



中华人民共和国国家标准

GB/T XXXXX—XXXX

医疗装备使用评估 第4部分：脑机接口

Evaluation for using of medical equipment—Part 4: Brain computer interfaces

(工作组讨论稿)

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

目 次

前 言	III
引 言	IV
1 范围	5
2 规范性引用文件	5
3 术语和定义	5
4 总体原则	7
5 评价指标体系	8
5.1 框架与层级	8
5.2 一级评价指标	8
5.3 二级评价指标	9
5.3.1.1 侵入式脑机接口医疗装备	9
5.3.1.1.2 非侵入式脑机接口医疗装备	9
5.3.1.2 数据采集方法和评价结果证明材料	10
5.3.1.3 量化评分规则	10
5.3.2.1.1 侵入式脑机接口医疗装备	11
5.3.2.1.2 非侵入式脑机接口医疗装备	11
5.3.2.2 数据采集方法和评价结果证明材料	11
5.3.1.3 量化评分规则	11
5.3.3.2 数据采集方法和评价结果证明材料	12
5.3.3.3 量化评分规则	12
5.3.4.2 数据采集方法和评价结果证明材料	13
5.3.4.3 量化评分规则	13
5.3.5.2 数据采集方法和评价结果证明材料	14
5.3.5.3 量化评分规则	15
5.3.6.2 数据采集方法和评价结果证明材料	15
5.3.6.3 量化评分规则	16
6 用于量化评分的数据源的取值规则	16
6.1 取值方法通则	16
6.2 等级/阈值	16
7 评估程序	16
7.1 确定评估目的和范围	16
7.2 确定评估指标	16
7.3 编制评估文件	17
7.4 确立评估方式和评估人员要求	17
7.5 确定参与评估的用户群	17
7.6 评估实施及记录	18
7.7 数据统计分析及结果输出	18
8 指标评估规则及数据处理	18
8.1 评估规则	18

8.2	评估计算	18
9	综合评估结果	21
9.1	等级划分	21
9.2	综合评估结果报告	21
C.1	统计数据法	48
C.2	试验/测量法	48
C.3	证据判断法	49
C.4	主观赋值法	50

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。

本文件由全国医疗装备产业与应用标准化工作组（SAC/SWG26）提出并归口。

本文件起草单位：浙江强脑科技有限公司、北京大学分子医学南京转化研究院、北京品驰医疗设备股份有限公司、首都医科大学附属北京天坛医院、机械工业仪器仪表综合技术经济研究所、重庆医疗器械质量检验中心、北京协和医院、首都医科大学附属北京友谊医院、西北工业大学、清云科技应用（昆明）有限公司、南京灵脑科技有限公司、南京鸿脑科技有限公司

本文件主要起草人：

引 言

随着神经工程和人工智能技术的不断发展，脑机接口（Brain-Computer Interface, BCI）在医疗领域中的应用日益广泛，特别是在神经和精神等疾病的辅助治疗和康复方面，展现出巨大的临床潜力。同时，脑机接口通过对大脑信号的实时采集、处理与解码，建立起人脑与外部设备之间的直接通信通道，也为无法进行常规运动、言语或意识交流的患者提供了一种全新的交互方式。

作为一类高度集成的智能医疗装备，其多模块、高风险、跨学科的特性，使得在实际医疗环境中，对其使用效果、安全性、稳定性和操作便捷性的系统性评估显得尤为重要。

当前，对于传统医疗设备已有较为成熟的评估体系，而针对脑机接口这类新兴技术装备的使用评估，尚缺乏统一标准。开展脑机接口医疗装备的使用评估，不仅有助于提升其使用效率、保障患者安全，也为制定相关政策、支撑采购遴选与上市后监管、规范产品与准入机制以及引领该领域产品高质量发展提供了实践依据。

本标准旨在建立一套针对脑机接口医疗装备科学、系统、可操作的多维度使用评估框架，为后续技术优化、推广应用、规范运行等提供数据支持与决策参考。

医疗装备使用评估 第4部分：脑机接口

1 范围

本文件确立了采用脑机接口技术的医疗装备（简称脑机接口医疗装备）在使用过程中开展评估的总体原则，规定了评价指标体系的框架和取值规则、评估结果的生成和报告，以及评估活动的组织、实施与管理要求。

本文件适用于按制造商的使用说明书中指出的预期在专业医疗机构和/或家庭护理环境中使用的脑机接口医疗装备（含侵入式、非侵入式技术类型）的使用评估。

本文件不适用于预期仅在YY 9706.112所述的紧急医疗服务环境或GB 9706.1（不包括YY 9706.112或YY 9706.111的附加内容）所述的专业医疗机构中使用的脑机接口医疗装备。

示例：预期用于家庭护理环境和专业医疗机构的脑机接口医疗装备。

注：家庭护理环境中的脑机接口医疗装备可能经常会被用于电源不稳及接地保护不可靠的地方。

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 9706.1—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

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

GB/T 16886.1 医疗器械生物学评价 第1部分：风险管理过程中的评价与试验（等同ISO 10993-1）

GB/T 42062—2022 医疗器械 风险管理对医疗器械的应用（等同ISO 14971:2019）

YY/T 1996—2025 采用脑机接口技术的医疗器械 具备闭环功能的植入式神经刺激器 感知与响应性能测试方法

3 术语和定义

YY/T 1987—2025界定的以及下列术语和定义适用于本文件。

3.1

脑机接口医疗装备 (brain-computer/machine interface medical equipment; BCI-ME/BMI-ME)

通过侵入或非侵入的方式测量神经系统产生的神经信号并实时解码，实现患者与外部设备的信息交互、控制或闭环调节，用于诊断、治疗、康复、监护或功能替代/辅助的有源医疗装备。

[来源：YY/T 1987—2025，3.1.1，有修改]

3.2

侵入式脑机接口医疗装备 (invasive BCI-ME/BMI-ME)

通过突破体表屏障的方式，将信号采集或执行组件放置于神经系统的内部、表面、腔隙或血管内记录神经活动的脑机接口医疗装备。

注1：侵入式脑机接口医疗装备通常包括半侵入式和完全侵入式。

注2：半侵入式与完全侵入式的分类判定标准为是否实质性侵入神经中枢。

示例：血管内电极、皮层脑电图（electrocorticography, ECoG）、皮层内微电极等。

[来源：YY/T 1987—2025, 3.1.2, 有修改]

3.3

非侵入式脑机接口医疗装备 (non-invasive BCI-ME/BMI-ME)

在不破坏体表屏障完整性的前提下，从人体表面或外部测量神经系统活动的脑机接口医疗装备。

注：神经活动包含但不仅限于常见技术包括头皮脑电图 (electroencephalography, EEG)、功能性近红外光谱 (functional near-infrared spectroscopy, fNIRS)、脑磁图 (magnetoencephalography, MEG) 以及表面肌电 (surface electromyography, sEMG) 等。

[来源：YY/T 1987—2025, 3.1.3, 有修改]

3.4

信息传输率 (Information Transfer Rate; ITR)

单位时间内从大脑向外部设备成功传输的平均信息量，常用公式为：

$$ITR = B \cdot C$$

单位：比特每分钟 (bits/min)。

式中：

B——单次选择的信息量，单位为bits/selection (bits/sel)；

C——单位时间的决策次数，单位为choices/min。

注：信息传输率是衡量脑机接口装备通信能力的综合指标，本条款中公式来源于Wolpaw定义。

3.5

校准时间 (calibration duration)

装备为达到可用控制性能所需的校准活动中，从校准开始到校准结束所经历的实际工作时长。

注：校准时间的单位为秒或分钟，符号为s或min。

3.6

生理伪迹 (physiological artifact)

由非目标神经活动引起的干扰信号。

示例：眼电 (electrooculography, EOG)、肌电 (electromyography, EMG)、心电 (electrocardiography, ECG)、呼吸伪迹 (respiratory artifact) 等。

3.7

电极间阻抗 (interelectrode impedance)

两个电极之间的等效阻抗。

注：电极间阻抗常用于表征ECG/EEG导联对的连接质量，或在植入式/神经刺激中表征电极触点之间的连通性与完整性。

3.8

用户 (user)

使用脑机接口医疗装备产品或服务的人员。

示例：患者、操作人员、护理人员、医疗保健人员、设备测试/维护人员。

3.9

患者 (patient)

接受脑机接口医疗装备检查、诊断、治疗和护理的生物（人或动物）。

注：患者可以是操作人员。

[来源：GB 9706.1—2020, 3.76, 有修改]

3.10

不良事件 (Adverse Event)

与脑机接口医疗装备有关的导致或可能导致患者、使用者或相关人员人体伤害的各种有害事件。

注：不良事件可能是一个不期望发生的或未预见的非预期疾病或损伤、任何不良临床体征（包括异常实验室检查结果），这些可能与脑机接口医疗装备的佩戴、植入、调整或操作过程有关联，也可能没有关联。

[来源：YY/T 0869.1—2016, 2.1；有修改]

3.11

风险管理 (Risk Management)

将管理方针、程序及其实践系统性地应用于分析、评价、控制和监视风险的活动。

[来源：GB/T 42062—2022, 3.24]

3.12

家庭护理环境 home healthcare environment

患者的住所或患者所在的其他地点，不包括只要患者出现就有经过培训的操作者在场的专业医疗机构。

示例：汽车、公交车、火车、轮船或飞机上，以及轮椅上或户外散步环境。

注1：专业医疗机构包括医院、医师办公室、独立外科中心、牙科诊所、独立妇幼保健院、专门护理机构、综合治疗机构和医疗急救机构。

注2：本部分中，疗养院也被认为是家庭护理环境。

注3：患者出现的其他地点包括步行或乘车所去的户外环境。

[来源：YY 9706.111—2021, 3.1]

4 总体原则

脑机接口医疗装备的使用评估宜遵循以下总体原则：

a) 科学性

评价指标体系基于科学的理论和方法构建，能够客观、准确地反映脑机接口医疗装备的应用效果，评价过程遵循科学的程序和方法。

b) 可操作性

评价指标具有明确的定义和评价得分计算方法，评价过程简便易行，便于实际操作和应用。

c) 范围全面

评估范围覆盖产品全使用周期，着重关注可能对用户安全、健康或治疗效果产生重大影响的风险点，充分考量影响评价结果客观性与准确性的各维度影响因素。核心评估维度涵盖安全性、临床有效性、人因可用性、技术可靠性、环境适应性、数据与网络安全。

注：可能产生重大影响的风险点包括但不限于误触发导致误操作、信号丢失引发治疗中断、长时间佩戴引发皮肤损伤等。

d) 指标体系合理

指标体系宜与产品特征、评估目的和实际应用场景相匹配，评价指标选取合理，各维度评价指标间逻辑关联清晰、内容无交叉重叠、因果关系无混淆。

注1：评价指标体系及其权重需根据实际使用场景（例如医院重症监护室、康复训练中心、家庭护理环境等）和产品的不同技术路线进行动态调整和修订，确保评估结果真实反映脑机接口医疗装备在实际使用场景中的临床表现。

注2：脑机接口医疗装备涵盖非侵入式与侵入式两类技术路线，包含EEG、功能性近红外光谱(Functional Near-Infrared Spectroscopy, fNIRS)、超声(Ultrasound, US)、脑磁图(Magnetoencephalography, MEG)等技术模式，以及闭环调控、混合BCI、自适应解码等新型交互范式。

e) 结果真实有效

所有评价指标，包括客观测量指标（例如信息传输率、误操作率）和主观评价指标（例如用户满意度、舒适度），具有设置合理和规范明确的判定准则、数据采集方法、量化评分规则和结果表达形式，评估结果具有可重复性、可比对性和可验证性。

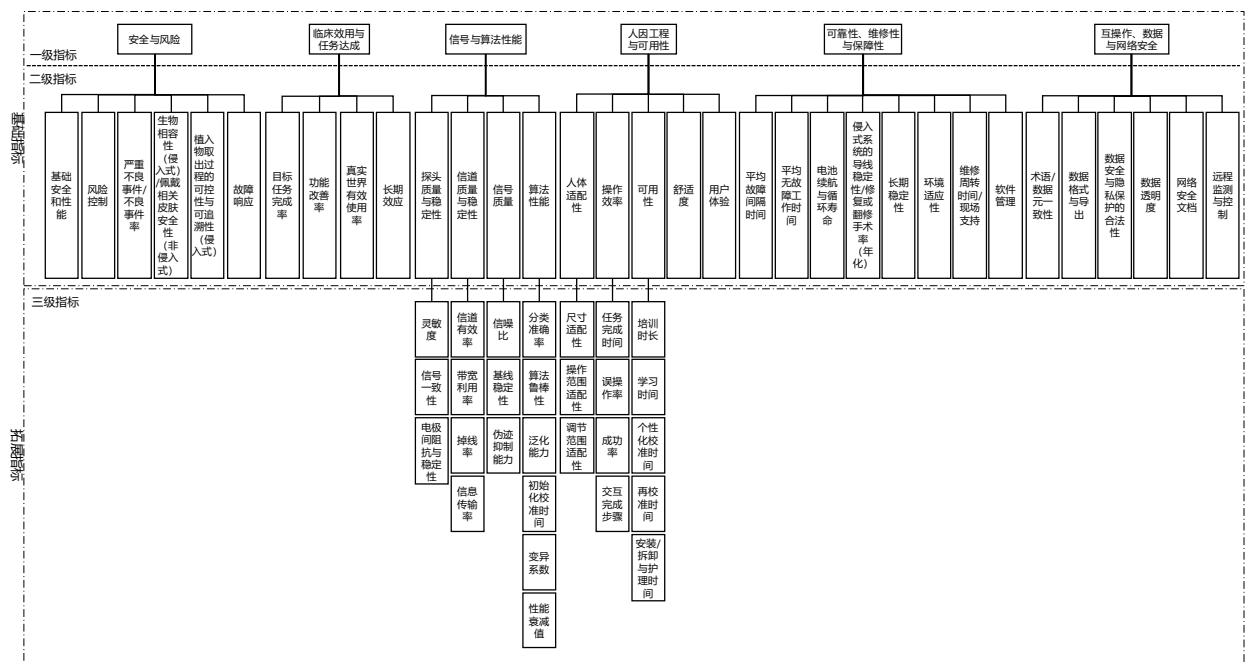
f) 动态性

评价活动具有动态性，评价指标体系把握原则和方向，评价方法能够随技术迭代、市场需求变化或政策调整而灵活更新，避免静态指标的滞后性。

5 评价指标体系

5.1 框架与层级

脑机接口使用评估指标体系（见图1）采用3级结构：一级评价指标6项；二级和三级评价指标按5.3中的原则选取和确定。



说明：侵入式脑机接口的安全与风险二级指标中的“生物相容性”在非侵入式脑机接口替换为“佩戴相关皮肤安全性”。

图1 脑机接口适用评估指标体系图

5.2 一级评价指标

本文件规定以下6项一级评价指标，评价目标分别为：

a) 安全与风险

评估产品在正常使用和可预见误用条件下对用户健康与安全以及环境安全的潜在危害控制能力。

- b) 临床效用与任务达成
评估产品在目标适应症或/和临床使用场景中实现预期诊疗功能的能力。
- c) 信号与算法性能
评估探头、信道、信号采集和算法处理的性能。
- d) 人因工程与可用性
评估产品的人体适配性，以及衡量用户在实际使用过程中对产品软、硬件操作的效率、有效性、舒适性的主观满意度。
- e) 可靠性、维修性与保障性
评估产品在预期使用寿命内的运行稳定性、故障率、维修便捷性及技术支持可及性。
- f) 互操作、数据与网络安全
评估患者神经数据的采集、存储、传输与隐私保护的合规性。

5.3 二级评价指标

5.3.1 安全与风险

5.3.1.1 评价内容

5.3.1.1.1 侵入式脑机接口医疗装备

评价内容如下：

- a) *基本安全和基本性能
应符合GB9706.1—2020中与评估目的和评估需求相关的基本安全相关指标的要求。
- b) *风险控制
——应具有GB/T 42062—2022全流程的风险管理文档；
——应符合GB/T 42062-2022中与评估目的和评估需求相关的风险控制措施的要求。
- c) *严重不良事件/不良事件率
依据《医疗器械不良事件监测和再评价管理办法》，按评估方的评价方法，产品严重不良事件/不良事件率应符合评估组织方的可接受准则要求。
- d) *生物相容性
——应符合GB/T 16886.1中与评估目的和评估需求相关的生物相容性要求。
- e) *植入物取出过程的可控性与可追溯性
——植入物应有永久性标识（如激光刻码等）或在植入位置附近应有影像可见的标记；
——植入物在植入期内不应自发断裂或遗留；
——需要移除时应能安全取出；
——制造商应能提供详细的取出手术指南和植入物难以完整取出的风险评估文件；
——制造商和医院应建立植入器械唯一编号与患者对应的追溯系统；
- f) *故障响应
应具有故障检测和响应机制以及设备出现故障时进入安全响应模式的等级设计。

5.3.1.1.2 非侵入式脑机接口医疗装备

评价内容如下：

- a) 基本安全和基本性能
同5.3.1.1.1 a)。

- b) 风险控制
同5.3.1.1.1 b)。
- c) 严重不良事件/不良事件率
同5.3.1.1.1 c)。
- d) 故障响应
同5.3.1.1.1 f)。
- e) *佩戴相关皮肤安全性
——应符合GB/T 16886.1中与评估目的和评估需求相关的皮肤安全性要求。

5.3.1.2 数据采集方法和评价结果证明材料

安全与风险相关二级评价指标宜按表1的要求开展数据采集和评价结果证明材料的收集。

表1 安全与风险相关二级评价指标

数据采集方法和评价结果证明材料一览表

序号	二级评价指标	数据采集方法	评价结果证明材料
1	基本安全和基本性能	证据收集法	第三方检验报告。 注：出具报告的检验机构应具有医疗器械检验检测资质。
2	风险控制	证据收集法	1.完整覆盖 GB/T 42062—2022 全流程的风险管理文档； 2.风险控制措施完成记录。
3	严重不良事件/不良事件率	统计数据法	严重不良事件/不良事件率记录。
4	生物相容性与植入相关并发症率	证据收集法 统计数据法	1.生物相容性评价报告； 2.植入装置在临床使用过程中引发的并发症的发生次数/月记录。
5	植入物取出过程的可控性与可追溯性	证据收集法	1.标识或标记的影像记录； 2.植入物自发断裂或遗留不良反应事件次数和严重程度评估报告、移除记录； 3.制造商提供的手术指南； 4.植入物难以完整取出的风险评估文件； 5.植入器械唯一编号及其对患者记录。
6	故障响应	证据收集法	1.故障检测和响应机制的验证记录； 2.各类故障及其进入安全响应对应模式的记录。
7	佩戴相关皮肤安全性	证据收集法	1.皮肤相关生物相容性评价报告； 2.真人佩戴研究和用户反馈报告。

5.3.1.3 量化评分规则

安全与风险相关二级评价指标宜依据评估组织方评估方案中的量化评分规则开展量化评分，相关示例见附录B中表B.1。

若该二级评价指标具有下属三级指标，则其最终得分由其下属第三级评价参数的量化评分值加权汇总计算得出。

5.3.2 临床效用与任务达成

5.3.2.1 评价内容

5.3.2.1.1 侵入式脑机接口医疗装备

评价内容如下：

a) *目标任务完成率

应符合评估组织方预设的在实验室或真实环境下的成功率阈值，以及其对数据完整性、统计显著性和多维度证据的要求。

b) *功能改善率

采用经过验证的功能独立性评定量表进行评估，计算术后90天及长期随访（如12个月）时评分相对于基线的变化率，该变化率应符合评估组织方对功能改善率的可接受准则要求。

c) *真实世界有效使用率

基于产品使用日志与患者日记，统计随访期内符合“有效使用日”标准的天数占总评估天数的比例，该比例应符合评估组织方对真实世界有效使用率的可接受准则要求。

d) *长期效应

通过在设定随访时间点统计仍持续使用产品的用户比例，并结合退出用户的结构化访谈记录开展综合评价，其评价结果应符合评估组织方对长期效应的可接受准则要求。

5.3.2.1.2 非侵入式脑机接口医疗装备

评价内容如下：

a) 目标任务完成率

同5.3.2.1.1 a)。

b) 功能改善率

c) 同5.3.2.1.1 b)，计算使用后90天及长期随访（如12个月）时评分相对于基线的变化率。真实世界有效使用率

同5.3.2.1.1 c)。

d) 长期效应

同5.3.2.1.1 d)。

5.3.2.2 数据采集方法和评价结果证明材料

临床效用与任务达成相关二级评价指标宜按表2的要求开展数据采集和评价结果证明材料的收集。

表2 临床效用与任务达成相关二级评价指标数据采集方法和评价结果证明材料一览表

序号	二级评价指标	数据采集方法	评价结果证明材料
1	目标任务完成率	试验/测量法	任务成功率、任务达成时间
2	功能改善率	证据收集法	功能改善率
3	真实世界有效使用率	统计数据法	真实世界有效使用率
4	长期效用	证据收集法	临床随访记录、退出用户访谈记录

5.3.1.3 量化评分规则

临床效用与任务达成相关二级评价指标宜依据评估组织方评估方案中的量化评分规则开展量化评分，相关示例见附录B中表B.2。

若该二级评价指标具有下属三级指标，则其最终得分由其下属第三级评价参数的量化评分值加权汇总计算得出。

5.3.3 信号与算法性能

5.3.3.1 评价内容

评价内容如下：

- a) *探头性能与稳定性
应符合评估组织方评估方案中探头性能与稳定性的可接受准则要求。
- b) *信道质量与稳定性
应符合评估组织方评估方案中信道质量与稳定性的可接受准则要求。
- c) *信号质量
应符合评估组织方评估方案中信号质量的可接受准则要求。
- d) *算法性能
应符合评估组织方评估方案中算法性能的可接受准则要求。

5.3.3.2 数据采集方法和评价结果证明材料

信号与算法性能相关二级评价指标宜按表3的要求开展数据采集和评价结果证明材料的收集。

表3 信号与算法性能相关二级评价指标数据采集方法和评价结果证明材料一览表

序号	二级评价指标	数据采集方法	评价结果证明材料	说明
1	探头性能与稳定性	证据收集法 试验/测量法	其下三级指标的检测报告。	三级指标宜包括但不限于灵敏度、信号一致性、电极间阻抗与稳定性等参数指标。
2	信道质量与稳定性	证据收集法 试验/测量法	其下三级指标的检测报告。	三级指标宜包括但不限于信道有效率、带宽利用率、掉线率、信息传输率等参数指标。
3	信号质量	证据收集法 试验/测量法	其下三级指标的检测报告。	三级指标宜包括但不限于信噪比、基线稳定性、伪迹抑制能力等参数指标。
4	算法性能	证据收集法 试验/测量法	其下三级指标的检测报告。	，三级指标宜包括但不限于分类准确率、算法鲁棒性、泛化能力、初始化校准时间、变异系数、性能衰减值等参数指标。

5.3.3.3 量化评分规则

信号与算法性能相关二级评价指标宜依据评估组织方评估方案中的量化评分规则开展量化评分，相关示例见附录B中表B.3。

若该二级评价指标具有下属三级指标，则其最终得分由其下属第三级评价参数的量化评分值加权汇总计算得出。

5.3.4 人因工程与可用性

5.3.4.1 评价内容

评价内容如下：

- a) *人体适配性
应符合评估组织方对人体适配性的可接受准则要求。

- b) *操作效率
应符合评估组织方对操作效率的可接受准则要求。
- c) *可用性
应符合评估组织方对可用度的可接受准则要求。
- d) *舒适度
应符合评估组织方对舒适度的可接受准则要求。
- e) *用户体验
应符合评估组织方对用户体验的可接受准则要求。

5.3.4.2 数据采集方法和评价结果证明材料

人因工程与可用性相关二级评价指标宜按表4的要求开展数据采集和评价结果证明材料的收集。

表4 人因工程与可用性相关二级评价指标数据采集方法和评价结果证明材料一览表

序号	二级评价指标	数据采集方法	评价结果证明材料	说明
1	人体适配性	统计数据法	其下三级指标的检测报告。	三级指标宜包括但不限于尺寸适配性、操作范围适配性和调节范围适配性等参数指标。
2	操作效率	证据收集法 试验/测量法	其下三级指标的检测报告。	三级指标宜包括但不限于任务完成时间、误操作率、成功率、交互完成步骤等参数指标。
3	可用性	证据收集法 试验/测量法	其下三级指标的检测报告。	三级指标宜包括但不限于培训时长、学习时间、个性化校准时间、再校准时间、安装/拆卸与护理时间等参数指标。
4	舒适度	主观赋值法	系统可用性量表（System Usability Scale, SUS）评分。	/
5	用户体验	主观赋值法	SUS 评分。	/

5.3.4.3 量化评分规则

人因工程与可用性相关二级评价指标宜依据评估组织方评估方案中的量化评分规则开展量化评分，相关示例见附录B中表B.4的要求。

若该二级评价指标具有下属三级指标，则其最终得分由其下属第三级评价参数的量化评分值加权汇总计算得出。

5.3.5 可靠性、维护与后勤保障

5.3.5.1 评价内容

评价内容如下：

- a) *平均故障间隔时间（Mean Time Between Failure, MTBF）
应符合评估组织方对平均故障间隔时间的可接受准则要求。
- b) *平均无故障工作时间（Mean Time To Failure, MTTF）
应符合评估组织方对平均无故障工作时间的可接受准则要求。
- c) *电池续航与循环寿命
——应符合评估组织方对电池续航与循环寿命的接受度；
——植入设备若不可更换电池，其寿命应不少于5年。

- d) *侵入式系统的导线稳定性/修复或翻修手术率（年化）
应符合评估组织方对年化翻修率的可接受准则要求。
注：翻修包括因导线故障、感染、信号衰减等需再次开颅或血管介入的操作。
- e) *长期稳定性
按评估方的评价方法，应符合评估组织方对使用寿命的可接受准则要求。
- f) *环境适应性
应符合评估组织方对由环境适应性引发事件的可接受准则要求。
- g) *维修周转时间/现场支持
抽查近6个月工单处理时效，应符合评估组织方对年化翻修率的可接受准则要求。
- h) *软件运维
应符合评估组织方对软件运行效率及软件管理效果的可接受准则要求。

5.3.5.2 数据采集方法和评价结果证明材料

可靠性、维护与后勤保障相关二级评价指标宜按表5的要求开展数据采集和评价结果证明材料的收集。

表5 可靠性、维护与后勤保障相关二级评价指标数据采集方法和评价结果证明材料一览表

序号	二级评价指标	数据采集方法	评价结果证明材料	说明
1	平均故障间隔时间	证据收集法 统计数据法	平均故障间隔时间记录，相关记录包括但不限于临床使用日志、远程监控平台数据。	<ul style="list-style-type: none"> 可靠性评价指标； 适用于可修复的外部设备； 评价时间以天或月为单位； 根据不同时间区间进行评分。
2	平均无故障工作时间	证据收集法 统计数据法	<ul style="list-style-type: none"> 正常使用时间调查表； 平均故障间隔时间记录。 	<ul style="list-style-type: none"> 可靠性评价指标； 适用于不可修复的植入部件或者一次性设备； 评价时间以天或月为单位； 根据不同时间区间进行评分； 需排除用户误操作导致的“假故障”。
3	电池续航与循环寿命	证据收集法 试验/测量法	<ul style="list-style-type: none"> 产品使用说明书； 第三方检测报告，或电池供应商证明材料。 	该指标为可靠性评价指标。
4	侵入式系统的导线稳定性/修复或翻修手术率（年化）	证据收集法 统计数据法	临床随访数据计算的年化翻修率统计数据。	该指标为可靠性评价指标。
5	长期稳定性	证据收集法	<ul style="list-style-type: none"> 产品使用说明书； 第三方检测报告，或加速老化试验验证报告。 	该指标为可靠性评价指标。
6	环境适应性	证据收集法 统计数据法	每月发生的由环境适应性引发事件次数及其严重程度的记录。	该指标为可靠性评价指标。
7	维修周转时间/现场支持	证据收集法 统计数据法	<ul style="list-style-type: none"> 售后服务记录； SLA 协议及用户满意度调查表。 	该指标为维修性与保障性评价指标。

序号	二级评价指标	数据采集方法	评价结果证明材料	说明
8	软件运维	证据收集法 统计数据法	包括但不限于以下记录： · 软件运行失效次数/月； · 软件升级及其风险告知次数/月； · 软件升级及其故障次数/月。	该指标为维修性与保障性评价指标。

5.3.5.3 量化评分规则

可靠性、维护与后勤保障相关二级评价指标宜依据评估组织方评估方案中的量化评分规则开展量化评分，相关示例见附录B中表B.5的要求。

若该二级评价指标具有下属三级指标，则其最终得分由其下属第三级评价参数的量化评分值加权汇总计算得出。

5.3.6 互操作、数据与网络安全

5.3.6.1 评价内容

评价内容如下：

a) 术语/数据元一致性

对业务术语与数据元的定义、命名、分类、标识和值域与使用规则的约定与执行宜统一、无歧义和可复用，且应与行业内推荐术语/数据元一致；

b) 数据格式与导出

产品相关数据的导出方式与存储格式应符合通用标准的要求。

c) 数据安全与隐私保护的合法性

数据收集和使用应符合个人信息保护相关法律法规的要求；

d) 数据透明度

应符合评估组织方对数据透明度的可接受准则要求。

e) 网络安全文档

应符合评估组织方对数据透明度的可接受准则要求。

f) 远程监测与控制

应符合国家相关法律法规、标准的要求。

5.3.6.2 数据采集方法和评价结果证明材料

互操作、数据与网络安全相关二级评价指标宜按表6的要求开展数据采集和评价结果证明材料的收集。

表5 互操作、数据与网络安全相关二级评价指标数据采集方法和评价结果证明材料一览表

序号	二级评价指标	数据采集方法	评价结果证明材料	说明
1	术语和数据元一致性	证据收集法	一致性验证报告。	/
2	数据格式与导出	证据收集法	符合性验证报告。	/
3	数据安全与隐私保护的合法性	证据收集法	符合性验证报告。	/
4	数据透明度	证据收集法	符合性验证报告。	验证内容宜包括但不限于测试协议和数据集。

序号	二级评价指标	数据采集方法	评价结果证明材料	说明
5	网络安全文档	证据收集法	<ul style="list-style-type: none"> 风险管理文档； 符合性验证报告。 	<ul style="list-style-type: none"> 验证内容宜包括但不限于SBOM、威胁建模、更新与补丁策略、漏洞沟通流程等 检验依据可包括但不限于： <ul style="list-style-type: none"> GB/T 44585.1-2024《接入IT网络的医疗器械风险管理 第1部分：联网医疗器械网络安全》； 《面向医疗器械的上市前网络安全指南（FDA2025 版）》的推荐要素（适用于联网/可更新设备）； IEC 81001-5-1。
6	远程监测与控制	证据收集法	<ul style="list-style-type: none"> 远程检测与控制记录； 符合性检验报告。 	/

5.3.6.3 量化评分规则

互操作、数据与网络安全相关二级评价指标宜依据评估组织方评估方案中的量化评分规则开展量化评分，相关示例见附录B中表B.6的要求。

若该二级评价指标具有下属三级指标，则其最终得分由其下属第三级评价参数的量化评分值加权汇总计算得出。

6 用于量化评分的数据源的取值规则

6.1 取值方法通则

用于量化评分的数据源取值，可采用统计数据法、试验/测量法、证据判断法、主观赋值法等适宜方法获取，且应在评估方案中考虑数据来源特征以及取值适用的时间范围、单位和计算方法。

注：本部分中部分数据源的取值方法示例详见附录C。

6.2 等级/阈值

等级/阈值应由组织评估方在评估方案中进行设定。设定等级/阈值时，应综合考虑产品的临床使用风险以及组织评估方的评估目的、范围、评估对象数量和特征。

7 评估程序

7.1 确定评估目的和范围

应根据临床使用情景和预期用途确定评估的目的和范围。

示例：可对评估对象在其主要临床使用情境下的主要预期用途的实现情况进行评估，以获得评估对象的临床使用安全性、有效性和可靠性的综合评估结果。也可对评价对象临床使用安全性、有效性和可靠性某方面的实现情况或在特定使用情境下的预期用途的实现情况进行评估，以获得评估对象特定方面的临床使用表现评估结果。

7.2 确定评估指标

依据不同评估对象和评估目的、评估范围及评估要求等，采用专家讨论或投票打分等方式选取并确定表1或表2中的评价指标。

7.3 编制评估文件

编制评估实施所需的相关文件，如评估方案、评估指导书、知情同意书、主观量表、数据记录表等。同时，应收集与评估工作有关的文件和资料。

评估方案内容一般应包括：

- a) 任务来源；
- b) 评估目的；
- c) 评估依据；
- d) 评估对象及其技术状态；
- e) 评估时间、地点；
- f) 评估内容，包括评估项目、评估方法、评估程序、参与评估的用户群等；
- g) 评估人员；
- h) 评估条件；
- i) 数据记录、分析及处理；
- j) 风险分析与控制；
- k) 质量控制；

l) 评估应急处置预案，且评估应急处置预案中应包括：不符合GB 9706.1—2020时，应立即通知评估组织方停止该评估对象的临床使用；

m) 技术保障条件；

n) 评估结论的出具规则及要求。

7.4 确立评估方式和评估人员要求

7.4.1 评估组织方应采用自主组织评估或委托独立检验检测机构的方式之一开展评估活动。

7.4.2 自主组织评估时，评估组织方应根据评估要求确立评估人员能力，应采取培训或演练的方式使每名参与评估的人员熟悉评估文件的内容和注意事项。

注：评估人员须具备人因工程、神经工程、临床康复/神经内外科、网络安全等相关能力。

7.4.3 委托独立检验检测机构评估时，该机构应具备医疗装备检验检测资质，宜具备评估对象的相关检验检测能力。

7.5 确定参与评估的用户群

7.5.1 参与评估用户群范围

需要用户参与的评估项目，应对评估对象的实际用户和潜在用户的范围进行确定，宜包括现有用户和新用户。评估前应确定参与评估的用户是否包括全部用户，宜考虑极端用户特性人群的需要。应识别可能会对实际评估效果产生重大影响的所有用户特征。

注1：如果用户类别不止一个，则根据评估目的对所有用户群或仅对某些用户群进行评估。

注2：如需评估可达性，参与评估的用户群需从最难实现指定目标（特别是与有效性相关的目标）的人群中挑选，而不是从一般性人群中选取。

7.5.2 确定参与评估用户群人数及其评估时间

根据评估目标和评估内容，确定预期的统计置信水平，评估用户样本量按照以下规定确定：

- a) 非侵入式：不少于20名用户，累计 ≥ 100 会话；
- b) 侵入式：不少于5~10名用户，随访时间应符合评估要求。

7.5.3 招募评估用户群

招募能够代表预期用户或实际用户的目标用户样本参与评估。用户样本的分布或属性应尽可能接近预期或实际用户群的用户特征。不宜从医疗机构或评估对象制造商内部挑选评估用户，如有，应尽量避免从熟悉评估对象的部门进行挑选。评估之后应对评估数据进行检查，若发现评估用户群内各子群之间的差异大于预期差异，则需要为每个子群选取新的用户样本，再次进行评估。

7.6 评估实施及记录

依据评估大纲，组织评估实施，并准确记录过程评估数据，主要包括：

- a) 评估对象的名称、编号、试验日期及起止时间；
- b) 真实世界数据（RWD）采集的伦理审批与隐私合规审查（若适用）；
- c) 评估前的状态数据；
- d) 评估环境条件数据；
- e) 物理测量统计数据；
- f) 主观反馈数据；
- g) 评估完成后的状态数据等。

7.7 数据统计分析及结果输出

统计数据分布以及具体问题与建议，采用相应分析方法，对统计的客观、主观数据等进行分析，最终给出评估结论，可采用评估报告、数据库等多种形式输出。

8 指标评估规则及数据处理

8.1 评估规则

依据评估大纲，各级指标应按以下规则实施评估：

- a) 一级指标按8.2.3公式计算确定；
- b) 二级指标首先采用专业技术测评、用户体验测评或两者相结合的方法实施测评，再按8.2.1的评分原则进行评分；
- c) 三级指标（若适用）选择采用专业技术测评、用户体验测评或两者相结合的方法实施测评，作为二级指标评分的依据。

8.2 评估计算

8.2.1 评分原则

专业技术测评和用户体验测评统一采用5分制评分法。针对测评结果偏离测评要求的程度，每个二级评价指标设置5个评分等级：优（5分）、良（4分）、一般（3分）、较差（2分）、差（1分）。

8.2.2 评价指标权重

评价体系中的指标权重确立基于专家德尔菲法和层次分析法，权重分配详见表7和表8。

表7 侵入式脑机接口评价指标权重

一级指标	一级指标 权重%	二级指标		二级指标 权重%
安全与风险	30	基础安全和性能		30
		风险控制		10
		严重不良事件/不良事件率		15
		生物相容性与植入相关并发症率		20
		植入物取出过程的可控性与可追溯性		10
		故障响应		15
临床效用与 任务达成	25	目标任务完成率		30
		功能改善率		30
		真实世界有效使用率		20
		长期效用		20
信号与算法性能	15	探头质量与稳定性		10
		通道质量与稳定性		20
		信号质量		20
		算法性能		50
人因工程与可用性	10	人体适配性		20
		操作效率		20
		可用性		20
		舒适度		20
		用户体验		20
可靠性、维修性与 保障性	15	可靠性	平均故障间隔时间	10
			平均无故障工作时间	10
			电池续航与循环寿命	10
			侵入式系统的导线稳定性/修复 或翻修手术率（年化）	10
			长期稳定性	10
			环境适应性	10
		维修性与保障性	维修周转时间/现场支持	20
			软件管理	20
互操作、数据与网 络安全	5	互操作	术语/数据元一致性	15
			数据格式与导出	15
		数据与网络安全	数据安全与隐私保护的合法性	20
			数据透明度	15
			网络安全文档	15
			远程监测与控制	20

一级指标	一级指标 权重%	二级指标	二级指标 权重%
注 1：评估组织方依据评价目的和评价对象特征等因素直接使用或修改使用本表，其中，修改使用仅限于减少评价系统中二级指标的数量。			
注 2：评估组织方修改使用本表时，对表中的权重值重新进行归一处理，即当指标体系中删除某项指标时，其他各项指标相应调整。			

表 8 非侵入式脑机接口评价指标权重

一级指标	一级指标 权重%	二级指标	二级指标 权重%	
安全与风险	25	基础安全和性能	30	
		风险控制	10	
		严重不良事件/不良事件率	20	
		故障响应	20	
		佩戴相关皮肤安全性	20	
临床效用与 任务达成	25	目标任务完成率	30	
		功能改善率	30	
		真实世界有效使用率	20	
		长期效用	20	
信号与算法性能	15	探头质量与稳定性	10	
		通道质量与稳定性	20	
		信号质量	20	
		算法性能	50	
人因工程与可用性	15	人体适配性	20	
		操作效率	20	
		可用性	20	
		舒适度	20	
		用户体验	20	
可靠性、维修性与 保障性	15	可靠性	平均故障间隔时间	15
			平均无故障工作时间	15
			电池续航与循环寿命	10
			长期稳定性	10
			环境适应性	10
		维修性与保障性	维修周转时间/现场支持	20
			软件管理	20
互操作、数据与网 络安全	5	互操作	术语/数据元一致性	15
			数据格式与导出	15
		数据与 网络安全	数据安全与隐私保护的合法性	20
			数据透明度	15

一级指标	一级指标 权重%	二级指标		二级指标 权重%
			网络安全文档	15
			远程监测与控制	20
<p>注 1：评估组织方依据评价目的和评价对象特征等因素直接使用或修改使用本表，其中，修改使用仅限于减少评价系统中二级指标的数量。</p> <p>注 2：评估组织方修改使用本表时，对表中的权重值重新进行归一处理，即当指标体系中删除某项指标时，其他各项指标相应调整。</p>				

8.2.3 一级指标得分

根据评分原则，对专业技术测评中的各个二级评价指标进行5级评分。

某一级指标得分为其下一层级所有二级评价指标的综合得分，按公式（1）计算：

$$E_i = 20 \sum_{j=1}^n a_{ij} w_{ij} \quad \dots\dots\dots (1)$$

式中：

E_i ——第*i*个一级指标得分；

a_{ij} ——第*i*个一级指标下一层的第*j*个二级指标得分；

w_{ij} ——第*i*个一级指标下一层的第*j*个二级指标的权重（%），且 $\sum_{j=1}^n w_{ij}=100\%$ 。

8.2.4 综合评估得分

综合评估得分按公式（2）计算：

$$X = \sum_{i=1}^6 E_i W_i \quad \dots\dots\dots (2)$$

式中： W_i 为一级指标权重（%）； $\sum W_i = 100\%$

X ——综合评估得分；

E_i ——第*i*个一级指标得分；

W_i ——第*i*个一级指标的权重（%），且 $\sum_{i=1}^6 W_i = 100\%$ ；

9 综合评估结果

9.1 等级划分

综合评估结果宜划分为如下5个等级：

优秀（★★★★★）： $X \geq 90$ ；

良好（★★★★）： $80 \leq X < 90$ ；

中等（★★★）： $70 \leq X < 80$ ；

一般（★★）： $60 \leq X < 70$ ；

差（★）： $X < 60$ 。

等级规则可由组织实施机构结合评估目的进行修订或细化。

9.2 综合评估结果报告

综合评估结果报告应包括但不仅限于以下内容：

- 评估对象基本信息，包括但不限于名称、型号、制造商、临床使用场景及核心功能；
- 评估目标和范围；
- 评估任务及其适用场景；
- 评价指标及其权重和评分原则；
- 评估方案，包括但不限于评价指标的符合性判定标准、评估方法和评估设备；
- 参与评估的用户群信息，包括但不限于用户特征、用户总数量、用户群选取原因、每个用户群的关键特征和能力、如何挑选参与评价用户以及他们是否能够代表用户群的基本特征、参与评价用户和用户群之间的区别、参与评价用户特征（列）表；
- 评估得分和评估结果。

附录 A
(资料性附录)
原理说明

A.1 特殊章和条款的原理说明

以下是本部分中的特殊章和条款的原理说明，与本部分正文的章和条款相对应。

条款 5.3.1.1.1 a) 基本安全和基本性能

评价在正常状态与单一故障状态下，产品通过合理设计与规范制造具备消除对患者、使用者及他人造成不可接受风险的能力。

条款 5.3.1.1.1 b) 风险控制

评价在全使用周期内，产品具备通过风险管理活动将综合剩余风险降至可接受水平以及持续开展风险监测与改进的能力。

条款 5.3.1.1.1 c) 严重不良事件/不良事件率

评价与脑机接口医疗装备有关的、导致或可能导致患者、使用者或相关人员人体伤害的各种有害事件不良事件。

条款 5.3.1.1.1 d) 生物相容性

评价材料在特定应用与接触条件下引发适宜的用户反应的能力。

并发症包括手术部位感染、出血、癫痫发作、组织反应、皮肤破溃、装置移位或失效等。

条款 5.3.1.1.1 e) 植入物取出过程的可控性与可追溯性

需要移除的情况包括使用寿命结束、感染或设备升级、故障或不良反应、有计划取出/置换，以及临床研究要求的受控取出等。

条款 5.3.1.1.1 f) 故障响应

故障检测和响应机制是指当远程监测系统失灵、通信中断或异常指令输入时，设备应能触发报警并自动进入安全状态。例如，若通信丢失，设备须有超时保护，停止刺激输出或维持在安全阈值以内。

故障报警和响应记录中须记录其在各种故障场景下是否按预期触发报警或进入安全模式，核查风险分析是否涵盖远程交互风险及相应缓解措施。

条款 5.3.1.1.2 e) 佩戴相关皮肤安全性

针对长时间佩戴可能引起的物理压迫、摩擦损伤、汗液积聚、过敏反应等问题，佩戴相关皮肤安全性用于考察用户在典型使用周期（如连续佩戴 2 h-8 h）内的皮肤耐受性。

条款 5.3.2.1.1 a) 目标任务完成率

目标任务完成率用于考察用户在标准化测试中完成规定目标任务的准确性与效率。

条款 5.3.2.1.1 b) 功能改善率

功能改善率的评价内容包含用户自我报告的健康状态、功能独立性及生活质量的改善程度。

条款 5.3.2.1.1 c) 真实世界有效使用率

真实世界有效使用率评价内容包含用户在家庭或社区环境中主动、有效使用系统的模式与持续性。

条款 5.3.2.1.1 d) 长期效应

长期效应评价内容包括系统在扩展时间范围内被持续使用的程度，以及停止使用的原因分析。

条款 5.3.3.1 a) 探头性能与稳定性

探头性能与稳定性用于考察脑机接口探头的各项基本性能，评价内容包括但不限于灵敏度、信号一致性、电极间阻抗与稳定性等参数。

条款 5.3.3.1 b) 信道质量与稳定性

信道质量与稳定性用于考察脑机接口设备对信道的使用能力，评价内容包括但不限于信道有效率、带宽利用率、掉线率、信息传输率等参数。

条款 5.3.3.1 c) 信号质量

信号质量用于考察信号中有效信息与无用干扰的相对强度、信号质量改善程度以及脑机接口设备的信号稳定性，评价内容包括但不限于信噪比、基线稳定性、伪迹抑制能力等参数。

条款 5.3.3.1 d) 算法性能

在实时操作中，算法正确识别用户意图或分类信号的比例，和在用户状态变化、存在环境干扰等情况下维持性能的能力，如分类准确率、算法鲁棒性、泛化能力、初始化校准时间、变异系数、性能衰减等。

条款 5.3.4.1 a) 人体适配性

人体适配性用于考察产品在尺寸、形态、力学、生理、行为等维度上与人体特征及其需求相匹配的程度。

条款 5.3.4.1 b) 操作效率

操作效率用于考察产品可被有效、高效地使用的程度。

条款 5.3.4.1 c) 可用性

可用性用于考察产品可被有效、高效且令用户满意使用的程度。

条款 5.3.4.1 d) 舒适度

舒适度用于考察用户在佩戴与使用期间的主观舒适感受与可耐受性，评价内容包括但不限于重量/压感、触感/夹紧/摩擦、热感/出汗、移动限制等维度。

条款 5.3.4.1 e) 用户体验

用户体验用于考察用户在产品体验过程中的综合主观感受。

附录 B
(资料性附录)
二级评价指标量化评分规则

B.1 安全与风险相关二级评价指标量化评分规则

脑机接口设备安全与风险相关二级评价指标的评价内容、数据采集方法和评价结果证明材料的要求分辨见5.3.1.1和5.3.1.2。安全与风险相关二级评价指标的量化评分规则见表B.1。

表B.1 安全与风险相关二级评价指标量化评分规则一览表

序号	二级评价指标	评分规则	说明	
1	基本安全和基本性能	• 符合评估方案中全部基本安全和基本性能指标要求；	5分	剩余风险的严重度由评估组织方在评估方案中的风险可接受性准则确定。
		• 非诊疗功能相关指标不符合评估方案要求，且存在补充风险控制措施后可消除的可忽略的或轻微的剩余风险；	4分	
		• 非诊疗功能相关指标不符合评估方案要求，且存在补充风险控制措施后不可消除的可忽略的或轻微的剩余风险；或，临床诊疗功能相关指标不符合评估方案要求，且存在补充风险控制措施后可消除的可忽略的或轻微的剩余风险；	3分	
		• 诊疗功能相关指标不符合评估方案要求，且存在补充风险控制措施后可消除的严重及其以上程度的剩余风险；	2分	
		• 诊疗功能相关指标不符合评估方案要求，且存在不可消除的严重及其以上程度的剩余风险；或无第三方检验报告。	1分	
2	风险控制	• 风险管理文档完整覆盖 GB/T 42062—2022 全流程，结果证明材料真实、准确和完整，经多学科评审仅存在可忽略的或轻微的剩余风险。	5分	
		• 结果证明材料真实且基本准确和完整，个别环节内容的准确性和完整性可修订，经多学科评审仅存在可忽略的或轻微的剩余风险。	4分	
		• 结果证明材料的真实性、准确性和完整性中任意一项存在严重缺陷，经多学科评审存在严重剩余风险，但可通过补充风险控制措施降低至可接受水平。	3分	
		• 结果证明材料的真实性、准确性和完整性中任意一项存在严重缺陷，经多学科评审仅存在不可接受的剩余风险。	2分	
		• 无结果证明材料。	1分	
3	严重不良事件/不良事件率	• ≤ 0.1 例/百台·月，且均为轻微可逆事件。	5分	严重不良事件是指指导致死亡、危及生命、需住院或延长住院时
		• 0.1 - 0.5 例/百台·月，无严重伤害。	4分	
		• 0.5 - 1.0 例/百台·月，偶有中度影响。	3分	

序号	二级评价指标	评分规则	说明
		<ul style="list-style-type: none"> • 1.0 - 2.0 例/百台·月，含用户焦虑或皮肤损伤等。 • >2.0 例/百台·月，或发生严重不良事件。 	2分 1分 间、导致永久性功能丧失或缺陷的事件。
4	生物相容性与植入相关并发症率	<ul style="list-style-type: none"> • 提供符合 GB/T 16886.1 的系列生物相容性评价报告，覆盖所有接触材料与预期接触时间。 • 生物相容性评价报告基本齐全，仅缺少一项测试，其等效替代证据的剩余风险经多学科评审判定为可接受。 • 仅提供部分生物相容性评价报告，且无等效替代证据，经多学科评审后长期接触风险可接受。 • 仅提供材料供应商声明，无实际测试数据，经多学科评审后长期接触风险可接受。 • 经对提供材料进行多学科评审后长期接触风险不可接受，或提供未经评估材料，或存在已知致敏/刺激案例。 	5分 4分 3分 2分 1分
5	植入物取出过程的可控性与可追溯性	<ul style="list-style-type: none"> • 取出过程安全便捷且追溯系统实时精准。 • 取出过程的安全便捷性以及追溯系统的实时性和精准性存在轻微缺陷，其导致的剩余风险经多学科评审可接受。 • 取出过程的安全便捷性以及追溯系统的实时性和精准性存在较严重缺陷，其导致的剩余风险经多学科评审可接受。 • 体系存在严重缺陷，即指南不清晰和/或追溯系统不完善，其导致的剩余风险经多学科评审不可接受。 • 存在根本性风险，即植入物无法安全取出或追溯系统失效。 	5分 4分 3分 2分 1分
6	故障响应	<ul style="list-style-type: none"> • 具有智能故障预测和早期提醒功能以及主动自愈能力，且故障发生率及其处理后修复能力满足评估方案要求。 • 不具有智能故障预测和早期提醒功能以及主动自愈能力，但故障发生率及其处理能力的鲁棒性满足评估方案要求。 • 故障发生时能自动进入预定义的安全状态，具有明确报警指示，且风险管理文件涵盖远程交互风险及其相应控制措施。 • 故障发生时无报警或报警不明确，产品故障状态无明确定义，且相关风险分析内容缺失。 • 故障时产品出现未定义的、随机的、或危险的非预期动作，且无报警。 	5分 4分 3分 2分 1分
7	佩戴相关皮肤安全性	<ul style="list-style-type: none"> • 采用化学与生物安全性良好、压力分布测试及长期人体佩戴研究结果优良、用户认可度高的材料。 	5分

序号	二级评价指标	评分规则	说明
		<ul style="list-style-type: none"> • 化学与生物安全性评价结果合格，且设计舒适度评价表现良好，以及工程测试数据可客观佐证其设计能有效分散压力、预防压疮风险。 	4分
		<ul style="list-style-type: none"> • 化学与生物安全性评价结果合格，且预期使用时长内无普遍性不适，以及设计上无已知的、会导致明显压迫疼痛的尖锐点或硬质边缘。 	3分
		<ul style="list-style-type: none"> • 使用过程中持续收到生物相容性和/或物理性风险相关的不良反应事件报告。 	2分
		<ul style="list-style-type: none"> • 生物相容性评价结果不合格，或出现与产品相关的需要医疗介入的物理性损伤。 	1分

B.2 临床效用与任务达成二级评价指标量化评分规则

脑机接口设备临床效用与任务达成相关二级评价指标的评价内容、数据采集方法和评价结果证明材料的要求分辨见5.3.2.1和5.3.2.2。临床效用与任务达成相关二级评价指标的量化评分规则见表B.2。

表B.2 临床效用与任务达成相关二级评价指标量化评分规则一览表

序号	二级评价指标	评分规则	说明
1	目标任务完成率	按评估组织方评估方案中的方法开展测试和统计，目标任务完成率为： <ul style="list-style-type: none"> • $\geq 95\%$; • 90% ~ 94%; • 80% ~ 89%; • 70% ~ 79%; • $< 70\%$。 	测试与统计方法示例 详见附录 X.1。 5分 4分 3分 2分 1分
		按评估组织方评估方案中的方法开展测试和统计，功能改善率为： <ul style="list-style-type: none"> • $\geq 90\%$; • 85% ~ 89%; • 80% ~ 84%; • 70% ~ 79%; • $< 70\%$。 	测试与统计方法示例 详见附录 X.2。 5分 4分 3分 2分 1分
2	功能改善率	按评估组织方评估方案中的方法开展测试和统计，功能改善率为： <ul style="list-style-type: none"> • 提升 $\geq 30\%$; • 提升 20% ~ 29%; • 提升 10% ~ 19%; • 提升 5% ~ 9%; • 提升 $< 5\%$。 	测试与统计方法示例 详见附录 X.3。 5分 4分 3分 2分 1分
		按评估组织方评估方案中的方法开展测试和统计，功能改善率为：	测试与统计方法示例 详见附录 X.4。

序号	二级评价指标	评分规则			说明
			<ul style="list-style-type: none"> • 提升$\geq 25\%$; • 提升 15% ~ 24%; • 提升 10% ~ 14%; • 提升 5% ~ 9%; • 提升$< 5\%$。 	5分 4分 3分 2分 1分	
3	真实世界有效使用率	侵入式	按评估组织方评估方案中的方法开展测试和统计，真实世界有效使用率为： <ul style="list-style-type: none"> • $\geq 95\%$; • 90% ~ 94%; • 80% ~ 89%; • 70% ~ 79%; • $< 70\%$。 	5分 4分 3分 2分 1分	测试与统计方法示例 详见附录 X. 5。
		非侵入式	按评估组织方评估方案中的方法开展测试和统计，真实世界有效使用率为： <ul style="list-style-type: none"> • $\geq 80\%$; • 60% ~ 79%; • 40% ~ 59%; • 20% ~ 39%; • $< 20\%$。 	5分 4分 3分 2分 1分	测试与统计方法示例 详见附录 X. 6。
4	长期效用	侵入式	按评估组织方评估方案中的方法开展测试和统计，12个月随访期内持续使用受试者比例为： <ul style="list-style-type: none"> • $\geq 95\%$; • 90% ~ 94%; • 80% ~ 89%; • 50% ~ 79%; • $< 50\%$。 	5分 4分 3分 2分 1分	测试与统计方法示例 详见附录 X. 7。
		非侵入式	按评估组织方评估方案中的方法开展测试和统计，6个月随访期内持续使用受试者比例为： <ul style="list-style-type: none"> • $\geq 80\%$; • 65% ~ 79%; • 50% ~ 64%; • 40% ~ 49%; • $< 40\%$。 	5分 4分 3分 2分 1分	测试与统计方法示例 详见附录 X. 8。

B.3 信号与算法性能相关二级评价指标量化评分规则

脑机接口设备信号与算法性能相关二级评价指标的评价内容、数据采集方法和评价结果证明材料的要求分辨见5.3.3.1和5.3.3.2。信号与算法性能相关二级评价指标的量化评分规则见表B.3。

表B.3 信号与算法性能相关二级评价指标量化评分规则一览表

指标	评分规则		说明
灵敏度	在目标工作段对“输入-输出”曲线做线性拟合，斜率 k 的归一化评分值： <ul style="list-style-type: none"> · ≥ 1.25 5分 · 1.10~1.25 4分 · 0.90~1.10 3分 · 0.70~0.90 2分 · < 0.70 1分 		斜率归一化公式： $k = k_{\text{测量值}} / k_{\text{标准均值}}$
信号一致性	CCC（一致性相关系数）： <ul style="list-style-type: none"> · ≥ 0.85 5分 · 0.75~0.85 4分 · 0.6~0.74 3分 · 0.4~0.59 2分 · < 0.4 1分 		基于 CCC（一致性相关系数）， $CCC = (2 \cdot \text{Cov}(x, y)) / (\text{Var}(x) + \text{Var}(y) + (\mu_x - \mu_y)^2)$ ，其中， μ_x 和 μ_y 为均值； $\text{Var}(x)$ 和 $\text{Var}(y)$ 为方差； $\text{Cov}(x, y)$ 为协方差。
EEG 电极的 电极间阻抗	通用 评估	湿电极 采用 10Hz 测试信号测量时，医用湿电极电阻值的中位数： <ul style="list-style-type: none"> · $\leq 5 \text{ k}\Omega$ 5分 · 5~10 (kΩ) 4分 · 10~50 (kΩ) 3分 · 50~100 (kΩ) 2分 · $\geq 100 \text{ k}\Omega$ 1分 	采用 10Hz 测试信号测量时，医用湿电极电阻值不大于 10k Ω 时被认为与人体皮肤接触良好。
		主动干电极 采用 10Hz 或 30Hz 测试信号测量时，主动干电极电阻值的中位数： <ul style="list-style-type: none"> · $\leq 100 \text{ k}\Omega$ 5分 · 100~200 (kΩ) 4分 · 200~500 (kΩ) 3分 · 500~1000 (kΩ) 2分 · $\geq 1000 \text{ k}\Omega$ 1分 	
		被动干电极 被动干电极电阻值的中位数： <ul style="list-style-type: none"> · $\leq 200 \text{ k}\Omega$ 5分 · 200~500 (kΩ) 4分 · 500~1000 (kΩ) 3分 · 1000~2000 (kΩ) 2分 · $\geq 2000 \text{ k}\Omega$ 1分 	

指标	评分规则			说明
专用评估	P300	湿电极的关键电极为 Fz, Cz, Pz, 其阻抗值需满足: <ul style="list-style-type: none"> · $\leq 5 \text{ k}\Omega$ · 5~10 (kΩ) · 10~50 (kΩ) · 50~100 (kΩ) · $\geq 100 \text{ k}\Omega$ 	5分 4分 3分 2分 1分	<ul style="list-style-type: none"> · 参考电极和地线阻抗应低于测量电极。 · 参考电极或地线阻抗超标,即为系统不可用,所有评分无效。 · 任一关键电极阻抗超标但其他指标合格,需总评分下调1档。 · 同侧电极间阻抗差应$\leq 50\%$。 · 对称电极(如 C3/C4)阻抗差应$\leq 30\%$,阻抗差$>50\%$,不满足需下调1档并提示检查佩戴情况。
	SSVEP	湿电极的关键电极为 O1, O2, Oz, 其阻抗值需满足: <ul style="list-style-type: none"> · $\leq 5 \text{ k}\Omega$ · 5~10 (kΩ) · 10~50 (kΩ) · 50~100 (kΩ) · $\geq 100 \text{ k}\Omega$ 	5分 4分 3分 2分 1分	阻抗测量应在佩戴稳定 5 min 后进行。
	RSVP	湿电极的关键电极为 Pz, POz, Oz, 其阻抗值需满足: <ul style="list-style-type: none"> · $\leq 8 \text{ k}\Omega$ · 8~15 (kΩ) · 15~80 (kΩ) · 80~150 (kΩ) · $\geq 150 \text{ k}\Omega$ 	5分 4分 3分 2分 1分	
	运动想象	湿电极的关键电极为 C3, C4, Cz, 其阻抗值需满足: <ul style="list-style-type: none"> · $\leq 8 \text{ k}\Omega$ · 8~15 (kΩ) · 15~80 (kΩ) · 80~150 (kΩ) · $\geq 150 \text{ k}\Omega$ 	5分 4分 3分 2分 1分	
	所有范式	关键电极为参考电极、地线,其阻抗值需满足: <ul style="list-style-type: none"> · $\leq 3 \text{ k}\Omega$ · 3 kΩ~6 (kΩ) · 6 kΩ~30 (kΩ) · 30~60 (kΩ) · $\geq 60 \text{ k}\Omega$ 	5分 4分 3分 2分 1分	

指标	评分规则			说明
EEG 电极的 阻抗稳定性	通用 评估	1 小时内阻抗变化率，取中位数： · ≤10% · 10%~20% · 20%~30% · 30%~50% · >50%		5 分 4 分 3 分 2 分 1 分 $Z = Z(60\text{min}) - Z(0) / Z(0) \times 100\%$
	专用 评估	湿电极 (凝胶)	1 小时内阻抗变化率，取中位数： · ≤10% · 10%~20% · 20%~30% · 30%~50% · >50%	5 分 4 分 3 分 2 分 1 分 $Z = Z(60\text{min}) - Z(0) / Z(0) \times 100\%$
		湿电极 (盐水)	1 小时内阻抗变化率，取中位数： · ≤15% · 15%~25% · 25%~35% · 35%~55% · >55%	5 分 4 分 3 分 2 分 1 分
		主动干 电极	1 小时内阻抗变化率，取中位数： · ≤15% · 15%~25% · 25%~35% · 35%~55% · >55%	5 分 4 分 3 分 2 分 1 分
		被动干 电极	1 小时内阻抗变化率，取中位数： · ≤20% · 20%~30% · 30%~40% · 40%~60% · >60%	5 分 4 分 3 分 2 分 1 分
		P300	session 内阻抗变化率，取中位数： · ≤10% · 10%~20% · 20%~30% · 30%~50% · >50%	5 分 4 分 3 分 2 分 1 分 $Z = Z(60\text{min}) - Z(0) / Z(0) \times 100\%$ ， 其中，典型 session 时长 20 min~40 min，推荐测试时长 1 h。

指标	评分规则			说明	
		SSVEP	session 内阻抗变化率，取中位数： <ul style="list-style-type: none"> · ≤15% · 15%~25% · 25%~35% · 35%~55% · >55% 	5分 4分 3分 2分 1分	$Z = Z(60\text{min}) - Z(0) / Z(0) \times 100\%$ ， 其中，典型 session 时长 10 min~20 min，推荐测试时长 1 h。
		RSVP	session 内阻抗变化率，取中位数： <ul style="list-style-type: none"> · ≤10% · 10%~20% · 20%~30% · 30%~50% · >50% 	5分 4分 3分 2分 1分	$Z = Z(60\text{min}) - Z(0) / Z(0) \times 100\%$ ， 其中，典型 session 时长 30 min~60 min，推荐测试时长 1 h。
		运动想象	session 内阻抗变化率，取中位数： <ul style="list-style-type: none"> · ≤5% · 5%~15% · 15%~25% · 25%~45% · >45% 	5分 4分 3分 2分 1分	$Z = Z(60\text{min}) - Z(0) / Z(0) \times 100\%$ ， 其中，典型 session 时长 30 min~60 min，推荐测试时长 2 h。
tES 电极	通用评估	tES 电极电流密度	所有电极的平均电流密度： <ul style="list-style-type: none"> · <50 $\mu\text{A}/\text{cm}^2$ · 50~80 ($\mu\text{A}/\text{cm}^2$) · 80~100 ($\mu\text{A}/\text{cm}^2$) · 100~200 ($\mu\text{A}/\text{cm}^2$) · >200 $\mu\text{A}/\text{cm}^2$ 	5分 4分 3分 2分 1分	电流密度=总电流/单电极有效面积
		tES 电极刺激电流稳定性	连续刺激 30 min，每分钟记录一次当前电流值，刺激结束后统计电流标准差，取所有电极的平均值： <ul style="list-style-type: none"> · ≤2% · 2%~5% · 5%~10% · 10%~15% · >15% 	5分 4分 3分 2分 1分	电流稳定性= 电流标准差 /目标电流×100%
		tES 电极皮肤不良反应率	连续刺激 30 min，期间若被试出现红斑、瘙痒、灼热感即为皮肤不适，统计出现皮肤不适的受试者数量，计算不良反应率： <ul style="list-style-type: none"> · ≤5% · 5%~15% · 15%~25% · 25%~35% · >35% 	5分 4分 3分 2分 1分	不良反应率=出现皮肤不适的受试者数量/被试总数量×100%

指标	评分规则			说明	
侵入式电极	通用评估	侵入式电极有效通道率	统计可记录单神经元活动的通道数量，计算有效通道率： · $\geq 90\%$ · 80%~90% · 70%~80% · 60%~70% · $<60\%$	5分 4分 3分 2分 1分	有效通道率：可记录单神经元活动的通道数量/总通道数量
		侵入式电极阻抗变化率	术后 6 - 12 个月测量阻抗变化率，取所有被试的平均值： · $\leq 20\%$ · 20%~30% · 30%~40% · 40%~50% · $>50\%$	5分 4分 3分 2分 1分	阻抗变化率= 当前通道阻抗-基线阻抗 /基线阻抗，其中：当前通道阻抗为当前所有通道阻抗的中位数；基线阻抗为植入后第 7 天所有通道阻抗的中位数。
		侵入式电极胶质瘢痕厚度	术后 6 - 12 个月测量电极胶质瘢痕厚度，取 MRI 或组织学切片测量电极周围 100 μm 范围内平均厚度： · $\leq 50 \mu\text{m}$ · 50~100 (μm) · 100~150 (μm) · 150~200 (μm) · $>200 \mu\text{m}$ 或明显神经元丢失	5分 4分 3分 2分 1分	瘢痕厚度为每个被试基于自身数据独立计算，最终取值结果为所有被试的平均值。
		侵入式电极长期信号衰减	在同一任务下计算 ITR 的均值： · $\leq 10\%$ · 10%~20% · 20%~30% · 30%~40% · $>40\%$	5分 4分 3分 2分 1分	信号衰减率=(1-当前 ITR/峰值 ITR)*100%，其中：当前 ITR=术后第 350 - 380 天内的 7 天平均 ITR；峰值 ITR=植入后第 14~60 天内连续 7 天最高单日 ITR。
信道有效率	通用评估	统计有效信道数量，计算信道有效率： · $\geq 90\%$ · 85%~89% · 80%~84% · 75%~79% · $<75\%$	5分 4分 3分 2分 1分	· 信道有效率=有效信道数/总信道数 $\times 100\%$ · 有效定义：阻抗合格且 SNR 达标	
	专用评估	P300 统计有效信道数量，计算信道有效率： · $\geq 95\%$ · 90%~94% · 85%~89% · 80%~84% · $<80\%$	5分 4分 3分 2分 1分	· 信道有效率=有效信道数/总信道数 $\times 100\%$ · 关键信道 (Fz/Cz/Pz) 必须为有效信道。	

指标	评分规则			说明	
		SSVEP	统计有效信道数量, 计算信道有效率: <ul style="list-style-type: none"> · $\geq 92\%$ · 88%~91% · 83%~87% · 78%~82% · $< 78\%$ 	5分 4分 3分 2分 1分	<ul style="list-style-type: none"> · 信道有效率=有效信道数/总信道数$\times 100\%$ · 枕区关键信道(O1/O2/Oz) 必须为有效信道。
		RSVP	统计有效信道数量, 计算信道有效率: <ul style="list-style-type: none"> · $\geq 93\%$ · 88%~92% · 83%~87% · 78%~82% · $< 78\%$ 	5分 4分 3分 2分 1分	<ul style="list-style-type: none"> · 信道有效率=有效信道数/总信道数$\times 100\%$ · 后部关键信道(Pz/P0z/Oz) 必须为有效信道。
		运动想象	统计有效信道数量, 计算信道有效率: <ul style="list-style-type: none"> · $\geq 90\%$ · 85%~89% · 80%~84% · 75%~79% · $< 75\%$ 	5分 4分 3分 2分 1分	<ul style="list-style-type: none"> · 信道有效率=有效信道数/总信道数$\times 100\%$ · 感觉运动区关键信道(C3/C4/Cz) 必须为有效信道。
带宽利用率	通用评估	所有被试带宽利用率的平均值: <ul style="list-style-type: none"> · $\geq 80\%$ · 70%~80% · 60%~70% · 50%~60% · $< 50\%$ 	5分 4分 3分 2分 1分	带宽利用率=实际 ITR/理论最大 ITR, 其中: 理论最大 ITR 为 100%准确率、无空闲时间下的理论值。	
掉线率	通用评估	连续记录不少于 1 小时, 统计数据丢失时长的占比: <ul style="list-style-type: none"> · $\leq 0.1\%$ · 0.1%~0.5% · 0.5%~1% · 1%~2% · $> 2\%$ 	5分 4分 3分 2分 1分	掉线率=数据丢失时长/总记录时长 $\times 100\%$	
信息传输率	通用评估	所有被试的平均信息传输率: <ul style="list-style-type: none"> · ≥ 60 bit/min · 40~59 (bit/min) · 20~39 (bit/min) · 10~19 (bit/min) · < 10 bit/min 	5分 4分 3分 2分 1分	$ITR = [\log_2 N + P \times \log_2 P + (1-P) \times \log_2 [(1-P)/(N-1)]] \times (60/T)$ 其中: N 为目标数, P 为准确率, T 为单次 Trial 所需时间(s), ITR 的单位为 bit/min	

指标	评分规则			说明	
专用评估	P300	所有被试的平均信息传输率： · ≥ 40 bit/min · 30~39 (bit/min) · 20~29 (bit/min) · 10~19 (bit/min) · < 10 bit/min	5分 4分 3分 2分 1分		
	SSVEP	所有被试的平均信息传输率： · ≥ 150 bit/min · 100~149 (bit/min) · 60~99 (bit/min) · 30~59 (bit/min) · < 30 bit/min	5分 4分 3分 2分 1分		
	运动想象	所有被试的平均信息传输率： · ≥ 20 bit/min · 15~19 (bit/min) · 10~14 (bit/min) · 5~9 (bit/min) · < 5 bit/min	5分 4分 3分 2分 1分		
信噪比	通用评估	信噪比： · ≥ 20 dB · 15~19.9 (dB) · 10~14.9 (dB) · 5~9.9 (dB) · < 5 dB	5分 4分 3分 2分 1分	$SNR(dB) = 10 \cdot \log_{10}(P_{\text{signal}}/P_{\text{noise}})$, 其中, P_{signal} 是信号的功率, P_{noise} 是噪声的功率	
	专用评估	P300	针对 Cz/Pz 电极, 计算不少于 30 次叠加后的峰值信噪比： · ≥ 15 dB · 12~14.9 (dB) · 9~11.9 (dB) · 6~8.9 (dB) · < 6 dB	5分 4分 3分 2分 1分	峰值 $SNR=20 \times \log(A_{\text{peak}}/\sigma_{\text{baseline}})$
		SSVEP	针对 O1/O2/Oz 电极, 计算刺激频率与邻近频段的信噪比： · ≥ 25 dB · 20~24.9 (dB) · 15~19.9 (dB) · 10~14.9 (dB) · < 10 dB	5分 4分 3分 2分 1分	$SNR=10 \times \log[P_{f_{\text{stim}}}/P_{f_{\text{neighbor}}}]$

指标	评分规则			说明	
		RSVP	针对 Pz/Oz 电极，计算峰值信噪比： <ul style="list-style-type: none"> · ≥ 13 dB · 10~12.9 (dB) · 7~9.9 (dB) · 5~6.9 (dB) · < 5 dB 	5分 4分 3分 2分 1分	峰值 $SNR=20 \times \log(ERP_{peak}/\sigma_{baseline})$
		运动想象	针对 C3/C4 电极，做 ERD/ERS 分析，计算 8-30Hz 波段的信噪比： <ul style="list-style-type: none"> · ≥ 15 dB · 12~14.9 (dB) · 9~11.9 (dB) · 6~8.9 (dB) · < 6 dB 	5分 4分 3分 2分 1分	$SNR=10 \times \log_{10}[P_{task}/P_{baseline}]$
		侵入式电极信噪比	信噪比： <ul style="list-style-type: none"> · ≥ 12 dB · 10~12 (dB) · 8~10 (dB) · 6~8 (dB) · < 6 dB 	5分 4分 3分 2分 1分	$SNR=20 \times \log_{10}$ (信号强度/噪声强度), 其中: 信号强度=所有通道神经元平均模板波峰-波谷平均幅度; 噪声强度为所有通道背景噪声标准差的均值。
基线稳定性	通用评估	测量 30 分钟静息态数据的线性趋势斜率： <ul style="list-style-type: none"> · ≤ 2 $\mu V/min$ · 2~5 ($\mu V/min$) · 5~8 ($\mu V/min$) · 8~12 ($\mu V/min$) · > 12 $\mu V/min$ 	5分 4分 3分 2分 1分	医用湿电极 (Ag/AgCl): 使用标准值 主动干电极: 标准 $\times 1.5$ 被动干电极: 标准 $\times 2$ 盐水电极: 标准 $\times 1.2$ 侵入式电极: 标准 $\div 2$	
	专用评估	P300	测量 30 分钟静息态数据的线性趋势斜率： <ul style="list-style-type: none"> · ≤ 1 $\mu V/min$ · 1~3 ($\mu V/min$) · 3~5 ($\mu V/min$) · 5~8 ($\mu V/min$) · > 8 $\mu V/min$ 	5分 4分 3分 2分 1分	
	专用评估	SSVEP	测量 30 分钟静息态数据的线性趋势斜率： <ul style="list-style-type: none"> · ≤ 2 $\mu V/min$ · 2~5 ($\mu V/min$) · 5~8 ($\mu V/min$) · 8~12 ($\mu V/min$) · > 12 $\mu V/min$ 	5分 4分 3分 2分 1分	

指标	评分规则			说明
		RSVP	测量 30 分钟静息态数据的线性趋势斜率： <ul style="list-style-type: none"> · $\leq 1 \mu\text{V}/\text{min}$ 5 分 · 1~3 ($\mu\text{V}/\text{min}$) 4 分 · 3~5 ($\mu\text{V}/\text{min}$) 3 分 · 5~8 ($\mu\text{V}/\text{min}$) 2 分 · $> 8 \mu\text{V}/\text{min}$ 1 分 	
		运动想象	测量 30 分钟静息态数据的线性趋势斜率： <ul style="list-style-type: none"> · $\leq 1 \mu\text{V}/\text{min}$ 5 分 · 1~3 ($\mu\text{V}/\text{min}$) 4 分 · 3~5 ($\mu\text{V}/\text{min}$) 3 分 · 5~8 ($\mu\text{V}/\text{min}$) 2 分 · $> 8 \mu\text{V}/\text{min}$ 1 分 	
伪迹抑制能力		通用评估	采用标准测试集（包含眨眼/肌电/心电各类伪迹），计算综合评分： <ul style="list-style-type: none"> · ≥ 90 分 5 分 · 80~89（分） 4 分 · 70~79（分） 3 分 · 60~69（分） 2 分 · < 60 分 1 分 	综合评分= $0.4 \times \text{ASR 分} + 0.3 \times \text{检测准确率分} + 0.3 \times \text{失真度分}$ 其中： <ul style="list-style-type: none"> · ASR 测量：使用标准伪迹数据集（推荐 EEGdenoiseNet）或诱发测试（受试者故意眨眼 20 次、咬牙 10 次），计算处理前后幅值比； · 检测准确率：使用专家标注的混合数据（纯净+伪迹段各 50%），计算分类准确率； · 失真度：使用仿真数据（已知 EEG+添加伪迹），计算处理后与真值的 RMSE； · 评分换算：ASR (dB) \rightarrow 分数：20dB=100 分，10dB=60 分，线性插值；检测率 (%) \rightarrow 分数：95%=100 分，80%=60 分；失真度 (%) \rightarrow 分数：5%=100 分，20%=0 分。
		专用评估	P300 综合评分，重点评估眨眼和眼动伪迹（额部电极）： <ul style="list-style-type: none"> · ≥ 88 分 5 分 · 78~87（分） 4 分 · 68~77（分） 3 分 · 58~67（分） 2 分 · < 58 分 1 分 	

指标	评分规则			说明	
		SSVEP	综合评分，重点评估肌电伪迹（频率混淆）： <ul style="list-style-type: none"> · ≥ 85 分 · 75~84（分） · 65~74（分） · 55~64（分） · < 55 分 	5分 4分 3分 2分 1分	
		RSVP	综合评分，重点评估眨眼伪迹（快速呈现易诱发）： <ul style="list-style-type: none"> · ≥ 90 分 · 80~89（分） · 70~79（分） · 60~69（分） · < 60 分 	5分 4分 3分 2分 1分	
		运动想象	综合评分，重点评估肌电伪迹（动作想象 vs 实际动作）： <ul style="list-style-type: none"> · ≥ 92 分 · 82~91（分） · 72~81（分） · 62~71（分） · < 62 分 	5分 4分 3分 2分 1分	
分类准确率	通用评估	F1 分数： <ul style="list-style-type: none"> · ≥ 0.85 · 0.75~0.84 · 0.65~0.74 · 0.55~0.64 · < 0.55 	5分 4分 3分 2分 1分	$F1 \text{ 分数} = 2 \times (\text{精确率}[\text{Precision}] \times \text{召回率}[\text{Recall}] / (\text{精确率}[\text{Precision}] + \text{召回率}[\text{Recall}]))$ 其中，精确率[Precision]=TP/(TP + FP)；召回率[Recall]=TP/(TP + FN)；TP 为真阳性计数；TN 为真阴性计数；FP 为假阳性计数；FN 为假阴性计数。	
	专用评估	P300	多分类平均 F1 或宏平均 F1： <ul style="list-style-type: none"> · ≥ 0.88 · 0.80~0.87 · 0.70~0.79 · 0.60~0.69 · < 0.60 	5分 4分 3分 2分 1分	
		SSVEP	多类别宏平均 F1： <ul style="list-style-type: none"> · ≥ 0.93 · 0.86~0.92 · 0.78~0.85 · 0.70~0.77 · < 0.70 	5分 4分 3分 2分 1分	

指标	评分规则			说明
		运动想象	双分类 F1 或多类宏平均 F1: · ≥ 0.78 · 0.68~0.77 · 0.63~0.67 · 0.58~0.62 · < 0.58	5 分 4 分 3 分 2 分 1 分
系统时延	通用评估		统计刺激/任务开始到输出决策的时间，取中位数: · ≤ 300 ms · 301~400 (ms) · 401~500 (ms) · 501~700 (ms) · > 700 ms	5 分 4 分 3 分 2 分 1 分
		P300	统计刺激序列结束到字符输出的时间，取中位数: · ≤ 200 ms · 201~300 (ms) · 301~400 (ms) · 401~500 (ms) · > 500 ms	5 分 4 分 3 分 2 分 1 分
		SSVEP	统计凝视结束到命令输出的时间，取中位数: · ≤ 150 ms · 151~250 (ms) · 251~350 (ms) · 351~500 (ms) · > 500 ms	5 分 4 分 3 分 2 分 1 分
		RSVP	统计目标图像呈现到标记输出的时间，取中位数: · ≤ 300 ms · 301~450 ms · 451~600 ms · 601~800 ms · > 800 ms	5 分 4 分 3 分 2 分 1 分
		运动想象	统计提示后到分类输出的时间（不含想象期），取中位数: · ≤ 400 ms · 401~600 (ms) · 601~800 (ms) · 801~1000 (ms) · > 1000 ms	5 分 4 分 3 分 2 分 1 分

指标	评分规则		说明	
算法鲁棒性	通用评估	计算干扰或状态变化条件下的性能保持率： <ul style="list-style-type: none"> · >90% · 80% ~ <90% · 70% ~ <80% · 60% ~ <70% · ≤60% 	5分 4分 3分 2分 1分 性能保持率=(干扰或状态变化条件下性能 (F1-score))/(无干扰条件下基线性能 (F1-score))×100%	
泛化能力	通用评估	计算跨域情况下的性能保持率： <ul style="list-style-type: none"> · >90% · 80% ~ <90% · 70% ~ <80% · 60% ~ <70% · ≤60% 	5分 4分 3分 2分 1分 性能保持率=(跨域性能 (F1-score))/(基线性能 (F1-score))×100%	
初始化校准时间	通用评估	统计从用户端数据采集到模型训练完成的总时间，包含数据收集和训练： <ul style="list-style-type: none"> · ≤8 min · 8~12 (min) · 12~18 (min) · 18~25 (min) · >25 min 	5分 4分 3分 2分 1分	
	专用评估	P300	统计从用户端数据采集到模型训练完成的总时间，包含收集训练数据和模型训练，至少拼写 5-10 个字符： <ul style="list-style-type: none"> · ≤5 min · 5~8 (min) · 8~12 (min) · 12~18 (min) · >18 min 	5分 4分 3分 2分 1分
		SSVEP	统计从用户端数据采集到模型训练完成的总时间，包含各频率数据采集和模型训练，每频率至少 10 秒×2 轮： <ul style="list-style-type: none"> · ≤3 min · 3~5 (min) · 5~8 (min) · 8~12 (min) · >12 min 	5分 4分 3分 2分 1分

指标	评分规则			说明	
		RSVP	统计从用户端数据采集到模型训练完成的总时间，包含收集各类目标的训练样本，至少 30 个目标： <ul style="list-style-type: none"> · ≤8 min · 8~12 (min) · 12~18 (min) · 18~25 (min) · >25 min 	5 分 4 分 3 分 2 分 1 分	
		运动想象	统计从用户端数据采集到模型训练完成的总时间，包含各类 MI 数据采集+训练，每类至少 20-40 次 trial： <ul style="list-style-type: none"> · ≤10 min · 10~15 (min) · 15~20 (min) · 20~30 (min) · >30 min 	5 分 4 分 3 分 2 分 1 分	
变异系数	通用评估	在连续使用过程中（如 30 分钟或 1 小时）测量关键指标（如信号幅度、分类准确率或时延）的变异系数： <ul style="list-style-type: none"> · ≤5% · 5%~10% · 10%~15% · 15%~20% · >20% 	5 分 4 分 3 分 2 分 1 分	变异系数=（标准差/均值）×100%	
性能衰减	通用评估	在模拟长期使用场景下（如连续操作 1 小时、4 小时或 24 小时）进行评估，记录起始和结束时的性能指标： <ul style="list-style-type: none"> · ≤5% · 5%~10% · 10%~15% · 15%~20% · >20% 	5 分 4 分 3 分 2 分 1 分	性能衰减=（基准值-目标值）/基准值×100%	
噪声电平	专用评估	脑电信号采集 各通道输入端短接，按照制造商的滤波设置，采集 10s 以上并回放，测量每个通道 10s 连续波形中最大噪声峰值： <ul style="list-style-type: none"> · ≤2 μVrms · ≤4 μVrms · ≤6 μVrms · ≤10 μVrms · >10 μVrms 	5 分 4 分 3 分 2 分 1 分	测量方法参考 GB 9706.226、YY/T 1996-2025	

B.4 人因工程与可用性相关二级评价指标量化评分规则

脑机接口设备人因工程与可用性相关二级评价指标的评价内容、数据采集方法和评价结果证明材料的要求分辨见5.3.4.1和5.3.4.2。人因工程与可用性相关二级评价指标的量化评分规则见表B.4。

表B.4 信号与算法性能相关二级评价指标量化评分规则一览表

指标	评分规则			说明
尺寸适配性	通用评估	SUS 评分： · ≥ 85 · 70~84 · 60~69 · 50~59 · < 50	5分 4分 3分 2分 1分	判定以“目标人群/解剖范围、配合器械清单、允许公差/间隙/过盈、清洁/灭菌/消毒兼容性”为边界，要求在临床最不利工况下仍满足功能与安全裕度。
操作范围适配性	通用评估	SUS 评分： · ≥ 85 · 70~84 · 60~69 · 50~59 · < 50	5分 4分 3分 2分 1分	判定以“可达空间、姿势与力矩、手部尺寸适配、人机界面（手柄/触发/踏板）”为边界，强调不同使用者群体（手型、体格、穿戴防护）的可操作性与疲劳风险。
调节范围适配性	通用评估	SUS 评分： · ≥ 85 · 70~84 · 60~69 · 50~59 · < 50	5分 4分 3分 2分 1分	判定以“临床任务需求、最差工况稳定性、互操作性与兼容性（接口/耗材/系统）”为边界，要求设定值可重复、保持性（抗漂移）与失效安全（超限告警/互锁）。
任务完成时间	通用评估	用户使用系统达成相应目标所需的平均时间： · $\leq 10s$ · 10~15 (s) · 15~20 (s) · 20~30 (s) · $> 30s$	5分 4分 3分 2分 1分	
误操作率	通用评估	用户在使用系统过程中出现操作错误的比例值： · $\leq 1\%$ · 1%~3% · 3%~5% · 5%~10% · $> 10\%$	5分 4分 3分 2分 1分	误操作率=误操作次数/总操作次数

指标	评分规则			说明
成功率	通用评估	用户使用系统达成相应目标的比例值： <ul style="list-style-type: none"> · ≥99% · 95%~99% · 90%~95% · 85%~90% · <85% 	5分 4分 3分 2分 1分	成功率=成功次数/总操作次数
交互步骤数量	通用评估	完成一次完整脑机指令或操作任务所需的平均步骤数： <ul style="list-style-type: none"> · ≤3步 · 3~5（步） · 5~8（步） · 8~10（步） · >10步 	5分 4分 3分 2分 1分	
培训时长	通用评估	用户在初次使用系统时，为掌握任务操作所需的总培训时间： <ul style="list-style-type: none"> · <5min · 5~<10（min） · 10~<20（min） · 20~<30（min） · ≥30min 	5分 4分 3分 2分 1分	
学习时间	通用评估	用户达到稳定操作性能所需的训练时间： <ul style="list-style-type: none"> <5min · 5~<10（min） · 10~<20（min） · 20~<30（min） · ≥30min 	5分 4分 3分 2分 1分	
个性化校准时间	专用评估	非侵入式	用户完成专属模型或信号参数校准所需时间： <ul style="list-style-type: none"> · <10min · 11~20（min） · 21~30（min） · 31~60（min） · >60min 	5分 4分 3分 2分 1分
		侵入式	用户完成专属模型或信号参数校准所需时间： <ul style="list-style-type: none"> · <5min · 5~10（min） · 11~20（min） · 21~30（min） · >30min 	5分 4分 3分 2分 1分

指标	评分规则			说明
再校准时间	通用评估	在用户使用过程中，设备算法发生性能衰减后重新校准到正常水平所需的时间： <ul style="list-style-type: none"> · <30 s · 30 s~<1 min · 1~<2 (min) · 2~<5 (min) · ≥5 min 	5分 4分 3分 2分 1分	若未发生需再校准的情况，按满分评价
安装/拆卸与护理时间	通用评估	安装或拆卸与护理总时长： <ul style="list-style-type: none"> · ≤5min · 5~10 (min) · 10~20 (min) · 20~30 (min) · >30min 	5分 4分 3分 2分 1分	
舒适度	通用评估	SUS 评分： <ul style="list-style-type: none"> · ≥85 · 70~84 · 60~69 · 50~59 · <50 	5分 4分 3分 2分 1分	采用问卷、深度访谈等型式开展 SUS 评分。
用户体验	通用评估	SUS 评分： <ul style="list-style-type: none"> · ≥85 · 70~84 · 60~69 · 50~59 · <50 	5分 4分 3分 2分 1分	采用问卷、深度访谈等型式开展 SUS 评分。

B.5 可靠性、维护与后勤保障二级评价指标量化评分规则

脑机接口设备可靠性、维护与后勤保障相关二级评价指标的评价内容、数据采集方法和评价结果证明材料的要求分辨见5.3.5.1和5.3.5.2。可靠性、维护与后勤保障相关二级评价指标的量化评分规则见表B.5。

表 B.5 可靠性、维护与后勤保障相关二级评价指标量化评分规则一览表

序号	二级评价指标	评分规则			说明
1	平均故障间隔时间	非侵入式	· >120 天	5分	MTBF = 总正常运行时间 / 故障总次数。
			· 91-120 (天)	4分	
· 61-90 (天)	3分				
· 31-60 (天)	2分				
· ≤30 天	1分				
外部设备为侵入式	· >150 天	5分			
	· 121-150 (天)	4分			

序号	二级评价指标	评分规则			说明
			<ul style="list-style-type: none"> • 91-120 (天) • 61-90 (天) • ≤60 天 	3分 2分 1分	
2	平均无故障工作时间	非侵入式	<ul style="list-style-type: none"> • >120 天 • 91-120 (天) • 61-90 (天) • 31-60 (天) • ≤30 天 	5分 4分 3分 2分 1分	<ul style="list-style-type: none"> • MTTF = 所有样本的总工作时间 / 样本总数; • 需排除用户误操作导致的“假故障”。
		植入部件为侵入式	<ul style="list-style-type: none"> • >7 年 • 6-7 (年) • 4-5 (年) • 2-3 (年) • <2 年 	5分 4分 3分 2分 1分	
3	电池续航与循环寿命		<ul style="list-style-type: none"> • 续航满足全天使用 (≥8h), 循环寿命≥500 次 (便携) 或设计寿命内无需更换 (植入); • 续航 6-8h, 循环寿命 300-500 次; • 续航 4-6h, 循环寿命 100-300 次; • 续航<4h 或 循环寿命<100 次; • 电池性能严重不足, 频繁充电或提前失效; 或植入部分寿命小于 5 年。 	5分 4分 3分 2分 1分	植入设备若不可更换电池, 需提供寿命预测模型, 证明其寿命≥5 年。
4	侵入式系统的导线稳定性/修复或翻修手术率 (年化)	年化翻修率为: <ul style="list-style-type: none"> • ≤1%; • 1-2%; • 2-5%; • 5-10%; • >10% 或 发生导线断裂/移位导致严重并发症。 	5分 4分 3分 2分 1分	需基于不少于 12 个月的临床随访数据计算年化翻修率。	
5	长期稳定性		<ul style="list-style-type: none"> • >9 年 • 8-9 (年) • 6-7 (年) • 4-5 (年) • ≤3 年 	5分 4分 3分 2分 1分	/
6	环境适应性		<ul style="list-style-type: none"> • 每月发生风险事件次数≤1, 且均为可忽略风险; • 每月发生风险事件次数≤3, 且均为可忽略风险; • 每月发生风险事件次数≤5, 且均为可忽略风险或轻微风险, 其中轻微风险事件次数≤1; • 每月发生风险事件次数≤10, 存在严重风险事件≤2; • 每月发生风险事件次数>10, 或存在严重风险事件≥2。 	5分 4分 3分 2分 1分	

序号	二级评价指标	评分规则		说明
7	维修周转时间/现场支持	<ul style="list-style-type: none"> 平均响应≤2h, 修复≤24h, 支持 7×24 服务; 响应≤4h, 修复≤48h; 响应≤24h, 修复≤5 天; 响应>24h 或 修复>7 天; 无明确服务承诺, 多次投诉未解决。 	5 分 4 分 3 分 2 分 1 分	/
8	软件管理	<ul style="list-style-type: none"> 全使用周期证据闭环, 风险可控, 更新合规; 基本闭环, 少量薄弱点单不影响安全有效; 关键环节到位, 但存在明显差距; 重要证据缺失或控制失效; 系统性失效。 	5 分 4 分 3 分 2 分 1 分	评分细则由评估组织方在评估方案中细化规定, 宜包括并不仅限于软件功能安全和软件更新相关评分要点。

B.6 互操作、数据与网络安全二级评价指标量化评分规则

脑机接口设备互操作、数据与网络安全相关二级评价指标的评价内容、数据采集方法和评价结果证明材料的要求分辨见5.3.6.1和5.3.6.2。互操作、数据与网络安全相关二级评价指标的量化评分规则见表B.6。

表B.6 互操作、数据与网络安全相关二级评价指标量化评分规则一览表

序号	二级评价指标	评分规则	
1	术语和数据元一致性	<ul style="list-style-type: none"> 与推荐术语和数据元一致; 单一要素不符合。 	5 分
			1 分
2	数据格式与导出	<ul style="list-style-type: none"> 符合通用标准要求; 单一要素不符合。 	5 分
			1 分
3	数据安全与隐私保护的合法性	<ul style="list-style-type: none"> 遵守法律要求; 单一要素不符合。 	5 分
			1 分
4	数据透明度	<ul style="list-style-type: none"> 提供详细的数据报告, 并验证报告证明符合评估组织方对数据透明度的可接受准则要求; 数据报告质量有严重缺陷和/或符合性验证报告中单一要素不符合。 	5 分
			1 分
5	网络安全文档	<ul style="list-style-type: none"> 符合性验证报告证明符合评估组织方对数据透明度的可接受准则要求; 符合性验证报告中单一要素不符合。 	5 分
			1 分
6	远程监测与控制	<ul style="list-style-type: none"> 符合性验证报告证明符合相关法律法规和标准要求; 符合性验证报告中单一要素不符合。 	5 分
			1 分

附录 C
(资料性附录)
直接取值方法示例

C.1 统计数据法

统计数据法主要依据系统记录、统计报表或跟踪调查所得数据，获取量化评分所需的数据源取值。

表C.1 基于统计数据法的取值方法示例

序号	指标名称	取值方法
1	不良事件/事件率	建立上市后监测方案，收集所有与设备相关的不良事件报告，按设备数量和使用时长计算发生率。
2	生物相容性与植入相关并发症率	在随访中记录所有与植入物相关的医疗并发症事件，统计每例设备或每例患者随访期内的并发症比例。
3	目标任务成功率	在代表性场景（如拼写沟通、基础光标控制、神经反馈训练）下，统计成功完成次数与总尝试次数。
4	真实世界有效使用率	基于设备使用日志与患者日记，统计在随访期内符合“有效使用日”。
5	持续使用率	通过为期12个月的随访，统计在随访期末仍坚持使用该系统的受试者比例。 持续使用率 = $\frac{\text{随访期内持续使用的受试者数}}{\text{入组受试者数}} \times 100\%$ 。
6	无故障使用天数	依据设备日志或服务记录开展统计。
7	误操作率	用户在使用系统过程中出现操作错误或未达成相应目标的发生频率。
8	交互步骤数量	完成一次完整脑机指令或操作任务所需的平均步骤数。
9	培训时长	用户在初次使用系统所需的总培训时间，时间单位为小时（h）。
10	学习时间	用户可稳定操作设备达成任务所需的累计训练时间，时间单位为小时（h）。
11	个性化校准时间	为使系统达到目标性能，训练、调参与验证所花费的时间，时间单位为小时（h）。
12	安装/拆卸与护理时间	部署/装配、拆卸/移除、维护/修理所需的时间，时间单位为小时（h）。

C.2 试验/测量法

试验/测量法主要通过特定条件下通过检验、检测、监测或测量手段获取量化评分所需的物理或性能数据。

表C.2 基于试验/测量法的取值方法示例

序号	指标名称	取值方法
1	ITR	$ITR = B \cdot C$ <p>其中，B为单次选择的信息量，C为单位时间的决策次数，详见条款3.4。</p>
2	电极间阻抗	会话开始前进行全通道电极间阻抗测量，统计中位数与95分位，评分阈值参考临床 EEG 指南。
3	电极间阻抗稳定性	$\text{稳定性} = \frac{\text{稳定时间}}{\text{总时间}} \times 100\%。$
4	信噪比	$SNR(dB) = 10 \cdot \log_{10} \left(\frac{P_{\text{signal}}}{P_{\text{noise}}} \right)$ <p>其中，P_{signal}是信号的功率，P_{noise}是噪声的功率。</p>
5	通道有效率	$\text{通道有效率} = \frac{\text{有效通道数}}{\text{总通道数}} \times 100\%。$
6	掉线率	$\text{掉线率} = \frac{\text{掉线次数/时间}}{\text{总连接次数/总在线时间}} \times 100\%。$
7	F1-score	$F1 - score = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$ <p>其中，$\text{Precision} = \frac{TP}{TP + FP}$，$\text{Recall} = \frac{TP}{TP + FN}$，TP为真阳性计数，TN为真阴性计数，FP为假阳性计数，FN为假阴性计数。</p>
8	性能保持率	$\text{性能保持率} = \frac{\text{某条件下性能 (F1-score)}}{\text{基线性能 (F1-score)}} \times 100\%。$
9	系统时延	检测用户动作或意图产生到系统反馈之间的时间，时间单位为毫秒(ms)。
10	初始化校准时间	系统从启动到完成首次有效模型或信号参数校准所需时间，时间单位为分钟 (min)。
11	再校准时间	系统在运行过程中出现性能下降、通道漂移或用户状态变化时，为恢复性能所需的再校准时长，时间单位为秒 (s)。

C.3 证据判断法

证据判断法主要依据文档、记录等形式的证据，通过专业判断给出符合性评价或等级判定结果。

表C.3 基于证据判断法的取值方法示例

序号	指标名称	取值方法
1	产品安全符合性	通过具有医疗器械检验检测资质的检测机构开展安全符合性检验，并按所有单项目检验结果的符合性综合情况分档赋分，赋分的分档依据为评估组织方认可的评估方案。
2	风险管理符合性	由专业人士对产品的风险管理文档与GB/T 42062—2022 要求的一致性开展核查，并按充分性分档赋分，赋分的分档依据为评估组织方认可的评估方案。

序号	指标名称	取值方法
3	软件运维	对本部分表1中的评价结果证明材料的真实性、准确性和完整性进行核查，并根据综合情况分档赋分，赋分的分档依据为评估组织方认可的评估方案。
4	植入物取出过程的可控性与可追溯性	
5	故障响应	

C.4 主观赋值法

主观赋值法主要基于用户、专家或其他主体的主观反馈或感知，按预定规则转换为定量评分以获取量化评分所需的数据源取值。

——患者报告结局（PRO）改善：采用经过验证的患者报告结局量表（如针对特定适应症的QUESI、SIS等）获得评估得分，并根据评估组织方认可的评估方案中的赋分规则进行赋分。

——用户体验：采用SUS获得评估得分，按表6评价，并记录样本量与信度。

表C.4 SUS得分赋分规则

序号	SUS得分	评分
1	≥ 85	5分
2	≥ 70 且 < 85	4分
3	≥ 60 且 < 70	3分
4	≥ 50 且 < 60	2分
5	< 50	1分

参考文献

- [1] 国家质量监督检验检疫总局, 国家标准化管理委员会. 医用电器环境要求及试验方法: GB/T 14710—2009[S]. 北京: 中国标准出版社, 2009.
- [2] 国家市场监督管理总局, 国家标准化管理委员会. 物联网 生命体征感知设备通用规范: GB/T 40687—2021[S].北京: 中国标准出版社, 2021.
- [3] 国家药品监督管理局. 医疗器械软件 软件生存周期过程: YY/T 0664—2020[S]. 北京: 中国标准出版社, 2020.
- [4] 国家药品监督管理局. 医疗器械 风险管理对医疗器械的应用: YY/T 1437—2023[S]. 北京: 中国标准出版社, 2023.
- [5] 国家药品监督管理局. 采用脑机接口技术的医疗器械 术语: YY/T 1987—2025[S]. 北京: 中国标准出版社, 2025.
- [6] 国家药品监督管理局. 医用电气设备 第1-2部分: 基本安全和基本性能的通用要求 并列标准: 电磁兼容 要求和试验: YY 9706.102—2021[S]. 北京: 中国标准出版社, 2021.
- [7] 国家药品监督管理局. 医疗器械 不良事件分级编码结构 第1部分:事件类型编码: YY/T 0869.1—2016[S]. 北京: 中国标准出版社, 2016.
- [8] 国家药品监督管理局. 医疗器械 风险管理对医疗器械的应用: GB/T 42062—2022[S]. 北京: 中国标准出版社, 2022.
- [9] International Organization for Standardization. Clinical investigation of medical devices for human subjects — Good clinical practice: ISO 14155:2020[S/OL]. (2020)[2026-1-13]. <https://www.iso.org/standard/71690.html>.
- [10] International Organization for Standardization. Medical devices—Application of risk management to medical devices: ISO 14971:2019[S/OL]. (2019)[2026-1-13]. <https://www.iso.org/standard/72704.html>.
- [11] International Electrotechnical Commission. Medical devices—Part 1: Application of usability engineering to medical devices: IEC 62366-1:2015+A1:2020 [S/OL]. (2020-06-17) [2026-1-13]. <https://webstore.iec.ch/publication/67220>