

## · 标准解读 ·

# 《世界卫生组织结核病整合指南模块 3: 诊断——结核病检测的快速诊断》第三版解读

王小敏 陈锦云 曾昱芹 马俊 孔星星 蒙建州 卢水华

**【摘要】** 2024 年 3 月 20 日,世界卫生组织发布了《世界卫生组织结核病整合指南模块 3: 诊断——结核病检测的快速诊断》(第三版)。指南主要针对 3 个方面提出建议: 耐药结核病的初始诊断检测方法、结核病的初始诊断检测方法(不包括耐药性检测)和结核病确诊后的耐药性检测。另外,指南首次增加靶向高通量测序(targeted next-generation sequencing, tNGS)技术可用于针对结核病确诊后耐药性检测的推荐意见。笔者对第三版指南的编制及内容要点加以解读,以促进该指南更好地实施与推广。

**【关键词】** 世界卫生组织; 指南; 结核; 药物耐受性; 诊断

doi: 10.19982/j.issn.1000-6621.20240221

**【中图分类号】** R52

**Interpretation of the third edition of WHO consolidated guidelines on tuberculosis: module 3: diagnosis: rapid diagnostics for tuberculosis detection** Wang Xiaomin, Chen Jinyun, Zeng Yuqin, Ma Quan, Kong Xingxing, Meng Jianzhou, Lu Shuihua. National Clinical Research Center for Infectious Disease, Shenzhen Third People's Hospital, Shenzhen 518112, China

Corresponding author: Lu Shuihua, Email: lushuihua66@126.com

**【Abstract】** On 20 March 2024, the World Health Organization (WHO) published the WHO consolidated guidelines on tuberculosis: Module 3: diagnosis: rapid diagnostics for tuberculosis detection. The third edition of the guidelines focuses on three areas: initial diagnostic tests for diagnosis of tuberculosis with drug-resistance detection, initial diagnostic tests for diagnosis of tuberculosis without drug-resistant detection, and follow-on diagnostic tests for detection of additional drug-resistance after tuberculosis confirmation. In addition, this current guidelines include for the first time a chapter on targeted next generation sequencing (tNGS) tests for the diagnosis of drug-resistant tuberculosis in people with diagnosed tuberculosis with or without rifampicin-resistant tuberculosis. The author now introduces the compilation and key points of the third edition of the guidelines in order to promote the implementation and dissemination of these guidelines.

**【Key words】** World Health Organization; Guidebooks; Tuberculosis; Drug tolerance; Diagnosis

**【Fund program】** National Natural Science Foundation of China (32394014); Shenzhen Science and Technology Program (Basic Research General Programme)(JCYJ20220530163212027)

自 2020 年以来,全球结核病患病例数逐年增加,仅 2022 年全球新确诊结核病患者就高达 750 万例<sup>[1]</sup>,扭转了过去 20 年大部分时间每年约 2% 的下降趋

势。同年,全球耐多药结核病(multidrug-resistant tuberculosis, MDR-TB)或利福平耐药结核病(rifampicin-resistant tuberculosis, RR-TB)患者约为 41 万例(95% UI: 37 万~45 万)<sup>[1-2]</sup>,仅有约 2/5 的耐药结核病(drug-resistant tuberculosis, DR-TB)患者接受了治疗。虽然自 2000 年以来,全球抗击结核病行动总共挽救了 7500 多万人的生命,但是结核病诊断和治疗方面距离全球结核病防治目标仍然存在巨大差距。世界卫生组织(World Health Organization, WHO)在《2023 年全球结核病报告》中,建议使用快速分子诊断检测作为所有存在结核病体征和症状者的初始诊断检测方法。

2023 年 9 月 22 日,联合国大会防治结核病高



开放科学(资源服务)标识码(OSID)的开放科学计划以二维码为入口,提供丰富的线上扩展功能,包括作者对论文背景的语音介绍、该研究的附加说明、与读者的交互问答、拓展学术圈等。读者“扫一扫”此二维码即可获得上述增值服务。

基金项目:国家自然科学基金(32394014);深圳市科技计划项目基础研究(面上项目)(JCYJ20220530163212027)

作者单位:国家感染性疾病临床医学研究中心/深圳市第三人民医院,深圳 518112

通信作者:卢水华,Email:lushuihua66@126.com

级别会议强化了 2018 年做出的承诺和确定的目标,并设定了 2023—2027 年的新目标,旨在推进全球在终止结核病流行方面的努力,其中两个目标就是使用 WHO 推荐的快速检测方法作为首选的结核病诊断方法和至少 90% 的结核病患者得到有质量保证的诊断和治疗<sup>[3-4]</sup>。WHO 在《终止结核病战略》中亦呼吁早期诊断结核病并普及药物敏感性试验(drug susceptibility testing, DST),强调了实验室在快速准确检测结核病和耐药性方面的关键作用<sup>[5]</sup>。

自 2021 年以来,WHO 一直在发布诊断结核病和 DR-TB 的综合指南。2024 年 3 月 20 日,WHO 发布了 *WHO consolidated guidelines on tuberculosis: Module 3: diagnosis-rapid diagnostics for tuberculosis detection, Third edition* (简称《2024 版指南》),以取代 2022 年 9 月 30 日发布的 *WHO consolidated guidelines on tuberculosis: Module 3: diagnosis: tests for tuberculosis infection* 提出的相关指南建议。《2024 版指南》基于最新循证医学证据,经系统评估后,提供了新型诊断技术的背景、证据和推荐建议。与 2022 年发布的第二版指南相比,第三版首次增加了靶向高通量测序(targeted next-generation sequencing, tNGS)技术可用于针对结核病确诊后耐药性检测的推荐意见。笔者对《2024 版指南》的编制过程和形成的 3 个核心方向的推荐意见加以解读,旨在为我国药物敏感结核病(drug-resistant tuberculosis, DS-TB)和 DR-TB 快速诊断工作的开展和推广提供参考依据。

### 《2024 版指南》编制过程

《2024 版指南》沿用既往结核病诊断技术的评价体系,即侧重于评估结核病诊断技术,而不是特定产品。指南的制定过程承袭了 WHO 指南严谨的循证过程,每项纳入该综合指南的建议都有其独立的指南制定小组(guideline development groups, GDG),通过对一系列已经发表或未发表的相关文献数据进行系统综述和荟萃分析,由 WHO 召集的指南指导小组审慎地采用国际证据推荐分级的评估、制定与评价方法,提出最终的推荐建议。

基于上述循证医学步骤,《2024 版指南》针对 3 个方向提出推荐意见:(1)DR-TB 的初始诊断检测方法;(2)结核病的初始诊断检测方法(不包括耐药性检测);(3)结核病确诊后的耐药性检测。WHO 指导小组依据研究对象、干预措施、对照和结

果/结局(population, intervention, comparator and outcome, PICO)原则,围绕推荐理由和诊断准确性等方面全面提出问题,并进行了细致的阐述,同时还从成本效果及用户视角方面对提出的建议进行了综合评估,最后还提出了未来的研究方向。

### 推荐意见一:耐药结核病的初始诊断检测方法

基于严谨的循证医学步骤,《2024 版指南》在 2022 版的基础上产生了 3 条核心推荐意见:(1)Xpert MTB/RIF 和 Xpert MTB/RIF Ultra(简称“Xpert Ultra”;Xpert MTB/RIF 和 Xpert MTB/RIF Ultra 统称为“Xpert”),用于 DR-TB 的初始诊断;(2)新的分子检测方法 Truenat MTB, MTB Plus(印度 Molbio Diagnostics 公司),适用于结核病的初始诊断;Truenat MTB-RIF Dx(简称“MTB-RIF Dx”;印度 Molbio Diagnostics 公司),适用于初始 Truenat MTB 阳性患者对利福平耐药性的检测;(3)中等复杂度自动化核酸扩增实验(nucleic acid amplification tests, NAAT;美国雅培公司),可用于从呼吸道样本中检测肺结核,以及对利福平和异烟肼的耐药性检测,而不是基于传统培养的表型 DST。指南针对不同结核病类型、不同人群及特定类型的检测推荐意见有所区别,证据等级和推荐程度也有所不同。在此,笔者将重点对 DR-TB 的 3 条初始诊断方法及相关循证依据等展开详细解读。

#### 一、Xpert MTB/RIF 和 Xpert Ultra 检测法

##### (一)推荐意见

《2024 版指南》针对不同结核病类型、不同人群及特定类型的检测从 4 个方面给出 17 条建议:

1. 关于将 Xpert MTB/RIF 和 Xpert Ultra 作为对有肺结核症状和体征的成人和儿童进行初始检测的建议:(1)对于有肺结核症状和体征的成人,应将 Xpert MTB/RIF 作为痰中结核分枝杆菌(*Mycobacterium tuberculosis*, MTB)和利福平耐药性检测的初始诊断检测,而不是涂片显微镜检查/培养和表型 DST(强烈推荐,准确性证据等级高,患者重要结局证据等级中)。(2)对于有肺结核症状和体征的儿童,应将 Xpert MTB/RIF 作为痰液(包括排出痰和诱导痰)、胃吸液、鼻咽吸液和粪便中 MTB 和利福平耐药性检测的初始诊断检测,而不是涂片显微镜检查/培养和表型 DST(强烈推荐,对于痰液,准确性证据等级中;对于胃吸液、鼻咽吸液和粪便,准确性证据等级低)。另外,该建议可推而广之,适用于感染 HIV 的儿童。(3)对于有肺结核症状和体征、

既往无结核病史( $\leq 5$ 年)或有远期结核病治疗史(治疗结束后 $> 5$ 年)的成人,应将 Xpert Ultra 用作结核病的初始诊断和痰液中利福平耐药性检测的方法,而不是涂片显微镜检查/培养和表型 DST(强烈推荐,准确性证据等级高)。(4)对于有肺结核症状和体征、既往有肺结核病史且在过去 5 年内结束治疗的成人,可将 Xpert Ultra 用作肺结核的初始诊断检测,并检测痰中的利福平耐药性,而不是涂片显微镜检查/培养和表型 DST(有条件推荐,准确性证据等级低)。(5)对于有肺结核症状和体征的儿童,应将 Xpert Ultra 作为肺结核的初始诊断检测,并检测痰液或鼻咽吸液中的利福平耐药性,而不是涂片显微镜检查/培养和表型 DST(强烈推荐,对于痰液,准确性证据等级低;对于鼻咽吸液,准确性证据等级极低)。

2. 关于将 Xpert MTB/RIF 和 Xpert Ultra 作为对有肺外结核症状和体征的成人和儿童进行初始检测的建议:(6)对于有结核性脑膜炎症状和体征的成人和儿童,应使用 Xpert MTB/RIF 或 Xpert Ultra 检测脑脊液(cerebrospinal fluid, CSF)作为结核性脑膜炎的初始诊断检测,而不是涂片镜检/培养(强烈推荐,对于 Xpert MTB/RIF 准确性证据等级中,对于 Xpert Ultra 准确性证据等级低)。(7)对于有肺外结核症状和体征的成人和儿童, Xpert MTB/RIF 可用于淋巴结穿刺、淋巴结活检、胸腔积液、腹腔积液、心包积液、滑膜液或尿液标本的初始诊断检测,而非涂片镜检/培养(有条件推荐,对于胸腔积液,准确性证据等级中;对于淋巴结穿刺、腹腔积液、滑膜液和尿液,证据等级低;对于心包积液和淋巴结活检,证据等级极低)。(8)对于有肺外结核症状和体征的成人和儿童,可在淋巴结穿刺和淋巴结活检中使用 Xpert Ultra 作为初始诊断检测,而不是涂片镜检/培养(有条件推荐,证据等级低)。(9)对于有肺外结核症状和体征的成人和儿童,应使用 Xpert MTB/RIF 或 Xpert Ultra 检测对利福平的耐药性,而不是培养和表型 DST 的测定(强烈推荐,对于 Xpert MTB/RIF 证据等级高;对于 Xpert Ultra 证据等级低)。(10)对于有播散性肺结核症状和体征的 HIV 阳性成人和儿童,可使用 Xpert MTB/RIF 检测血液样本,作为播散性肺结核的初始诊断检测(有条件推荐,准确性证据等级极低)。此外, GDG 认为不宜将此建议推广到其他患者群体。

3. 关于对有肺结核症状和体征的成人和儿童进行 Xpert MTB/RIF 和 Xpert Ultra 重复检测的建

议:(11)对于有肺结核症状和体征的成人,如果初次检测结果为 Xpert Ultra 微量阳性,则可以不再重复进行 Xpert Ultra 检测(有条件推荐,准确性证据等级极低)。(12)对于有肺结核症状和体征的儿童,若验前概率低于 5%且初始检测的 Xpert MTB/RIF 结果为阴性,则不得使用 Xpert MTB/RIF 对痰液、胃液、鼻咽吸液或粪便标本进行重复检测(有条件推荐,对于痰液,准确性证据等级低;对于其他标本,准确性证据等级极低)。(13)对于有肺结核症状和体征的儿童,若验前概率 $\geq 5\%$ 且初始 Xpert MTB/RIF 检测结果为阴性,则可使用 Xpert MTB/RIF 对痰液、胃液、鼻咽吸液和粪便标本进行重复检测(共两次)(有条件推荐,对于痰液,准确性证据等级低;对于其他标本,准确性证据等级极低)。(14)对于有肺结核症状和体征的儿童,若验前概率低于 5%且 Xpert Ultra 初始检测结果为阴性,则不得使用 Xpert Ultra 对痰液或鼻咽吸液标本进行重复检测(有条件推荐,准确性证据等级极低)。(15)对于有肺结核症状和体征的儿童,若验前概率 $\geq 5\%$ 且初始 Xpert Ultra 检测结果为阴性,可在痰液和鼻咽吸液标本中重复进行一次 Xpert Ultra 检测(共两次)(有条件推荐,准确性证据等级极低)。

4. 关于将 Xpert MTB/RIF 和 Xpert Ultra 作为出现肺结核症状和体征或胸部 X 线摄片(简称“胸片”)显示肺部异常或两者兼有的成人肺结核初始检测方法的建议:(16)对于有肺结核症状或体征或胸片显示肺部异常或两者兼有的成年人, Xpert MTB/RIF 或 Xpert Ultra 可取代培养作为肺结核的初始检测方法(有条件推荐,对于 Xpert MTB/RIF 准确性证据等级低;对于 Xpert Ultra 准确性证据等级中)。该建议仅适用于无法使用或使用辅助检测会造成延误诊断,需要立即决定是否对患者进行治疗时。(17)对于肺结核症状筛查呈阳性或胸片显示肺部异常或两者均呈阳性的成人,可使用一次 Xpert Ultra 检测而非两次 Xpert Ultra 检测作为肺结核的初始检测方法(有条件推荐,准确性证据等级极低)。该建议仅适用于 Xpert Ultra 用于临床病例管理的情况下。

## (二)推荐理由和证据

这是 WHO 制定的用于诊断成人和儿童肺部和肺外结核初始检测的分子测定使用最新指南进程的一部分。WHO 指导小组基于 PICO 原则围绕推荐理由、诊断准确性全面提出问题,并进行了详细的阐述,同时也从成本效果和用户视角进行了综合

评价。

#### 1. PICO 问题:包括 PICO 1~6。

PICO 1:在有肺结核体征和症状的成年人中,在卫生保健机构寻求治疗时,Xpert MTB/RIF 和 Xpert Ultra 是否应作为诊断肺结核和利福平耐药性的初始检测?

WHO 指导小组从以下两个方面分析了这个问题:(1)评估使用 Xpert MTB/RIF 诊断策略与使用涂片显微镜诊断策略对患者重要结局的影响?考虑到以下结局:全因死亡率、治疗前失访、治愈、诊断时间和开始治疗时间。WHO 指导小组发现,Xpert MTB/RIF 组的治疗时间[中位数(四分位数)]为 4(1~10) d,涂片镜检组的治疗时间为 5(1~15) d。与涂片镜检组相比,使用 Xpert MTB/RIF 并没有导致全因死亡率的显著降低,但可降低治疗前失访的风险,估计 RR 为 0.59(95%CI:0.42~0.84),会增加治愈的概率(OR=1.09,95%CI:1.02~1.16)。(2)与微生物参考标准(microbiological reference standard, MRS)相比,Xpert MTB/RIF、Xpert Ultra 对肺结核和利福平耐药的诊断准确性如何?WHO 指导小组纳入随机试验、横断面研究和队列研究,使用呼吸道标本评估 Xpert MTB/RIF、Xpert Ultra 或两者,对照培养检测结核病和基于培养的 DST 或 MTBDRplus 检测利福平耐药性的参考标准,发现 Xpert MTB/RIF 在诊断肺结核和利福平耐药性方面具有较高的敏感度和特异度,这与之前报道的结果一致。Xpert MTB/RIF 对涂片阳性者比涂片阴性者更敏感。在头对头比较研究中,Xpert Ultra 对结核病检测的敏感度(90%)高于 Xpert MTB/RIF(83%),特异度(96%)低于 Xpert MTB/RIF(99%)(6 项研究)。Xpert MTB/RIF 和 Xpert Ultra 对利福平耐药的敏感度和特异度相似。综上,Xpert MTB/RIF 和 Xpert Ultra 可提供准确的结果,从而快速启动对 MDR-TB 的治疗。

PICO 2:在有肺结核体征和症状的儿童中,医疗机构是否应使用 Xpert MTB/RIF 或 Xpert Ultra 初始诊断肺结核和利福平耐药性?

敏感度差异可能部分归因于临床环境的差异(更常见的是住院胃标本采集)和参比标准品质量的差异。Xpert Ultra 对痰液的总敏感度高于 Xpert MTB/RIF,但对鼻咽标本的总敏感度不变;痰和鼻咽吸出物的特异度相似(间接比较)。Xpert Ultra 在提高儿童结核病的检出率方面效果显著。Xpert MTB/RIF 的敏感度(由培养定义)对于多个标本比

单个标本更高,并且不同标本类型相似。同一类型的多个标本行 Xpert MTB/RIF 或 Xpert Ultra,可能会增加测试敏感度(由培养定义)。

在使用复合参考标准时,Xpert MTB/RIF 和 Xpert Ultra 仅可识别约 30%的患者。因此,需要进行更多的研究,以确定一种改进的参考标准,以准确定义儿童结核病。现有诊断方法的敏感度都不是最佳的,这也突出了开发新工具以正确诊断更高比例的儿童结核病患者的持续紧迫性。理想情况下,新工具应是快速、负担得起、可行,并为儿童和他们的父母所接受的。

PICO 3:在有肺外结核体征和症状的成年人中,医疗机构是否应使用 Xpert MTB/RIF 或 Xpert Ultra 初始诊断肺外结核和利福平耐药性?

由于获得肺外标本困难,以及传统细菌学方法辅助诊断技术的局限性,各种非肺标本和复合参考标准常被用于评估肺外结核新诊断技术的性能。对于肺外结核的检测,纳入了 65 项研究。39 项(60%)研究发生在结核病高负担国家,41 项(63%)研究发生在 MTB/HIV 双重感染高负担国家。共有 63 项研究(13 144 例受试者)评估了 Xpert MTB/RIF,其中 5 项也同时评估了 Xpert Ultra。

Xpert MTB/RIF 合并敏感度(由培养定义)在不同类型的标本中有所不同(胸膜液为 50%,骨或关节液为 97%)。Xpert MTB/RIF 在淋巴结液、淋巴结组织、尿液和骨或关节液中的总敏感度为 80%或更高。Xpert MTB/RIF 合并特异度(由培养定义)在不同标本中的变化小于敏感度(淋巴结组织为 79%,胸膜液为 99%)。Xpert MTB/RIF 合并特异度在脑脊液、胸膜液、尿液和腹膜液中为 96%或更高。间接比较显示,对于结核性脑膜炎,Xpert Ultra 的综合敏感度(84%)高于 Xpert MTB/RIF(70%),而综合特异度(88%)低于 Xpert MTB/RIF(97%)。Xpert MTB/RIF 和 Xpert Ultra 对利福平耐药的敏感度和特异度相似。

WHO 指导小组指出,应对包括 Xpert Ultra 在内的不同测试进行比较,揭示哪些测试(或策略)具有更高的诊断准确性,还应纳入儿童和 HIV 感染人群,同时须认识到将培养作为少菌标本参考标准的问题,进一步考虑解决这一限制的方法。另外,研究小组应该把重点放在开发诊断测试和策略上,使用现成的临床标本,如尿液,而不是需要侵入性程序收集的标本。

PICO 4:对于在医疗机构就诊,有肺外结核体

征和症状且对利福平耐药的儿童,是否应使用 Xpert MTB/RIF 或 Xpert Ultra 初始诊断肺结核和利福平耐药性?

关于这个问题,指南纳入了有淋巴结结核或结核性脑膜炎体征或症状的儿童 Xpert MTB/RIF 诊断准确性的研究。Xpert MTB/RIF 合并敏感度和特异度分别为 54.0% (27.8%~78.2%) 和 93.8% (84.5%~97.6%) (6 项研究,262 例受试者;敏感度证据等级极低;特异度证据等级低)。目前,尚无评估 Xpert Ultra 在检测淋巴结结核或结核性脑膜炎方面准确性的研究。

PICO 5:对于有肺结核体征和症状的人群,在随后的样本中重复 Xpert MTB/RIF 或 Xpert Ultra 检测作为诊断肺结核和利福平耐药性的初始检测,与单一初始检测相比,是否会增加敏感度和特异度?

对于具有初始 Xpert Ultra 微量结果的成年人,三项研究被纳入分析:Mishra 2019a (4 例受试者), Piersimoni 2019 (4 例受试者) 和 Dorman 2018 (42 例受试者)。在 Mishra 2019a 和 Piersimoni 2019 研究中,Xpert Ultra 的敏感度和特异度均为 100%。Dorman 2018 中 Xpert Ultra 敏感度为 69% (95% CI: 39%~91%), 特异度为 66% (95% CI: 46%~82%)。由于研究数量有限,未进行荟萃分析,证据等级被判定为敏感度和特异度均极低。

对于儿童,使用多个痰标本与使用单个痰标本相比,Xpert MTB/RIF 的总敏感度和特异度差异分别为 12.8% (95% CI: -6.78%~32.3%) 和 -0.34% (95% CI: -1.09%~0.41%) (5 项研究,1925 例受试者进行多次 Xpert MTB/RIF 检测,1939 例受试者进行一次 Xpert MTB/RIF 检测);敏感度证据等级低;特异度证据等级高)。使用多个粪便标本与使用一个粪便标本相比,Xpert MTB/RIF 的敏感度和特异度差异分别为 10.3% (95% CI: -20.8%~41.4%) 和 0.02% (95% CI: -1.21%~1.25%) (1 项研究,247 例受试者进行多次 Xpert MTB/RIF 检测,236 例受试者进行一次 Xpert MTB/RIF 检测;敏感度证据等级低)。多个鼻咽标本与单个鼻咽标本相比,Xpert 对 MTB/RIF 的敏感度和特异度差异分别为 16.7% (95% CI: -11.1%~44.5%) 和 -1.89% (95% CI: -6.34%~2.57%) (1 项研究,130 例受试者;敏感度证据等级极低;特异度证据等级低)。总结起来:多次 Xpert MTB/RIF 或 Xpert Ultra 试验与单次 Xpert MTB/RIF 试验诊断肺结核的敏感度比较,Xpert MTB/RIF 痰液:

12.8% (95% CI: -6.78%~32.3%),  $P=0.20$ ; Xpert Ultra 痰液:10.7% (95% CI: -13.2%~34.6%),  $P=0.38$ ; Xpert MTB/RIF 胃液标本:13.2% (95% CI: -4.64%~31.1%),  $P=0.15$ ; Xpert MTB/RIF 鼻咽吸痰:13.5% (95% CI: -9.50%~36.5%),  $P=0.25$ ; Xpert Ultra 鼻咽吸痰:16.7% (95% CI: -11.1%~44.5%),  $P=0.25$ ; Xpert MTB/RIF 粪便:10.3% (95% CI: -20.8%~41.4%),  $P=0.52$ 。

因此,对于儿童,Xpert MTB/RIF 的敏感度(由培养定义)对于多个标本比单个标本更高,并且不同标本类型相似。同一类型的多个 Xpert MTB/RIF 或 Xpert Ultra 标本,可能会增加测试敏感度(由培养定义)。

PICO 6:在有结核病体征和症状或胸片显示肺部异常提示肺结核或两者兼有的成年人中,是否应单独使用 Xpert MTB/RIF 或 Xpert Ultra 来定义活动性肺结核患者?

这个问题旨在评估 Xpert MTB/RIF 和 Xpert Ultra 对普通人群中成人(年龄 $\geq 15$ 岁)肺结核的诊断准确性。本研究纳入了 4 项具有全国代表性和两次国家级的活动性结核病流行病学调查的数据,这些调查采用横断面设计,使用痰样本对照结核分枝杆菌培养参考标准对 Xpert MTB/RIF 或 Xpert Ultra 进行评估。与 MRS 相比,Xpert MTB/RIF 和 Xpert Ultra 的敏感度证据等级低,特异度证据等级中。两次 Xpert Ultra 和一次 Xpert Ultra 的 95% CI 较宽,敏感度证据等级极低,特异度证据等级中。

2. 成本效益分析:围绕成本效益分析的关键问题是确定实施 Xpert MTB/RIF 和 Xpert Ultra 的相对成本、可负担性和成本效益。WHO 项目小组纳入 28 项符合标准的研究,主要关注非洲国家门诊环境中的 Xpert MTB/RIF,但也考虑了其他国家的门诊和住院患者。由于研究采用了不同的建模方法、人群和环境,直接比较具有挑战性。对于高结核病患病率、HIV 感染者和经验治疗率低的人群来说,Xpert MTB/RIF 的成本效益更为显著。此外,Xpert MTB/RIF 的成本效益受到多种因素的影响,如 GeneXpert 机器的位置、测试量、结核病流行率、经验治疗水平等。在推广成本效益时,需考虑当地实施条件和环境,并进行本地实施研究以评估可能影响患者发现、长期结果和成本效益的因素。尽管在诊断结核病的门诊患者中有大量经济证据表明实施和扩大 Xpert MTB/RIF 具有成本效益,但由于

实施方法和环境的不同,也存在一些例外情况。另外,目前尚无研究评估 Xpert Ultra 的成本效益。

3. 用户视角:用户视角在新技术的推广和实施中尤为重要。WHO 指导小组主要讨论了 Xpert MTB/RIF 和 Xpert Ultra 技术在 DR-TB 诊断中的应用及其对患者诊断可行性、可及性、公平性的影响。研究表明,这些技术能显著提升对 DR-TB 的诊断能力,但在实际应用中存在一些挑战,包括:临床治疗的广泛性、结果周转时间的延长、技术实施的可行性,以及在新发现结核病患者方面的效果尚不明确。

此外,尽管 Xpert 检测流程已得到改进和简化,使得实验室工作更便利和自动化,但并不是所有需要此项技术的患者都能实施该检测。实施这一技术的可行性还取决于多种因素,如基础设施和稳定电力的保障、药盒和实验室服务的供应、对专业知识投资、有效的维修服务、监测能力强的工作人员、样品运输的正常化、可持续的资金筹集方式、透明的捐赠协议及简单的诊断流程。

最后,尽管 Xpert 提高了实验室操作的便利性和自动化,但在某些情况下,由于对技术的偏好可能会导致对传统显微镜检测方法的忽视,一旦 GeneXpert 设备出现故障,可能会需要恢复使用显微镜。此外,临床医生虽然对 Xpert 结果持较高信心,但由于上述挑战,有时可能会犹豫是否进行 Xpert 检测。

### (三)未来的研究方向

《2024 版指南》提出的研究重点可以概括为:(1)评估 Xpert Ultra 检测对患者重要结局的影响:主要包括治愈率、死亡率、诊断时间和开始治疗时间等。(2)评估 Xpert Ultra 对儿童肺结核和肺外结核的诊断准确性:评估应用 Xpert Ultra 检测胃液或粪便标本,用以诊断儿童肺结核和肺外结核的准确性。(3)评估不同标本类型的组合效果:目前有限的研究表明,非侵入性标本的组合与传统的胃液标本或诱导痰标本效果相当。(4)最低侵入性标本收集的最佳方法:研究额外的操作和定性研究,确定最低侵入性标本收集的最佳方法,尤其适用于低技能或低资源环境。(5)鼻咽吸痰方法的实施研究:针对适用于低技能或低资源环境的鼻咽吸痰方法的实施予以研究。(6)粪便作为诊断标本的研究:开展广泛的操作性研究,将粪便作为诊断标本纳入正常的诊断临床路径,评估粪便检测对患者重要结局的影响。(7)改进参考标准:确定一种改进的参考标准,准确定义儿童结核病和少菌标本,以提高现有诊断方法的准确

性。(8)开发新工具:提高儿童结核病的诊断正确率,新工具应该是快速、经济、可行并且为儿童及其父母所接受的。(9)比较不同的检测方法:包括 Xpert MTB/RIF 和 Xpert Ultra,确定哪些检测或策略具有更高的诊断准确性。(10)开发肺外结核快速即时诊断方法:应重点开发使用尿液等非侵入性程序采集标本的诊断方法和策略。(11)业务研究:进行业务研究,以确保检测在预定用途的环境中得到最佳使用。

## 二、Truenat MTB、MTB Plus 和 MTB-RIF Dx 检测法

### (一)推荐意见

(1)对于有肺结核症状和体征的成人和儿童,可将 Truenat MTB 或 MTB Plus 作为肺结核的初始诊断检测,而不是涂片显微镜检查/培养(有条件推荐,准确性证据等级中)。该建议包括涂片阴性的患者,也有可能推广至 HIV 感染者(people living with HIV, PLHIV)。(2)对于有肺结核症状和体征且 Truenat MTB 或 MTB Plus 阳性的成人和儿童,可将 MTB-RIF Dx 作为肺结核的初始诊断检测,而不是培养和表型 DST(有条件推荐,准确性证据等级极低)。

### (二)推荐理由和证据

#### 1. PICO 问题:包括 PICO 1。

PICO 1:在有肺结核体征和症状并在医疗保健机构寻求治疗的人群中,是否应使用 Molbio Truenat MTB、MTB Plus 和 MTB-RIF Dx(3 种方法统称“Truenat”)初始诊断肺结核和利福平耐药性?

《2024 版指南》评估了 Truenat MTB、MTB Plus 和 MTB-RIF Dx 的临床研究结果。在研究中,通过创新诊断基金委员会(Foundation for Innovative New Diagnostics, FIND)在 4 个国家的 19 个临床中心和 7 个参考实验室进行多中心前瞻性临床评估,旨在评估 Truenat 在成人肺结核和利福平耐药性诊断方面的效果,并以微生物确认(培养)作为参考标准,评估 Truenat 在预期使用环境(显微镜中心)中的诊断准确性。同时,该研究还将 Truenat 与参考实验室的 Xpert MTB/RIF 或 Xpert Ultra 进行了直接比较。由于敏感度相关证据不精确和不一致,证据等级被判定为低,但特异度证据等级高。对于 Truenat MTB Plus,因敏感度的不精确性,证据等级被判定为低,但特异度证据等级较高。对于 MTB-RIF Dx,186 例受试者的结果在敏感度评估方面,存在间接性和不一致性,证据等级被判定为极

低。由于利福平耐药患者数量少且大多数来自印度和秘鲁,研究者们对特异度的间接性表示高度关注。

2. 成本效益分析:《2024 版指南》对基于分子的活动性结核病诊断进行了经济评价,特别关注了新型 Molbio Truenat MTB 检测。目的在于总结目前的经济证据,评估这些分子检测的成本、成本效益和可负担性。目前,仅有一项研究评估了 Truenat MTB 的成本效益,结果表明,在印度医疗点实施该检测可能具有成本效益,但该研究依赖于几个关键的模型假设,需要在实际试验中进行评估。另外,进行成本效益和经济评估时,应谨慎考虑当地实施条件和环境,或许有助于评估对患者发现、长期结局和成本效益的可能影响。

3. 用户视角:《2024 版指南》在该部分主要讨论了实施 Truenat MTB、MTB Plus 和 MTB-RIF Dx 系统对患者诊断价值、可行性、可及性和公平性的影响。现有研究主要从患者和政策制定角度出发,基于 Xpert MTB/RIF 和 Xpert Ultra 的定性结果。虽然大部分 Xpert MTB/RIF 和 Xpert Ultra 的定性证据可适用于 Truenat 检测,但仍需谨慎,因为技术特征(如诊断准确性和适用的特定患者群体情况)可能有所不同。此外,相关的供应链和维护方面也可能不同。大多数实验室工作人员认为 Truenat 是可以接受和可行的,但也有人指出,相比 Xpert MTB/RIF,该方法操作更复杂。

### (三)未来的研究方向

WHO 指导小组指出,进行业务研究是确保检测在预期用途环境中得到最佳利用的关键步骤。另外,需要评估 Truenat 在特定患者群体中的诊断准确性,这些群体包括 HIV 感染者、曾经患肺结核的患者,以及患有肺外结核的成人和儿童患者。

三、用于诊断结核病和利福平、异烟肼耐药性的中等复杂程度的自动 NAAT

随着各国技术的不断发展,新的技术不断被研发出来,快速检测结核病和利福平耐药性的希望也越来越大。然而,在全球范围内,估计有 13.1% (95%CI: 9.9%~16.9%) 的新发患者和 17.4% (95%CI: 0.5%~54.0%) 的既往治疗患者发生异烟肼耐药、利福平敏感结核病。中等复杂程度的自动化 NAAT 有可能填补这个技术缺口。该类检测技术相对于表型培养的 DST 和线性探针测定(line probe assays, LPA)更快,操作更简单,在样品制备步骤之后基本实现了自动化。可作为同时检测结核病和对两种一线抗结核药物(利福平和异烟肼)耐药

性的初始检测。它们有可能快速提供准确的结果(对患者很重要),并在每日需要大量检测的情况下提高检测效率(对规划很重要),适用于人口密度高和样本需要快速转诊的地区。

### (一)推荐意见

对于有肺结核症状和体征的患者,可在呼吸道样本中使用 NAAT 检测肺结核,以及对利福平和异烟肼的耐药性,而不是培养和表型 DST(有条件推荐,准确性证据等级中)。

### (二)推荐理由和证据

1. PICO 问题:包括 PICO 1~3。

PICO 1:与培养相比,是否应该对有肺结核体征和症状患者的呼吸道样本使用中等复杂程度的自动 NAAT 来检测肺结核?

以分枝杆菌培养作为结核病检测研究的参考标准,评估了 29 项研究(13 852 份标本)。其中 16 项(55%)研究在纳入研究前检测了样本,采用了方便抽样或未报告参与者的选择方法,偏倚风险高或不明确。由于存在偏倚风险,证据被降级一级。综上,敏感度证据等级中,特异度证据等级高。

PICO 2:与基于培养的表型 DST 相比,是否应该在有肺结核体征和症状患者的呼吸道样本中使用中等复杂程度的自动 NAAT 检测利福平耐药性?

以表型 DST 作为抗性检测参考标准,采用表型 DST 和测序结果(基因型 DST)的复合参考标准,评估了 18 项研究(2874 份标本)。18 项研究中,有 8 项(44%)因未报告参与者选择或纳入研究前检测的标本,存在较高或不明确的偏倚风险。对利福平耐药性的总体敏感度为 88%~100%,特异度为 98%~100%;合并敏感度为 96.7% (95%CI: 93.1%~98.4%),合并特异度为 98.9% (95%CI: 97.5%~99.5%)。对于利福平耐药性检测,中等复杂程度的自动 NAAT 与表型 DST,以及复合参考标准,三者的诊断检测准确率是相近的。

PICO 3:与基于培养的表型 DST 相比,是否应该在有肺结核体征和症状患者的呼吸道样本中使用中等复杂程度的自动 NAAT 检测异烟肼耐药性?

以表型 DST 作为抗性检测参考标准,采用表型 DST 和基因型 DST 的复合参考标准评估了 18 项研究(1758 份标本)。18 项研究中,有 8 项(44%)因未报告参与者选择或纳入研究前检测的标本,存在较高或不明确的偏倚风险。对异烟肼耐药性的总体敏感度为 58%~100%,特异度为 94%~100%。合并敏感度为 86.4% (95%CI: 82.1%~89.8%),合

并特异度为 99.8% (95% CI: 98.3% ~ 99.8%)。对于异烟肼耐药性检测,中等复杂程度的自动 NAAT 与表型 DST,以及复合参考标准,三者的诊断检测准确率是相似的。

2. 成本效益分析:《2024 版指南》回顾了 2010 年 1 月 1 日至 2020 年 9 月 17 日期间发表的相关研究,还联系了专家和相关制造商以获取未发表的研究数据,评估了中等复杂程度的自动化 NAAT 的成本、成本效益和可负担性。

就资源需求(成本)而言,两种中等复杂程度的自动 NAAT 的单元测试成本范围在 18.52(13.79 ~ 40.70)美元和 15.37(9.61 ~ 37.40)美元之间。设备成本是造成成本差异的主要因素,并且会因实验室网络和操作的不同而有所不同。在单向敏感度分析中,每年的检测量从少于 5000 次到超过 25 000 次不等。当每年检测量较小时,单次检测费用对检测量非常敏感;但当每年检测量达到 5000 ~ 10 000 次时,单位测试成本开始稳定;当每年检测量超过 10 000 次时,单位成本估算是可靠的。此外,如果设备可以复用,并以高容量运行,则每次测试的成本可以降至最低。

关于资源需求(成本)证据的证据等级,可用的每项测试成本数据未公布,但成本又确实包括间接费用、设备、建筑、人员和消耗品成本。另外,由于研究中心、国家和实施方法之间存在重要差异,评估这些差异的证据也是有限的。

目前,尚无评估中等复杂程度的自动 NAAT 的成本效益的研究。由于涉及到标准护理、护理级联和相关成本、操作条件、测试量和诊断准确性等的差异,推断成本效益是不合适的。此外,实施方面的因素(如检测安排、实验室网络和项目快速启动治疗的能力)也可能影响单元测试成本和成本效益。因此,需要在不同设置中进行经济建模,以了解中等复杂程度的自动 NAAT 的成本效益概况和在不同操作标准下的变化情况。

3. 用户视角:在实施这些技术时,用户对诊断技术的价值、可行性、可用性和可接受性的看法很重要。如果不考虑实验室人员、临床医生、患者和结核病规划人员的观点,这些技术就有可能无法被目标受众获得和使用。中等复杂程度的自动 NAAT 符合临床医生和实验室工作人员的一些偏好和价值,因为它们比基于培养的表型 DST 更快,具有自动化优势,且能提供更多与临床相关的耐药性信息。使用中等复杂程度的自动 NAAT 检测结核病,以及对

利福平、异烟肼耐药性时应考虑当地耐药流行情况、成本估算、供应商支持服务和维护的可用性和及时性、实验室认证和质量管理体系等因素。

### (三)未来的研究方向

(1)特定患者群体(如儿童、HIV 感染者和有肺外结核症状和体征的患者)和非痰样本的诊断准确性。(2)在所有患者人群中,诊断技术对患者重要的临床决策和结局(例如治愈率、死亡率、诊断时间和开始治疗的时间)的影响。(3)特定突变对 DR-TB 患者治疗结局的影响。(4)在整个检测和护理领域中使用、整合和优化诊断技术、诊断路径和算法。(5)对不同诊断技术的成本、成本效益和成本效果进行评估的经济学研究。(6)评价不同诊断技术的公平性、可接受性、可行性和最终用户价值的定性研究。(7)不可采取行动的结果(不确定、非确定或无效)对诊断准确性和对患者重要结局的影响。(8)对中等复杂程度的自动 NAAT 中各技术优缺点的操作研究。(9)中等复杂程度自动化 NAAT 在促进疾病规划之间的协作和整合方面的作用。(10)检测 *katG* 耐药性以识别可能因没有 RRDR 突变而被遗漏的 MDR-TB 克隆(例如 Eswatini MDR-TB 克隆同时具有 *katG* S315T 和非 RRDR *rpoB* 1491F 突变)。

## 推荐意见二:结核病的初始诊断检测方法 (不包括耐药检测)

《2024 版指南》第二条推荐意见包含了 2 种商业化的检测方法,一种是基于环介导等温扩增技术(LAMP)的 Loopamp™ 结核分枝杆菌复合群(MTBC)检测试剂盒,称为 TB-LAMP。另一种是侧流尿脂阿拉伯甘露聚糖测定法(Alere LAM, Alere Determine™ TB LAM Ag assay),称为 LF-LAM。

### 一、TB-LAMP

#### (一)推荐意见

(1)对于有肺结核症状和体征的成人,TB-LAMP 可替代痰涂片镜检诊断肺结核(有条件推荐,证据等级极低)。(2)对于有肺结核症状和体征的成人,TB-LAMP 可作为痰涂片镜检的后续检测,尤其是在需要进一步检测痰涂片阴性标本时(有条件推荐,证据等级极低)。

备注:(1)这些建议适用于可以进行传统痰涂片显微镜检查的环境。(2)TB-LAMP 不应取代检测结核病和利福平耐药性的快速分子检测方法,特别是在有 MDR-TB 风险的人群中。(3)对于有体征和症状的 PLHIV,该检查与传统痰



涂片显微镜检查相比,额外诊断价值有限。(4)这些建议仅适用于检测与肺结核体征和症状一致患者的痰标本。(5)这些建议是基于成人数据的归纳,尽管收集儿童痰标本困难,但仍推断可在儿童中使用 TB-LAMP。

## (二)推荐理由和证据

1. PICO 问题:WHO 指导小组在制定本指南时基于 PICO 原则对商业 TB-LAMP 手动检测方法进行了证据质量等级的综合评估。

1)主要围绕以下 PICO 问题:(1)当 TB-LAMP 作为痰涂片显微镜检查的替代试验时,与培养作为参考标准相比,TB-LAMP 检测成人肺结核的诊断准确性是多少?(结果按 HIV 感染状况进行分层)。(2)当痰涂片显微镜检查阴性后的 TB-LAMP 作为参考标准时,TB-LAMP 检测成人肺结核的诊断准确性如何?(3)以分枝杆菌培养为参照,TB-LAMP 和 Xpert MTB/RIF 在所有成人中检测肺结核的诊断准确性有什么差异?(4)当 TB-LAMP 用于所有成人和 HIV 感染成人检测肺结核时,不确定或无效结果的比例是多少?

2)分析和结论:考虑到各研究进行的培养数量不同和参与者可获得的培养结果的数量不同,WHO 指导小组使用了 3 个基于培养的分级参考标准来评估诊断的准确性。

标准 1 包括:结核病:至少 1 份阳性培养样本通过分型检测鉴定为 MTBC。非结核病:两种不同痰液样本无阳性和至少两次培养阴性。

标准 2 包括:结核病:至少 1 份阳性培养样本通过分型检测鉴定为 MTBC。非结核病:至少 1 份痰液样本无阳性和至少两次培养阴性。

标准 3 包括:结核病:至少 1 份阳性培养样本通过分型检测鉴定为 MTBC。非结核病:无阳性和至少 1 次培养阴性。

在循证过程中,使用了来自 7 项研究的 1810 例受试者(符合标准 1)的数据,其中 TB-LAMP 的敏感度达到 77.7%(95%CI:71.2%~83.0%),比涂片镜检的敏感度[63%(95%CI:56.0%~69.0%)]高约 14.7%。然而,TB-LAMP 的特异度为 98.1%(95%CI:95.7%~99.2%),比涂片镜检的特异度[100.0%(95%CI:97.0%~100.0%)]低约 1.9%。

另外,使用了 1349 例疑似结核病患者(符合标准 1)的数据,在 7 项研究中,TB-LAMP 在涂片显微镜检查阴性的假定结核病患者中的合并敏感度为 42.1%(95%CI:30.0%~55.3%),特异度为 98.4%(95%CI:95.9%~99.4%)。对于涂片阴性患者,

TB-LAMP 的阳性总百分比(total positive percentage, TPP)目标为 70.0%,而即使是使用 Xpert MTB/RIF 检测,其 TPP 为 68.0%,还是达不到 TB-LAMP 的性能目标。

在循证过程中,根据纳入该综述的 4 项研究(271 例 HIV 感染者,符合标准 2)的数据,TB-LAMP 的敏感度为 63.8%(95%CI:49.0%~76.4%)。然而,涂片镜检的敏感度异常地达到了 62.0%(95%CI:34.0%~89.0%),这引出了一个问题:TB-LAMP 的敏感度是否因涂片阳性患者比例异常高而被人为提高。TB-LAMP 的特异度为 98.8%(95%CI:85.1%~99.9%),涂片镜检则为 99.0%(95%CI:95.0%~100.0%)。WHO 指导小组判断,在可能患有肺结核的 HIV 阳性成人中,TB-LAMP 与涂片镜检相比没有额外的价值。

2. 成本效益分析:WHO 指导小组开展了一项自下而上的微观成本分析,旨在确定、测量和评估马拉维和越南的周边实验室提供 TB-LAMP 和 Xpert MTB/RIF 作为常规诊断检测相关的所有资源。将两种 TB-LAMP 策略(作为痰涂片镜检的替代试验和痰涂片镜检阴性患者的附加检测)与基础病例算法(对涂片镜检结果阴性的患者进行痰涂片镜检后再进行临床诊断)进行比较。结果显示,这两种 TB-LAMP 策略的成本均比 Xpert MTB/RIF 的成本更低,而且都提高了病例检出率,但 TB-LAMP 替代痰涂片镜检策略的成本有所增加。尽管如此,与 WHO 的支付意愿阈值水平相比,这两种策略都具有成本效益。以上这些结论在所进行的一系列敏感度分析中没有变化,与 HIV 阳性患者群体相比,没有单独对 HIV 阴性患者群体进行成本效益分析。成本效益分析结果表明,TB-LAMP 有可能成为痰涂片镜检联合临床诊断这一基本方案的成本效益替代方案。

3. 实施注意事项:在实施的过程中,需要注意的事项主要包括:(1)操作相关的问题:TB-LAMP 的制造商建议了相关的储存温度、电力需求、储存空间、废物处理、库存监测及温度控制(通常为 30℃)。(2)TB-LAMP 已被评估可用于检测痰标本中的 MTB,尚未对其他样本(如尿液、血清、血浆、脑脊液或其他体液)的适用性进行充分评估。(3)采用 TB-LAMP 并不能消除涂片显微镜检查的需求,后者可以监测药物敏感结核患者的治疗。然而,在 TB-LAMP 完全或部分取代传统涂片显微镜检查的情况下,对后者的需求可能会减少。(4)TB-LAMP

不应取代 Xpert MTB/RIF, 因为后者可以同时检测 MTB 和利福平耐药性, 而且是自动化的, 执行起来相对简单。(5) 在无法采用 Xpert MTB/RIF 检测的情况下(例如, 由于电力供应不足, 或温度、湿度、灰尘过高), TB-LAMP 可能是一种合理的替代方案。

### (三) 未来的研究方向

(1) 对不同流行病学、地理环境及患者群体中的诊断算法进行评估。(2) 采用更高质量的参考标准进行更严格的研究(包括多种标本类型和肺外标本), 以提高特异性评估的可信度。(3) 明确培训需求, 评估能力和质量。(4) 收集更多影响结核病治疗启动、发病率和死亡率的证据。(5) 针对特定国家进行成本效益和成本效果分析, 以评估在不同项目背景下有针对性地使用 TB-LAMP 的经济效益。(6) 在未来研究中符合诊断准确性研究报告标准(standards for reporting diagnostic accuracy studies, STARD)。

## 二、LF-LAM

### (一) 推荐意见

(1) 在住院环境中, 对于以下情况: 有结核病(肺部和/或肺外)症状和体征的患者、HIV 感染晚期疾病(advanced HIV disease, AHD)患者或重症患者、无论是否有结核病症状和体征但是  $CD4^+$  T 细胞计数低于  $200$  个/ $mm^3$  的患者, 强烈推荐使用 LF-LAM 协助诊断 HIV 阳性成人、青少年和儿童的活动性结核病(强烈推荐, 有关干预效果的证据等级中)。(2) 在门诊环境中, 对于以下情况: 有结核病(肺部和/或肺外)症状和体征的患者或病情严重的患者(有条件推荐, 准确性证据等级低)和无论是否有结核病症状和体征但是  $CD4^+$  T 细胞计数低于  $100$  个/ $mm^3$  的患者(有条件推荐, 准确性证据等级极低), 推荐使用 LF-LAM 协助诊断 HIV 阳性成人、青少年和儿童的活动性结核病。(3) 在门诊环境中, 对于以下情况: 未评估结核病症状的患者(强烈推荐, 准确性证据等级极低)、无结核病症状且  $CD4^+$  T 细胞计数未知或无结核病症状但  $CD4^+$  T 细胞计数  $\geq 200$  个/ $mm^3$  的患者(强烈推荐, 准确性证据等级极低)、无结核病症状但  $CD4^+$  T 细胞计数为  $100 \sim 200$  个/ $mm^3$  的患者(有条件推荐, 准确性证据等级极低), 不建议使用 LF-LAM 协助诊断 HIV 阳性成人、青少年和儿童的活动性结核病。

备注: (1) 上述建议仅适用于 Alere LAM 的使用, 因为其他基于 LAM 的内部分析方法还没有得到充分验证, 或仅

在有限的研究场景中使用。(2) 所有存在肺结核体征和症状且能够产生痰的患者应提供至少一份痰标本进行 Xpert MTB/RIF 或 Xpert Ultra 检测, 作为其初始诊断, 包括能够提供痰液样本的 HIV 感染的儿童和青少年。(3) 以上这些推荐建议也适用于青少年和感染 HIV 的儿童, 其依据是来自成年人的综合数据, 即使关于这些人口群体的数据采集非常有限。(4) LF-LAM 检测应作为辅助临床判断的工具, 与其他诊断测试联合使用, 而不应作为替代手段或用于初步筛选的独立检测。

### (二) 推荐理由和证据

#### 1. PICO 问题: 包括 PICO 1~4。

PICO 1: 对于有结核病体征和症状的 HIV 阳性成人的结核病诊断, Alere LAM 的诊断准确性如何?

在住院环境中, 敏感度为  $52\%$  ( $95\%CI: 40\% \sim 64\%$ ), 特异度为  $87\%$  ( $95\%CI: 78\% \sim 93\%$ ); 在门诊环境中, 敏感度为  $29\%$  ( $95\%CI: 17\% \sim 47\%$ ), 特异度为  $96\%$  ( $95\%CI: 91\% \sim 99\%$ ); 在所有环境中, 敏感度为  $42\%$  ( $95\%CI: 31\% \sim 55\%$ ), 特异度为  $91\%$  ( $95\%CI: 85\% \sim 95\%$ )。

PICO 2: 对于 HIV 阳性成人的结核病诊断, 无论结核病的体征和症状如何, Alere LAM 的诊断准确性如何?

在住院环境中, 敏感度为  $62\%$  ( $95\%CI: 41\% \sim 83\%$ ), 特异度为  $84\%$  ( $95\%CI: 48\% \sim 96\%$ ); 在门诊环境中, 敏感度为  $31\%$  ( $95\%CI: 18\% \sim 47\%$ ), 特异度为  $95\%$  ( $95\%CI: 87\% \sim 99\%$ ); 在所有环境中, 敏感度为  $35\%$  ( $95\%CI: 22\% \sim 50\%$ ), 特异度为  $95\%$  ( $95\%CI: 89\% \sim 98\%$ )。而对于 HIV 阳性儿童的结核病诊断, 在所有环境下(包括所有儿童), 一项在门诊进行的研究显示: 敏感度为  $42\%$  ( $95\%CI: 15\% \sim 72\%$ ), 特异度为  $94\%$  ( $95\%CI: 73\% \sim 100\%$ ); 在住院环境中进行的研究显示: 敏感度为  $56\%$  ( $95\%CI: 21\% \sim 86\%$ ), 特异度为  $95\%$  ( $95\%CI: 90\% \sim 98\%$ ); 一项在住院和门诊进行的研究显示, 敏感度为  $43\%$  ( $95\%CI: 23\% \sim 66\%$ ), 特异度为  $80\%$  ( $95\%CI: 69\% \sim 88\%$ )。

PICO 3: 对于 AHD 成人的结核病诊断, 无论结核病的体征和症状如何, Alere LAM 的诊断准确性(可用数据有限)如何?

在住院环境中,  $CD4^+$  T 细胞计数  $\leq 200$  个/ $mm^3$ , 敏感度为  $64\%$  ( $95\%CI: 35\% \sim 87\%$ ), 特异度为  $82\%$  ( $95\%CI: 67\% \sim 93\%$ ) (一项研究); 在门诊环境中,  $CD4^+$  T 细胞计数  $\leq 200$  个/ $mm^3$ , 敏感度为  $21\%$  ( $95\%CI: 8\% \sim 48\%$ ), 特异度为  $96\%$  ( $95\%CI:$

89%~99%);在所有情况下,CD4<sup>+</sup> T 细胞计数 $\leq$ 200 个/mm<sup>3</sup>,敏感度为 26%(95%CI:9%~56%),特异度为 96%(95%CI:87%~98%);在住院环境中,CD4<sup>+</sup> T 细胞计数 $\leq$ 100 个/mm<sup>3</sup>,敏感度为 57%(95%CI:33%~79%),特异度为 90%(95%CI:69%~97%);在门诊环境中,CD4<sup>+</sup> T 细胞计数 $\leq$ 100 个/mm<sup>3</sup>,敏感度为 40%(95%CI:20%~64%),特异度为 87%(95%CI:68%~94%);在所有情况下,CD4<sup>+</sup> T 细胞计数 $\leq$ 100 个/mm<sup>3</sup>,敏感度为 47%(95%CI:30%~64%),特异度为 90%(95%CI:77%~96%)。

PICO 4:在 HIV 感染的成年人中使用 Alere LAM 能否降低与 AHD 相关的死亡率?

对于使用 Alere LAM 来降低与 AHD 相关的死亡率(两项随机试验):死亡率的合并风险比为 85%(95%CI:76%~94%);绝对效应为每 1000 人减少 35(95%CI:14~55)人死亡。

2. 成本效益分析:要实施和扩大一个新的诊断方法需要更全面的经济证据。有研究对 Alere LAM 检测 HIV 阳性患者活动性结核病的经济评估进行了系统的回顾,旨在汇集并综合现有证据,以理解 Alere LAM 在 HIV 阳性人群中结核病诊断实施的成本、成本效益和预算影响下的可负担性。共纳入了 6 项全部来自撒哈拉以南非洲地区的研究,研究方法和人群是异质性的,评估了一系列诊断算法,其中只有 4 项研究评估了成本效益。结果表明,LF-LAM 在非洲成人 HIV 感染者中(特别是在住院患者中)很有可能具有成本效益。2018 年的一项研究进行了详细的 Alere LAM 微观成本分析,报告中显示 Alere LAM 的单位检测成本为 23.55 美元,比之前估计的 3~3.99 美元高出数倍,这在很大程度上是由于纳入了人力成本等临床成本。如果成本效益模型低估了实施 Alere LAM 的真实成本,那么增量成本效益比(incremental cost effectiveness ratio, ICER)可能不会那么有吸引力,项目的成本可能更高。然而,随着规模的继续扩大,规模经济和其他效益可能会提高。

3. 用户视角:关于 LF-LAM 的用户视角定性研究,主要采取了半结构化访谈的方法。在 2019 年 2 月和 3 月, Nora Engel 和 Muthoni Mwaura 采访了肯尼亚、南非和乌干达的临床医生、护士、项目官员、实验室工作人员和患者权益倡导者,旨在更广泛地了解使用 LF-LAM 在 PLHIV 中诊断结核病的经验,并了解用户对 LF-LAM 的偏好。研究结果显

示,本研究的参与者认为 LF-LAM 是一种易于使用、快速的检测方法,更重要的是它检测的是更安全更容易获得的尿液标本,而不是依赖痰标本。然而,还存在以下问题:(1)根据测试使用的具体情境和测试应用的能力,样品的感知好处、周转时间、用户友好性、成本和维护需求也可能构成挑战。(2)对基础设施的要求较少,但是在尿液标本容器、存货和保质期方面仍然存在挑战。(3)尽管周转时间理论上只有 25 min,但在很多情况下,治疗直到第二天才开始。

总的来说,考虑到对这一特定患者群体缺乏可行的诊断替代方案,LF-LAM 的普及利大于弊。此外,关注 LF-LAM 的具体操作才是促进其实施的关键。

### (三)未来的研究方向

(1)开发基于 LAM 的简单、更准确的检测方法,并有可能用于 HIV 阴性人群。(2)评估 LF-LAM 在无结核病症状和体征 PLHIV 中的应用。(3)评估 LF-LAM 在感染 HIV 的儿童和青少年中的应用。(4)评估 LF-LAM 和快速定性 CD4<sup>+</sup> T 细胞计数系统组合的应用。(5)对常规临床环境中 LF-LAM 的接受程度、推广和影响进行实施研究。(6)对 LF-LAM 的可行性、可及性和公平性问题的用户视角进行定性研究。(7)开展将 LF-LAM 纳入 HIV 护理包的实施研究。(8)随着 HIV 流行的发展和更多的患者住院接受病毒载量抑制治疗,评估在此过程中 LF-LAM 的表现。(9)评估 LF-LAM 的成本效益。(10)评估其他基于 LAM 的快速测试,如 FujiLAM(Fujifilm SILVAMP TB LAM)。

### 推荐意见三:结核病确诊后的耐药性检测

一、检测对异烟肼和二线抗结核药物耐药性的低复杂程度的自动 NAAT

#### (一)推荐意见

(1)对于经细菌学确诊的肺结核患者,推荐使用低复杂程度的自动 NAAT 对痰液进行异烟肼和氟喹诺酮类药物耐药性的初始检测,而不是基于培养的表型 DST(有条件推荐,准确性证据等级中)。(2)对于经细菌学确诊和对利福平耐药的肺结核患者,推荐使用低复杂程度的自动 NAAT 对痰液进行乙硫异烟胺耐药性的初始检测,而不是对 *inhA* 启动子进行 DNA 测序(有条件推荐,准确性证据等级极低)。(3)对于经细菌学确诊和对利福平耐药的肺结核患者,推荐使用低复杂程度的自

动 NAAT 对痰液进行阿米卡星耐药性的初始检测,而不是基于培养的表型 DST(有条件推荐,准确性证据等级低)。

## (二)推荐理由和证据

### 1. PICO 问题:包括 PICO 1~4。

PICO 1:对于有肺结核体征和症状患者的痰标本,与基于培养的表型 DST 相比,是否应该将低复杂程度的自动 NAAT 用于检测对异烟肼的耐药性,而不考虑对利福平的耐药性?

首先,在所纳入的研究中,异烟肼耐药发生率中位数为 67.2%,高于全球异烟肼耐药发生率的估值。因此,对于异烟肼耐药发生率较低的环境,该诊断方法的适用性存在一定的不确定性。其次,异烟肼单耐药菌株和耐多药菌株的突变存在差异,有研究表明,在单耐药菌株中发现的突变更多样化。最后,尽管这个 PICO 问题的人群是“不考虑利福平耐药性”,但研究的入组标准意味着纳入研究的大多数参与者患有 RR-TB。由于这些担忧,证据的等级在敏感度和特异度方面因其间接性而被降低一级,皆被评为证据等级中。

PICO 2:对于有肺结核体征和症状患者(无论对利福平是否耐药)的痰液,与基于培养的表型 DST 相比,是否应该使用低复杂程度的自动 NAAT 检测对氟喹诺酮类药物的耐药性?

3 项包含 1337 例受试者的多国研究提供了用于评估氟喹诺酮类药物耐药性检测的数据,且研究的参考标准都是基于培养的表型 DST。多国研究中的每个研究中心都作为单独的研究进行了分析,特异度估计值不一致,孟买为 84%,新德里为 91%,其他研究中心均超过 96%。因此,证据等级降低一级,最终敏感度证据等级被判定为高,特异度证据等级被判定为中。

PICO 3:对于有肺结核体征和症状且对利福平耐药患者的分离培养物,与 *inhA* 启动子的基因型测序相比,是否应该将低复杂程度的自动 NAAT 用于检测乙硫异烟胺的耐药性?

一项纳入 434 例受试者的多国研究中,以 *inhA* 启动子的 DNA 测序分析作为参考标准。由于该研究采用参考标准未包括正确分类参考标准的所有位点(即 *ethA*、*ethR* 和 *inhA* 启动子),因此存在非常严重的偏倚风险。与表型 DST 相比,合并敏感度下降至 51.7%(95%CI:33.1%~69.8%)。由于敏感度和特异度均存在偏倚风险,证据等级降低两级。另

外,敏感度和特异度的 95%CI 都很宽,存在不精确性,证据等级降低一级。最终,敏感度和特异度证据等级均被判定为极低。

PICO 4:对于有肺结核体征和症状且对利福平耐药患者的痰液,与基于培养的表型 DST 相比,是否应该采用低复杂程度的自动 NAAT 检测阿米卡星的耐药性?

一项纳入 490 例受试者的多国研究中,以基于培养的表型 DST 为参考标准。由于敏感度的 95%CI 很宽,且很少有对阿米卡星耐药的参与者参与这项敏感度分析,因此,证据等级降低两级。最终,敏感度证据等级均被判定为低,特异度证据等级均被判定为高。

2. 成本效益分析:目前,有两种低复杂程度的自动 NAAT:即 MeltPro MTB/RIF(厦门致善生物科技股份有限公司)和 Xpert MTB/XDR(美国赛沛公司)。与 Xpert MTB/RIF 一样,Xpert MTB/XDR 可用于检测未经处理或浓缩的痰液。然而,缺乏可直接证明低复杂程度自动 NAAT 成本或成本效益的已发表研究。通过与 Xpert MTB/XDR 制造商(美国赛沛公司)直接沟通了解到,中低收入国家 Xpert MTB/XDR 的出厂价预计为 19.80 美元,运输和海关费用将由订购国或组织承担。此外,Xpert MTB/XDR 的成本效益受到诊断准确性等多个因素的影响。总成本也将因检测量和人口中广泛耐药结核病发生率的不同而不同。此外,对预算的影响还取决于当前的治疗标准和相关的资源使用情况。

## (三)未来的研究方向

(1)提高其在特定患者群体(如儿童、HIV 感染者和有肺外结核体征和症状的患者)和非痰样本中的诊断准确性。(2)在所有患者人群中,评估诊断技术对患者重要的临床决策和结局的影响(例如治愈、死亡率、诊断时间和开始治疗的时间)。(3)评估特定突变对 DR-TB 患者治疗结果的影响。(4)在整个检测和护理过程中,使用、整合和优化诊断技术,以及诊断路径和算法。(5)对不同诊断技术的成本、成本效益和成本效果进行评估的经济学研究。(6)定性研究不同诊断技术的公平性、可接受性、可行性和最终用户价值。(7)评估低复杂程度的自动 NAAT 是否可用于结核病的初始检测,并将其作为有结核体征和症状的患者、儿童和 HIV 感染者的一种后续检测。(8)评价 *katG* 耐药性检测的潜在效用,以识别可能因没有 RRDR 突变而被遗漏的 MDR-TB 克隆(例如,同时具有 *katG* S315T 和非

RRDR *rpoB* I491F 突变的 Eswatini MDR-TB 克隆)。

## 二、一线 LPA

### (一) 推荐意见

对于痰涂片阳性标本或培养分离出 MTBC 的患者,推荐将商用分子 LPA 作为初始检测方法,而非基于培养的表型 DST,以检测对利福平和异烟肼的耐药性(有条件推荐,准确性证据等级中)。

备注:(1)这些建议适用于使用 LPA 检测痰涂片阳性标本(直接检测)和从肺/肺外部位培养的 MTBC 分离株(间接检测);(2)不建议将 LPA 直接用于痰涂片阴性标本的检测;(3)这些建议适用于 MTBC 的检测和 MDR-TB 的诊断,但检测利福平和异烟肼耐药性的准确性不同,因此,降低了 MDR-TB 诊断的准确性;(4)这些建议并没有消除对传统的基于培养 DST 的需求,因为 DST 对于确定患者对其他抗结核药物的耐药性和监测更多耐药性的出现是必要的;(5)当 LPA 结果未检测到异烟肼耐药性时,传统的基于培养的异烟肼 DST 仍可用于评估患者。

### (二) 推荐理由和证据

1. PICO 汇总:(1)在对来自具有结核病症状和体征患者的标本进行直接检测和对培养分离物进行间接检测时,LPA 是否能够指导临床决定使用利福平/异烟肼?(2)是否应使用 LPA 诊断有结核病症状和体征的 MDR-TB 患者?(3)对于有结核病症状和体征但痰涂片结果为阴性的患者,LPA 是否能够用于诊断结核病?

将 LPA 与基于表型培养的 DST 参考标准,以及将基因测序结果与基于表型培养的 DST 结果相结合的复合参考标准进行比较。表型 DST 是所有分析中适用于所有参与者的主要参考标准。这些分析是分层进行的,首先,按对利福平或异烟肼(或两者)的敏感度或耐药性进行分层;其次,按 LPA 检测类型(间接检测或直接检测)进行分层。一些研究关注了敏感度(无真阳性和假阴性)或特异度(无真阴性和假阳性),但缺乏两者同时关注的研究。对于这些研究,研究人员分别对敏感度或特异度的估计值进行了单变量、随机效应的荟萃分析,以最佳地利用数据,但可用于荟萃分析的研究太少。最终,判定准确性证据等级中。

2. 实施注意事项:《2024 版指南》指导小组指出,采用 LPA 检测利福平和异烟肼耐药性并不能消除对传统培养和 DST 能力的需求。培养和基于表型培养的 DST 在监测患者对治疗的反应和检测对二线抗结核药物的耐药性方面发挥着至关重要的作用。

### (三) 未来的研究方向

(1)提高对使用基于培养的 DST 检测耐药突变与患者预后之间相关性的理解。(2)审查证据,以确认或修订基于培养的 DST 方法中使用的不同临界浓度。(3)确定 LPA 检测异质性耐药的检测限。(4)确定培训需求,评估能力并确保质量。(5)评估 LPA 对 MDR-TB 患者启动适当治疗对死亡率的影响。(6)符合未来诊断研究的 STARD 标准。(7)对不同背景下使用 LPA 的成本效益和成本效果进行国别分析。

## 三、二线 LPA

MTBDRsl(1.0 版)是首个用于检测二线抗结核药物耐药性的商用二线 LPA。2015 年,制造商开发了 MTBDRsl 2.0 版本,可检测氟喹诺酮类和二线注射药物(second-line injectable drug, SLID)耐药性相关突变及其他突变。一旦确诊为 MDR/RR-TB,就可以使用二线 LPA 检测其对二线药物的耐药性。

### (一) 推荐意见

(1)对于确诊的 MDR/RR-TB 患者,推荐使用二线 LPA 作为初始检测方法,而不是基于培养的表型 DST,以检测对氟喹诺酮类药物的耐药性。(2)对于确诊的 MDR/RR-TB 患者,推荐使用二线 LPA 作为初始检测,而不是基于培养的表型 DST,以检测对 SLID 的耐药性。

备注:(1)这些建议适用于使用二线 LPA 检测痰标本(直接检测)和来源于肺/肺外样本的分离培养物(间接检测)。直接检测痰标本可以更早地开始适当的治疗;(2)无论涂片结果如何,这些建议适用于直接检测 MDR/RR-TB 患者的痰标本;另外,该方法在涂片阴性痰标本中的不确定率高于涂片阳性痰标本;(3)传统表型 DST 对于确认患者对其他药物的耐药性和监测更多耐药性的出现是必要的;(4)传统的表型 DST 仍可用于评估二线 LPA 阴性结果的患者,特别是在对氟喹诺酮类药物或 SLID(或两者)耐药概率较高的人群中;(5)这些建议是基于对成人数据的归纳,但也适用于在确诊为 MDR/RR-TB 的儿童中使用二线 LPA;(6)二线 LPA 检测到的耐药性相关突变与对氧氟沙星、左氧氟沙星和 SLID 的表型耐药性高度相关;(7)考虑到二线 LPA 检测氟喹诺酮类药物和 SLID 耐药的高特异度,其阳性结果可用于指导实施适当的感染控制措施。

### (二) 推荐理由和证据

1. PICO 汇总:是否应使用 MTBDRsl 检测来指导临床决策,以便在确诊为 MDR/RR-TB 的患者中使用氟喹诺酮类药物/SLID?

共确定了 29 项独特的研究,其中,26 项评估了 MTBDRsl 1.0 版,3 项研究评估了 2.0 版。采用基

于表型培养的 DST 参考标准,对 1.0 版和 2.0 版 MTBDRsl 试验数据分别进行分析。首先,根据对特定药物的敏感度或耐药性对这些分析进行分层,其次,根据二线 LPA 检测类型(间接检测或直接检测)对这些分析进行分层。然而,关于 2.0 版的研究太少,导致数据过于稀少或过于分散,无法进行荟萃分析,也无法对间接检测和直接检测进行比较分析。三项研究评估了 MTBDRsl 2.0 版对 562 例患者的检测结果,其中包括 111 例确诊的对氟喹诺酮类药物耐药的结核病患者。以基于培养的表型 DST 作为参考标准,其敏感度估计值介于 84%~100% 之间,特异度介于 99%~100% 之间。

2. 实施注意事项:二线 LPA 只能用于检测确诊为 MDR/RR-TB 患者的标本。尽管二线 LPA 在检测氟喹诺酮类药物和 SLID 的耐药性方面具有良好的特异度,但要完全排除对这些药物以及其他二线药物的耐药性,还需要进行培养和表型 DST 的检测。

### (三)未来的研究方向

(1)提高对耐药突变检测与表型 DST 结果和患者预后之间相关性的理解。(2)进一步了解二线 LPA 检测到的特定突变与氟喹诺酮类药物和 SLID 的最低抑菌浓度之间的相关性。(3)确定二线 LPA 检测异质性耐药的检测限。(4)评估 MTBDRsl 对 MDR-TB 患者启动适当治疗对死亡率的影响。(5)遵循 STARD 标准中的建议,提高报告质量。(6)对不同环境下使用二线 LPA 的成本效益和成本效果进行国别分析。

### 四、用于检测吡嗪酰胺耐药性的基于反向杂交的高复杂度 NAAT

吡嗪酰胺具有强大的抗 MTB 作用,以及与其他抗生素的协同作用,是治疗 DS-TB 和 DR-TB 的重要抗生素。对吡嗪酰胺产生单一耐药性的情况很少见,然而,吡嗪酰胺耐药性与 MDR/RR-TB 密切相关,约有 30%~60% 的 MDR/RR-TB 患者也对吡嗪酰胺产生耐药性。因此,对于确诊为 RR-TB 的患者来说,检测是否存在吡嗪酰胺耐药性非常重要,可以指导临床医生决定是否将吡嗪酰胺纳入治疗方案。基于反向杂交的高复杂度 NAAT 可用于诊断患者分离物对吡嗪酰胺的耐药性,然而,实现这种检测需要适当的基础设施和熟练的工作人员。

#### (一)推荐意见

对于经细菌学确诊的结核病患者,可对分离培养出的 MTB 培养物,使用基于反向杂交的高复杂

度 NAAT 检测吡嗪酰胺耐药性,而不是基于培养的表型 DST(有条件推荐,准确性证据等级极低)。

#### (二)推荐理由和证据

1. PICO 问题:包括 PICO 1~2。

PICO 1:与基于培养的表型 DST 或复合参考标准相比,是否应将基于痰液的反向杂交的高复杂度 NAAT 用于诊断微生物学确诊的肺结核患者对吡嗪酰胺的耐药性(无论是否对利福平耐药)?

3 项研究(122 例受试者)为评价 NAAT 检测吡嗪酰胺的耐药性提供了数据,其中 2 项研究(101 例受试者)采用基于培养的表型 DST 作为参考标准,1 项研究(21 例受试者)采用基因型 DST 作为参考标准。研究和参与者的数量被认为不足以对基于反向杂交的高复杂度 NAAT 的诊断准确性做出结论。

PICO 2:与基于培养的表型 DST 相比,是否应该使用基于分离培养物的反向杂交高复杂度 NAAT 用于诊断微生物学确诊的肺结核患者对吡嗪酰胺的耐药性(无论是否对利福平耐药)?

共有 7 项研究(964 例受试者)提供了评估 NAAT 检测吡嗪酰胺耐药性的数据。这些研究存在选择偏倚,因为它们选择了具有多种不同 *pncA* 突变的分离株,而不是从群体中选取具有代表性的菌株,证据等级被降低一级。另外,因纳入的研究未直接解决综述问题,证据因间接性而降低一级。与其他研究相比,Burhan 研究和 Rienthong 研究的敏感度为离群值,证据因不一致性被降低一级。综上,敏感度证据等级被判定为极低,特异度证据等级被判定为低。

2. 成本效益分析:尚未发现已发表的研究对使用市售的基于反向杂交的高复杂度 NAAT(Genoscholar PZA-TB II,日本尼普洛株式会社)成本或成本效益进行评估。Genoscholar PZA LPA 是为与日本尼普洛株式会社自动 MultiBlot 设备一起使用而开发的,然而,最近一项未发表的试验表明,Hain Lifescience 公司的 Twincubator 平台可以成功地与这种 LPA 配合使用,使得 Genoscholar PZA LPA 更容易在已使用 Hain Lifescience 设备的特定环境中实施。

关于所需资源总额,没有从已发表的研究中找到直接证据。所需资源包括购买检测试剂盒(Genoscholar PZA LPA:16 美元/单位检测包)和设备(14 000 美元)。日本尼普洛株式会社希望,当 Genoscholar PZA-TB II 产品在全球销售时,能够进一步降低检测成本。然而,间接数据表明,Genos-

cholar PZA-TB II 的单位检测总成本可能比 16 美元的单位检测包成本高出数倍。总成本将因检测量、符合检测条件的人数和人群中吡嗪酰胺耐药性的流行程度而异。对预算的影响将取决于当前的治疗标准、诊断和治疗途径及相关的资源使用情况。

3. 用户视角:关于最终用户观点定性证据的综合和分析,已在上文“中等复杂程度的自动 NAAT 的用户视角”一节中讨论。

### (三)未来的研究方向

(1)在有结核病症状和体征、对利福平有或无耐药性的人群中,评估基于反向杂交的高复杂度 NAAT 对痰和非痰样本间接检测的诊断准确性。(2)在所有患者人群中,评价诊断技术对临床决策和患者重要结局的影响(例如治愈率、死亡率、诊断时间和开始治疗的时间)。(3)特定突变对 DR-TB 患者治疗结局的影响。(4)分析诊断技术的成本、成本效益和成本效果。(5)定性研究以评估诊断技术的公平性、可接受性、可行性和最终用户价值。(6)比较分析基因型 DST、表型 DST 和基于反向杂交的高复杂度 NAAT。

### 五、tNGS

tNGS 技术将选定基因的扩增与 NGS 技术结合起来,只需一次检测就能发现多种药物的耐药性。同时,由于 tNGS 可以检测整个基因,确定与耐药性相关的特定突变,因此基于该技术的检测可能比现有的 WHO 推荐的诊断检测方法(WHO-recommended diagnostic tests, WRD)更准确。此外,基于 NGS 的新检测方法还能检测到目前任何其他分子检测方法未纳入的新药和重新利用的药物的耐药性。因此,基于 tNGS 的检测在提供与现代治疗方案相匹配的全面耐药性检测方面具有巨大潜力。

#### (一)推荐意见

(1)对于经细菌学确诊的肺结核患者,可在呼吸道样本中使用 tNGS 技术来诊断对利福平、异烟肼、氟喹诺酮类药物、吡嗪酰胺和乙胺丁醇的耐药性,而不是基于培养的表型 DST(有条件推荐,对于异烟肼和吡嗪酰胺的证据等级中,对于利福平、氟喹诺酮类药物和乙胺丁醇的证据等级低)。

备注:(1)应优先考虑对一线抗结核药物耐药风险较高的患者,包括以下人员:在治疗 2 个月或更长治疗后,涂片或培养仍呈阳性,或出现治疗失败;以前接受过抗结核治疗;与已知对抗结核药物具有耐药性的患者有过接触;可能对利福平、异烟肼或氟喹诺酮类药物耐药的人群。(2)对于经细菌学确诊的利福平耐药肺结核患者,可在呼吸道样本中使用

tNGS 技术检测对异烟肼、氟喹诺酮类药物、贝达喹啉、利奈唑胺、氯法齐明、吡嗪酰胺、乙胺丁醇、阿米卡星和链霉素的耐药性,而不是基于培养的表型 DST(有条件推荐,对于异烟肼、氟喹诺酮类药物和吡嗪酰胺证据等级高,对于乙胺丁醇证据等级中,对于贝达喹啉、利奈唑胺、氯法齐明和链霉素证据等级低,对于阿米卡星证据等级极低)。备注:应优先考虑用于有较高耐药风险的患者,包括以下人群:经过 2 个月或更长时间的治疗后,涂片或培养仍呈阳性,或已经历治疗失败;以前接受过抗结核治疗,包括使用新药和重新利用的药物;对氟喹诺酮类药物耐药的广泛耐药结核病患者。

#### (二)推荐理由和证据

1. PICO 问题:包括 PICO 1~2。

PICO 1:tNGS 是否应作为细菌学确诊肺结核患者的初始检测方法来诊断耐药性?

分析中包含的现有证据因药物而异,12 项研究(1440 例受试者)涉及异烟肼的敏感度,3 项研究(269 例受试者)涉及吡嗪酰胺的特异度。因为所有研究都对利福平耐药进行了富集,导致适用性问题,所有药物均因敏感度和特异度的间接性而降低一级。此外,对于利福平、左氧氟沙星和吡嗪酰胺,特异度因不精确性,被进一步降级。对于乙胺丁醇,由于用于指标和参考测试的样本不同,存在偏倚风险而被降级。检测准确性证据的总体证据等级范围为中度至极低。经测定,该检测对所有纳入评估的药物均准确,对异烟肼、莫西沙星和乙胺丁醇的合并敏感度至少为 95%,对利福平和左氧氟沙星的合并敏感度超过 93%,对吡嗪酰胺的合并敏感度为 88%。对于所有药物的合并特异度至少为 96%。

以基于培养的表型 DST 作为异烟肼、左氧氟沙星和莫西沙星的对照标准,以表型 DST 和全基因组测序(whole genome sequencing, WGS)联合作为利福平、吡嗪酰胺和乙胺丁醇的对照标准。检测结果不确定的百分比范围为 9%(左氧氟沙星和莫西沙星)~18%(吡嗪酰胺),在细菌载量较低(半定量类别低或极低)的样本中,不确定率更高。

PICO 2:对于细菌学确诊的利福平耐药肺结核患者,是否应该使用 tNGS 来诊断耐药性?

不同药物的可用证据各不相同,关于异烟肼敏感度的研究有 12 项(1440 例受试者),而贝达喹啉敏感度的研究仅有 3 项(31 例受试者)。随后,对汇总的数据采用多因素混合效应模型进行分析。结果表明,其中一些药物的总体证据等级较高。左氧氟沙星由于不一致性被降低一级。由于耐药样本数量低于阈值设置,并且置信区间较宽,因此,贝达喹啉和利奈唑胺因敏感度不精确,被降低两级。氯法齐

明也因敏感度不一致(因为两项研究是离群值)和不精确性(因为置信区间较宽)被降低两级。因为对大部分样本使用了超出 WHO 推荐浓度的临界浓度,阿米卡星的敏感度和特异度降低一个级别。由于不一致性和不精确性,阿米卡星的敏感度进一步被降低两个级别。乙胺丁醇由于偏倚风险降低了一个级别,因为用于指标检测和参考检测的样本不同。由于不一致性和不精确性,链霉素的特异度被降级两个级别。因此,不同药物的检测准确性证据的总体证据等级范围为从高到极低。

在 RR-TB 患者中,针对异烟肼、左氧氟沙星、莫西沙星、乙胺丁醇和链霉素的检测结果准确,总敏感度 $\geq 95\%$ ;针对吡嗪酰胺(90%)、贝达喹啉(68%)、利奈唑胺(69%)、氯法齐明(70%)和阿米卡星(87%)的检测敏感度均在可接受范围内。另外,除链霉素(75%)外,所有药物的合并特异度均 $\geq 95\%$ 。除乙胺丁醇和吡嗪酰胺采用表型 DST 和 WGS 相结合的参考标准,其余药物的参考标准均为基于培养的表型 DST。检测结果不确定的百分比从 9%(左氧氟沙星和莫西沙星)~21%(乙胺丁醇);在细菌载量较低(半定量类别低或极低)的样本中,不确定率更高。

2. 成本效益分析:关于使用 tNGS 或 WGS 诊断 DR-TB 的成本和成本效益的系统综述检索了 3 个数据库:PubMed、Embase 和 Scopus。然而,来自系统综述的成本数据有限。因此,与制造商和创新诊断基金会(Foundation for Innovative New Diagnostics, FIND)协商后,进行了实证单位成本计算。在开展这项工作时,只获得了 Deeplex Myc-TB 试剂盒的定价,并将其用于估算该类的成本。单位成本包括耗材、设备、人员和间接费用(如有);此外,假定对所有药物进行 tNGS 检测。根据实证分析法,估算出 tNGS 的成本为:南非 134~257 美元,格鲁吉亚 120~198 美元,印度 121~175 美元。这些费用还取决于患者数量、批次和每个 tNGS 试剂盒的谈判成本。

3. 用户视角:鉴于没有找到直接证据,指南指导小组注意到 2022 年发表的 Cochrane 定性证据综述,该综述研究了相关间接性证据。结核病患者重视诊断准确性,避免诊断延误和保持与诊断相关的低成本,而医疗服务提供者重视准确性和由此产生的对低复杂程度 NAAT 结果的信心,快速周转时间和寻求诊断者的低成本。为了弥补直接证据方面的差距,WHO 委托进行了一项额外的定性横断面研

究,包括半结构式访谈,访谈对象主要是在 3 个 FIND 试验点的实验室工作人员和直接参与实施 tNGS 的管理人员,以及参与结核病治疗和诊断的 3 名全球专家。共有 17 名受访者,于 2022 年 9—10 月接受了访谈。目的是探讨 tNGS 技术实施者对 tNGS 可接受性、可行性,以及价值、偏好和公平性的看法和经验。

### (三)未来的研究方向

1. 临床研究方面:(1)开展临床试验,评估 tNGS 对患者重要结局的影响。(2)在利福平或其他药物耐药性流行范围内,评估已确诊肺结核患者群体中 tNGS 的准确性及其对患者重要结局的影响。(3)在不同的地理和流行病学环境中,评价 tNGS 对新药和重新利用的药物耐药性的准确性和对患者重要结局的影响。(4)评估 tNGS 分析肺外样本,包括脑膜炎的 CSF、儿童的非痰样本(如鼻咽抽吸物、胃抽吸物或粪便),以及成人和儿童的其他样本类型(如舌拭子)的准确性和对患者重要结局的影响。(5)开展更多定性和定量研究,进一步了解最终用户和临床医生对使用 tNGS 的可接受性和可行性的看法。

2. 实施研究方面:(1)将 tNGS 整合到实验室网络和优化算法中,开发和评估有效和高效的实施模型,以提高检测和治疗的及时性,改善患者疗效。(2)制定策略以提高 tNGS 检测的效率,包括样品处理和浓缩技术,在进行 tNGS 检测之前确定初始细菌负荷的最佳阈值,采用分子运输介质进行环境储存和将样品转移到测试点。(3)定期更新 WHO 突变目录,纳入更多的突变靶点,包括新药(例如普瑞马尼),以提高 tNGS 的敏感度和特异度。(4)探索技术进步,简化测试过程,实现自动化(特别是文库制备),开发分散的 tNGS 解决方案,并研究与现有初始测试(例如使用剩余的 DNA 或涂片阳性载玻片)的潜在协同作用。(5)对各国的测序能力进行全面测绘,并进行诊断网络优化工作。该技术的应用应考虑多种疾病的测序需求,促进机器的跨学科使用和成本分担。(6)采集汇总 tNGS 技术在其他疾病(如新型冠状病毒感染)中应用的经验教训,为其在结核病中的应用制定更有效的实施策略。

3. 监测和评估方面:(1)对不同 tNGS 技术报告结果的术语进行标准化,以便纳入医疗信息数据系统。(2)确保单独记录真正的失败和未分类的突变,并监测随时间推移的趋势,这是结果报告的重要组成部分。(3)定期监测数据,包括总体耐药率、特定



药物或靶点的耐药率及周转时间(包括总时间和实验室内时间)。(4)纳入质量监测措施,如跟踪不确定率、测序覆盖率和深度,以及参与外部质量保证计划。(5)建立外部测序质量保证计划,涵盖所有相关靶标。(6)将测序生成的数据纳入现有监测系统,以有效监测耐药性的流行情况和趋势。共享数据,更新 WHO 突变目录。(7)收集成本数据,以解决一些重要问题,如在不同环境中引入和推广 tNGS 的相关成本、周转时间与批次之间的权衡,以及在不同环境中的最佳平衡。(8)评估多疾病检测对计划运作和成本的影响,包括特定疾病的检测量、周转时间、成本计算、资源共享和资源需求。(9)评估 tNGS 对开始或改变治疗的时间、治疗效果和总体成本效益的影响。

## 总 结

针对高危人群开展 MTB 感染和耐药性早期快速诊断,以确保结核病患者得到有质量保证的快速诊断和治疗,对于有效控制结核病具有非常重要的意义,也是实现 2035 年全球“终止结核病流行”战略目标和联合国可持续发展目标的有效手段。《2024 版指南》整合了最新研究证据,提供了很多新的可供选择的分子检测手段,但多数推荐意见为条件性推荐,其证据等级低或极低,需要在今后更新的研究中

进一步评估以更新推荐意见。笔者认为,学习借鉴国际指南的同时,需结合我国现阶段的结核病诊疗现状和流行病学研究数据,针对不同应用场景、不同技术的特点及其实际可行性和可及性等综合分析,制定适合我国的结核病快速诊断的指南,以选择适宜的检测技术或组合。

**利益冲突** 所有作者均声明不存在利益冲突

**作者贡献** 王小敏、陈锦云、曾昱芹和马佳:设计、起草文章;孔星星和蒙建州:审阅文章;卢水华:对文章进行批评性审阅、指导、获取研究经费

## 参 考 文 献

- [1] World Health Organization. Global tuberculosis report 2023. Geneva: World Health Organization, 2023.
- [2] World Health Organization. WHO consolidated guidelines on tuberculosis: Module 3: diagnosis-rapid diagnostics for tuberculosis detection. Geneva: World Health Organization, 2024.
- [3] United Nations General Assembly. Resolution 73/3: Political declaration of the high-level meeting of the General Assembly on the fight against tuberculosis. New York: United Nations General Assembly, 2018.
- [4] United Nations. Resolution 78/L. 4. Political declaration of the high-level meeting of the General Assembly on the fight against tuberculosis. New York: United Nations, 2023.
- [5] World Health Organization. Implementing the End TB Strategy: the essentials. Geneva: World Health Organization, 2015.

(收稿日期:2024-05-31)

(本文编辑:郭萌)