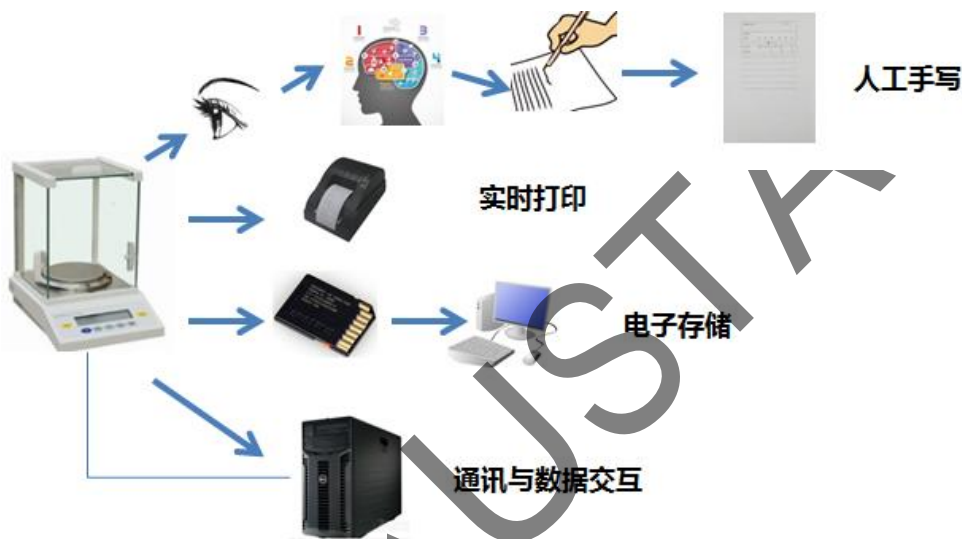
中国药品数据管理规范（征求意见稿）描述如下，

确保数据准确的控制措施至少包括：

- (一) 产生数据的设备应当经过校准、确认和维护；
- (二) 产生、储存、分配、维护及归档电子数据的计算机化系统应当经过验证；
- (三) 分析方法和生产工艺应当经过验证或确认；
- (四) 数据应当经过审核；
- (五) 偏差、异常值、超标结果等应当经过调查；
- (六) 应当建立完善的工作流程减少差错的发生。

从六个不同层面阐述了常见控制措施。

为了便于大家日常应用理解，我们做了如下简单的示意图来分析说明：



按照上述示意图所示，实际上真正的原始数据应该是当物料放到电子天平稳定后的读数，这个读数客观反映被称量物料的事实真相（“质量”），而质量的“理论真值”是不被任何人所知道的，通过校准溯源的方式我们能得到一个参考的读数（与标准器的比对），用其来代表该称量物料的质量。所以这里“仪表校准”是保证“读数准不准”的第一步。经过校准的天平读数产生之后，当我们移开物料拿去使用，而该读数就消失了，它是不会自己保留下来的，这个“瞬息”的数据并不能产生“耐久性”，无法作为原始数据保留下来，所以就需要通过人眼观察记录，打印机实时打印至纸张上，SD卡存储为电子数据或计算机化系统通讯传输处理等方式被保留下来。因此保证数据的准确性，需要从如下几个方面着手及思考：

1. 需要考虑数据源头捕获与采集的方式（感官？传感器？摄影？其他？）
2. 需要考虑数据保留和存储的方式（纸质？电子？录像？）
3. 需要考虑数据转换迁移的过程（通讯传输？USB拷贝？打印输出？备份与恢复？）

对于感官指标数据，读数经过人眼观察、大脑记忆、手工书写的程序。首先记录人必须符合基本要求并经过培训（比如：灯检岗位对视力有要求，红绿色盲患者无法用标准比色卡的方式去测pH值，记录人必须是懂阿拉伯数字的意义和书写，读取液体体积时视线必须和液面水平后读取凹陷处数值），此外由于人的主

观能动性（疲劳、情绪波动、有意作假），因此保证感官指标准确可靠，就加入了双人复核的要求，包括中国GMP附录《计算机化系统》中也对人工往系统录入关键数据需要双人复核也有类似规定一样。事实上，“大脑就是一个隐形的便签纸/复写纸”，所以导致纸质记录容易被操纵而且不留下痕迹。

对于实时打印的数据，数据从天平上产生，最后通过连接线到打印机输出至打印纸上，由于中间过程没有任何的“打断与干预”，一旦确认打印纸上的数据和原天平的数据一致之后，即可认为打印纸是准确可靠的。因此，此类数据依赖打印功能的确认。此外，除了重量读数的准确性，还应包括时间戳的准确性与锁定受控。

对于本机电子存储的数据，数据从天平上产生后以电子方式留存下来，至最后使用和保管。由于数据以电子方式存储，若不加以管理，该电子数据有可能被“干预”（如，存储为Excel格式，则可能被人为改动后再打印输出作为放行报告一部分），不像实时打印到一个纸张固化下来。因此要保证数据的准确可靠，就必须满足电子数据的管理要求，详细内容可参考21 CFR Part 11 Sub Part B电子记录部分。例如：存储的数据是否为不可更改或特定平台才可访问的数据，倘若数据本身为可更改的方式，是否有措施保证更改过程可溯或其他安全的方式避免被更改，如审计跟踪、权限、数据锁定/加密等。

对于通讯/交互处理的数据，数据从天平上产生后以一种网络协议方式传输并存储记录于一个服务器、硬盘、虚拟服务器或云存储的某个空间，为了保证最终数据的准确，则需要对传输过程进行确认，即对产生、处理、存储该数据的计算机化系统进行验证。中国GMP附录《计算机化系统》第九条要求：“数据转换格式或迁移时，应当确认数据的数值及含义没有改变”。

因此，不同的数据类型，在保证数据准确性方面的做法不尽相同。

另外，我们需要从更广义的视角理解准确性：

例如：经过清洁验证之后，则有充足证据与说服力地证明在此经过验证的清洁方法之下的清洁操作，能够达到残留交叉污染水平可接受（清洁方法是“准”的）；经过工艺验证之后，则有充足证据与说服力地证明在此经过验证的工艺规程之下的工艺操作，能够始终如一的生产符合质量要求的产品（工艺操作是“准”的）；通过HVAC的确认，我们知道各个机组在什么样的参数状态下（如各风机的频率、回风温湿度设定等），能够达到送风区域的环境指标符合要求（风量、压差、温湿度等）（HVAC设定值是“准”的，有理论依据的）。也即是，通过确认验证收集证据证明既定方法或指标是准确可靠的。

例如：技术转移实现研发到生产（或A到B）的知识传递过渡的准确性，从而保证GMP生产“准确可靠”；仿制药一致性评价的数据来用以证明仿制药与原研药在疗效、安全性和质量的一致，通过一致性评价的药品被认为“准确可靠”；国家局新颁布的药品追溯体系建立，保证整个药品供应链下，最终患者用药的“准确可靠”。

最后，关于偏差、异常值、超标结果的调查，这是当前FDA、欧盟认证检查缺陷项目的“常客”；“偏差调查不能识别根本原因”、“出现重复性偏差”“无效OOS率过高”……屡见各个缺陷信中。无论如何，只有经过充分的、识别出根本原因的调查与处理纠偏，才能从根本上解决问题，才能真正做到“准确可靠”。