



中华人民共和国国家标准

GB/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

立项草案稿

XXXX—XX—XX 发布

XXXX—XX—XX 实施

国家市场监督管理总局
国家标准化管理委员会 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并归口。

本文件起草单位：

本文件主要起草人：

引 言

采用脑机接口技术的医疗器械经常调用人工智能算法对脑电信号进行处理，以实现产品的预期用途。人工智能算法的训练、调优、测试均需要依托高质量的脑电数据集。目前，不同机构在脑电数据采集、传输、编码、标注等方面存在差异，导致数据集质量参差不齐。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10章围绕脑电数据集的各种技术要素，分别提出了质量要求。第11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/T 30524-2014 科技平台 元数据注册与管理

YY/T 1833.1-2022 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

3 术语和定义

YY/T 1833.1-2022、YY/T 1833.2-2022界定的以及下列术语和定义适用于本文件。

3.1

脑电图机 electroencephalography

一种记录大脑电活动的电生理监测系统。

注：脑电图机可能是侵入式的或非侵入式的，通常指非侵入式的，即头皮脑电。

3.2

脑机接口 brain-computer interface

一种能够在大脑和外部设备间发送和接收信息的系统，从而实现二者的直接通讯。

注1：脑机接口允许在大脑和外部设备间单向或双向通讯，从而实现控制或反馈功能。

注2：脑机接口系统可在人与人之间或人与动物之间建立。

3.3

数据集 data set

具有一定主题，可以标识并可以被计算机化处理的数据集合。

[来源：WS/T 305-2009，3.1.2]

3.4

数据 data

信息的可再解释的形式化表示，以适用于通信、解释或处理注：可以通过人工或自动手段处理数据。

[来源：GB/T 5271.1-2000，01-01-02]

3.5

数据元 data element

用一组属性规定其定义、标识、表示和允许值等的数据单元。

[来源：WS/T 305-2009，3.1.6，有修改]

3.6

元数据 metadata

定义和描述其他数据的数据。

[来源：GB 18391.1-2009，3.2.16]

3.7

采样率 frequency of sample

单位时间从脑电信号中提取并组成离散信号的采样个数。

3.8

电极 electrode

用于头皮或插入脑部区域用于检测点活动的传感器。

[来源：GB/T 9706.225-2021，定义201.3.208，有修改]

3.9

导联 lead

电极间的电压。

[来源：GB/T 9706.227-2021，定义201.3.206]

3.10

标签 label

在脑电数据中反映特定行为、事件与状态的标记。

3.11

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。

3.12

范式 paradigm

脑机接口范式是指在脑机接口系统中用于诱发特定神经活动或指导使用者主动产生特定神经活动的预设执行流程，在脑机接口系统中用于从脑信号中提取和传输信息的特定方法或策略。

3.13

通道 channel

从头皮上记录EEG信号的单个电极放置位置。每个通道对应一个特定位置，并提供反映该点大脑电活动的独特信号。

3.14

电极位置规范 electrode position specification

电极位置规范是一套标准化的系统，用于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。

3.15

脑电信号 EEG signal

脑电信号大脑神经元放电产生的，反应大脑活动的生物电信号。

4 电极信息说明要求

4.1 电极类型信息

数据集说明应提供数据采集使用的电极类型信息，根据电极与人体组织的接触方式和放置位置，可将电极分为：

- a) 侵入式电极：
注：侵入式电极通过外科手术直接植入大脑皮层或更深的脑组织中，与大脑细胞直接接触。
- b) 半侵入式电极：
注：半侵入式电极通过小型手术植入颅骨下，与大脑皮层或硬脑膜直接接触，提供比非侵入式电极更清晰的信号。
- c) 非侵入式电极：
注：非侵入式电极直接放置在头皮上，不穿透皮肤或颅骨，通过头皮采集大脑的电活动，无需进行任何形式的手术。

4.2 电极形态和尺寸信息

数据集说明应提供数据采集使用的电极形态和尺寸信息，包括：

- a) 电极整体信息
应提供使用电极的整体形状、尺寸信息，包括长度、直径和厚度（如适用）。
- b) 电极触电信息
明确各个电极触点（即电极与神经组织接触部分）的形状、数量、尺寸和电极接触点之间的物理距离等信息。
- c) 电极材料信息
数据集应说明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
数据集应包括信号采集时电极阻抗的测试方法与测试结果信息。

4.3 电极安全性信息

数据集应详细记录与电极安全性相关的信息，包括：

- a) 注册信息
应提供电极的注册信息，包括注册的国家或地区、注册编号、注册日期、型号、生产厂家、以及电极获得批准的具体医疗用途。
- b) 电极安全性试验结果
对于未进行正式注册的电极，应详对电极安全性试验的结果进行描述。

5 数据采集条件要求

5.1 电极布局记录

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。

注：SEEG电极应根据立体定向脑电图技术规范植入，覆盖目标脑区，间距合理。ECoG电极应按网格或条带形式植入大脑皮层表面，密度符合临床需求。记录局部场电位（LFP）和锋电位（Spikes）的电极阵列或深部电极植入应针对特定神经元群体或单个神经元。侵入式且需要按照ECoG、sEEG、深部电极、电极阵列有所区分。EEG电极建议遵循国际通用的10/20或10/10等规范，应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。

5.2 接触阻抗信息

对于非侵入式脑机接口，应提供脑机接口实验前后所用通道的头皮电极接触阻抗，及测量接触阻抗所用的电流频率和峰峰值。应包括采集时电极阻抗的测试方法与测试结果，采集开始前应确保电极与皮肤接触的阻抗在可接受范围内。

5.3 电极植入信息

对于侵入式、半侵入式脑机接口，应记录电极植入的相关信息：电极植入位置（含立体定向坐标、接触深度）、手术方法（含导航系统）、植入时长、电极材料与参数（含触点设计）、术后阻抗测试结果及安全性评估（含并发症记录）。

具体内容包括：

- 位置：立体定向坐标（如 Talairach/MNI）、接触深度（如硬膜下 3mm）。
- 方法：手术类型（如开颅术）、导航工具（如 ROSA 机器人）。
- 时长：植入持续时间（如 72 小时）及信号采集窗口。
- 材料参数：导电层材料（如铂铱合金）、触点间距（如 0.5mm）。
- 阻抗：术后测试值（如 $<10\text{k}\Omega@1\text{kHz}$ ）及方法。
- 安全性：注册证编号、术后并发症（如感染）、组织相容性评估。

5.4 信号类型信息

5.4.1 Electroencephalography (EEG)

对非侵入式脑机接口，应记录 EEG 信号采集的种类及其来源，包括所使用电极的类别（如湿电极、盐水电极或半干电极）及其相应的厂商与型号信息，同时注明脑电放大器的制造商和具体型号。

注：EEG 信号是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。

5.4.2 Electrocorticography (ECoG)

如适用，应采集硬膜下电极记录皮层表面电活动（幅值高 10–100 倍），用于癫痫定位和精准控制。

注：ECoG 是通过直接放置在大脑皮层表面的电极记录的脑电活动。与 EEG 不同，ECoG 的信号具有更高的空间和时间分辨率，能够更精确地反映局部神经元群体的电活动。ECoG 记录的信号通常比 EEG 更清晰，因为它直接接触硬脑膜或脑皮层，减少了头皮和颅骨对信号的衰减。

5.4.3 局部场点位 (LFP)

如适用，应采集植入电极记录神经元群体低频电活动（ $<300\text{Hz}$ ），兼具时空特异性，支持复杂意图解码

LFP 是指在大脑局部区域内记录的电位变化，通常由大量神经元的同步活动产生。LFP 反映了局部神经元群体的电活动，一般是通过植入大脑组织内的电极记录的，能够提供关于局部神经网络活动的详细信息。

5.4.4 Spike banding power (SBP)

如适用，应采集单个神经元高频脉冲信号（ $>300\text{Hz}$ ），直接反映神经编码，用于意念控制设备（如Neuralink芯片）。

注：通常定义为300到6000Hz的范围内的神经元尖峰活动功率，是一种与单个神经元或小群神经元的放电有关的神经信号特征。

5.4.5 尖峰信号（Spike）

如适用，应采集单个神经元高频脉冲信号（ $>300\text{Hz}$ ），直接反映神经编码，用于意念控制设备（如Neuralink芯片）。

注：尖峰信号是由单个神经元或小群神经元产生的短暂、高幅值的电位变化。这些信号通常具有尖锐的波形特征，频率较高，持续时间较短。尖峰信号是神经元动作电位的直接反映，通常通过插入大脑组织内的微电极记录。

5.4.6 背景脑电信息

对于非侵入式脑机接口，数据集应包含脑机接口实验前后被试处于睁眼、闭眼和眨眼状态时脑电图机采集的原始背景脑电数据。其中，睁眼时长应 $\geq 10\text{s}$ ，闭眼时长应 $\geq 10\text{s}$ ，眨眼时长应 $\geq 10\text{s}$ ，眨眼间隔约为2s。

对于侵入式/半侵入式脑机接口，宜采集电极术后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）。

具体参数说明如下：

- 时间点：术后急性期（24h内）、组织适应期（3-7天）、移除前。
- 时长：每次 ≥ 30 分钟，静息态 ≥ 15 分钟（睁眼/闭眼）。
- 环境控制：关闭电子设备、使用屏蔽电缆，侵入式需排除脑脊液伪影。
- 标注要求：记录阶段（如“急性期”）、患者状态（如“静息闭眼”）、刺激参数（如“视觉刺激10Hz”）。

5.4.7 系统噪声信息

数据集制造责任方应提供脑机接口系统的噪声数据和声称噪声有效值指标。在不实施电刺激的情况下，对于sEEG系统，噪声有效值应不大于 $5\mu\text{Vrms}$ ，对于ECoG系统，噪声有效值应不大于 $3\mu\text{Vrms}$ ，对于记录Spike的装置，噪声有效值应不大于 $10\mu\text{Vrms}$ ，且尖峰信噪比应不小于4:1。

对于记录EEG的装置，噪声有效值应 $\leq 0.5\mu\text{Vrms}$ @ 0.5-100 Hz（待确认）。

5.5 采样率信息

应记录数据采集使用的采样率。采样率应为目标信号最高频率的4倍以上。对于LFP信号的装置，采样率应不低于500Hz；对于采集Spike信号的装置，采样率应不低于20kHz；对于采集EEG信号的装置，采样率应不低于1kHz。

5.6 信号分辨率

应记录采集时信号模数转换的位数。对于LFP信号，最小可检测信号应 $\leq 1\mu\text{V}$ ；对于Spike信号，最小可检测信号应 $\leq 5\mu\text{V}$ ；对于EEG信号，最小可检测信号应 $\leq 0.1\mu\text{V}$ 。

5.7 滤波器设置

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）。

具体参数说明如下：

- 类型：明确滤波器功能（如带通 8-32Hz 提取 μ/β 波）。
- 参数：记录截止频率（如 0.1Hz HPF）、滚降（如 24dB/oct 四阶巴特沃斯）、带宽（如 16-20Hz 子频段）。
- 边缘处理：说明是否添加无刺激时段（如首尾各 2s）或采用零相位滤波。
- 方法：注明滤波器设计类型（如切比雪夫）及实现方式（硬件滤波更稳定，软件滤波可灵活调整）。

5.8 信噪比

应记录在体环境的信噪比，应该明确描述采集数据的信噪比计算方法、公式，建议在数据使用与共享前提供信噪比结果作为信号质量的参考信息。

6 数据传输条件要求

6.1 丢包率

应记录脑电采集过程中的丢包情况并计算丢包率。对于有线传输的情况，丢包率应为 0；对于无线传输的情况，丢包率应 $\leq 0.01\%$ 。

6.2 传输稳定性信息

应记录数据传输过程中的稳定性指标，并保存数据校验信息。

6.3 传输实时性信息

应记录系统的数据传输方案、系统延迟等参数，保证数据传输的实时性。

6.4 传输同步性信息

应明确数据传输过程中多设备/信号之间时钟是否对齐，记录数据同步机制、同步精度等信息。

6.5 传输速率

对于多通道数据传输，传输速率应能够支持至少 X Mbps 速率。

7 数据采集实验要求

7.1 实验范式类型

应记录实验范式的类型，脑机接口的实验范式可以分为以下几种类型：

a) 脑电诱发范式：

——视觉诱发范式：通过视觉刺激引发脑电活动，如 SSVEP 闪烁、P300 矩阵等；

——听觉诱发范式：通过听觉刺激引发脑电活动，如声音、语音等；

——触觉诱发范式：通过触觉刺激引发脑电活动，如 20Hz 振动、触摸等；

——其他感觉诱发范式：通过其他感觉刺激引发脑电活动。

——反应任务：执行特定时间反应任务（如对特定类型的图片做出按键反应）并记录脑电信号。

b) 自主节律范式：

——运动想象范式：通过受试者想象运动激活对侧皮层 ERD/ERS，引发脑电信号的变化；

——注意力调节范式：通过冥想（增强 α 波）、放松或集中注意力改变脑电活动，用于运动意图解码和专注力训练；

——情绪调节范式：通过控制情绪状态（如快乐、愤怒）来调节脑电信号。

c) 任务相关范式：

——认知任务：执行特定的认知任务（如计算、记忆、n-back 认知任务等）并记录脑电信号；

——运动任务：执行实际运动任务（如手部动作、混合 SSVEP+运动想象等）并记录脑电信号；

——语言任务：执行语言处理任务（如阅读、说话、言语合成和多模态融合等）并记录脑电信号；

d) 闭环反馈范式：通过 Spike 信号等实时控制植入与外周设备接收发放交互信号，进行神经可塑性训练；

e) 其他实验范式。

7.2 实验范式内容

7.2.1 一般要求

应记录实验范式的具体内容，包括使用的刺激类型、刺激参数（如闪光刺激的尺寸、强度、频率等）、任务要求、任务逻辑及条件设计、范式中的引导语或引导标识等，确保可复现性。

刺激参数、任务逻辑及条件设计举例如下：

- 刺激参数：SSVEP 矩阵布局，频率，对比度高/低；听觉范式组合声音强度（如 30-70dB）与频率（如 1-4kHz）。

- 任务逻辑：SSVEP 要求 600ms 内识别目标，运动想象需在提示后数秒内编码意图，触觉实验保持肢体固定以排除干扰。

7.2.2 任务要求

应记录实验中的具体要求，如动作的时间限制或空间限制。如在 600 毫秒内完成点接触摸屏的动作或在运动过程中保持肘部位置不变等。

7.2.3 条件设置

应记录实验中的条件(condition)集合，即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。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。

7.2.4 条件顺序

应明确实验为区组设计(block design)或逐试次设计(trial-by-trial design)，即每个试次的条件有固定顺序或是随机选取。

7.3 采集过程

7.3.1

应记录实验使用的刺激类型、刺激参数和刺激时间。

- a) 刺激类型，如视觉、听觉、触觉、力学、电刺激等；
- b) 刺激格式，如图像、声音、视频、文字、力场等；
- c) 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；
- d) 刺激时间，如出现和持续的时间、间隔时间等信息。

注：刺激方法、实验的引导和实验步骤对所有BCI系统均适用，但侵入式BCI、半侵入式BCI可能需要特殊刺激方案，如在双向闭环的侵入式BCI中，可能用不同频率和强度的皮层内微电流刺激模拟触觉反馈信息。

7.3.2

应描述实验前和实验中的引导信号。

- a) 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。
- b) 实验中引导：实验中各步骤的引导信号（引导语或引导标识）。

7.3.3

应记录实验中的步骤，如刺激呈现、行为反应、反应测量和反馈等。

7.4 实验过程

7.4.1 一般要求

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

7.4.2 交互方式

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。

7.4.3 行为评估

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）。

7.4.4 行为反馈

应记录每个试次结束后的反馈信息及实验对象是否获得总结性反馈。

7.5 受试者特征记录

7.5.1 性别（涉及到匿名化问题，应为可选项）

应记录受试者的性别信息。

7.5.2 年龄

应记录受试者的年龄信息。

7.5.3 相关个体特征（是否展开？）

根据研究目的和实验设计，应记录相关的个体特征，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。

7.5.4 有/无 BCI 经验

应记录受试者是否具有 BCI 经验。

8 数据集内容要求

8.1 数据文件完备性

8.1.1 脑电数据

数据集制造责任方应确保脑电数据结构能够为脑电数据的解析、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应为连续数据，即在实验任务中硬件设备直接采集得到的整段脑电信号，推荐表示为二维数组。

注：导联数量，采样点数建议直接规定二维数组维度顺序。

8.1.2 数据标签

标签内容由标签值、标签说明、标签时刻组成，反映脑电数据在特定时刻的关键行为、事件与状态。其中，标签值用于数据记录过程中标定不同关键行为、事件与状态的出现，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关键行为、事件与状态的具体情况，其数据类型可以为文本。标签时刻用于标定不同关键行为、事件与状态出现的时刻，其数据类型可以为数字、时间、文本。

给出常见实验范式的推荐记录标准，使用标准化的事件标记

对于非侵入式脑机接口数据集，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。例如，对于一个包含两种稳态视觉诱发电位的脑电数据集，一个完善的标签内容可以如下表所示：

标签值	标签说明	标签时刻
1	对应事件为：由 10Hz 正弦波调制的视觉刺激开始闪烁	10285
2	对应事件为：由 15Hz 正弦波调制的视觉刺激开始闪烁	12285
1	对应事件为：由 10Hz 正弦波调制的视觉刺激开始闪烁	14285
2	对应事件为：由 15Hz 正弦波调制的视觉刺激开始闪烁	16285
...

其中标签时刻的数据类型为数字，表示刺激开始闪烁时对应的脑电数据采集点索引。

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。

注2：标签值是否规定常见标签范围？

注3：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。实验范式产生标签的实际时刻与脑电图机记录到标签的时刻之间的误差应 \leq ? ms

8.1.3 元数据

数据集应包含全面的元数据，提供关于数据集的详细背景信息和描述，包括但不限于电极、信号采集和实验信息，推荐以结构体的形式存储。

元数据应包含电极信息、数据采集信息、数据传输信息、实验人员、地点、时间、机构等信息。其符号、数据形式举例见附录 A。

8.1.4 其他数据

数据集文件应包含采集过程中记录的背景脑电数据、噪声数据和指标，允许包含眼电、肌电等其他模态生理数据，阻抗数据，范式说明数据，license协议文件以及额外的数据补充信息。

数据集应提供应用示例，示例需要包括数据使用和数据分析的具体步骤和分析结果等。数据集应明确数据文件的格式和文件命名标准，以支持数据的有效存储、管理和检索。应明确说明数据集中各数据文件的格式（如 BDF、EDF、RHD、CSV 等）。

8.2 数据集合规性信息

8.2.1 合规性陈述

数据集说明应提供数据来源的合规性陈述。涉及动物、人体实验数据集应提供伦理审批机构的名称和批准号等信息。

8.2.2 实验安全

8.2.2.1 不良影响

应描述(并记录)实验中可能(以及出现过的)对实验对象造成的不良影响,如眩晕或疲劳。

8.2.2.2 保护提示

应描述对实验对象的保护提示,如正确率阈值、反应时间阈值、实验时间阈值等。如正确率降低至60%以下、平均反应时间超过300毫秒、实验时间超过2小时则结束实验。

8.2.2.3 退出机制

如适用,应描述实验对象主动退出当前实验区间的方式。退出时应执行安全程序并记录退出时间/原因(含预设时长结束、保护机制触发、主动退出三种情形),确保实验过程可追溯。

注:实验对象有权通过口头、书面或系统操作主动退出实验,研究者可能因健康风险、违规或安全原因终止参与。

8.2.2.4 范式安全性信息

对侵入式脑机接口、半侵入式脑机接口,应描述范式安全性信息。例如侵入式脑机接口中,如果使用皮层内微电流刺激作为刺激方案,需要将电流对人体的影响和可能诱发的不良反应如癫痫等列入安全风险,描述风险等级、剩余风险和控制措施。

9 数据处理和标注要求

9.1 数据脱敏处理

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、数据匿名化等。适当时,数据集说明文档应描述数据去标识化或者数据匿名化的规则。对参与者的身份信息进行匿名化处理,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。对于无法去除或加密的数据进行通过降低数据的精度来保护隐私。

9.2 数据去噪处理

9.2.1 噪声来源分析与记录

如适用,应描述数据处理时使用的去噪方法及其参数设置。噪声类型包括但不限于以下情形:

- 生理噪声:包括肌电(EMG, 50-200Hz, 额区显著)、眼电(EOG, 0.1-10Hz, 眨眼干扰)及心电(ECG, 1-5Hz 周期性尖峰, 额颞区明显),分别通过电极优化、ICA及陷波滤波处理。
- 设备与环境噪声:工频干扰(50/60Hz, 陷波滤波)、电极接触噪声(<0.1Hz 基线漂移, 阻抗监测+多项式拟合)。
- 侵入式特有噪声:脑脊液波动(0.05-0.2Hz 低频伪影, 带通滤波+空间插值)、胶质瘢痕高频尖峰(>500Hz, 小波变换剔除)。

注：侵入式电极记录面临特殊的噪声来源，如电极微动、组织反应等，去噪尤为重要。

9.2.2 去噪算法与参数设置

应描述去噪算法的全称（如“三阶巴特沃斯带通滤波”）、参数（如陷波频率50Hz）及运行软件环境版本（如MATLAB EEGLAB 2023a），附信噪比评估方法（信号功率/噪声功率比）。

选用去噪算法时，一般按信号类型进行区分处理，例如：

- 通用方法：非侵入式EEG采用0.1-100Hz带通（三阶巴特沃斯，滚降 $\geq 18\text{dB/oct}$ ）+ FastICA（ ≤ 200 次迭代）剔除生理伪影；侵入式ECoG/sEEG扩展至0.01-500Hz，新增QRS波模板匹配剔除心电干扰。

- 特殊处理：植入式电极通过PCA校正运动伪影（前3主成解释释方差 $>80\%$ ），Spike信号结合阈值检测（3倍标准差）与波形筛选，确保信噪比（侵入式Spike $\geq 4:1$ ，非侵入式EEG $\geq 2:1$ ）。

9.3 数据标注

9.3.1 一般要求

如果对数据进行标注，应保存标注文件，并在数据集说明文档中描述数据集标注的内容、方法、标准、参与人员等信息。

9.3.2 标注对象与粒度

标注内容应针对脑电采集过程中出现的关键事件、行为或状态。例如：

- 刺激呈现（如视觉刺激 A、声音提示 B）；
- 被试反应（如按键、眨眼、运动想象）；
- 状态变化（如注意、疲劳、癫痫发作）。

标注粒度应根据具体研究任务确定，可为：

- 离散时刻标注：如事件触发器，记录某一时刻的事件；
- 连续区间标注：如情绪状态、睡眠分期，标注某一时间段的状态标签。

核心标注要素包括：

- 时间戳：精确到信号采样精度（如1000Hz采样对应0.001秒分辨率），标注事件触发时刻（如“视觉刺激开始于记录第15.200秒”）。
- 事件类型：采用统一分类体系（如“SSVEP-10Hz 闪烁”“左手运动想象”）。
- 任务阶段：区分基线期（刺激前）、执行期（刺激呈现）、恢复期（刺激结束后）。

侵入式数据的特殊标注对象包括：

- 神经-行为关联：标注Spike/LFP信号与临床指标的对应关系（如“杏仁核Spike放电频率升高-学习能力表现增加”），结合视频记录或神经心理量表。
- 电极定位标签：记录每个通道的脑区坐标空间（如MNI空间 $x=-24$, $y=-56$, $z=12$ 对应初级运动皮层）。

9.3.3 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：

- c) 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，例如
 - 事件检测：基于信号特征自动识别刺激触发点（如 P300 峰值检测），人工校正准确率 $\geq 95\%$ 。
 - Spike 分类：通过波形聚类算法（如模板匹配）区分不同神经元放电模式，标注动作电位类型。
- d) 手动标注：人工对原始数据进行逐条判读与标注，例如：
 - 双人独立标注：两名标注员分别完成事件标记，交叉验证分歧率 $< 5\%$ ，否则提交仲裁；
 - 多模态对齐：同步标注 EEG、视频、生理信号（如呼吸、心电），确保时间误差 ≤ 5 毫秒，使用工具或硬件进行跨模态时间对齐。
 - 临床标注特殊流程：治疗性 BCI 需标注不良反应与神经信号的时间关联（如“电极移位发生于术后第 7 天，伴随海马区 θ 波功率下降 50%”），参照《Clinical Neurophysiology》安全性标注指南。

应明确标注结果的编码规范，例如：

- 时间格式：采用 ISO 8601 时间戳（如“2025-04-01T14:23:15.000+08:00”）。
- 事件编码：建议设置标准事件编码。

执行标注任务时，应建立质量验证机制并形成记录，例如：

- 内部信度：计算标注者间一致性（Kappa 系数 ≥ 0.8 ）。
- 外部效度：通过机器学习模型验证标注准确性（如标注后的运动想象数据分类正确率 $\geq 90\%$ ）。

9.3.4 参与人员与复核机制

人工标注或专家判断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。

注1：建议使标准化标注：使用统一的标注规范（如时间戳、事件类型、任务阶段）确保不同实验或数据集的标注方式一致。建立标注验证流程，确保标注的准确性和一致性。

注2：将神经活动与状态、行为或症状关联的标注对侵入式BCI数据至关重要，特别是在治疗性应用中

9.4 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。

9.5 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。

9.6 数据分段

数据分段通常基于实验目的或脑电活动特点进行，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。

将连续记录分割成有意义的片段对于分析神经状态变化及其与症状相关性很重要

9.7 数据清洗

使用数据集前应对数据进行清洗，根据后续算法和数据处理需要选择排除坏导、坏段等清洗操作并对清洗方法、标准进行记录（如有）。

9.8 归一化

不同通道、不同时段的数据需要标准化

9.9 重参考

共平均参考？

9.10 频谱分析

Bode图？

9.11 编解码

如数据集包含编解码前后的信息，应描述并验证编解码方法，具体要求见附录B。

10 数据存储与访问

10.1 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。

10.2 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

10.3 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，一般情况下建议使用无损压缩方式

10.4 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(原始数据文件编辑记录?)

10.5 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。

10.6 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。

10.7 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说明文件和使用示例,确保数据集在不同版本间的可追溯性和一致性

10.8 访问控制权限

数据集管理过程中应根据使用者的角色和需求设置不同的访问权限。研究人员、开发人员、医生或其他相关人员应仅能访问与其工作相关的数据和信息。数据集应拒绝任何未经授权访问请求。

10.9 备份策略

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

10.10 访问可追溯性

数据集管理过程中应记录所有数据访问操作,并保留详细的日志。日志内容应至少包含:访问者信息、访问时间、访问终端、访问过程中进行的操作。

11 质量符合性测试方法

11.1 电极说明信息检查

检查数据集制造责任方提供的电极说明信息或随附文件,判断是否满足第4章的要求。

11.2 数据采集条件验证

检查数据集制造责任方提供的数据采集记录和数据集说明文档,判断是否包含条款5.1-5.4、5.9要求的全部信息;

以病例为统计单元,对脑电数据序列进行均匀随机抽样检验,抽样比例不低于10%;使用以下公式,分别计算各个定量指标。

请起草组补充各个定量指标的计算公式

如抽样样本全部满足数据集制造责任方声称的指标,则判定符合条款5.5-5.8、条款5.10的要求,反之亦然。

11.3 数据传输条件验证

根据数据集制造责任方提供的数据传输方案,任选脑电通道,使用信号发生器输入参考波形,对脑电数据传输过程进行复现,记录传输后的波形。

注:参考波形的幅度、频率由具体的脑电放大器决定,一般为正弦信号。

按照下列公式计算丢包率、系统延时、传输速率、稳定性指标，判断是否满足第6章的要求。

请起草组补充各个定量指标的计算公式

11.4 数据采集实验记录检查

检查数据集制造责任方提供的实验记录或数据集随附文件，判断内容是否全面、清晰、准确，判断是否符合条款7.1的要求；

检查脱敏后的受试者信息或记录，判断是否符合条款8.2的要求。

11.5 数据集内容检查

以病例为统计单元，对脑电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，判断是否满足条款8.1的要求；

检查合规性证明文件和安全文档，判断内容是否全面、清晰，判断是否满足条款8.2的要求；

11.6 数据处理与标注检查

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，判断是否满足条款9.1的要求；

检查数据集制造责任方的数据去噪工具和记录，判断是否满足条款9.2的要求；

检查数据标注结果、过程记录和标注人员信息，判断是否满足条款9.3的要求；

检查不同模态数据之间的时间戳或对齐记录，判断是否满足条款9.4的要求；

检查数据集制造责任方提供的筛选记录，对数据集进行10%均匀随机抽样，验证抽样样本是否全部满足入组条件，判断是否满足条款9.5的要求；

请起草组围绕条款9.6-9.11的测试方法按照以上风格进行补充

11.7 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、数据备份记录和策略、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10章的要求。

附 录 A
(资料性)
元数据格式

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
实验日期	ExpDate	日期格式
实验目的	ExpPurpose	文本描述
受试者编号	SubjID	字符串
性别	Gender	枚举字符串
年龄	Age	数字
健康状况	HealthStatus	文本描述
采集通道数	ChNum	数字
电极类型	ElectrodeType	字符串
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。
滤波范围	Fcut	采集过程中允许通过的频率范围, 用成对的阿拉伯数字表示。包括三种滤波方式: 带通: 允许通过频率下限与上限之间的信号, 频率范围包含上下限值, 表示为(数字, 数字)Hz; 低通: 允许通过从 0 Hz 开始至指定上限频率之间的信号, 表示为(0, 数字) Hz; 高通: 允许通过指定下限频率至频率上限无限制之间的信号, 表示为(数字, ∞) Hz。

导联名称	ChanName	采集过程中使用导联的名称，用字符串表示。
导联位置	ChanLoc	采集过程中使用导联的三维坐标，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。
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）。
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

附录 B (规范性) 对脑电数据编解码的补充要求

B.1 数据解码要求

B.1.1 锋电位检测

spike detection

B.1.2 峰电位排序

spike sorting

B.1.3 数据解码算法

应对使用的解码算法、算法的重要参数以及算法使用方法等信息进行详细的描述（如有）。

B.1.3.1 算法信息

应详细说明所使用的解码算法的类型、基本原理和实现方式，包括算法的名称、架构以及其应用领域（例如运动想象、情感调节等）。此外，还应描述算法背后的数学模型或计算框架，以及算法的来源，无论是公开的标准算法还是自定义开发的算法。同时，需要阐明该算法的理论基础和先前研究的支持，确保该算法具有科学性和可信度。

B.1.3.2 模型参数

应列出解码算法中使用的关键参数，并说明选择这些参数的依据。

B.1.3.3 模型更新机制

应阐述解码模型如何进行更新，包括训练数据的选择方式、训练过程中的参数调整方法以及模型如何在实际应用中进行实时更新。如果模型需要周期性重新训练，应该说明训练频率、数据使用方式等。详细描述模型更新机制有助于评估解码算法的适应性和长期有效性，特别是在动态或个性化的脑机接口系统中，及时更新模型以应对输入数据的变化，是确保解码精度的关键

B.1.4 解码效果评估

数据解码算法的评估是确保脑机接口（BCI）系统性能的重要环节。首先，解码算法应基于公开或自建的数据集进行离线开发和评估，确保其在不同情境下的准确性和稳定性。在离线评估达到一定性能指标后，才应进入在线开发和实时评估阶段。此过程不仅确保算法的基础性能，还可以验证其在动态、实时环境中的表现。

B.1.4.1 准确率

准确率（Accuracy, A）是衡量 BCI 系统解读用户意图的正确率的关键指标。在 BCI 系统中，目标数量（N）表示用户可以选择的不同选项或命令数量，而选择时间（T）则是系统完成一次选择所需的时间。准确率通常是通过比较系统输出的命令与用户实际意图的匹配程度来计算的。在大多数情况下，BCI 系

统的准确率应达到一定标准，以确保系统的有效性和用户体验的可靠性。

B.1.4.2 召回率

即真阳性率 (Recall)，可以理解为有多少阳性样本被正确地预测。召回率高则说明漏检率低。

$$Recall = \frac{TP}{TP + FN} \quad (4)$$

其中，FN 为预测为阴性但实际为阳性的样本数量或比率。

B.1.4.3 精确度

即阳性预测值 (Precision)，可以理解为被预测为阳性的样本中有多少是正确的。精确度尽可能保证预测的正确率。

$$Precision = \frac{TP}{TP + FP} \quad (3)$$

其中，TP 为预测为阳性且实际也为阳性的样本数量或比率，FP 为预测为阳性但实际为阴性的样本数量或比率。

B.1.4.4 F1-Score

定义为精确率和召回率的调和平均值，范围为 0~1，综合考虑了精确度和召回率的影响。

$$F1 = \frac{2 \cdot Precision \cdot Recall}{Precision + Recall} \quad (5)$$

B.1.4.5 信号传输率

准确率(A): BCI 系统解读用户意图的正确率。

目标数量(N): 用户可以选择的不同选项或命令数量。

选择时间(T): 系统进行一次选择所需的时间。

高 ITR 表示系统既能快速传递信息，又能保持较高的准确性。在大多数情况下，BCI 系统的 ITR 应该至少达到 X 比特每分钟，以确保系统的基本可用性。在对于特定的 BCI 系统完成特定的任务需要 ITR 不低于 X。

B.1.4.6 解码时间

解码时间是指从输入脑电信号到输出解码结果所需的时间。解码算法的处理时间应小于神经特征数据的时间，确保系统的实时响应能力。解码时间的测量应使用高精度时间同步设备，准确标记脑电信号进入解码模块的时刻与解码结果输出的时刻。多次测量并取平均值，作为系统的响应时间。测量误差应控制在 $\pm[X]\mu\text{s}$ 以内，以确保精确度和一致性。

B.1.4.7 泛化性能

泛化性能指解码算法在不同任务、不同用户和不同实验条件下的表现能力。良好的泛化性意味着解码算法不仅在训练数据集上表现良好，还能在新、未见过的数据上保持高效的性能。为测试泛化性，可以进行跨任务、跨用户、跨数据集等不同情境下的评估。

B.1.4.8 解码鲁棒性

鲁棒性指解码算法在面对噪声、干扰和其他不确定因素时保持稳定和高效的能力。BCI 系统中可能会遇到不同类型的信号干扰，如电源噪声、运动伪迹或用户状态的变化等。因此，算法应具备处理这些不确定因素的能力，保证在各种复杂环境下的稳定性和高准确率。鲁棒性评估可以通过模拟干扰和噪声来测试算法的耐受性和适应能力。

B.1.4.9 时间稳定性

连续多天进行相同任务实验，计算性能指标，评估解码算法性能稳定性（如有）。

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