

ICS 11.040.99

CCS C30

中华人民共和国医药行业标准

YY/T XXXXX—XXXX

采用脑机接口技术的医疗器械 用于人工智能算法的脑电数据集质量要求与评价方法

Medical device adopting BCI technology—Quality requirement and evaluation
methods of EEG datasets for artificial intelligence

征求意见稿

XXXX—XX—XX 发布

XXXX—XX—XX 实施

发 布

前 言

本文件按照GB/T 1.1—2020《标准化工作导则 第1部分：标准化文件的结构和起草规则》的规定起草。

请注意本文件的某些内容可能涉及专利。本文件的发布机构不承担识别专利的责任。

本文件由国家药品监督管理局提出。

本文件由人工智能医疗器械标准化技术归口单位归口。

本文件起草单位：

本文件主要起草人：

引 言

本文件根据当前技术现状，提出了脑电数据集的专用质量要求和评价方法。由于脑机接口医疗器械还处于快速发展阶段，充分考虑适用性和风险后提出的新的质量要求和方法不受本文件制约。本文件第4章-第7章围绕脑电数据集的技术要素分别提出了质量要求。第8章提出了质量符合性评价方法。

采用脑机接口技术的医疗器械 用于人工智能算法的脑电数据集

质量要求与评价方法

1 范围

本文件规定了采用脑机接口技术的人工智能医疗器械专用脑电数据集的质量要求和试验方法。

本文件适用于由人工智能算法对脑电信号进行后处理、采用脑机接口技术的医疗器械研发、生产、质控等环节使用的数据集。

2 规范性引用文件

下列文件中的内容通过文中的规范性引用而构成本文件必不可少的条款。其中，注日期的引用文件，仅该日期对应的版本适用于本文件；不注日期的引用文件，其最新版本（包括所有的修改单）适用于本文件。

GB 9706.226-2021 医用电气设备 第2-26部分：脑电图机的基本安全和基本性能专用要求

XX/T XXXX.1-XXXX 人工智能医疗器械 质量要求和评价 第1部分：术语

YY/T 1833.2-2022 人工智能医疗器械 质量要求和评价 第2部分：数据集通用要求

XX/T XXXX-XXXX 采用脑机接口技术的医疗器械 术语

3 术语和定义

GB 9706.226-2021、XX/T XXXX.1-XXXX、YY/T 1833.2-2022、XX/X XXXX-XXXX界定的以及下列术语和定义适用于本文件。

3.1

标签 label

在脑电数据中反映特定行为、事件与状态的标记。

注1：行为（Behavior）：指受试者主动产生的、有意图或任务反应性的动作或认知活动，如按键、言语、选择等。

注2：事件（Event）：指在数据采集过程中被动产生的、由系统或环境引发的客观现象，可能包括实验程序中的刺激呈现，也包括如肌肉收缩、电生理反射等非意图性生理事件。

注3：状态（State）：指受试者在数据记录期间持续处于的一种心理或生理状态，如清醒、注意、疲劳、病理状态等。

3.2

标注 annotate

为脑电数据添加标签的行为。

3.3

交互 interaction

通过测量、解码用户的脑电活动，以实现用户对计算机软件或其他外部设备的控制或信息交流的过程。

3.4

电极位置规范 electrode position specification

确保在神经电生理学研究中电极的放置具有一致性和可重复性的标准化方法。

3.5

神经电生理 neuroelectrophysiology / neuroelectrophysiological

通过放置在头皮上的电极记录的大脑皮层神经元群电活动的总和，反映了大脑皮层神经元同步放电的宏观电活动，通常需要大量神经元的同步活动才能在头皮上产生可检测的信号。

3.6

区组设计(block design)

将相同或相似类型的刺激或任务集中在一起，然后将这些区组以随机或伪随机的方式呈现给被试者的设计方法。

3.7

逐试次设计(trial-by-trial design)

将每个独立的刺激呈现或任务执行视为一个单独的“试次”，并在每个试次中独立地记录和分析被试的脑活动和行为反应的设计方法。

3.8

共平均参考(common average reference, CAR)

一种脑电信号重参考方法。通过计算所有或部分头皮电极的平均电位，并将该平均电位作为参考值，从每个通道的原始信号中减去，以减少共模噪声。

3.9

参考电极标准化技术(Reference Electrode Standardization Technique, REST)

通过数学方法将脑电数据重参考至理论无限远处，从而近似实现一个理想中性参考点，从而提高脑电图数据分析的准确性和可比性的重参考方法。

4 总则

数据集应包含数据集说明文档，元数据和数据文件三个部分。

数据集说明文档用于介绍数据集内容，数据采集实验范式和流程，重要参数及参数内容，给出数据内容、文件结构、实验的整体设计、数据流情况以及元数据定义。

元数据描述脑电数据集数据的组织、数据域及其关系，建立数据内容和数据值之间的配对关系。在脑机接口应用中元数据用于结构化描述神经信号采集、处理、分析过程等信息。

数据文件记录相关数据完整内容单独存储的数据文件。为脑电数据的解析、复现和使用提供足够的信息。数据集中脑电数据的存储方式宜为连续数据，即在实验任务中硬件设备直接采集得到的整段脑电信号。

注1：本文件各章节内容与数据集说明文档，元数据和数据文件的对应关系见附录A。

注2：元数据符号、数据形式举例见附录B。

5 数据采集要求

5.1 电极信息

5.1.1 电极类型信息

元数据应提供数据采集使用的电极类型信息。

注：根据电极与人体组织的接触方式和放置位置，将电极分为：

- a) 非侵入式电极；
- b) 侵入式电极。侵入式通常包括半侵入式、介入式及植入式等形式。

5.1.2 电极结构信息

元数据应提供数据采集使用的电极结构信息，包括：

- a) 电极整体信息

使用电极的几何形状、尺寸信息（包括长度、直径和厚度）信息。

- b) 电极触点信息

各个电极触点（即电极与神经组织接触部分）的形状、数量、尺寸和电极触点之间的物理距离等信息。

- c) 电极材料信息

用于制造电极的关键材料，包括电极导电层、涂层、绝缘层的材料类型。

- d) 电极阻抗信息

出厂时电极阻抗的测试方法与测试结果信息。

5.1.3 电极安全性信息

数据集说明文档应记录与电极安全性相关的信息，包括：

- a) 注册信息

对于已注册上市的电极，应给出注册证信息。

- b) 电极安全性试验结果

对于未进行正式注册的电极，应对电极安全性试验的结果进行描述。

5.2 电极采集条件

5.2.1 电极布局信息

数据集说明文档应记录所覆盖的目标脑区空间信息、间距、密度、对应的神经元群体或单个神经元。

非侵入式电极宜详细描述电极的类型、数量、放置位置、遵循的电极位置规范、采集时的参考与接地方案与对应电极位置；

侵入式电极宜按照立体定向脑电图（sEEG）电极、脑皮层电图（ECoG）电极、局部场电位（LFP）电极和锋电位（Spike）电极有所区分；立体定向脑电图电极宜考虑立体定向脑电图技术规范，明确目标脑区，间距；脑皮层电图电极宜按描述网格或条带形式，结合临床需求明确密度。局部场电位和锋电位的电极阵列或深部电极植入宜记录所针对的特定神经元群体或单个神经元。

5.2.2 接触阻抗信息

对于非侵入式电极，数据集说明文档应提供使用前后、采集过程中（如适用）所用电极的接触阻抗、测量接触阻抗所用的电流频率和峰峰值。

5.2.3 电极植入信息

对于侵入式电极，元数据应提供电极植入的相关信息。如电极植入位置（含立体定向坐标、接触深度）、手术方法（手术类型、导航系统）、植入时长（在体电极信号持续时间及信号采集窗口）、术后阻抗测试方法及结果。

5.2.4 信号类型信息

数据集说明文档应记录数据集所采集信号的种类及其来源，包括但不限于如下类型：

- 脑电图（EEG）；
- 脑皮层电图（ECoG）；
- 局部场电位（LFP）；
- 峰电位信号（Spike）。

5.2.5 背景脑电信息

对于非侵入式脑机接口，数据集说明文档应包含测试前后被试者处于睁眼、闭眼和眨眼状态时采集的原始背景脑电数据。

对于侵入式脑机接口，数据集说明文档应采集植入后24小时内、3—7天及移除前的背景脑电（每次建议 ≥ 30 分钟，含静息态），并标注记录阶段、患者状态及环境参数（如屏蔽措施、设备状态）。

5.2.6 系统噪声

脑机接口系统的噪声有效值应满足数据集制造责任方的宣称。

元数据应提供脑机接口系统的噪声数据和声称噪声有效值指标。

5.2.7 采样率

对于脑电图、立体定向脑电图、脑皮层电图和局部场电位信号的采集，采样率应不低于目标信号最高频率的两倍；对于尖峰信号的采集，采样率应不低于15kHz。

元数据应记录数据采集使用的采样率。

5.2.8 信号分辨率

元数据应记录采集的信号分辨率信息。

5.2.9 滤波器设置

元数据应记录数据采集使用的滤波器类型（低通/高通/带通/陷波）、截止频率、滚降特性（dB/oct）、阶数、带宽及相移处理方式（如零相位滤波），并注明设计方法（如巴特沃斯）及滤波实现方式（硬件/软件）。

5.2.10 尖峰检测

如适用，元数据应记录单个通道信号处理所使用的尖峰检测设定值，并要说明滤波设置、阈值大小、以及窗大小。

5.2.11 尖峰排序

如适用，元数据应记录尖峰检测、特征提取方法、分类算法的详细说明与关键参数，过程如涉及人工校对应描述人工校对方法。

5.3 数据传输条件

5.3.1 传输稳定性信息

数据集说明文档宜记录数据传输过程中的稳定性指标，并保存数据校验信息。如丢包率，延迟的方差分布，持续稳定工作时间等。

5.3.2 传输实时性信息

数据集说明文档应记录系统的数据传输方案、系统延迟等参数。

5.3.3 传输同步性信息

数据集说明文档应记录数据传输过程中多设备/信号之间时钟是否对齐；如果是，应记录数据同步机制、同步精度等信息。

5.3.4 传输速率

数据集说明文档应记录数据传输速率。对于多通道数据传输，数据集的传输速率应与其临床功能相匹配。

5.4 实验条件

5.4.1 实验范式类型

数据集说明文档应记录实验范式的类型，脑机接口的实验范式包括但不限于分为以下几种类型：

- 1) 主动类型范式（需要用户主动产生特定脑电信号来控制设备的系统）
 - 运动想象范式；
 - 注意力调节范式；
 - 情绪调节范式；
 - 认知任务；
 - 运动任务；
 - 语言任务。
- 2) 被动类型范式（通过监测用户的自然脑活动来执行操作，用户不需要主动产生特定的脑电信号）
 - 状态监测（注意力、情绪、疲劳等）；
 - 脑疾病诊断。
- 3) 反应类型范式（依赖于用户对外部刺激的反应来引发特定的脑电信号，用户间接调节信号）
 - 视觉诱发范式；
 - 听觉诱发范式；
 - 触觉诱发范式；
 - 其他感觉诱发范式；

——反应任务。

4) 混合类型范式（同时含有主动、被动、反应式范式中两种或两种以上的脑机接口实验范式）

——闭环反馈任务

5.4.2 实验范式设计

5.4.2.1 任务要求

元数据应记录实验中的具体要求。

注：如动作的时间限制或空间限制、采集环境（温湿度、光照）、屏蔽情况等信息。

5.4.2.2 条件设置

元数据应记录实验中的条件集合。

注1：如按照视觉刺激的对比度分为5组或按照听觉刺激的强度分为10组等。

注2：5.4.2.2~5.4.2.4中的条件指的是根据研究目的和一定规则穷举实验相关参数的不同组合或模式。

5.4.2.3 条件顺序

元数据应明确实验为区组设计或逐试次设计，应明确实验是否伪随机，是否盲法。

5.4.2.4 试次数设计

元数据应说明范式中各条件、区组的试次数量，实验天数设计，重复次数设计的依据。

5.4.3 实验刺激呈现

5.4.3.1 刺激信号

数据集说明文档应记录实验使用的刺激类型、刺激格式、刺激参数和刺激时间。

——刺激类型，如视觉、听觉、触觉、力学、电刺激等；

——刺激格式，如图像、声音、视频、文字、力场等；

——刺激参数，如视觉刺激的颜色、形状、尺寸、对比度；电刺激的强度和频率等；

——刺激时间，如出现和持续的时间、间隔时间等信息。

5.4.3.2 引导信号

数据集说明文档应描述实验前和实验中的引导信号。

实验前引导：实验开始前实验对象获得的引导信息，如任务说明和注意事项等。

实验中引导：实验中各步骤的引导信号，如引导语或引导标识。

5.4.4 行为信息记录

5.4.4.1 一般要求

数据集说明文档应记录实验过程中采集对象执行的具体任务、实验流程对应的时序信息以及实验过程中出现的其他情况。

5.4.4.2 交互方式

数据集说明文档应记录实验对象与实验系统的交互方式，如点接触摸屏、移动光标、眼动跟踪等。

5.4.4.3 行为评估

数据集说明文档应记录对实验行为结果的分类（正确/错误）和判断标准（阈值或选择）。

5.4.4.4 行为反馈

数据集说明文档应记录每个试次结束后的反馈信息或奖励（reward）及实验对象是否获得总结性反馈。

5.4.5 受试者特征记录

数据集说明文档应记录受试者性别、年龄、相关个体特征（如受试者的左/右利手信息、视力情况、行动能力和有关病史等）以及是否具有BCI经验等。

6 数据处理和标注要求

6.1 数据去标识化

数据集说明文档应描述用于保护受试者隐私的技术手段，如数据去标识化等。适当时，数据集说明文档应描述数据去标识化的规则，去除或加密所有能够直接或间接识别个人的字段，如姓名、出生日期、地址等。

6.2 数据筛选

数据集说明文档应描述筛选数据时采用的条件，包括入组条件和排除条件，以及使用的筛选方法。

注：数据集制造责任方声称数据筛选的依据和指标，如信噪比、伪影分布、个体信息完整性等。

6.3 数据清洗

数据集说明文档应描述使用对数据进行清洗的方法和要求（如根据后续算法和数据处理需要选择排除的坏导、坏段等），并对清洗方法、要求进行记录。

6.4 数据标签

数据集说明文档应描述标签内容。标签内容由标签值、标签说明、标签时间（序号）组成，反映脑电数据在特定时刻或时间段内的关键行为、事件或状态。其中，标签值用于标定数据记录过程中出现的不同行为、事件或状态，其数据类型可以为数字、字母、符号、文本。标签说明用于解释标签值对应的行为、事件或状态的具体情况，其数据类型可以为文本。标签时间用于标定不同关键行为、事件与状态出现的时刻或时间段，其数据类型可以为数字、时间、文本。

示例：推荐在数据采集记录过程中使用数字或字母的数据类型来记录标签，使用对应采样点索引的数字来记录标签时刻。

表 1 包含两种行为的数据标签示例

标签值	标签说明	标签时间（序号）
1	对应行为为：受试者按下按钮1	00246
2	对应事件为：受试者按下按钮2	022196
1	对应事件为：受试者按下按钮1	04231
2	对应事件为：受试者按下按钮2	06249
...	...	...

表 2 包含两种事件的数据标签示例

标签值	标签说明	标签时间（序号）
1	对应事件为：由10Hz正弦波调制的视觉刺激开始闪烁	00216
2	对应事件为：由15Hz正弦波调制的视觉刺激开始闪烁	02216
1	对应事件为：由10Hz正弦波调制的视觉刺激开始闪烁	04216
2	对应事件为：由15Hz正弦波调制的视觉刺激开始闪烁	06216
...	...	...

表 3 包含两种状态的数据标签示例

标签值	标签说明	标签时间（序号）
1	受试者处于清醒状态	00100-89100
2	受试者处于疲劳状态	90449-145200

注1：推荐在数据采集记录过程中使用数字数据类型来记录标签，采用ISO 8601时间戳。

注2：标签时刻应精确对应于行为、事件与状态产生或存在的时间点。

6.5 数据标注

6.5.1 一般要求

如果数据集需要对数据进行标注，应添加相关的标签，并应在数据集说明文档中描述标注对象、标注粒度、标注依据、方法和工具、参与人员与复核机制等信息。

6.5.2 标注对象

标注的对象应为与数据集主题相关的行为、事件或状态，不应出现与数据集主题无关的标签。

注：标注对象的举例见表2。

表 2 标注对象的示例

实验范式类型 标注对象	行为	事件	状态
主动类型范式	用户执行运动想象动作（如想象手部运动）、主动调节注意力（如专注特定目标）、尝试调节情绪（如努力唤起积极情绪）、开展认知任务（如记忆任务操作）、进行运动任务（如实际简单肢体动作）、执行语言任务（如默念词语）	运动想象任务开始/结束、注意力调节任务启动、情绪调节指令发出、认知任务阶段切换、运动任务动作起始、语言任务语句输入	运动想象时大脑激活状态、注意力高度集中状态、情绪调节中的情绪倾向状态、认知任务中的专注度状态、运动任务执行中的运动意图强度、语言任务里的语言表征激活状态

被动式脑机接口范式	无（因是监测自然脑活动，用户无主动控制行为）	注意力监测周期开始、情绪监测数据采集、疲劳监测启动、脑疾病诊断扫描开始	注意力水平状态（如注意力分散/集中）、情绪状态（如焦虑/平静）、疲劳状态（如疲劳/清醒）、脑疾病相关病理状态（如异常脑电特征呈现）
反应式脑机接口范式	用户对视觉刺激（如看到闪烁图案后产生脑电反应）、听觉刺激（如听到特定声音后有脑信号变化）、触觉刺激（如感受到触碰后引发脑反应）、其他感觉刺激（如闻到气味后产生脑活动）做出反应，执行反应任务操作（如对刺激序列做判断回应）	视觉刺激呈现事件、听觉刺激播放事件、触觉刺激施加事件、其他感觉刺激给予事件、反应任务指令发布	对视觉刺激的脑电响应状态（如P300波是否出现）、对听觉刺激的脑信号反馈状态、对触觉刺激的神经反应强度状态、完成反应任务时的脑活动同步状态
混合式脑机接口范式	用户主动产生脑信号（如结合运动想象）同时依据系统被动监测反馈调整行为（如根据疲劳反馈调整任务强度），执行闭环反馈任务中的主动交互动作（如根据实时脑状态调整控制指令）	主动交互阶段开始、系统反馈触发、闭环反馈任务循环节点	主动脑信号产生强度与准确性状态、结合被动反馈后的任务调整状态、闭环反馈中脑机协同工作的适配状态

6.5.3 标注粒度

标注的粒度应根据具体的标注对象决定。标注粒度可以为时刻级标注、时间段级标注或全数据集级标注。

- 注1：时刻级标注 用于标记在某一特定时间点上发生的行为或事件。典型用途包括：刺激出现、按键、眨眼、电刺激发生等。
- 注2：时间段级标注 用于标记在一段连续时间内持续存在的状态或特征。典型用途包括：疲劳状态、闭眼状态、癫痫发作间期等。
- 注3：全数据集级标注 用于标记整个数据集（或数据集中某个文件）仅包含单一类型的状态或事件，无需进一步细化标注粒度。典型用途包括：特定病理期记录（如整段均为癫痫发作或清醒期），或任务唯一、时序无关的数据。

6.5.4 标注依据、方法和工具

应明确数据集中标签标注的依据，例如：

- 实验设计与事件触发信号；
- 受试者行为触发；
- 专家判读与标注。

应明确数据集标注的方法和执行要求，例如：

辅助标注：使用软件或算法工具根据触发器（包含手动触发）、刺激日志等条件辅助产生标签，然后由人工复核与修改。

手动标注：人工对原始数据进行逐条判读与标注。

执行标注任务时，应建立质量验证机制并形成记录。

6.5.5 参与人员与复核机制

应记录参与人员的基本信息（如编号或身份），并建立标注结果的复核机制，建议至少由两人进行交叉核对。若存在争议，需记录争议及最终确认过程。

6.6 数据对齐处理

如适用，数据文件应保证不同模态间数据时间上的一致性，应将神经记录与刺激输出、行为测量或生理指标精确对齐，尤其是在闭环系统中。

6.7 数据分段

如适用，数据集说明文档应描述数据分段时数据分段的含义、分段使用的方法与参数，分段后的数据应按照标准化格式进行存储。

6.8 重参考

应根据实际情况决定是否进行重参考。

注：重参考方法如共平均参考、参考电极标准化技术、拉普拉斯滤波等。

7 数据存储与访问

7.1 存储管理

数据集制造责任方应对数据集存储文件进行管理和维护，包括备份、更新和优化等方面。

7.2 数据安全

数据存储和传输过程应采取必要的安全措施，确保数据的完整性、机密性和可用性，防止数据泄露、丢失或篡改。

7.3 数据压缩

宜使用无损压缩方式进行数据压缩。如对数据进行了压缩，数据集说明文档应说明压缩和解压缩方法与工具。

7.4 数据存储信息

数据集说明文档应描述数据存储的内容、存储方式与存储路径等信息，原始数据文件应当保留、处理后的数据文件应当说明处理方法与参数。

7.5 数据命名规范

数据集说明文档应提供数据命名的规则、字段代表的具体信息。

7.6 数据备份

应规划数据的备份方式并定期对数据进行备份，准备数据恢复的流程方法，保护数据不受意外删除、丢失、损坏或其他灾害影响。

7.7 数据更新和版本控制

数据集在每次更新时，应记录数据集的更新信息，包括更新时间、更新的内容和管理方式。每次版本更新应同步更新相关数据集说明文档和使用示例，确保数据集在不同版本间的可追溯性和一致性。

7.8 访问控制权限

应根据使用者的角色和需求设置不同的访问权限。数据集应拒绝任何未经授权访问请求。

7.9 访问可追溯性

应记录数据访问操作信息，包含访问者信息、访问时间、访问终端、访问过程中进行的操作等。

8 质量符合性测试方法

8.1 总则

以数据集为统计单元，对数据序列进行均匀随机抽样检验，抽样比例不低于10%；对脑电数据进行解析，对数据集说明文档、元数据字段进行检查，对数据集包含的其他数据进行预览，对抽样样本的数据标签进行检查，判断是否满足第4章的要求。

8.2 数据采集要求

8.2.1 电极信息

通过审查数据集制造责任方提供的元数据文件、数据集说明文档确认符合性。

8.2.2 电极采集条件

通过审查数据集制造责任方提供的元数据文件、数据集说明文档确认符合性。其中，

8.2.2.1 系统噪声信息

通过以下方法确认符合性。

a) 工频噪声

提取基波（50/60 Hz）和谐波频域的峰值幅值或计算其功率谱密度（PSD）。

b) 基底噪声

取半采样频率附近区间内的信号进行基底噪声测试(应尽量避开工频噪声干扰和信号源频率范围)，计算该区间范围内经过带通滤波的信号RMS值作为量化指标。

8.2.3 数据传输条件

通过审查数据集制造责任方提供的元数据文件、数据集说明文档确认符合性。

8.2.4 实验条件

通过审查数据集制造责任方提供的元数据文件、数据集说明文档确认符合性。

8.3 数据处理与标注要求

8.3.1 总体符合性测试方法

检查脑电数据、元数据的字段设置和数据集制造责任方的脱敏记录，查看脱敏措施是否已生效，判断是否满足第6章的要求。各子条款的测试细节描述如下

8.3.2 标注粒度

8.3.2.1 时刻级标注

从数据集中提取时刻级标注信息(如刺激出现、按键等行为/事件的标注时刻，记为 $T_{\text{标注}}$)，同时查找数据采集时同步记录的系统日志里实际行为/事件发生的参考时刻(若数据采集系统有记录，记为 $T_{\text{实际参考}}$)。

计算时间差 $\Delta T = |T_{\text{标注}} - T_{\text{实际参考}}|$ ，对多个不同的时刻级标注事件，进行均匀随机抽样检验，抽样比例不低于10%，统计 ΔT 的最大值、最小值和平均值，判断是否满足预设的误差要求。

8.3.2.2 时间段级标注

对时间段级标注数据进行均匀随机抽样检验，抽样比例不低于10%，并提取其时间段级标注信息，包括起始时刻（ $T_{\text{起标注}}$ ）、结束时刻（ $T_{\text{止标注}}$ ）及对应的状态描述（如疲劳状态、闭眼状态等）。结合数据集中脑电数据的特征（如闭眼状态时脑电的alpha波特征、疲劳状态下的脑电节律变化等），得到 $T_{\text{起特征}}$ 和 $T_{\text{止特征}}$ 或人工判读（有经验的脑电分析师查看脑电数据）的方式，确定实际状态对应的大致时间区间。

计算标注起始时刻与特征起始时刻的差值 $\Delta T_{\text{起}} = |T_{\text{起标注}} - T_{\text{起特征}}|$ 、标注结束时刻与特征结束时刻的差值 $\Delta T_{\text{止}} = |T_{\text{止标注}} - T_{\text{止特征}}|$ ，以及标注时长（ $D_{\text{标注}} = T_{\text{止标注}} - T_{\text{起标注}}$ ）与特征分析得到的实际时长（ $D_{\text{实际}} = T_{\text{止特征}} - T_{\text{起特征}}$ ）的偏差率 $\eta = \left| \frac{D_{\text{标注}} - D_{\text{实际}}}{D_{\text{实际}}} \right| \times 100\%$ 。

对抽样的标注数据进行统计，判断标注的时间段是否能合理反映状态持续过程。

8.3.2.3 全数据集级标注

查看数据集的标注说明和数据内容，确认全数据集级标注的数据集（或数据集中某个文件）对应的标注描述，检查整个数据集（或数据集中某个文件）内的脑电数据是否在特征上趋于一致，符合标注的单一类型状态/事件。

8.3.3 数据对齐处理

对于不同采集系统采集的数据，可根据事件标记进行时间同步对齐计算，具体计算方式为：

- （1）选择需要计算的不同数据里共有的事件标记序列，如试次开始、结束、关键点等事件标记。
- （2）通过作差计算不同数据中每个事件标记对应的时间差，得到时间误差序列。
- （3）求时间误差序列的均值，基于误差均值对齐数据，误差序列均值和方差应满足任务需求。

8.4 数据存储与访问检查

对数据存储管理规程、数据安全措施确认记录、数据存储信息和日志、数据命名规范、数据备份记录 and 策略、数据更新控制规程和记录等文档进行检查；对数据压缩工具、访问权限管理工具进行实际操作，判断是否满足第7章的要求。

附录 A 数据集内容

(规范性)

本标准条款与数据集说明文档、元数据、数据文件对应关系如下。

数据内容	存储方案	备注/解释
5.1.1 电极类型信息	元数据	无
5.1.2 电极结构信息	元数据	复杂的电极设计信息可以通过图片辅助在数据集说明文档中介绍，或者使用单独的电极设计技术文件进行保存
5.1.3 电极安全性信息	数据集说明文档	无
5.2.1 电极布局信息	数据集说明文档	通过文字、影像图片、三维建模结果等方式辅助介绍
5.2.2 接触阻抗信息	数据文件	无
5.2.3 电极植入信息	元数据	无
5.2.4 信号类型信息	数据集说明文档	无
5.2.5 背景脑电信息	数据文件	无
5.2.6 系统噪声信息	元数据	对市面上常用设备，披露设备型号即可，设备介绍中一般会包含相关信息。对于自研设备，相关应在元数据中记录
5.2.7 采样率信息	元数据	无
5.2.8 信号分辨率	元数据	对市面上常用设备，披露设备型号即可，设备介绍中一般会包含相关信息。对于自研设备，相关应在元数据中记录
5.2.10 尖峰检测	元数据	无
5.2.11 尖峰排序	元数据	无
5.3 数据传输条件	数据集说明文档	无需严格要求在数据集中体现，如出现丢包、稳定性等问题，应在数据集说明文档中进行说明
5.4.1 实验范式类型	数据集说明文档	无
5.4.2 实验范式设计	元数据	范式中各条件、区组的试次数量等信息建议在元数据中体现，具体的任务要求、设计等在数据集说明文档中体现
5.4.3 实验刺激呈现	数据集说明文档	无
5.4.4 行为信息记录	数据集说明文档	实验的名称、编号、实验日期、实验起止时间、实验人员信息、受试者编号及受试者相关信息应在元数据中体现
6.1 数据脱敏处理	数据集说明文档	无
6.2 数据筛选	数据集说明文档	无
6.3 数据清洗	数据集说明文档	无
6.4 数据标签	数据集说明文档 其中具体的标签与事件的映射关系建议在元数据中体现	实验范式的具体标签说明，建议在元数据中使用字典格式对实验范式中的各个标签进行注释说明，并建立标签与实际事件之间的一一映射关系
6.5 数据标注	数据文件	无
7.1 存储管理	数据集说明文档	无
7.2 数据安全	数据集说明文档	无
7.3 数据压缩	数据集说明文档	无
7.4 数据存储信息	数据集说明文档	无
7.5 数据命名规范	数据集说明文档	无
7.6 数据备份	数据集说明文档	无
7.7 数据更新和版本	数据集说明文档	无

控制		
7.8 访问控制权限	数据集说明文档	无
7.9 访问可追溯性	数据集说明文档	无
注 1：数据集说明文档：参考了 YY/T 1833.2 数据说明要求的内容，主要帮助数据使用者了解数据内容、文件结构、实验的整体设计、数据流情况并介绍元数据及其定义。元数据：体现元数据内容及其细节，这里不对元数据的定义进行重复介绍，只体现名称和对应数值/字符串。数据文件：数据量较大，数据结构较为复杂的数据单独存储，如脑电信号、视频、阻抗测试结果、实验记录信息等。 注 2：部分通过数字和简单文字较难完整、清晰描述的内容，可以通过图片等形式体现在数据集说明文档中辅助说明，或者单独作为技术文件存储，这部分内容不宜做过于细致的要求。 注3：如果有多种不同范式/实验流程，应针对不同范式/实验流程分别准备元数据。		

附 录 B 元数据格式
(规范性)

元数据内容如下：

元数据名称	标识符	定义及描述方式
实验名称	ExpName	字符串
实验编号	ExpID	字符串
实验开始时间	ExpStarttime	日期格式
实验结束时间	ExpEndtime	日期格式
实验目的	ExpPurpose	文本描述
实验人员编号	ExperimenterID	字符串
采集设备	Device	硬件设备名称型号，用字符串表示。
数据采集单位	Institute	负责实施采集任务的学校或研究机构等组织，用字符串表示。
数据采集时间	DateTime	采集数据集的精确时间记录，应符合 GB/T 7408.1-2023 要求（采用标准日期和时间格式）。
受试者编号	SubjID	字符串
性别	Gender	枚举字符串
年龄	Age	数字
受试者特征	SubjectCharacteristics	受试者相关个体特征，如利手（左/右）、视力、行动能力、病史、BCI 经验等，用文本描述。
采集通道数	ChNum	数字
电极类型	ElectrodeType	字符串
电极结构	ElectrodeStructure	文本描述
电极外径	ElectrodeOuterDiameter	数字
电极内径	ElectrodeInnerDiameter	数字
电极引线长度	ElectrodeLeadLength	数字
电极厚度	ElectrodeThickness	数字
电极材料	ElectrodeMaterial	文本描述
电气性能参数	ElectricalPerformanceParameters	文本描述
定位标准	PositioningStandard	文本描述
电极植入	ElectrodeImplantation	文本描述

系统噪声	SystemNoise	文本描述
延迟	Latency	文本描述
ADC 分辨率	ADCResolution	数字，单位为 bit
输入动态范围	InputDynamicRange	数字，单位为 mV
数据传输速率	DataTransferRate	数字
同步精度	SynchronizationPrecision	文本描述
标签	Label	信号的类别信息、时序信息等的标识，包含标签值和对应的采样点位置索引。
采样率	Fs	信号的采集频率，单位为 Hz。
信号分辨率	SignalResolution	采集时信号模数转换的位数和信号动态范围，用数字表示。
滤波器设置	FSettings	记录数据采集使用的滤波器类型、截止频率、滚降特性、阶数、带宽及相移处理方式、设计方法及滤波实现方式。文本描述
阈值穿越率（尖峰检测、尖峰排序）	TCR	单个通道信号处理所使用的阈值穿越率设定值，并说明滤波设置，阈值大小，以及窗大小。文本描述
条件顺序	ConditionalOrder	字符串
试次数量	Ntrials	数字
通道名称	ChanName	采集过程中使用通道的名称，用字符串表示。
通道位置	ChanLoc	采集过程中使用通道的三维坐标，用（x:数字,y:数字,z:数字）表示。
编码描述	Info	对编码过程的简要说明，用字符串表示。
数据采集时间	DateTime	采集数据集的精确时间记录，应符合 GB/T 7408.1-2023 要求（采用标准日期和时间格式）。
参考电极	Reference	参考电极所采用的通道名称，用字符串表示。
接地电极	Ground	接地电极所采用的通道名称，用字符串表示。
去噪算法	DenoiseAlgo	字符串
标注人员	Annotator	字符串
标注时间	AnnotateTime	日期时间格式
电极植入脑区	Area	字符串
其他数据	OtherData	同时记录的其他数据类型，如肌电、眼电、眼动、行为、音频、视频、被试主观报告及相应与神经数据对齐的标签，用字符串表示。
备注	Note	与该实验相关必要的备注和提醒，有助于数据分析。文本描述

参 考 文 献

- [1] GB 9706.1—2020 医用电气设备+第1部分：基本安全和基本性能的通用要求
- [2] GB 9706.226—2021 医用电气设备 脑电图机的基本安全和基本性能专用要求
- [3] GB 9706.227—2021 医用电气设备 心电监护设备的基本安全和基本性能专用要求
- [4] GB 9706.225—2021 医用电气设备 心电图机的基本安全和基本性能专用要求
- [5] GB/T 39725-2020 信息安全技术 健康医疗数据安全指南
- [6] GB/T 25000.12-2017 系统与软件工程 系统与软件质量要求和评价（SQuaRE） 第12部分：数据质量模型