

附件

# 真实世界数据用于医疗器械临床评价技术 指导原则（试行）

本指导原则旨在初步规范和合理引导真实世界数据在医疗器械临床评价中的应用，为申请人使用医疗器械真实世界数据申报注册以及监管部门对该类临床数据的技术审评提供技术指导。本指导原则中提及的医疗器械包括体外诊断试剂。

本指导原则是供申请人和审查人员使用的技术指导文件，不作为法规强制执行，应在遵循相关法规的前提下使用本指导原则。真实世界数据和真实世界研究处于快速发展阶段，本指导原则基于现有认知水平制定，需根据科学发展不断完善和修订。

## 一、概述

### （一）真实世界数据与证据

本指导原则所述真实世界数据是指传统临床试验以外的，从多种来源收集的各种与患者健康状况和/或常规诊疗及保健有关的数据。

围绕相关科学问题，综合运用流行病学、生物统计学、循证医学等多学科方法技术，利用真实世界数据开展的研究统称为真实世界研究。真实世界研究通过系统性收集真实世界数据，运用合理的设计和分析方法，开展前瞻或回顾性研究。

真实世界证据指的是，通过分析真实世界数据，形成医疗器

械使用、风险/收益相关的临床证据，可能作为有效的科学证据用于监管决策。由于真实世界数据来源不同，数据质量可能存在较大差异，并非所有的真实世界数据都能产生有效的真实世界证据。

## （二）真实世界研究的优势与局限性

相比于传统临床试验，一般来说，真实世界研究在现实环境下开展，对纳入患者限定相对更少，样本量可能较大，更可能获得长期临床结局，研究结果的外推性可能较好。真实世界研究可使用多种数据，如医院病历数据、登记数据、医疗保险数据等。真实世界研究还可用于观察罕见严重不良事件，回答罕见疾病诊疗相关问题，评价临床结局在不同人群、不同医疗环境、不同使用方法之间的差异等。

真实世界研究的局限性包括但不限于，真实世界数据来源众多，数据质量有待评价；真实世界研究通常存在较多的偏倚和混杂（包括选择偏倚、信息偏倚、混杂等），研究结论可能存在挑战。

## 二、常见真实世界数据来源

常见的真实世界数据包括但不限于登记数据、医院病历数据、区域健康医疗数据、医疗保险数据、健康档案、公共监测数据、患者自报数据、移动设备产生的数据等。此外，真实世界数据还可包括在医疗器械生产、销售、运输、存储、安装、使用、维护、退市、处置等过程中产生的数据（如验收报告、维修报告、使用者反馈、使用环境、校准记录、运行日志、影像原始数据等）。

真实世界数据依其来源及特征，包括但不限于以下情形：

（一）产生于健康医疗服务的提供和付费过程，基于管理目的生成，如医院电子病历数据、医保数据、健康档案等。

（二）基于数据库建立时的研究目的，设立统一的数据标准和数据收集模式，在常规临床实践中形成并建立的数据资源，如器械登记数据等。

### 三、真实世界数据质量评价

良好的真实世界数据质量是开展真实世界研究的基础，直接影响真实世界研究生成的证据强度。真实世界数据质量评价，在遵循伦理原则，符合法规要求，保障数据安全的基础上，需关注数据的相关性和可靠性。数据的相关性，指的是数据是否可充分回答与研究目的相关的临床问题，包括数据是否涵盖研究人群数据，是否能形成相对统一或标化的干预/暴露，是否可设置可比的对照，是否包含研究所需的结局变量及测量结果，是否可获得混杂因素的相关数据。数据的可靠性，指的是数据采集的准确性，包括采集前确定采集范围、采集变量，制定数据词典、规定采集方法、采集数据的流转方式、储存介质格式等，充分保障数据的真实性和完整性等。评价真实世界数据质量，具体可从以下方面进行考虑：

#### （一）代表性

数据所包含的人群是否涵盖研究的目标人群。

#### （二）完整性

数据被收集和获取的程度，即相对于研究目的，数据是否完

整，如研究变量的缺失是否影响研究结局的评估，样本量及随访时间是否足以回答研究问题等。

### （三）准确性

数据对患者健康状况、诊疗及保健反映的准确程度，如患者年龄、使用器械、手术类型是否准确。准确性评价包括原始数据记录的准确性，数据采集的准确性（如是否建立规范统一的数据采集方法，是否核查不同来源数据的准确性等），以及数据治理的恰当性（如是否建立规范统一的数据治理流程，包括数据安全性处理、数据链接、数据清洗、数据编码、数据结构化、数据传输等，是否核查数据治理算法的正确性）。

### （四）真实性

医疗器械可被唯一标识以及唯一标识被记录的程度，以识别和分析该器械的全部使用过程。

### （五）一致性

数据采集遵循相同的过程和程序的程度，包括统一的数据定义和稳定的病例报告表或版本受控的其他数据收集表。

### （六）可重复性

变量可重复的程度。例如，对同一患者，结局变量测量和分类的一致性。

## 四、真实世界研究设计常见类型及统计分析方法

真实世界数据用于医疗器械临床评价时，应基于具体研究目的，进行策划和设计，遵循伦理原则，符合法规要求，保障数据安全。

研究策划包括明确研究问题，确定数据来源及收集方式，以及组建研究团队等。研究设计包括确定设计类型，明确研究对象和研究变量，识别混杂及偏倚的来源并制定相应措施进行合理控制，以及事先制定统计分析计划等。

### （一）真实世界研究设计常见类型

真实世界研究设计类型主要分为试验性研究和观察性研究。

#### 1. 试验性研究

实用性临床试验是在常规或接近常规的临床实践中开展的临床试验，实效性随机对照试验是实用性临床试验的一种重要类型。

实用性临床试验关注干预措施在常规临床实践中的效果，其研究对象是在常规临床实践中应用干预措施的患者群体，可能存在多种合并症；干预措施由于与常规临床实践保持较好一致，从而受干预者技能和经验的影响。因此，研究设计需基于其特点进行全面考虑。

实用性临床试验通常选用常规治疗、标准治疗或公认有效的治疗措施作为对照，观察指标通常选择对患者或研究结果的使用者具有重要临床意义的指标，根据研究目的不同，可包括安全性、有效性、治疗依从性、卫生经济等方面，因其注重评价远期结局，随访时间较长，随访频率通常与常规临床随访一致。

#### 2. 观察性研究

观察性研究包括队列研究、病例-对照研究、横断面研究、病例系列等设计类型。申请人可根据研究目的，选择恰当的研究

设计。由于观察性研究更可能出现偏倚及混杂，需预先进行全面识别，并采取有效的控制措施。

### 3.其他

在单组试验中，使用真实世界数据作为外部对照，是形成临床证据的一种特殊设计类型。外部对照需充分考虑试验组和对照组的可比性，如研究人群、临床实践、诊断标准、测量和分类等。

#### （二）统计学分析方法

在真实世界研究中，研究者需要根据研究目的、数据以及设计类型，选择合理的统计学方法，常见的统计分析方法见附录。

试验性研究的统计分析方法与传统临床试验相似，其统计分析计划包括数据集定义、分析原则与策略、缺失数据处理、分析指标与分析方法、亚组或分层分析、敏感性分析、补充分析和结果报告等。统计分析的基本原则亦为意向性治疗分析原则。观察性研究由于更容易产生偏倚和混杂，数据分析的关键是采用统计分析技术最大限度的控制混杂产生的偏倚，可用的分析技术除传统的分层分析、多变量分析外，还包括倾向性评分等。

## 五、可考虑将真实世界证据用于医疗器械临床评价的常见情形

基于真实世界数据形成的真实世界证据可支持医疗器械全生命周期临床评价，涵盖上市前临床评价及上市后临床评价。真实世界证据用于医疗器械临床评价的常见情形如下：

#### （一）在同品种临床评价路径中提供临床证据

同品种临床评价路径主要基于同品种医疗器械的临床数据

开展临床评价，需要的临床数据包括同品种产品的临床数据和/或申报产品的临床数据。

对于同品种产品的临床数据，真实世界数据是其重要来源，其有助于确认产品在常规临床实践中的安全有效性；识别产品的潜在风险（如罕见的严重不良事件）；甚至通过获知同类产品在不同人群中的实际疗效，明确最佳使用人群；通过知晓同类产品的行业水平，为申报产品的上市前风险/收益评价提供信息。申报产品合法使用获得的真实世界数据，可用于确认申报产品与同品种器械间的差异，不对申报产品的安全性有效性产生不利影响。

### （二）用于支持产品注册，作为已有证据的补充

由于全球法规尚待进一步协调以及产品上市策略等因素影响，部分医疗器械尚未实现全球同步上市。注册申请人可综合考虑产品设计特点及适用范围，已有的临床证据，各监管国家或地区对于临床证据要求的差异等情况，在已上市国家或地区收集真实世界数据并形成真实世界证据，作为已有临床证据的补充，支持在中国的注册申报，可避免在原有临床证据不足时在中国境内开展临床试验。

### （三）临床急需进口器械在国内特许使用中产生的真实世界数据，可用于支持产品注册，作为已有证据的补充

根据国家相关规定，在部分区域指定医疗机构内，特许使用的临床急需进口医疗器械，按照相关管理制度和临床技术规范使用产生的真实世界数据，经过严格的数据采集和系统处理、科学

的统计分析以及多维度的结果评价，可用于支持产品注册，作为已有证据的补充。特别是通过境外临床试验进行临床评价，有证据表明/提示将境外临床试验数据外推至中国人群可能受到境内外差异的影响时，可考虑使用该类数据作为支持。

#### （四）作为单组试验的外部对照

在单组临床试验设计中，可从质量可控的真实世界数据库中提取与试验组具有可比性的病例及其临床数据，作为外部对照。外部对照通常来源于具有良好质量管理体系的登记数据库，其可接受申办者和监管方等的评估，以确认其数据的相关性和可靠性。建议采用同期外部对照，如使用历史数据进行对照，将因为时间差异引入多种偏倚，降低临床试验的证据强度。

#### （五）为单组目标值的构建提供临床数据

目标值是专业领域内公认的某类医疗器械有效性/安全性评价指标所应达到的最低标准，包括客观性能标准和性能目标，是在既往临床数据的基础上分析得出，用于试验器械主要评价指标的比较和评价。真实世界数据可作为构建或更新目标值的数据来源。

#### （六）支持适用范围、适应症、禁忌症的修改

医疗器械上市后，基于所在国家或地区的相关法规，在合法使用的前提下，获得的真实世界数据可用于支持适用范围、适应症及禁忌症的修改。可能的情形包括发现额外的疗效、潜在的获益人群、慎用人群、产品远期安全性确认等。

#### （七）支持在说明书中修改产品的临床价值

医疗器械上市后的真实世界证据，可用于支持修改说明书中



修改产品的临床价值。例如，对于测量、计算患者生理参数和功能指标的医疗器械，部分生理参数和功能指标在上市前评价时主要关注测量和计算的准确性，未充分发掘其临床价值。真实世界数据可用于构建生理参数和功能指标，或者基于其做出的临床治疗决定与临床结局之间的因果推断，从而修改说明书中产品的临床价值。

#### （八）支持附带条件批准产品的上市后研究

对用于治疗罕见病、严重危及生命且尚无有效治疗手段的疾病和应对公共卫生事件等急需的医疗器械，附带条件批准上市后，可利用真实世界数据开展上市后研究，以支持注册证载明事项的完成。

#### （九）用于高风险植入物等医疗器械的远期安全性和/或有效性评估

高风险植入物等医疗器械，特别是市场上首次出现的高风险植入物，在上市前临床评价中，难以确认产品的远期疗效和风险，识别罕见严重不良事件。可利用真实世界数据进行该类产品的上市后研究，评估产品的远期安全性和/或有效性，完成产品的全生命周期临床评价。

#### （十）用于治疗罕见病的医疗器械全生命周期临床评价，加快其上市进程，满足患者需求

真实世界数据可在多维度支持治疗罕见病的医疗器械快速上市。如拟开展上市前临床试验，真实世界数据可作为单组试验的外部对照，或者用于构建目标值；附带条件批准后，真实世界

数据可用于确认产品的有效性，识别产品风险，进行产品风险/收益的再评价。

#### （十一）上市后监测

产品的上市后监测，涉及不良事件监测、产品安全有效性再评价等方面，是医疗器械全生命周期临床评价的重要组成部分。真实世界数据在上市后监测中应当发挥重要作用，如通过收集、提取风险信号，开展不良事件归因分析，及时发现和控制已上市医疗器械的使用风险，同时促进生产企业对已上市产品的设计改进，推动新产品研发。

附：医疗器械真实世界研究常见统计分析方法

附

## 医疗器械真实世界研究常见统计分析方法

### 一、实效性随机对照试验的统计方法

与传统随机对照临床试验相比，实效性随机对照试验（以下简称 pRCT）在现实医疗环境中开展，患者个体差异可能较大，接受干预的标准化程度可能降低，患者依从性可能较差，临床专业人员的医疗技术可能存在不同，研究失访可能增加。pRCT 的统计分析需遵循事先制定的统计分析方案，考虑因素包括但不限于以下情形：

（一）意向性分析是常用的统计分析方法，需重视对患者失访的处理，预先明确失访患者的处理办法并说明原因；

（二）pRCT 的研究人群、临床环境等存在较大异质性，研究结果检验效能可能较低，应谨慎使用非劣效设计。

（三）pRCT 在随机后可能仍会出现混杂，例如患者接受的干预发生变化，不同组别患者的依从性不同等。研究者需根据研究问题与研究假设，采取适当统计方法调整随机后混杂的影响。

（四）pRCT 如果来自于多个中心，需要对中心效应进行控制，当主要结局变量是连续性指标时，可采用协方差分析方法；当主要结局变量是分类指标时，可采用考虑 Cochran-Mantel-Haenszel 方法。当除中心效应外还有其他协变量需要考虑时，可采用随机效应模型。

(五) 在 pRCT 统计分析中, 建议重视敏感性分析, 以评估统计推断的稳健性。

## 二、观察性研究常用的统计分析方法

在观察性研究中, 数据分析的关键是采用统计分析技术最大限度的控制混杂因素造成的偏倚。可用的分析方法包括但不限于以下情形:

### (一) 分层分析

分层分析是指将数据按可能的混杂因素分为多层, 每层的内部数据有较好的同质性, 是常用的识别和控制混杂造成的偏倚的方法之一。Mantel- Haenszel 法是常用的分层分析方法, 来评估混杂因素对结果的影响。该分析可判断外来因素是混杂还是效应修饰作用, 或以哪种作用为主, 以及确定混杂的大小和方向或效应修饰的大小。但是分层分析只能控制少数混杂因素, 若混杂因素数过多可能导致过度分层, 使层内样本量少; 对连续性变量只能用等级分层法, 常引起不合理的分组。

### (二) 多变量回归模型

多变量回归模型是最常见的控制混杂因素的统计分析方法, 根据结局变量的特点选择 logistic 回归、线性回归、Poisson 回归和 Cox 比例风险回归等。值得注意的是, 大多数回归模型用于估计相对效应值。针对存在层次结构的数据可考虑多水平模型, 针对存在重复测量的数据可考虑广义线性混合效应模型和广义估计方程。但在应用这些模型的时候, 仍需考虑其模型的模型假设以及模型适用性。

### （三）倾向性评分分析方法

倾向性评分分析是目前观察性研究中因果推断常用的分析方法，是一种针对较多混杂因素的调整方法，尤其适用于暴露因素常见而结局事件罕见的研究，或者有多个结局变量的研究。常见的倾向性评分应用方法包括倾向性评分匹配法，倾向性评分分层法，逆概率加权法，以及将倾向性评分作为唯一协变量纳入统计模型进行调整分析的方法。其中，倾向性评分的匹配和分层法在医疗器械临床评价的真实世界研究中已有较为成熟的应用。

值得注意的是，若使用倾向性评分方法，应首先在统计分析计划中，预先指明用于建立倾向评分模型的变量，以及对模型拟合优度和预测效果进行判断的标准；在对基线指标建立倾向评分模型时，应保持对结局指标的“盲态”，避免根据结局指标重新调整倾向评分模型，从而获得“理想”或“预期”结果的情况。

应用倾向性评分进行效应估计时，需判断倾向性评分接近的患者在不同处理组间的协变量分布是否均衡，报告使用倾向性评分之前和之后的结果，考虑倾向性评分处理后可能对研究结果造成的影响，例如，用倾向性评分匹配后可能导致的估计精度降低（因样本量下降）；或使用倾向性评分加权时，个别极大权重的样本可能对分析结果造成较大影响等。

倾向性评分方法仅能处理可观测到的混杂因素，不能控制研究中未采集的混杂因素可能带来的潜在影响，建议研究中针对评价结果进行合理的解读和讨论，并开展可能的定量分析。