

了解
和利用
结核病
数据

了解 和利用 结核病 数据



World Health
Organization

世界卫生组织书库图书在版编目数据：

Understanding and using tuberculosis data.

- 1.Tuberculosis – epidemiology. 2.Population Surveillance.
- 3.Data Interpretation, Statistical. I.World Health Organization.

ISBN 978 92 4 154878 6

（NLM 分类代码：WF 205）

© World Health Organization 2014

保留所有权利。世界卫生组织的出版物可从世界卫生组织网站（www.who.int）获取或者从世界卫生组织出版物（Press）部（地址：20 Avenue Appia, 1211 Geneva 27, Switzerland；电话：+41 22 791 3264；传真：+41 22 791 4857；电子邮件：bookorders@who.int）购买。

若要复制或翻译世界卫生组织出版物（不论是为了销售抑或为了非商业性质的发布），需要在世界卫生组织网站（www.who.int/about/licensing/copyright_form/en/index.html）上向世界卫生组织出版物部申请许可。

本出版物所采用的内容和形式并不表明世界卫生组织就国家、地区、城市或区域或相应当局的法律地位发表任何意见，也不表明世界卫生组织就它们的边疆或边界的划定发表任何意见。地图中的虚线代表大致的边界，相关方面对这些边界的划定可能尚未达成共识。

文中对具体公司或制造商产品的提及并不表明世界卫生组织认可或推荐它们，而不认可或推荐其它同类的公司或产品。专利产品的名称由英文单词首字母大写来标识（除非有错误或疏漏）。

世界卫生组织已适当地采取各种必要措施来检验本出版物中信息的准确性。然而，本文档的出版不附带任何明确的或隐含的担保。与本文档的解释和使用相关的责任由读者承担。对于因使用本文档而导致的任何损失，世界卫生组织均不承担责任。

卢森堡印刷

WHO/HTM/TB/2014.09

设计：Rania Spatha

目录

致谢	viii
简介	x
缩略语	xii
第 1 章	
聚合结核病通报数据的分析	15
1.1 聚合通报数据：什么是聚合通报数据？	16
1.2 聚合结核病通报数据质量评估和质量保障	18
1.3 聚合数据的分析	21
1.4 趋势分析示例	22
1.5 聚合通报数据的局限性	40
1.6 小结	41
参考文献	43
附录 1 结核病监测数据质量标准，带有示例	44
第 2 章	
病例相关结核病通报数据的分析	49
2.1 病例相关通报数据：什么是病例相关通报数据以及它们为什么很重要	50
2.2 制定分析计划	52
2.3 制作数据集	53
2.4 数据分析：开展和解释描述性分析	62
2.5 数据分析：开展和解释更为复杂的分析	71
2.6 传播分析结果	73
2.7 结论	75
参考文献	76
附录 2 分析计划示例	77
附录 3 用于评估失访因素的分层和多元分析的示例	80

第 3 章	
利用基因分型数据开展疫情暴发调查	83
3.1 基因分型数据：概述	84
3.2 数据准备	86
3.3 分析疫情暴发	88
3.4 分析大型群集	98
3.5 基因分型数据的局限性	103
3.6 在高结核病负担地区开展基因分型的特殊考虑事项	104
3.7 结论：利用基因分型数据改善大众健康	106
参考文献	107
第 4 章	
分析加剧结核病流行情况的因素	109
4.1 生态分析	110
4.2 结核病发病率	110
4.3 利用生态分析了解结核病流行情况	112
4.4 生态分析的概念框架	114
4.5 准备分析所需数据	116
4.6 个案研究	117
4.7 结论	120
参考文献	121
附录 4 结核病生态分析中应当调查哪些类型的数据？	122
附录 5 示例：关于各因素如何影响结核病负担的详细概念框架	130
第 5 章	
耐药结核病：负担和防治措施分析	131
5.1 方法	132
5.2 耐药结核病负担估计和时间趋势分析	139
5.3 监测规划有效性	141
5.4 结论	165
参考文献	166

第 6 章	
艾滋病相关结核病：负担和防治措施分析	169
6.1 艾滋病相关结核病简介	170
6.2 规划数据的分析	170
参考文献	181
第 7 章	
利用人口动态登记和死亡率调查数据估算结核病死亡率	183
7.1 死亡率数据的来源	184
7.2 监测 HIV 阴性个体的结核病死亡率	186
7.3 监测艾滋病病毒感染者的结核病死亡率	190
7.4 死亡数与通报数的比值	192
7.5 耐多药结核病死亡率	192
参考文献	194
第 8 章	
整合监测数据和调查数据来估测结核病负担	195
8.1 结核病发病率	196
8.2 结核病现患率	199
8.3 结核病死亡病例和病死率	201
参考文献	204
结语	205

致谢

本手册由世界卫生组织“全球结核病影响衡量”任务组制定。加强结核病监测是该任务组的三大战略工作领域之一。本文件的制定始于 2012 年，完成于 2014 年，旨在为各国深入评估、了解和系统性地利用结核病监测数据提供参考依据。

本手册的核心编写人员包括：Laura Anderson（英格兰公共卫生署）、Lori Armstrong（美国疾病预防控制中心）、Emily Bloss（美国疾病预防控制中心）、Anna Dean（世界卫生组织总部）、Julia Ershova（美国疾病预防控制中心）、Dennis Falzon（世界卫生组织总部）、Philippe Glaziou（世界卫生组织总部）、Susan van den Hof（荷兰皇家防痨协会结核病基金会）、Irwin Law（世界卫生组织总部）、Ellen Mitchell（荷兰皇家防痨协会结核病基金会）、Ikushi Onozaki（世界卫生组织总部）、Charalambos Sismanidis（世界卫生组织总部）、Deanna Tollefson（美国疾病预防控制中心）、Rachel Yelk Woodruff（美国疾病预防控制中心）和 Matteo Zignol（世界卫生组织总部）。编写组的主管和协调人是 Philippe Glaziou 和 Charalambos Sismanidis。全球结核病项目的督导与评估工作团队协调人 Katherine Floyd 负责提供总体指导。

本手册的制定离不开许多人员的贡献。核心编写组特别要感谢 Annabel Baddeley（世界卫生组织总部）、Charlotte Colvin（美国国际发展署）、Carole Mitnick（美国哈佛大学医学院）、Olivia Oxlade（加拿大麦吉尔大学）、Alberto Piubello（尼日尔达米恩行动）、Hazim Timimi（世界卫生组织总部）和 Armand van Deun（比利时安特卫普热带医学研究所），他们帮助起草或审阅了文件的内容。

作者还要感谢（通过出版文献或者与核心编写组的个人通讯）提供了监测数据的 30 个国家的国家结核病防治项目，这些数据在本手册中被用作示例。这些国家是：孟加拉国、白俄罗斯、博茨瓦纳、巴西、柬埔寨、中国、爱沙尼亚、埃塞俄比亚、格鲁吉亚、加纳、印度、印度尼西亚、伊拉克、日本、肯尼亚、拉脱维亚、莫桑比克、荷兰、尼日尔、尼日利亚、秘鲁、菲律宾、罗马尼亚、南非、泰国、乌干达、英国、美国、乌兹别克斯坦和越南。

本报告的研究结果和结论是作者得出的，不一定体现上述任何机构的官方立场。文中如有任何错误，将由核心编写组承担责任。

美国国际发展署（USAID）的卫生、传染病与营养办公室（HIDN）全球健康局通过其结核病 CARE I 项目（合作协议编号 AID-OAA-A-10-00020）为本指南的制定提供了经费支持。

本文件由 Inis Communication 编辑。

作者声明其在编写本文件的过程中不存在财务和个人的利益冲突。

简介

国家卫生信息系统提供丰富的数据源，可以了解结核病导致的疾病负担和致力于减轻这种负担的项目规划的有效性，这两方面对于公共卫生行动都很重要。不过，可用的数据通常没有被充分利用，或者根本未被利用。这至少在一定程度上可能反映了缺乏有关分析这类数据推荐方法的明确指南。本手册致力于解决这个缺口，针对结核病监测数据的分析提供了详细的和实用的示例，尤其是结核病通报数据、抗结核病治疗耐药监测数据、国家人口动态登记系统内编制的死亡率数据等。本手册从最基本的分析开始，然后逐渐过渡到更具有挑战性的主题，例如利用多个证据源估计疾病负担，包括专门调查数据。

本手册有七大主要目标：

1. 描述和解释如何分析结核病通报数据以便了解结核病流行病学，包括疾病的地域分布、按年龄和性别的分布和特定人群内的分布（[第 1 章](#)、[第 2 章](#)和[第 3 章](#)）。
2. 描述和解释如何分析结核病通报数据以便评估规划绩效（包括病例发现和治疗成功）和数据质量（[第 1 章](#)和[第 2 章](#)）。
3. 描述和解释如何利用基因分型数据调查疫情暴发（[第 3 章](#)）。
4. 描述和解释如何在国家层面上分析推动结核病流行情况的各种因素，以及这些因素如何（积极地或消极地）影响结核病通报和基本结核病发病率的趋势（[第 4 章](#)）。
5. 描述和解释如何分析耐药结核病和艾滋病相关结核病带来的疾病负担，以及利用监测数据的相关项目规划防治措施（[第 5 章](#)和[第 6 章](#)）。
6. 描述如何分析人口动态登记系统的数据以便估计结核病发病率（[第 7 章](#)）。
7. 展示如何将专门调查数据与监测数据结合起来，以便了解结核病疾病负担的趋势（[第 8 章](#)），推导出结核病导致的疾病负担的估计值。

本手册讨论和利用的典型数据源包括：

1. 病例相关或聚合结核病通报和治疗转归数据。
2. 机构审计或记录数据质量评估的结果。
3. 实验室数据。
4. 耐药监测（包括耐药调查）的结果。
5. 人口动态统计数据（包括死因数据）的户籍记录。
6. 死亡率调查的结果。
7. 结核病现患率调查的结果。
8. 病例存量研究的结果，衡量结核病漏报情况，以及特定情况下估计发病率。

本手册展示如何利用这些数据源，介绍现有数据质量分析工具，描述结核病负担和相关趋势估计方法。贯穿手册始终，重点还放在利用各国示例阐述如何开展分析和解释结果。

第1章至第2章的内容应当面向绝大多数（如果不是所有）读者，旨在提供有关如何开展特定分析的指南。**第3章至第8章**的内容应当主要概述更高级的监测数据分析类型，但不会提供有关如何开展这些分析的详细指南。

本手册致力于帮助广泛的目标受众，包括国家结核病防治规划管理人员、督导与评估官员、研究人员（包括流行病学家和统计学家），以及技术、金融和发展领域机构的工作人员。

Philippe Glaziou

高级流行病学家

世界卫生组织

全球结核病项目

结核病督导与评估团队

结核病疾病负担估计负责人

Charalambos (Babis) Sismanidis

统计学家

世界卫生组织

全球结核病项目

结核病督导与评估团队

增强结核病监测负责人

缩略语

ACF	主动病例发现
BMI	身体质量指数
BMU	基本管理单元
CDC	美国疾病预防控制中心
CI	置信区间
DR-TB	耐药结核病
DST	药敏试验
HIV	人类免疫缺陷病毒（艾滋病病毒）
LGA	地方政府区
MDR-TB	耐多药结核病
MIRU-VNTR	分枝杆菌散在分布重复单位 – 可变串联重复序列
NTIP	国家结核病指标项目
NTP	国家结核病防治规划
NTR	国家电子结核病登记簿
OR	优势比
PPM	医防合作
RFLP	限制性片段长度多态性
RVCT	确诊结核病修订报告
SOP	标准操作规程
TB	结核病

第 1 章

聚合结核病通报数据的分析

目标受众：

普通读者，不过重点是在国家结核病防治规划（NTP）内各个管理层面开展工作的督导与评估官员。

预期成果：

本章结束时，读者应当能够了解：

- 聚合结核病通报数据对于描述结核病流行情况和给国家结核病防治规划提供参考依据的重要性；
- 如何按照人物、地点和时间来分析、报告和解释聚合结核病数据；
- 如何评估聚合结核病通报数据的质量；
- 聚合结核病通报数据的优势和局限性。

作者：

Emily Bloss、Julia Ershova、Susan van den Hof、Irwin Law、Ellen Mitchell、Deanna Tollefson

1.1 聚合通报数据：什么是聚合通报数据？

自二十世纪九十年代中期以来，全球范围内就一直在采用一种标准化的纸质记录和报告系统，用于记录和报告被诊断出结核病的患者的人数以及这些患者的治疗转归。在该系统内，结核病数据以聚合形式进行报告（即病例总数相当于一个基本记录单元）

（1,2）。通常而言，卫生保健人员在个体患者结核病治疗卡上记录有关患者治疗史的信息。根据国际认可的统一记录和报告框架，收集结核病病例（个人结核病发作）的人口统计学、临床学和细菌学信息。^a然后将这些数据转录到结核病登记簿中，结核病登记簿罗列着在特定卫生保健机构和/或基本管理单元（BMU）内接受治疗的所有病例的信息。^b接下来，特定地理区域内所有登记簿中的全部病例的通报和治疗转归数据经过汇编后聚合到相应的报表中（2）。这些报告被提交给更高的管理层面（最终抵达国家层面），通常是每个季度提交一次；报告可以采取纸面形式，也可以采用电子形式（[文字框 1](#)）。在国家层面上，国家结核病防治规划会汇报这些聚合数据，作为年度报告的分析基础。聚合数据不提供个体病例的详细信息。

如果通报数据通过电子形式收集，务必要适当地选择一种数据管理软件，创建的数据库应当能够将地方层面上诸多季度文件合并到国家层面上一个数据集中。大量的 Excel 电子数据表虽然在基本管理单元层面上可以轻松用于数据收集，但在国家层面上却不能有效地用于监测数据的监控和分析。应当采用统一的平台收集通报数据，这样有助于数据分析。[文字框 1](#) 阐述了将 Epi Info™ 软件用于国家聚合监测数据库的示例。

结核病监测数据的分析对于规划的评估很重要，可以帮助指导有关病例管理和政策的决策。这些数据可以推动国家结核病防治规划监测全国范围内结核病病例的数量和分布趋势。这有助于国家结核病防治规划汇报国内结核病流行情况，以及在实现规划总目标和具体目标方面取得的进展。这还有助于国家结核病防治规划制定有针对性的国家策略和筹资计划。

-
- a. 世界卫生组织病例相关结核病监测数据标准和基准核查清单 (<http://www.who.int/tb/publications/standardsandbenchmarks/en/>, 2014 年 7 月 19 日检索) 也建议收集和分析一组最少变量（年龄或年龄段、性别、登记年份、细菌学检测结果、既往治疗史和疾病类型）。
 - b. 基本管理单元（BMU）是按照管理、监督和监控职责来界定。国家结核病防治规划的一个基本管理单元可能包含多个治疗机构、一个或多个实验室以及一个或多个医院。界定关键是有有一名管理人员或协调人员负责监督基本管理单元的结核病控制活动，以及维护接受治疗的所有结核病患者的主登记簿。该登记簿用于监控规划的实施情况和向较高管理层面汇报指标的完成情况。通常而言，基本管理单元对应着政府的二级地方行政区域，例如可能被称为区或县。国际社会建议一个基本管理单元覆盖 5 万到 15 万人口，或者在大型城市可以覆盖多达 30 万人口。（资料来源：*Compendium of indicators for monitoring and evaluating national tuberculosis programmes*（国家结核病防治规划督导与评估指标大纲）） (http://www.who.int/tb/publications/tb_compendium_of_indicators/en/, 2013 年 12 月 10 日检索)

在本章中，描述了分析聚合结核病监测数据的不同方法。这些分析方法主要侧重于评估地方区域不同人群（例如按年龄和性别划分）内病例通报率（每 10 万人口中的病例数量）的分布和趋势。为了开展这些分析，建议采用最少的一组变量，包括年龄（或年龄段）、性别、登记年份、细菌学检测结果、既往治疗史、疾病类型和地理区域（4）。一些监测系统只使用聚合结核病数据，讨论这些监测系统专门采用的数据质量指标和数据验证方法时，还应讨论聚合数据的局限性。

文字框 1： 电子聚合结核病监测数据库（尼日利亚的示例）

在尼日利亚，已经开发和安装了国家聚合结核病监测系统，用于在地方政府区和国家层面收集结核病监测数据。数据库包括多个部分，分别与世界卫生组织推荐的报表指南相对应（1）。数据库界面在视觉上遵循数据录入纸质表格的设计和顺序。数据库中包含所有推荐的变量 – 包括年龄段、性别、时间段、细菌学结果、既往治疗史、疾病类型和地理区域。尼日利亚地方政府区、州和国家层面上可以有效地、及时地监测和分析结核病病例的数量和分布趋势。

完整的数据库包括三张电子页面。下面举例说明地方政府区层面上数据库的第一个页面。

Make/Edit View: C:\NATSS\NATSS.MDB:Surveillance
File Edit View Insert Format Tools Help

国家聚合结核病监测系统 (NATSS)

NATSS 编号

地区

区域 州 地方 覆盖人群

填报

填报日期 季度 年份

机构类型

结核病指标

方框 1. 登记的所有结核病病例

新涂阳结核病病例		其他病例		肺部涂阴病例		肺外病例	
既往治疗	复发 ☐ 失败 ☐ 失访后复归 ☐	涂阳 ☐ 涂阴 ☐		< 15 岁 ☐ ≥ 15 岁 ☐		< 15 岁 ☐ ≥ 15 岁 ☐	

病例总数 CV/TS 管理的结核病病例 ☐

方框 2a. 按性别和年龄段划分的新涂阳肺结核病例

性别 x 年龄段									
男性 0-4 ☐	男性 5-14 ☐	男性 15-24 ☐	男性 25-34 ☐	男性 35-44 ☐	男性 45-54 ☐	男性 55-64 ☐	男性 >65 ☐	男性总数	
女性 0-4 ☐	女性 5-14 ☐	女性 15-24 ☐	女性 25-34 ☐	女性 35-44 ☐	女性 45-54 ☐	女性 55-64 ☐	女性 >65 ☐	女性总数	

1.2 聚合结核病通报数据质量评估和质量保障

决定数据质量的因素各不相同（例如准确性、精确性、真实性、一致性和有效性），不过数据的质量还受到捕获和报告这些数据的系统的限制。在结核病监测的情况下，如果监测系统收集的信息不完整、不一致或不正确，则数据无法准确地描述国内当前的结核病负担。因此，评估数据质量实际上是评估用于收集数据的各种方法的质量。所以，评估通报数据的质量会凸显结核病监测系统的优势和劣势，揭示这些数据能在多大程度上衡量国家结核病负担及其发展趋势。

表 1
结核病监测和人口动态登记系统标准和基准核查清单中采用的各项标准

数据质量
1. 病例定义吻合世界卫生组织指南。
2. 结核病监测系统旨在捕获所有结核病报告病例的一组最少变量。
3. 国家层面已经收到并处理了所有预定的定期数据提交内容。
4. 季度报告（或同等文件）中的数据具有准确性、完整性和内部一致性等特点（仅针对纸质系统）。
5. 国家数据库中的数据具有准确性、完整性、内部一致性和没有重复性等特点（只针对电子病例相关或患者相关系统）
6. 结核病监测数据具有外部一致性特点。
7. 结核病报告病例数在不同时间段都保持内部一致性。
覆盖面
8. 所有结核病诊断病例都被报告给上级层面。
9. 所有人都能很好地获得卫生保健服务。
人口动态登记
10. 人口动态登记系统具有较高的国内覆盖面和质量。
耐药结核病、结核/HIV 合并感染和儿童结核病
11. 监测数据直接衡量新病例中的耐药结核病。
12. 监测数据直接衡量结核病病例中的艾滋病病毒感染现患率。
13. 儿童结核病报告病例监测数据具有可靠性和准确性等特点，所有被诊断出的儿童结核病病例都被报告给上级层面。

世界卫生组织结核病监测和人口动态登记系统标准和基准核查清单提供一种系统化的方法，用于评估国家结核病监测系统的质量，包括分析通报数据质量的方法等（表 1）。有关如何开展结核病监测系统评估的详细信息可以在随附的用户指南中找到。（附录 1 提供了有关应用核查清单的结核病监测数据质量标准的示例）。

数据录入时的数据验证

不同的方法可以用来在数据收集，聚合、录入和核实程序中验证数据。数据录入时的一种数据验证方法是在将数据录入到聚合表格之前人工检验结核病登记簿。这包括检查同一名患者在不同机构的重复登记、记录内部的数据一致性（例如涂阴患者的治疗转归为治愈）、每条记录所有必填字段的完整性等。此外，数据录入表可以包括电子检查（参见文字框 2）。

文字框 2

电子数据录入时的数据验证（尼日利亚的示例）

为了加快数据收集时的电子数据录入和验证，可以创建大量数据检查（或检查代码）并添加到数据录入表中。尼日利亚电子聚合结核病监测数据库（参见文字框 1）就是拥有这些自动数据检查的一个系统。

例如，在本系统中，只要按性别录入不同年龄段的病例数目，就能自动计算男性和女性的涂阳病例总数（文字框 1、方框 2a 和方框 1）。如果计算出的总数不同于录入的涂阳病例总数（将方框 1 与方框 2a 进行比较），就会生成下面的错误信息：

“新涂阳结核病病例的数量（方框 1）应当 = 男性总数 + 女性总数（方框 2a） - 请检查数据！”

尼日利亚的系统在数据录入时验证数据的另一个方法是确保特定字段都已经录入了数据。也就是说，数据录入人员只有在必填字段中录入数据之后才能移动到下一个字段。在尼日利亚数据库中，这些必填字段是聚合数据的州、地方政府区、年份和季度。同样，系统采用预先定义的下拉列表用于区域、州、地方政府区和机构类型的选择，避免打字错误并确保录入有效值。

c. 核查清单和用户指南可以在世界卫生组织“全球结核病影响衡量”任务组的网站上找到：
<http://www.who.int/tb/publications/standardsandbenchmarks/en/>

数据录入后的数据验证

数据录入过程中可能会出现一些人为因素导致的错误，而录入时的数据验证可以最大限度地减少错误和帮助改正错误，确保聚合数据的正确性和真实性。区域或国家层面上可以开展额外验证。这应当包括检查缺失的数据（例如是否所有基本单元都在当年的每个季度提交了聚合数据？）。只有开展了这些初步数据验证步骤后，才能开展更为详细的数据质量分析。聚合数据验证的示例（[文字框 3](#)）包括确保结核病病例总数与下面几个总和之间的一致性：i）新病例和复治病例的总和；ii）男性和女性结核病病例的总和；iii）按患者年龄分组的结核病病例的总和。在具有电子聚合数据库的系统中，数据录入时和数据录入后的计算机辅助检查可以用来发现和纠正数据质量问题。

文字框 3 数据清理和分析准备

聚合数据被录入到数据库后，可以采用多种额外的数据质量检查手段，从而确保数据是完整的且可用于分析（例如确定某个机构的重复录入项和缺失或不一致的数值）。例如：

- **检查重复记录。**如果发现同一个机构或地区存在聚合数据的重复记录，应当仔细地调查涉及到的病例，方法是检查每项重复记录的数据，包括纸质表格。找出存在重复记录的原因，修正导致重复记录的错误。不过在有些系统中，可以在数据录入期间发现是否存在重复记录。
- **检查缺失值。**在依赖聚合数据的系统，缺失信息通常是由数据录入时的人为错误或纸质表格中的缺失数据所导致。务必要记住，零病例（零报告）的聚合数目有别于缺失值，应当相应地对此作出说明。如果可能的话，原始纸质数据源应当被用来检查和填补缺失值。
- **检查电子数据的一致性。**使用现有电子数据分析工具检查各年龄段的结核病病例总数是否与结核病病例总数相等。使用原始数据源修正发现的任何不一致。请注意，在一些系统中，绝大多数不一致的地方可以在数据录入期间被发现（[文字框 2](#)）。

1.3 聚合数据的分析

趋势分析的理论依据

监测数据的完整分析涉及调查通报率在不同时间段的变化，以及尝试了解出现这些变化的根本原因。结核病负担的趋势解释是一项重要的内容，可以评估一个国家在结核病控制和预防领域的进展情况。

当趋势被认为不一致时（即通报率出现快速变化或非预期变化），国家结核病防治规划管理人员应当查看可能的原因。^d应当调查病例报告中的快速变化或不一致趋势的原因。例如，非预期不一致是否能够代表结核病流行情况的真实变化，或者非预期不一致是否由其他结核病决定因素（例如城镇化、社会经济情况、健康保险方案的实施或特定结核病控制活动）所导致？在其他考量中，趋势的这些非预期不一致是否主要由病例定义的变化或记录和报告系统的变化所导致（例如通报系统的结构、覆盖面或绩效变化）？在寻找导致不一致趋势的原因时，务必要首先检查数据中的错误并在所有层面上做出更正。这有助于确保趋势不是由不准确的数据所导致。这就是为什么验证所收集的数据的质量也显得很重要。不过，评估趋势也是揭示数据质量方面可能存在的问题的一种方式。下面一节内容会探讨与患者（或人）、地点和时间相关的趋势分析。

国家结核病防治规划汇报特定年份的结核病病例数量时，通常是报告全国范围内一个历年（从一月份到十二月份）的数据。在有些情况下，地方层面（例如省或州）会报告相同的数据。国家层面每年将按患者年龄和性别、治疗史、疾病部位和细菌学结果划分的数据报告给世界卫生组织全球结核病项目，这样有助于了解全球结核病流行情况的当前概况。^e

不过，从一个国家的角度来看，还务必要追踪国家和地方层面上的结核病病例数量和分布趋势，这涉及到比报告给世界卫生组织的国家数据更为详细的分析。这可以展示结核病流行情况是否得到改善、继续恶化还是没有任何变化。因此，国家结核病防治规划应当定期评估结核病趋势是否存在任何不一致，以及是否需要调整项目规划策略来解决这些趋势。同样，数据必须达到最高质量，从而能够对当前流行情况做出最可靠的解释。

d. 不一致趋势的示例可能包括病例通报数出现超过 4% 的变化，或者在性别和年龄子集中出现突然的和非预期的变化。

e. 每年都会制作和发布年度报告，年度报告提供有关全球、区域和国家层面上结核病流行情况和结核病防治进展的最新信息和分析结果。年度报告主要基于世界卫生组织成员国在各年度全球结核病数据收集报告中报告的数据。在 2013 年，共有 197 个国家和地区报告了这类数据，这些国家和地区的结核病病例数占全球总数的 99% 以上（<http://www.who.int/tb/country/en/index.html>，2014 年 7 月 14 检索）。

不同时间段报告的结核病病例数量^f（例如国家和地方层面上聚合数据的时间序列通报数）可以提供非常重要的信息。就像病例总数一样，这类数据还可以表明谁患有结核病（例如男性或女性）、患有何种类型的结核病（例如肺结核或肺外结核）、在人生的哪个阶段患有结核病（例如年轻或年老）、生活在什么地方（例如病例的地域分布）。按人、地点和时间描述聚合数据是非常关键的活动，有助于描述一个国家的流行病学情况，以及这些分布方面是否存在相当大的趋势。

1.4 趋势分析示例

按时间划分的通报数

解释病例通报数的趋势要求寻找具体变化的原因。为了展示如何分析和解释不同时间段的通报率，下文探讨了英格兰和威尔士、孟加拉国和印度尼西亚的示例。总体而言，这些示例强调为了正确解释一个国家的流行病学情况，务必要深入了解国家的具体情况、国家结核病防治规划及其项目规划方法、结核病监测系统等。

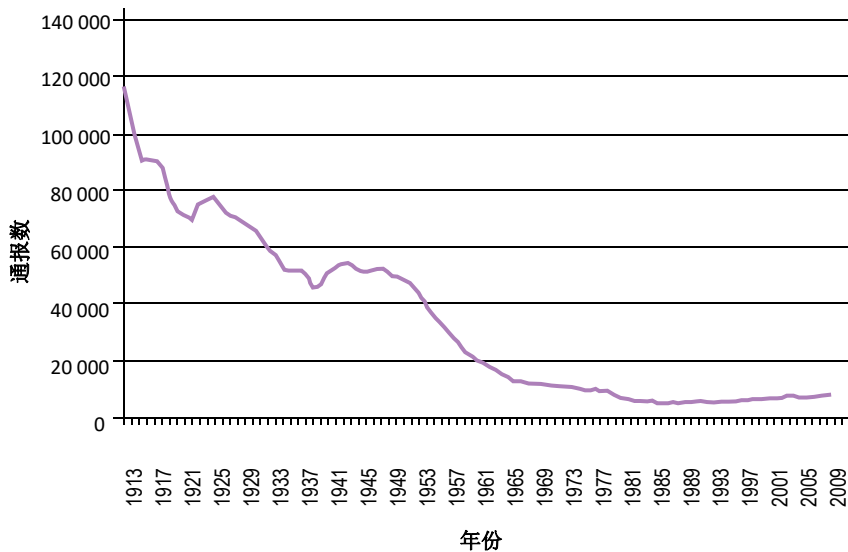
英格兰和威尔士

英格兰和威尔士已经书面记录了将近 100 年的通报率，收集了高质量的、可靠的数据（图 1）。绘制了每年报告的结核病病例数量的分布图，有助于国家结核病防治规划观察结核病病例数量在过去一个世纪的逐渐下降。1945 年后出现了每年减少约 10% 的最快下降，主要归功于生活条件的改善和快速的经济增长，以及有效的化学疗法、卡介苗（BCG）疫苗、巴氏杀菌和卫生保健普及服务的引入等。^g

f. 在本章中，各示例利用了所有结核病通报病例的国家数据，除非另有说明。一些监测系统可能只有非常有限的数据可供利用，因此可能应当只采用新病例和复发通报数，甚至只采用涂阳结核病通报数。

g. 1913–1982 年传染病法定通报数（NOIDS）：2010–2011 增强型结核病监测（ETS）。编制单位：英国科林达健康保护局结核病防治部。

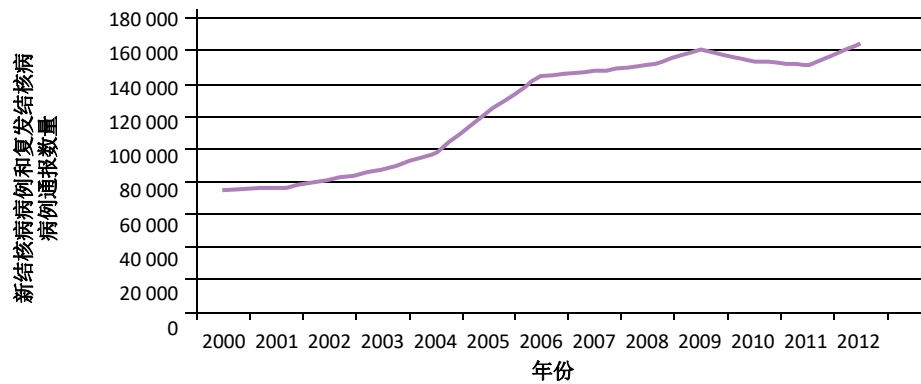
图 1
结核病病例（所有类型）通报，英格兰和威尔士（1913–2011 年）
 资料来源：英格兰公共卫生署（结核病防治部）



孟加拉国

孟加拉国从 2000 年到 2012 年的新结核病病例和复发结核病病例总通报数显示病例报告数呈上升趋势（图 2）。具体而言，2004 年到 2006 年出现了病例数的大幅上升，随后几年基本上没什么变化。国家结核病防治规划报告 2004 年到 2005 年的年度总通报数出现了约 25% 的增长。这对应着 24 000 名患者的绝对增长数。图 2 显示 2005 年到 2006 年出现了 18% 的增长，对应着 22 000 名患者。再往前看，五年期（2000–2004 年）内观察到了相同程度的变化。病例数量在不同时间段的这种上升是否能真实地反应孟加拉国的结核病流行情况，或者是否有其他原因导致了这种趋势？与国家结核病防治规划展开讨论后发现，在该十年的中间阶段，国家结核病防治规划大力改善了病例发现和落实了相关措施，促进了其他部门（例如私营部门）的卫生保健提供方参与到结核病防治规划病例发现和治疗中来。因此，结核病病例通报数的大幅增长可以归功于项目规划调整，而不是反映该国结核病发病率的真正增长。

图 2
孟加拉国国家结核病防治规划向世界卫生组织报告的新结核病病例和复发结核病病例通报总数（2000–2012 年）
资料来源：世界卫生组织 2013 年全球结核病报告



印度尼西亚

在全国范围内有效落实 DOTS 也可以促进病例通报数的增长。这一点在印度尼西亚已得到证实，该国的数据显示二十世纪九十年代后期在全国范围内落实 DOTS 后出现了通报数的典型上升，当时正好开始实施 DOTS 项目（图 3）。此外，继 2006 年到 2008 年病例发现数量出现停滞不前的困境后，2009 年到 2011 年观察到了上升趋势。研究人员考查了这种现象的潜在原因，发现病例发现数量的上升主要归功于下列原因：i）接受结核病症状调查的患者人数处于上升趋势；ii）医院-DOTS 衔接的增强使得医院和肺部诊所的通报数不断上升；iii）用于发现涂阳结核病病例的镜检服务的覆盖面增加，以及卫生所和医院提供的一线治疗的覆盖面增加。省级层面的数据进一步支持了通报数的上升，而且过去十年间接受评估的有症状个人的人数也有所增加（图 4）。

上述每个示例都强调为了通过常规数据正确解释一个国家的流行病学情况，务必要深入了解国家结核病防治规划及其结核病监测系统。

图 3

印度尼西亚国家结核病防治规划向世界卫生组织报告的新结核病病例和复发结核病病例通报总数（1990–2012 年）

资料来源：世界卫生组织 2013 年全球结核病报告

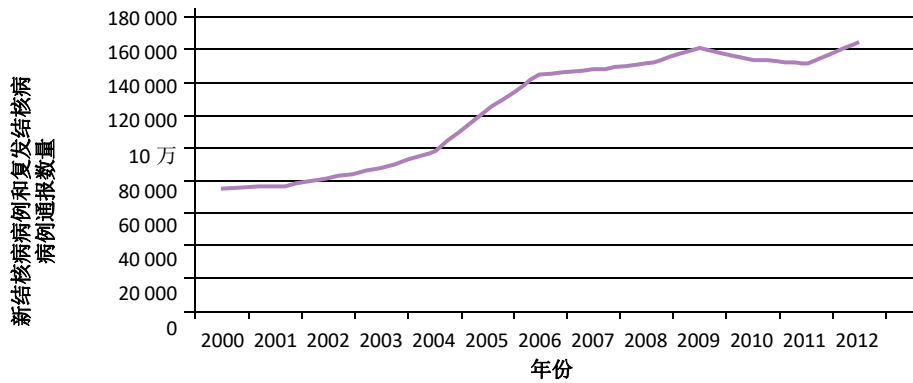


图 4

1990 年到 2011 年间疑似患者评估率的省级和国家（图 100_印度尼西亚）层面时间序列

资料来源：印度尼西亚国家结核病防治规划



按年龄划分的通报数

众所周知结核病发病率会随着年龄而异。在非洲，结核病主要影响到青少年和青壮年群体。不过，在结核病已经从较高发病率下降到较低发病率的国家（例如美国），结核病主要是在老年人中发病，或者是在免疫功能低下者中发病。

因此，在描述结核病流行病学情况时，务必要观察结核病病例在各年龄段的分布。这些信息有助于让国家结核病防治规划了解哪个年龄段正经历着最高的结核病负担，以及应当面向谁开展干预。下文介绍来自不同国家的多份个案研究。

南非

在南非，根据 2012 年所有结核病通报病例和估计人口计算了每个年龄段的病例通报率（图 5）。在本示例中，我们将各年龄段的结核病通报病例绝对数量与各年龄段的通报率（每 10 万人口中的病例数量）进行比较。两个指标都是公共卫生官员用来了解疾病负担的重要衡量手段。在南非，病例绝对数量的年龄分布与通报率的年龄分布略有不同。例如，最高的病例绝对数量出现在 25–34 岁年龄段，而最高的通报率出现在 35–44 岁年龄段。而且，65 岁以上年龄段的通报率非常高，但该年龄段的病例绝对数量最低。考虑到各国有着不同的人口年龄结构，这样的调整在必要时也有助于开展国家之间的比较。

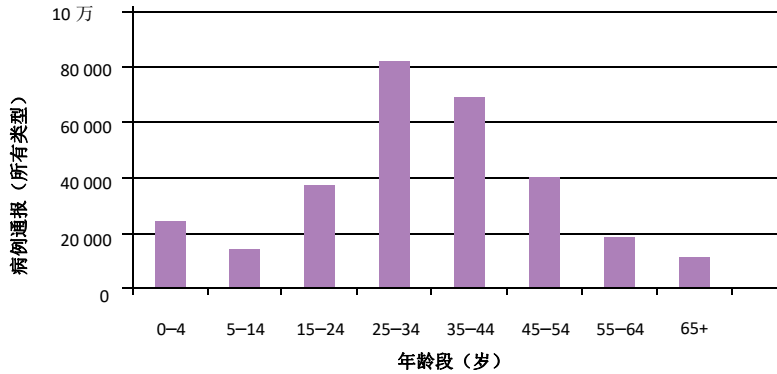
此外，如果我们按年龄段细分国家病例通报率（南非 2012 年为每 10 万人口中 567 名患者），可以明显看出各年龄段之间在通报率方面存在着较大的差异。通报率在婴儿阶段较高，然后在 5–14 岁年龄段有所下降；然后在 35–44 岁年龄段达到最高点，并在后面各年龄段再次下降。这种模式在结核病高度流行的国家非常典型，这些国家的绝大多数病例出现在一般劳动人口中（在南非，一部分原因是该年龄段有着最高的艾滋病病毒现患率）。这种模式也反映了国家结核病防治规划诊断和通报儿童结核病病例的能力。

图 5

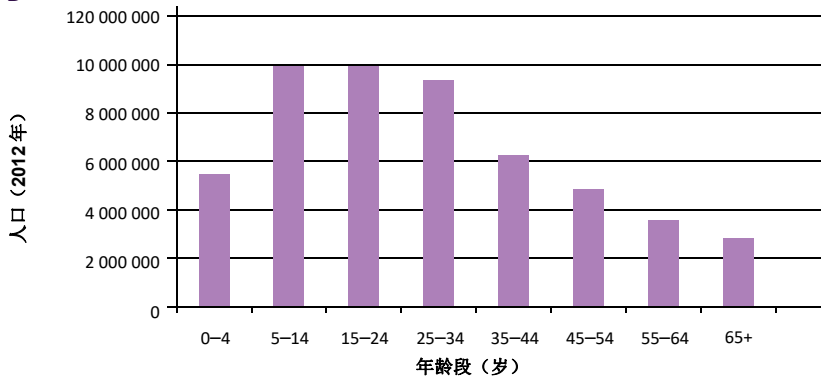
南非各年龄段病例通报率的估测（2012 年）：A）各年龄段的病例绝对数量；B）人口分布；C）病例通报率（每 10 万人口中的病例数量）

资料来源：世界卫生组织 2013 年全球结核病报告和《世界人口展望》2012 年修订版（联合国人口司）

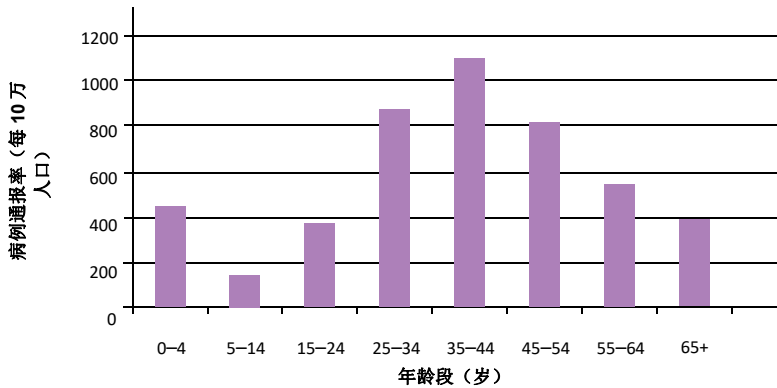
A



B

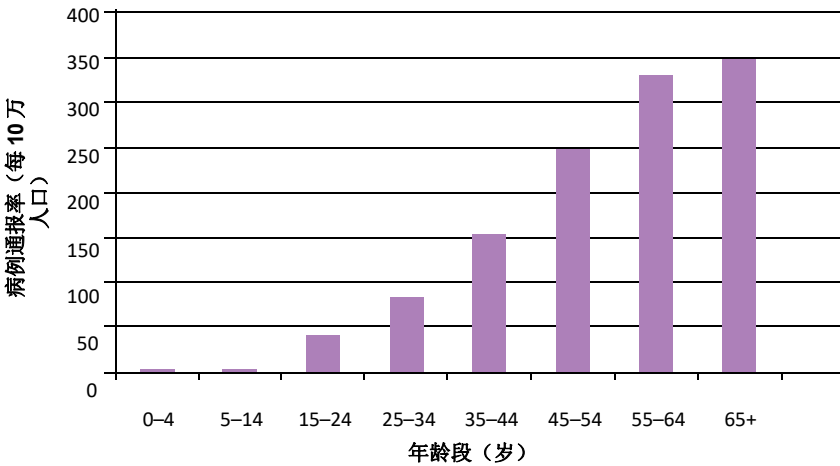


C



通过比较，我们可以看到柬埔寨的通报率随着年龄增长，最大年龄段有着最高通报率（图 6）。随着结核病预防、诊断技术和治疗不断得到完善，结核病最高通报率一般会从较小年龄段向较长年龄段转移，这是因为感染率逐渐下降，促使较小年龄段中通报率也下降。不过，0–4 岁和 5–14 岁年龄段的零病例通报也可能反映出国家结核病防治规划在诊断或报告儿童结核病病例方面存在问题。

图 6
柬埔寨国家结核病防治规划各年龄段的病例通报率（所有类型）（每 10 万人口），2012 年
资料来源：世界卫生组织 2013 年全球结核病报告



日本

病例相关监测系统的聚合数据^h还可以提供有关结核病负担的重要信息。日本就是一个很好的例子，随着结核病预防、关怀和治疗服务的扩展和改善，结核病通报率的最高点从较小年龄段转移到了较长年龄段（图 7）。从 1962 年到 2011 年，日本的结核病通报率在所有年龄段都有显著下降，而且下降率的变化随着年龄增长逐渐变小。这意味着更大比例的结核病病例随着时间的推移从较小年龄段转移到了较长年龄段。预期人口老龄化已经使得儿童结核病病例少于成年人结核病病例，这些数据明确强调了从出生到 14 岁的儿童群体中通报率的快速下降。该观察结果代表着社群中结核病传播速度的减缓，因此提供证据说明结核病防治规划的改善也能减少社群中结核病的传播。

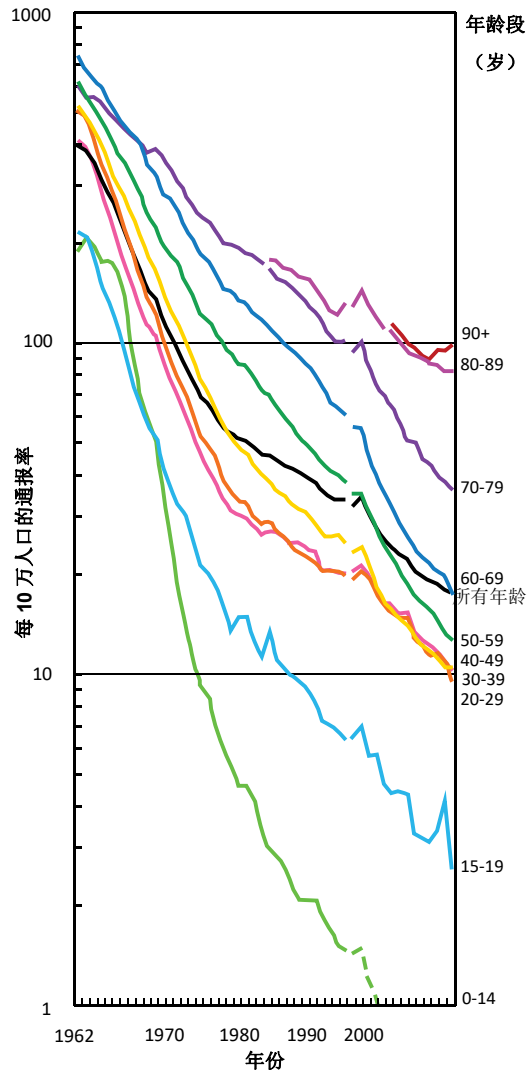
h. 病例相关监测数据包含相应的记录，其中个体结核病病例和相关治疗信息属于分析单元（参见第 2 章）。

聚合数据还捕获日本结核病监测系统的特定变化。自 1987 年以来，国家结核病防治规划开始采用“80 岁以上”类别，并在 2003 年引入了“90 岁以上”类别。

图 7

日本的结核病病例通报率（1962–2011 年）。Y 轴的对数标尺用来展示结核病通报率的快速下降。

资料来源：日本国家结核病防治规划 2012 年年度报告。

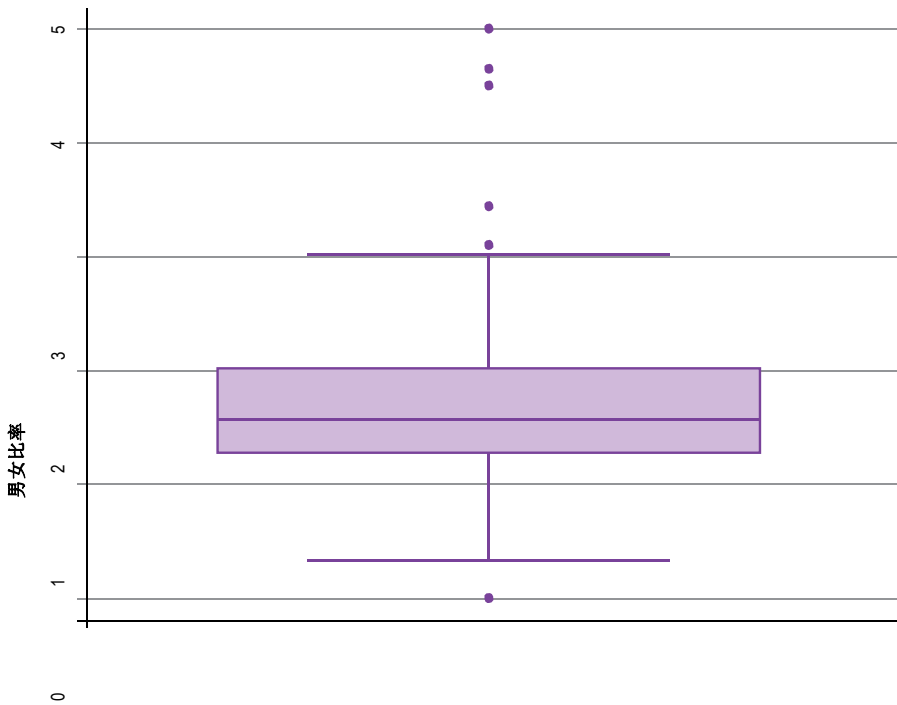


此外，1998 年对结核病病例定义进行了修订，排除了被诊断出感染非结核病分枝杆菌的个人，因此只有被诊断出**结核分枝杆菌**感染的个人被归类为病例。1999 年观察到了结核病通报数趋势的另一个显著变化，主要是国家结核病防治规划宣布国内结核病疫情属于紧急事件，因此在一般人群和卫生服务提供方中加强了宣传教育，从而诊断和通报了更多病例。

按性别划分的通报数

根据当前全球结核病通报数（2012），所有新病例（所有结核病类型）的全球男女比率是 1.6，各国之间存在着很大的差异（图 8）。虽然绝大多数国家的新病例性别比率都大于 1（意味着被通报的男性多于女性），还是有少数国家（例如阿富汗、巴基斯坦、巴布亚新几内亚）的新病例总体男女比率低于或等于 1（意味着被通报的男性少于女性）。

图 8
具有分性别聚合数据的 193 个国家的新结核病通报病例（所有类型）男女比率，2012 年
资料来源：世界卫生组织 2013 年全球结核病报告



如果在国家层面上观察不同时间段的男女比率，有助于评估国家结核病防治规划监测系统的内部一致性ⁱ。“内部一致性”世界卫生组织结核病监测和人口动态登记系统标准和基准核查清单中的数据质量标准之一，可以用来评估国家结核病防治规划常规监测系统的数据（附录 1）。为了评估监测系统的内部一致性，应当在不同的时间段考查一组最少变量，查看是否有需要解释的任何实质差异。结核病病例男女比率是推荐接受考查的指标之一。例如，乌干达的国家结核病防治规划分析了结核病病例从 2008 年到 2012 年的男女比率变化（表 2）。每年被诊断出结核病且接受治疗的男性的比例都有一定增长。由于不同年份之间的变化比较小，该项分析指出数据具有良好的内部一致性。不过，在五年的时间跨度中，男女比率的变化可以被认为比较显著：从 1.47 上升到了 1.64，即增长了 11.6%。因此，国家结核病防治规划应当进一步探讨该比率为什么随着时间的推移不断增长，这样会带来相应的益处。

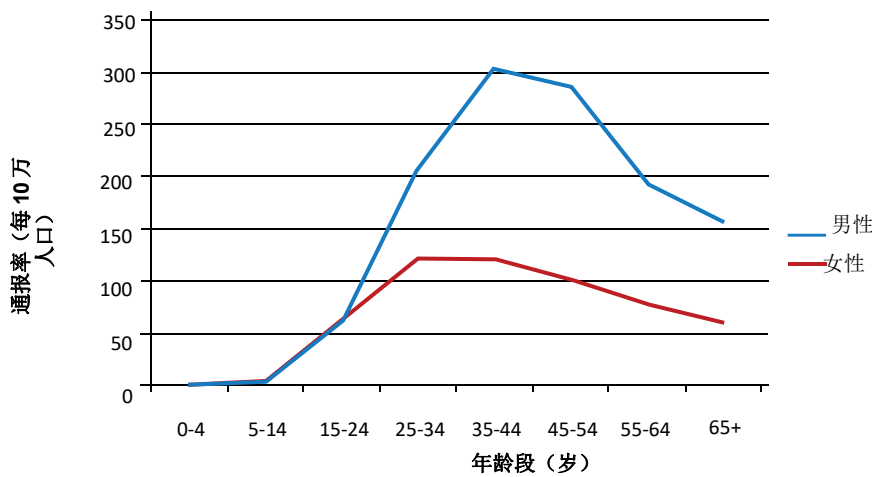
表 2
乌干达的结核病病例男女比率，2008–2012 年

年份	男女比率（所有新病例和复发病例）
2008	1.47
2009	1.52
2010	1.58
2011	1.63
2012	1.64

当我们进一步按年龄段和性别对乌干达的通报数据（2012 年）进行分层时（图 9），我们可以看到虽然男童和女童（0–14 岁）的通报率相同，但在儿童和青少年后面各年龄段中可以观察到更为明显的性别差异。男女病例的最大差异出现在 35–54 岁年龄段。关于这种情况的可能假设包括：i）女性随着年龄的增长更不可能寻求卫生保健服务，或即使寻求卫生保健服务也被漏诊；ii）男性更有可能寻求卫生保健服务，因此在结核病病例中占有更大比例；或 iii）从流行病学的角度来看，乌干达多年来确实出现了男性结核病病例的增长（或女性结核病病例的下降^j）。

i. “内部一致性”指在不同时间段可以复制和比较的监测系统指标（附录 1）。
j. 在乌干达，女性（15–49 岁）中的艾滋病病毒现患率要高于男性（联合国艾滋病规划署 2014 年），不过男性中的结核病现患率要比女性高。可以假设乌干达感染艾滋病病毒的女性更有可能来自最富裕的五分之一人口，因此更有可能接受抗逆转录病毒治疗，这样即便具有较高的艾滋病病毒现患率，也不太可能患有结核病。参见第 4 章和第 6 章，了解有关加剧结核病流行情况的各种因素的更详细信息。

图 9
乌干达按年龄段和性别划分的所有类型结核病病例通报率（每 10 万人口），2012 年
资料来源：乌干达国家结核病防治规划



按地点划分的通报数

分析监测数据的另一种有效方法是考查一个国家内不同地域之间通报病例数的差异。正如我们迄今为止已经看到的一样，聚合国家数据可以用来描述总体流行病学情况，但结核病负担通常并不是在人群内均匀分布。因此，地方数据还可以有策略性地用来给政策、项目和具体活动提供参考依据，从而帮助将资源分配到最需要它们的地方。

考查各地方层面（例如省级或区县层面）的病例通报率尤其有助于了解一个国家内结核病负担的差异。为了开展这类评估，各地方层面的病例通报总数会按照特定地区的人群进行划分。地方通报率被制成分布图，以便更清晰地呈现，采用不同的颜色来指示不同层面的病例通报率，这样可以非常明确地显示全国范围内的差异。^k例如，在柬埔寨，东北部省份的结核病通报率很显然要比国内其他地区低（图 10）。这些通报率可能反映了区域内的真实结核病负担，或者表明人口密度较低地区的卫生保健服务可及性较低。相反，人口密集地区有着较高的结核病通报率，可能是由于结核病感染率和患病率的实际差异，或者是因为这些地区的卫生保健服务可及性较高。中国的情况正好相反，其中东部人口密度较大且经济比较富庶的城镇地区有着较低的通报率，而西部卫生保健服务可及性较低且人口密度较小的农村地区有着较高的通报率（图 11）。

k. 这些分布图采用 Quantum GIS 制作。Quantum GIS 是一个免费的跨平台开放源（GIS）应用程序，面向桌面计算机，提供数据查看、编辑和分析功能（<http://www.qgis.org>）。

在泰国，主要城市和跨境人口流动率高的省份也有着较高的通报率（图 12）。同样，在柬埔寨，与越南和泰国接壤的省份也有着较高的结核病通报率。在莫桑比克，与斯威士兰和南非接壤的南部人口高度密集地区同样有着较高的通报率，此外海岸沿线的较大城市也有着较高的通报率（图 13）。

其他常规收集的数据可以通过可视化手段呈现，并在地方层面上进行比较。在莫桑比克，具体例子包括 15 岁以下儿童中结核病通报数所占的比例（图 14），以及肺外结核通报数所占的比例（图 15）。即使在同一个国家内，我们也可以看到不同省份之间的巨大差异，这可能归功于临床诊断能力、卫生保健服务可及性和省级监测系统捕获病例的质量等，而不是儿童结核病负担和肺外疾病负担的实际地区差异。

虽然超出了本手册的范围，还是务必要注意到如果具有人口和社会经济数据，进一步开展分析有助于确定人口统计学可变性与结核病负担之间的关系（例如：如何在富裕城镇、贫困城镇、富裕农村和贫困农村之间比较结核病病例通报数？）。

图 10
柬埔寨各省的结核病通报率（所有类型），2011 年
资料来源：柬埔寨国家结核病防治规划 2012 年年度报告

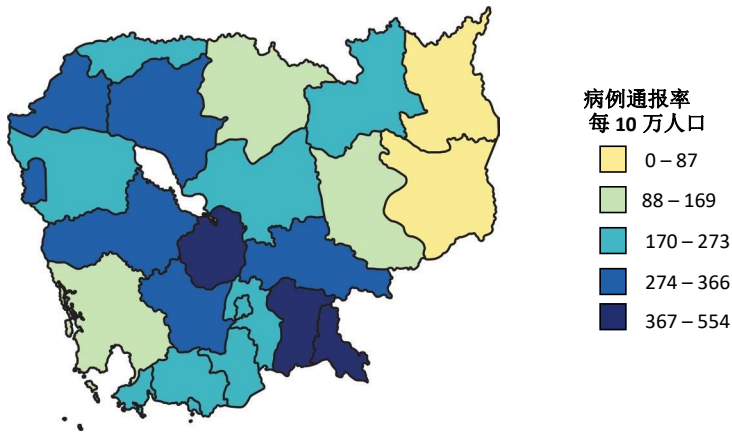


图 11
中国各省份的结核病通报率（所有类型），2010 年
资料来源：中国卫生部，中国十年国家结核病防治规划评估报告（2001-2010 年），
2011 年，军事医学科学院，北京

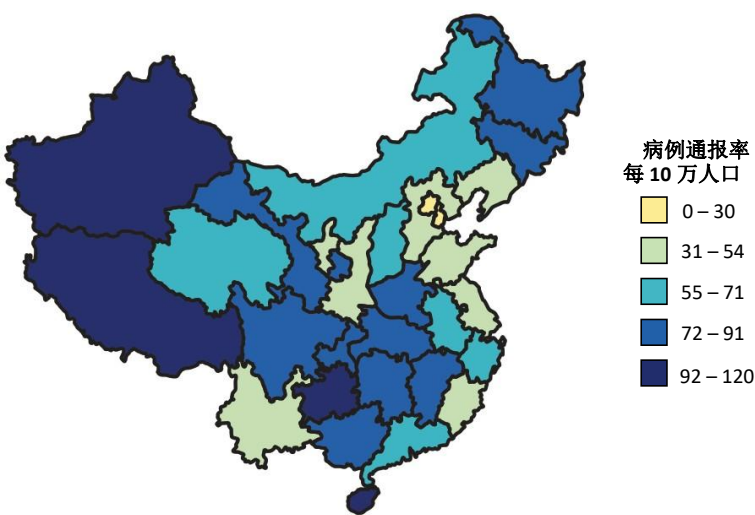


图 12
泰国各省份结核病通报率（所有类型），2012 年
资料来源：泰国国家结核病防治规划

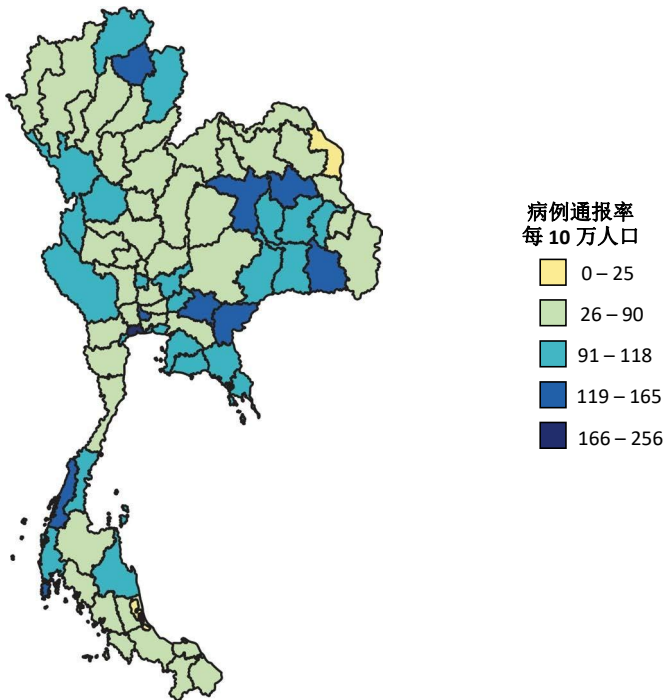


图 13
莫桑比克各地区结核病通报率（所有类型，新病例），2012 年
资料来源：莫桑比克国家结核病防治规划

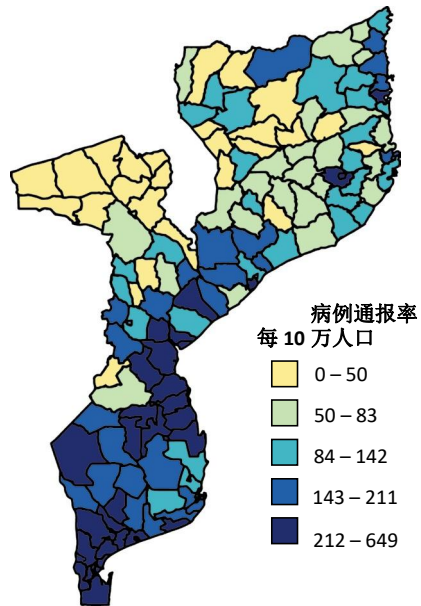


图 14
莫桑比克各省份通报病例总数中儿童结核病通报病例所占的比例
（所有类型，新病例），2012 年
资料来源：莫桑比克国家结核病防治规划

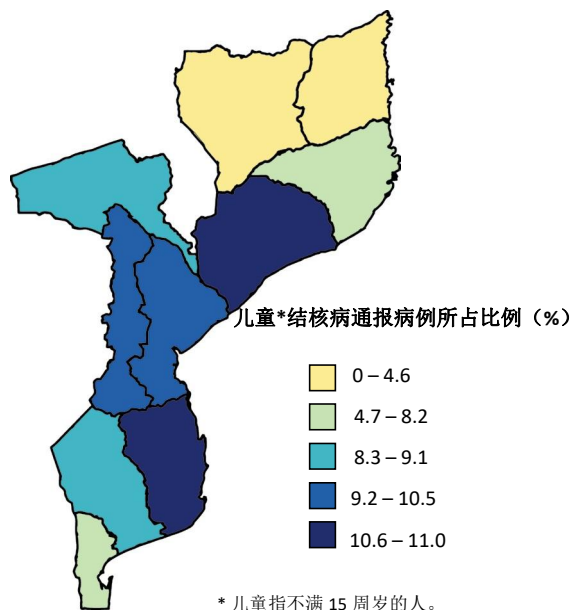
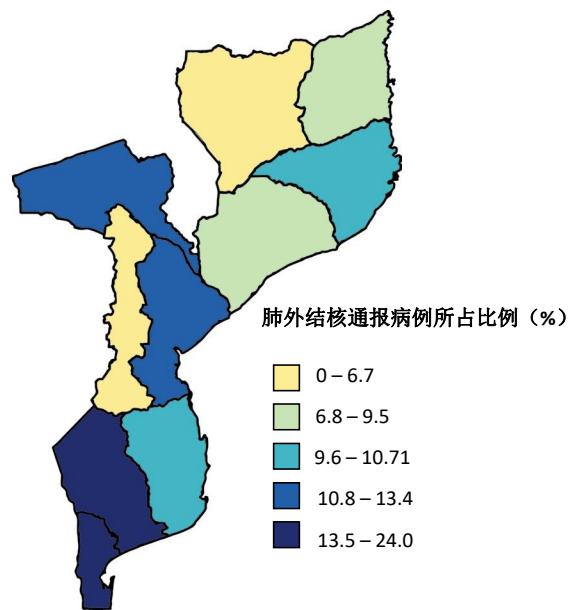


图 15
莫桑比克各省份通报病例总数中肺外结核通报病例（新病例）所占比例，
2012 年
资料来源：莫桑比克国家结核病防治规划



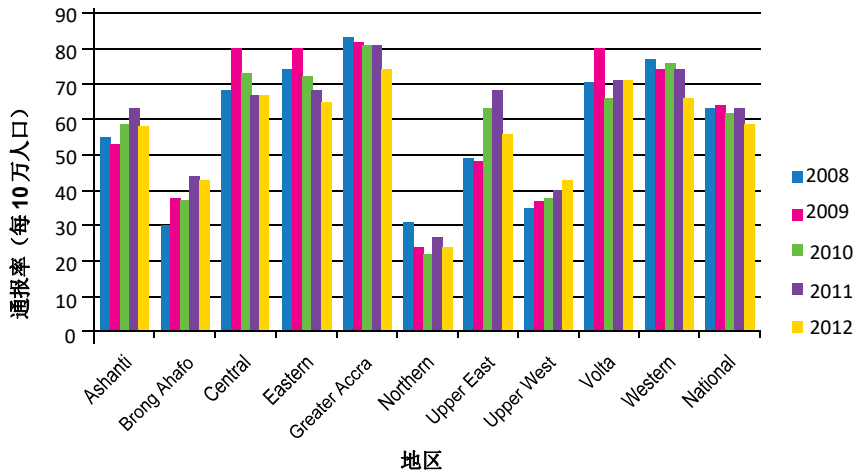
按地点和时间划分的通报数

可以按不同地区、不同的人和时间观察趋势。例如，我们可以评估地方区域内各年龄段和各性别通报率的趋势。如果开展该项评估时结合常规记录和报告系统的详细分析，则可以让国家结核病防治规划改善对结核病流行病学的了解，以及解决常规结核病监测的不足之处。

加纳

考查加纳从 2008 年到 2012 年的地方结核病通报率可以发现不同地区之间存在较大的差异（图 16）。平均而言，大阿克拉地区有着最高的通报率，而北部地区一直有着最低的通报率。如果只是观察了全国通报率，这类差异可能会被错过，结核病流行情况会被夸大。

图 16
加纳各区域结核病病例（所有类型）通报率，2008–2012 年
 资料来源：加纳国家结核病防治规划



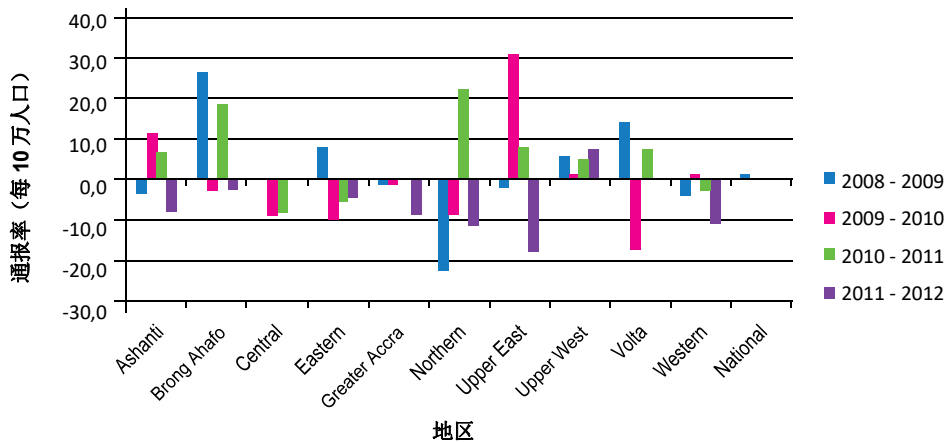
还可以观察到各地区内从 2008 年到 2012 年的差异。看起来布朗阿哈福地区从 2008 年到 2012 年出现了结核病通报率的非预期大幅增长，上东部省从 2009 年到 2010 年也出现了结核病通报率的非预期大幅增长。如果计算随后数年间的结核病通报率百分比变化，则这些差异还能发挥更大的作用（图 17）。可以采用下列公式完成计算，其中“*i* 年”对应着感兴趣的第一年，“*i* 年+1”对应着随后的一年：

$$\text{每年的百分比变化} = \frac{\text{通报率}(i \text{ 年}) - \text{通报率}(i \text{ 年} + 1)}{\text{通报率}(i \text{ 年})} \times 100$$

计算百分比变化属于数据质量的一项内部一致性检查。各年度之间一旦观察到较大的百分比变化，则务必要找出这些差异的可能原因。这些波动是否由自然变化所导致，或者是否存在人为影响监测数据的项目规划或人工错误？差异较大（例如变化超过 10%）的薄弱系统更有可能是与报告和记录错误相关的问题，而不是地区内病例数量的实际变化。¹

1. 在绩效较高的监测系统中，最初结核病病例较少的地区可能会出现较大的年度波动。

图 17
加纳各地区结核病病例（所有类型）通报数：各年度的百分比变化，
2008–2012 年
资料来源：加纳国家结核病防治规划

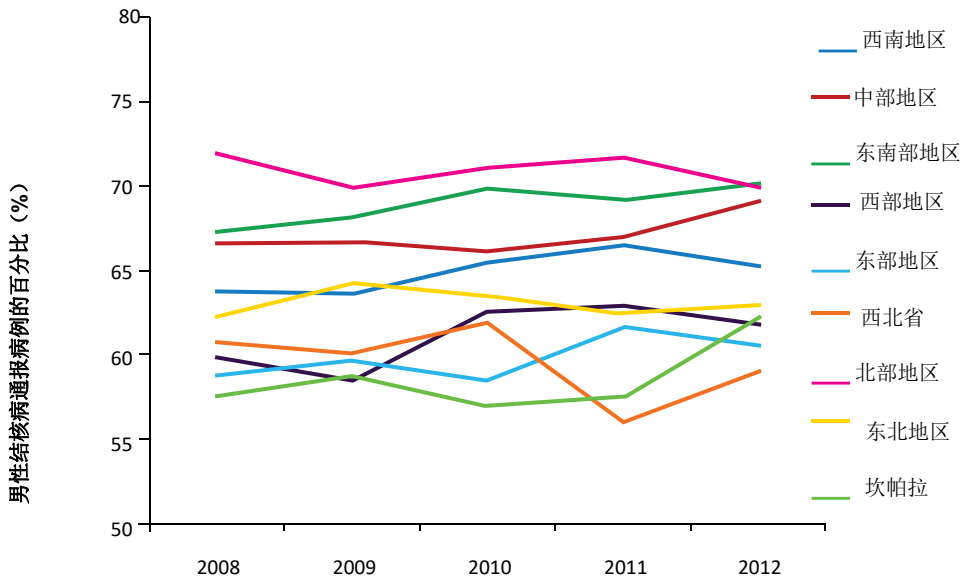


采用世界卫生组织标准和基准核查清单对加纳的国家结核病防治规划监测系统开展了评估，结果发现季度报告中的数据存在内部不一致的问题，有些情况下还存在不准确和不完整等问题。大量的数据质量审计发现，在四个地区选出的 29 个基本管理单元中，只有 11 个单元在季度国家结核病防治规划报告中的结核病病例总数与基本管理单元结核病登记簿中的病例数量相匹配。另外有 11 个基本管理单元被发现存在漏报现象，7 个基本管理单元存在多报现象。这类差异与特定地区观察到的通报率百分比变化的大幅波动有关。后来，国家结核病防治规划尝试着提高了总体数据质量，具体包括更加重视已修订的数据收集工具的正确使用、推动更为频繁的培训、落实标准操作规程、安排更经常的监测考查和季度数据质量审计。

乌干达

乌干达于 2008 年到 2012 年间在省级层面上考查了各性别的结核病病例数量，结果发现男性与女性之间存在较大的差异（图 18）。所有地区都是男性报告病例多于女性报告病例，不过所有男性病例的比例在不同地区之间也有差异。另外看起来好像还存在一些需要解决的不一致问题，从而提高监测系统的数据质量。例如，西北省和坎帕拉在这类报告中发现了更大的不一致，分别表现为 2010–2011 年和 2011–2012 年男性结核病通报病例百分比的大幅波动。需要找出这些不一致的原因，还应在其他变量中寻找类似的 inconsistency 模式，从而了解提高数据质量需要解决的各种问题。

图 18
乌干达各结核病报告地区男性结核病通报病例的比例，2008–2012 年
 资料来源：乌干达国家结核病防治规划



通报率在不同时间段出现变化的原因

通报率可能会因为很多原因出现变化，包括自然原因（例如人口流动的加剧或地方疫情的蔓延）、项目规划原因（例如将重点转移到发现儿童结核病病例上）或未知原因。通常，通报率的变化应当归根于监测系统和数据质量方面的问题，除非证明另有其他原因。回答下列问题有助于确定数据出现实质性差异的原因：

- 社区中是否有任何重大事件会自然地影响了结核病负担（例如难民涌入）？
- 社区中是否有任何重大事件积极地或消极地影响了卫生保健服务的可及性（例如战争、自然灾害、普遍健康保险项目的引入、新卫生所或医院的建立）？
- 报告期内监测系统是否出现了任何重大变化（例如报表的调整或人员的变动）？这些变化可能已经导致数据质量出现问题（例如对定义的不同解释或迟迟不采纳新的表格，或者由于新定义或新报告方法的广泛适应而导致人为增加或减少结核病指标）。

- 报告期内结核病防治项目或健康外展项目是否出现了任何重大变化（例如卫生工作者的增加）？结核病健康交流项目是否有所增加或者是否更加关注儿童结核病？是否做出新的努力来覆盖已知面临着结核病风险的格外脆弱群体？

如果研究结果发现观察到的变化与报告系统有关，并不能真正地反映结核病病例数的增加，则务必要找出问题的根源和制定相应的解决方案。

1.5 聚合通报数据的局限性

聚合报告在病例相关通报数据的收集方面有着多项局限性（1）。主要局限性包括：

1. 聚合报告通常基于多份季度纸质记录的人工病例数目。这是一个很容易导致错误的强化过程，然而通常没有嵌入式机制可以用来发现和更正错误。因此，在区县、省市和国家层面上所有卫生机构收集聚合数据并报告给结核病数据管理单元是一项非常耗时的任务，会在编制、分析和确定结核病通报病例和治疗转归的趋势方面造成延迟。而且，人工处理数据通常会在将信息从一个表格转移到另一个表格的过程中导致转录错误。
2. 由于将病例相关信息转录和编制为聚合报告属于劳动密集型工作，这些系统只能收集非常有限的信息。收集的数据包括：性别、年龄段、治疗史、疾病部位（肺部或肺外）、诊断是的细菌学状况。治疗转归通常只按照治疗史收集。
3. 聚合数据的分析比较复杂，而且报告选项也很有限。随着耐药结核病诊断和治疗相关综合信息的收集和报告需求越来越大，数据的分析和报告显得尤其重要。
4. 有关个体病例或患者的数据在卫生机构层面之上并不是现成可用。因此，项目负责人和政策制定者获得病例相关数据的途径有限，而他们需要开展有据可依的项目规划评估（例如患者的诊断和治疗管理）。
5. 不可能将聚合数据链接到其他数据库。这在最大程度上减少了系统收集的结核病数据的范围、潜能和效用，减少了如果使用病例相关数据可能会实现的研究、项目规划和政策衔接。

总体而言，病例相关结核病监测系统可以更详细、更广泛和更及时地收集和分析信息。因此，多个国家（包括印度尼西亚、肯尼亚和越南）正在从纸质聚合结核病监测系统转换为电子病例相关系统。有些国家需要创建这些系统，而其他一些国家只需将这些系统整合到面向其他传染病的现有电子报告系统中。世界卫生组织*结核病防治电子记录和报告指南（1）*中提供了这些系统的示例，以及有关如何选择、设计、实施和维护这类系统的详细信息。

1.6 小结

国家结核病防治规划被鼓励常规分析各自的数据。本章解释和展示了相关方法，国家结核病防治规划可以通过这些方法使用结核病监测系统常规收集的聚合数据。简单的描述性数据分析假定数据具有高质量，这类分析有助于国家结核病防治规划描述社区内结核病病例的“谁、什么、何时和何地”等要素。然后，这些分析可以生成相应的假设和进一步的探索性研究，从而回答当前现象“如何”和“为什么”会存在的问题。国家结核病防治规划还可以利用各自的数据和特定数据分析，评估通过监测系统收集的数据的质量，确定这些数据是否真正代表了国内当前的结核病流行情况。虽然本章的重点是病例通报数据，不过其他指标也可以用于类似分析，例如聚合死亡率数据或成功结核病治疗转归。尽管聚合数据有一定的效用，但这种效用在很多方面非常有限。务必要认识到这些局限性，如果可能的话，应当尽量开发或改编病例相关监测系统，从而开展更为可靠的分析。

文字框 4**国家结核病防治规划可以开展的数据分析**

- 收集国家通报数据和绘制各年度病例通报数分布图（所有类型，涂阳，取决于数据可用性）。
- 对于最近一年，绘制各年龄段病例通报数（所有类型或涂阳结核病病例）分布图。是否还可以绘制前面五年的分布图？描述已经绘制的内容。这意味着什么？
- 通报数据是否按年龄划分？如果是，创建通报率并按年龄段绘制分布图。是否还可以绘制前面五年的分布图？描述已经绘制的内容。这意味着什么？
- 所有结核病通报数中 15 岁以下的儿童占多大比例？该比例与前几年的比例是否相同？
- 通报数据是否按性别划分？如果是，创建通报率并按性别绘制分布图。是否还可以绘制前面五年的分布图？是否还可以按年龄段绘制分布图？描述已经绘制的内容。这意味着什么？
- 地方层面（例如省级）上是否具有通报和人口数据？如果是，创建地方通报率并进行比较。描述通报率之间的差异。同样，按年龄或性别计算地方通报率，然后比较分布和时间趋势。这些意味着什么？
- 是否也可以对其他指标（例如死亡率或成功治疗转归数据）开展同样的分析？
- 确定解释这些信息的最佳方式。

1. Electronic recording and reporting for tuberculosis care and control. Geneva: World Health Organization; 2011 (WHO/HTM/TB/2011.22). 网址: http://whqlibdoc.who.int/publications/2012/9789241564465_eng.pdf
2. Definitions and reporting framework for tuberculosis—2013 revision. Geneva: World Health Organization; 2013 (WHO/HTM/TB/2013.2). 网址: http://apps.who.int/iris/bitstream/10665/79199/1/9789241505345_eng.pdf
3. Checklist of standards and benchmarks for case-based TB surveillance data. Geneva: World Health Organization; 2014. 网址: http://apps.who.int/iris/bitstream/10665/112673/1/9789241506724_eng.pdf?ua=1
4. Worker's manual on the management of tuberculosis and leprosy diseases - 5th edition. Federal Ministry of Health, Department of Public Health, Nigeria. 网址: http://www.who.int/hiv/pub/guidelines/nigeria_tb.pdf

附录 1

结核病监测数据质量标准，带有示例

世界卫生组织结核病监测和人口动态登记系统标准和基准核查清单包含 13 项标准，其中许多标准可以用来评估国家结核病防治规划常规监测系统的数据质量。尤其应当定期考虑下列标准：

- **病例定义（标准 1）。**评估数据质量时，首先必须要考虑的事情就是结核病监测中采用的病例定义和采用这些定义的统一性。普遍适用的标准化病例定义显得非常重要，有助于提供有关流行病学趋势和防治项目绩效的一致性信息，还可以推动国家、区域和全球层面上的准确比较。不明确的、不一致的或非标准化的病例定义会给这类监测带来不利影响，导致聚合数据的分析变得困难或完全无法开展。因此，应当评估国家监测系统采用的病例定义的一致性，这对于解释系统收集的数据非常关键。

*示例：*尼日利亚国家结核病和麻风病防治项目已经开发和发放了《医务人员结核病和麻风病管理手册》（4）。手册对结核病病例给出了明确的定义，包括实验室确认病例与临床病例之间的差别、新病例与既往治疗病例之间的差别、肺部病例与肺外病例之间的差别等。这些病例定义属于各国结核病管理常规培训的一项重要内容。因此，尼日利亚的结核病病例定义与世界卫生组织指南保持了一致，尼日利亚的监测数据也满足这项标准。正是因为这些努力，国家结核病防治规划非常有信心，坚信全国范围内都根据这些定义统一地记录相关数据，这是获得高质量数据的第一步。

- **一组最少报告变量（标准 2）。**世界卫生组织已经确定了一组最少变量，代表着评估结核病监测系统数据质量所需的根本要素。^m 因此，应当给所有结核病病例收集和报告下列数据：年龄、性别、登记年份、细菌学结果、既往治疗史、疾病解剖部位。确保在监测系统中收集这些变量的数据，从而更好地评估数据质量（例如外部一致性和内部一致性）。

m. 年龄或年龄段、性别、登记年份、细菌学结果、既往治疗史、疾病解剖部位，病例相关系统还需要患者识别号（例如数字编号）。

示例：在尼日利亚，结核病监测系统在机构层面上采用纸质形式。患者数据被收集后会记录到机构结核病登记簿中。地方政府区层面负责编制聚合报告，然后通过州和地区结核病监督员提交给国家结核病防治规划，结核病监督员在转发报告之前会对数据进行验证。**2013** 年开发和安装了一个聚合电子数据库，致力于在地方政府区和州级层面上收集数据和分析趋势。常规收集下列变量的数据：年龄、性别、登记年份、细菌学结果、既往治疗史、疾病解剖部位。此外，地方政府区监督员在机构层面给每名患者分配一个识别号；国家指南提供了有关如何生成识别号的指南。根据评估内容，给所有结核病病例捕获一组最少变量，因此该项标准在尼日利亚得到了落实。这意味着尼日利亚常规收集足够多变量的信息，可以评估收集的数据的质量。

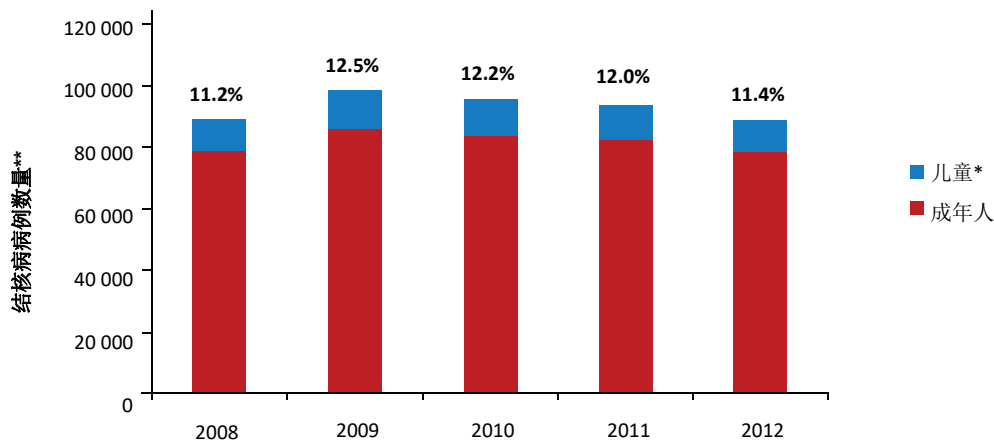
- **收到和处理所有数据提交内容（标准 3）。**遵循世界卫生组织指南的纸质系统依赖于结核病基本管理单元编制的季度报告，季度报告将一组卫生机构的结核病登记簿聚合起来。这些季度报告按行政级别层次往上提交，每个级别都有可能对数据进行聚合处理。国家层面的数据聚合应当基于国内 **100%** 基本管理单元（例如形成聚合的最低层面）在评估年度提交的报告。如果不能采用 **100%** 的预期季度报告或数据文件，则可能导致许多病例未被包括在国家统计数据中。

示例：乌干达在国家层面上可以了解每年提交报告的地区的数量，以及这些地区占预期地区总数的比例。**2011** 年，国家层面上收到了全部 **468** 份预期报告。

- **内部一致性（标准 4 针对聚合数据，标准 5 针对病例相关数据）。**卫生保健机构在源文件中会按年龄段、性别和病例类型收集结核病病例数据，这些病例数量的小计加起来应当与基本管理单元报告的结核病报告病例的总数相等。
- **外部一致性（标准 6）。**将国家结核病防治规划的结核病监测数据与全球结核病流行病学数据进行比较，这是衡量监测系统质量的一种方法，也是评估通过监测系统收集的数据的一种方法。对于国家艾滋病防治规划来说，外部一致性的评估方法是计算规划内被诊断为结核病的儿童的百分比，然后将该百分比与全球平均值进行比较（作为全球平均值，低收入和中等收入国家的儿童结核病病例（**15** 岁）应当占有结核病病例的 **5–15%**，该比例在高收入应低于 **10%**）。

示例：肯尼亚麻风病、结核病和肺部疾病监测系统收集的数据在 2008 年到 2012 年间具有外部一致性。儿童结核病病例的百分比在五年内位于中低收入国家的可接受范围（5–15%），而且各年度之间的差异非常小（图 A1）。ⁿ

图 A1
肯尼亚儿童结核病病例占有所有结核病通报病例的比例，2008–2012 年



* 患者 15 岁。所有儿童病例都被认为是新病例。
** 病例包括涂阳、涂阴、肺外和涂片为开展的病例。该项分析排除了复治病例。

- **内部一致性（标准 7）。**最后，务必要评估规划内部的数据的一致性。众所周知，结核病病例数和死亡数在不同时间段的变化应当最小，因为即使是在最好的情况下，全国结核病发病率的下降一直都没有超过每年 10%。因此，可以在国家级、省级或其他行政层面上考查不同时间段的一组最少变量（上文有相关描述），如果出现任何限制差异，则可能提示数据报告和记录的不一致，或者需要做出其他解释。有关分析的具体建议包括肺部通报病例与肺外通报比例之间的比率变化、男性病例与女性病例之间的比率变化、儿童结核病病例占有所有病例的比例、所有类型结核病的病例通报率，新涂阳结核病的病例通报率、结

n. 英国和荷兰的数据已经记录了全国范围内发病率的最大下降，即 1945 年后达到了 10%。当前绩效最佳和负担较高的国家是柬埔寨和中国，自 2000 年以来记录的发病率下降速度为 3–5%。

核病疑似病例与通报病例之间的比率。最后，如果具有死亡率数据，国家结核病报告病例数量的变化应当与国家结核病死亡率的变化保持一致。

第 2 章

病例相关结核病通报数据的分析

目标受众：

普通读者，不过重点是在国家结核病防治规划（NTP）内各个管理层面负责病例相关监测数据的督导与评估官员和流行病学家。

预期成果：

本章结束时，读者应当能够了解：

- 病例相关结核病通报数据对于描述结核病流行情况和给国家结核病防治规划提供参考依据的重要性；
- 如何评估病例相关结核病通报数据的质量？
- 如何分析、报告和解释病例相关数据？
- 病例相关数据的分析如何促成实践和政策的变化。

作者：

Emily Bloss、Rachel Yelk Woodruff、Deanna Tollefson、Susan van den Hof、Lori Armstrong、Laura Anderson

2.1 病例相关通报数据：什么是病例相关通报数据以及它们为什么很重要

病例相关监测数据包含相应的记录，其中个体结核病病例和相关治疗信息属于分析单元（1）。病例相关系统中的数据经过整合后可以生成聚合数据，不过仍然保留着每份结核病病例的单独记录，相比只具备聚合数据的系统，单独记录有助于开展更为详细的分析。利用病例相关通报数据的监测系统是一项比较理想的选择，不过许多高结核病负担国家目前在采用基于聚合数据的系统。国家结核病防治规划可以分析病例相关结核病监测数据，找出容易让人面临更高疾病风险和较差转归的人口特征，以便更好地制定适合当地和国内具体情况的干预。国家结核病防治规划还可以使用病例相关数据，发现疾病暴发和指导及时的公共卫生行动，确保结核病病例和接触者的适当管理，评估结核病控制领域取得的进展情况，并与国内和国际目标值进行比较，从而给政策提供参考依据（2）。不过，根据分析的目的或具体目标，国家结核病防治规划开展的分析类型可能会有差异。

对于病例相关结核病监测数据，每个病例都代表着系统内一个单独的录入项，治疗期间会给录入项添加更多信息。系统内通常可以捕获唯一识别号，该识别号可以将与结核病病例相关的不同信息片段衔接起来。在一些国家，国家结核病防治规划使用唯一个人识别号（例如社会保障号）促进不同疾病发作的衔接，确定同一个疾病发作的重复录入项，并与其他国家数据库挂钩。偶尔情况下，只有特定区县或卫生保健机构可能会使用该识别号。这使得很难确定既往结核病发作，尤其很难了解患者既往是否在其他地方接受过治疗。^a

虽然与维护一个采用聚合数据的纸质系统相比，建立一个电子病例相关监测系统可能需要更多资源和协调，但电子病例相关监测系统还是更优越，原因如下（1）：

- 记录可以用于患者的临床管理，确保高质量关怀和监测治疗转归。
- 从每名患者的记录中编制每个结核病病例的病史，这样有利于开展轻松的病例相关分析（例如分析结核病患者亚群的特征或患者特征对治疗转归的影响）。
- 可以将感兴趣的变量添加到监测系统，例如结核病的风险因素（例如吸烟状况或糖尿病），并且开展分析，从而更好地了解结核病流行情况。

a. 在患者相关记录中，个人（而不是结核病病例）属于基本记录单元，这样只要系统覆盖到全国范围，就可以了解患者的既往治疗史，系统还能自动利用患者的既往结核病发作来识别复发结核病患者。不过，许多系统并不满足这些标准，而是捕获病例相关数据。

- 数据质量检查可以轻松得到实施，帮助确保数据的准确性和有效性（例如发现重复录入项或不合理的数据组合，例如一名男性患者怀有身孕）。
- 记录可以被链接到其他数据库，以便检索更多信息（例如来自实验室信息系统的药敏试验结果）。这可以提高数据的质量和扩展数据使用的范围。
- 报告在任何时候都可以自动生成，而不需要等待季度聚合报告出来再生成。
- 数量更多的数据可被国家结核病防治规划工作人员用于分析、解释和使用。

有些地区的结核病监测系统继续依赖于通过区县登记簿编制的季度纸质表格，而考虑到上述益处，国家结核病防治规划可能需要细致地考查电子病例相关系统的各种益处。

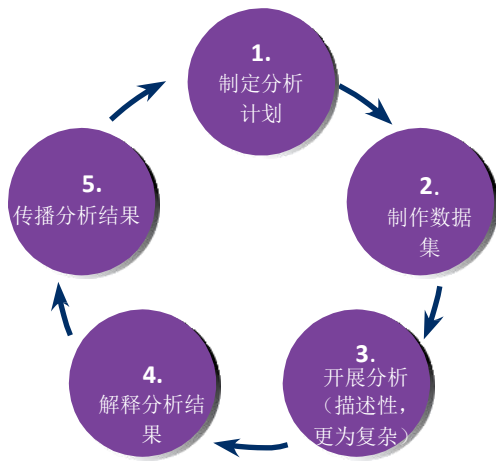
本章描述分析病例相关数据的标准方法，读者可以根据这些内容规划和开展类似的分析。这包括如何制定一份分析计划、准备用于分析的数据、分析和解释常规监测数据等。本章最后简要介绍了如何开展更为深入的分析。为了了解描述的相关概念和分析方法的更多信息，请参见世界卫生组织的出版物《基本流行病学》（3），该出版物提供有关各种流行病学概念的详细定义和分步骤解释。

病例相关数据分析的步骤

病例相关数据可以拓展数据分析的机会。聚合和病例相关系统中结核病监测数据的记录和报告都是基于国际认可的通用原则（4），绝大多数都在各自监测系统中收集有关结核病病例的人口统计学、临床学和细菌学信息。事实上，世界卫生组织建议收集和分析一组最少变量，病例相关结核病监测数据标准和基准核查清单（例如年龄/年龄段、性别、登记年份、细菌学结果、既往治疗史、疾病类型和唯一编号）中列出了这些变量（5）。不过，根据国内具体情况、流行病学现状和结核病控制重点工作，许多国家会常规收集和分析额外的变量，尤其是采用病例相关系统的国家。例如，美国和欧盟（EU）的国家结核病防治规划常规分析和报告有关出生地的数据（出生在特定国家内或出生在特定国家以外）。地方层面（例如州、区县或地区）开展风险因素，耐药结核病和趋势分析的现象也很普遍。

分析病例相关数据的过程涉及到很多步骤，其中一些步骤是在开始数据分析之前完成（图 1）。这些初步步骤对于工作规划和确定数据质量很重要，并会进一步影响到分析结果如何给结核病防治规划提供参考依据，或者分析结果如何以其他方式促进结核病研究工作。

图 1
分析病例相关数据的步骤



2.2 制定分析计划

在对常规或研究活动开展分析之前，国家结核病防治规划指定负责开展分析工作的人员（通常是一名流行病学家）应当制定一份分析计划。分析计划可以在两个方面带来益处。首先，分析计划有助于确定开展病例相关结核病监测数据分析时所需的关键步骤、技能和软件（例如采用什么统计检验，以及完成分析需要哪些额外的信息、工具或技能）。其次，分析计划提供一份蓝图，确保流行病学家有针对性地、高效地开展分析工作。流行病学家可以开展分析的过程中根据需要修订分析计划。分析计划应当包括：

- 分析目标和研究问题；
- 研究人群，以及纳入和排除标准；
- 要使用的变量和关键定义；
- 表格和样本图；
- 有关如何处理缺失值的指南；
- 要开展的分析；
- 要采用的统计检验。

实施分析计划之前必须选择好统计软件。流行病学家可以将病例相关数据导入到各种用于分析的专业统计软件包中，例如 R、SAS、Epi Info、SPSS 和 STATA。流行病学家还可以使用地理信息系统（GIS）来绘制分布图和开展空间分析。虽然这些类型的分析可以在地方和国家层面上开展，但还是应当雇佣一名熟悉统计程序且经验丰富的流行病学家（或数据管理员或统计学家）来帮助支持数据分析。

2.3 制作数据集

流行病学家一旦制定了分析计划，接下来应当制作数据集。这涉及多个步骤，第一步是一般数据清理，最后一步是干净的、“被冻结的”数据集，以便用于分析。

数据清理

监测数据的记录和报告有时会出现错误。数据在被分析之前应当妥善清理，因此必须具备标准规程最大限度地减少数据集中的记录错误，以及发现和更正这类记录错误。这可以确保数据分析产出更容易解释的结果，而且出现的偏倚会更少；它还有助于持续地评估和修订既定的数据收集流程。相对于聚合数据系统来说，病例相关监测系统的数据清理要轻松得多，因为后者可以更轻松地评估和更正数据问题。电子系统也优于纸质系统，因为电子系统可以实施自动化的验证检查。电子数据库还可以轻松检索所以更有利于评估数据的质量。

为了确保数据清理得当，流行病学家应当遵循下面的一般原则：

- 收集和录入过程中最大限度地保证数据质量；
- 系统性地检查数据，找出缺失的、不准确的、不一致的、分类不当的和不合理的数据值和重复记录；
- 以系统化的方式解决错误的数据点。

首先，流行病学家务必要采取措施最大限度地提高数据收集和录入时的质量。为了做到这一点，可以制定结核病防治规划每个行政管理层面（例如地方、区县、省级和国家）上开展数据收集、录入和查看的标准操作规程。这样可以推动所有相关方采取统一的数据收集和录入方法，提升这些流程涉及的所有人对数据质量的认知。制定这类标准操作规程后，关键还是要给省级和地方层面提供培训，介绍和审阅标准操作规程，以及提高数据收集和管理能力。给结核病数据记录和报告流程中涉及的所有人提供标准操作规程的书面手册或海报，这样有助于提高利用率和确保一致性。除了标准

操作规程之外，流行病学家还务必要确保设计的数据库能最大限度地减少数据录入错误。^b例如，应当尽量减少数据录入中的自由文本字段，帮助确保录入有效的数据值；相反，应当使用预先定义的下拉菜单或取值范围。同样，应当制定数据验证规则，并整合到数据库中（例如诊断日期 治疗开始日期）。最后，流行病学家应当确保数据库的设计不会让重要的数据值被遗漏。可以设计相应的错误信息，防止数据录入人员在录入相应数据值之前进行相关处理。积极落实这些步骤可以确保流行病学家在数据清理期间面临较少的错误。

示例： 采取措施确保高质量数据的使用

在英国，英格兰公共卫生署（PHE）管理着一个基于网络的结核病监测系统。该系统被设计来最大限度地提高数据收集和录入过程中的数据质量。在数据录入屏中，自由文本字段的数量有限，绝大多数字段采用预先定义的下拉菜单，这样是为了确保录入有效的数据值。还有自动数据验证检查功能，防止出现数据录入错误，例如“症状出现日期”必须在“第一次到诊所就诊日期”之前，而“第一次到诊所就诊日期”必须在“治疗开始日期”之前。同样，“出生日期”不能是未来的某一天，而且治愈或完成治疗方案的“治疗转归”数据不能记录在“通报日期”后的六个月内。除了这些措施之外，许多地区还有关于发现和更正错误的季度数据报告。病例各个变量之间的不兼容会用红旗标记，然后通过病例最初就诊的门诊跟进。

数据录入后，流行病学家应当系统地、全面地查看缺失的、不准确的、不一致的、分类不当的、不合理的或重复的数据值。流行病学家应当开展多项不同的检查，以便找出这些数据值。例如，流行病学家应当检查看看是否同样的患者在系统中出现了不止一次（例如查找重复病例）。还应检查类别变量（例如涂片状况）的答案，确保数据值在为相应变量指定的类别范围中（例如针对变量“涂片状况”的阳性、阴性或涂片为开展等选项）。同样，应当考查持续变量（例如年龄和体重）的范围（例如查找超过正常或合理范围的答案）。最后，应当检查数据点，看看是否符合逻辑和具有一致性（例如男性病例不可能怀有身孕；患者的年龄应当正确对应出生日期）。

b. 通常不建议对病例相关结核病监测数据进行双重数据录入。

示例： 数据质量控制

荷兰国家结核病防治规划在国家层面上积极开展重复录入查找活动。他们采用出生日期、诊断年份和卫生保健机构代码的组合，必要时还采用性别、诊断结果或邮政编码等变量。如果出现重复记录，就会联系当地工作人员核实。如果当地工作人员确认存在重复录入，市级卫生保健团队再次确认后会清除重复数据。

流行病学家确保数据清理得当的最后一步是以系统化的方式解决错误的数据点。应当具备明确的方案用于解释如何处理错误数据。还应当制定标准规程在记录中出现错误时发现错误和（如果适当的话）修正错误。例如，为了核实错误数据点已被删除，所有不准确或缺失数据都应当在被发现时得到妥善处理。请注意，应当采用单一的规则来处理缺失或不准确数据，并且将此单一规则应用于所有数据。理想的规程是首先追溯到源数据（例如地方层面上保留的数据），以便核实和更正录入项。

示例： 必要时修正错误

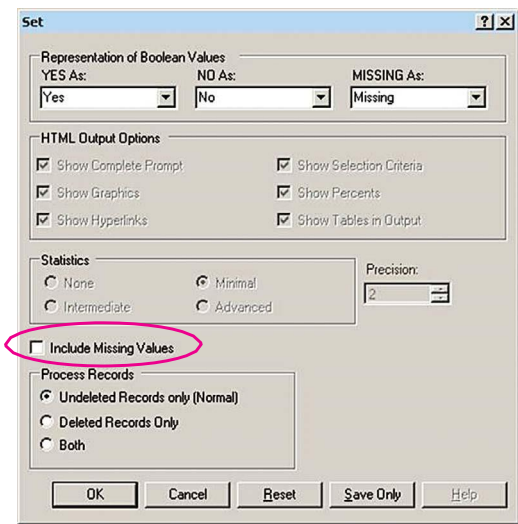
在英国，英格兰公共卫生署会跟进区域层面上（通过标准季度报告）发现的数据不一致，方法是联系最初提交数据的诊所。在美国，每个州的公共卫生官员都被鼓励根据患者的病历交叉检查数据库中缺失或明显不准确的记录。

在开始病例相关监测数据分析之前，许多方法可以用来查看、发现和解决缺失的或错误的数据。相比数据收集数月（或数年）后再尝试开展数据清理步骤，定期落实数据清理步骤可以捕获更多错误，而且可以更快和更轻松地完成数据清理流程。而且，开展定期数据审计和分析，举办有关数据质量保障的专门培训，将数据质量检查整合到常规监督中，遵循数据收集和录入的标准操作规程等，这些都能帮助确保最大限度地减少数据集中的错误，即使出现了错误，也能更为轻松地解决错误。

解决缺失的数据

清理数据集时，流行病学家务必要考虑如何解决缺失的数据值。监测数据通常会出现信息缺失现象，原因是患者可能不知道或可能拒绝提供所需的详细信息，或者服务提供方没有填写必要的信息。虽然如此，流行病学家还是可以轻松找出缺失的数据，方法是运行“frequency”（频率）程序，查找出现缺失数据的病例的数量。绝大多数软件程序提供一个选项，让用户决定是否希望将缺失值包括在统计计算中（图 2）。如果许多病例在特定变量中存在缺失值，则采用该变量的分析的结果可能会出现偏倚。

图 2
Epi-Info 7 的“包括缺失值”特点



一旦流行病学家已经确定了缺失数据，流行病学家接下来需要解决出现缺失值的病例。理想情况下，流行病学家应当首先检查原始数据源，确认缺失值。如果在原始数据源找到缺失值，流行病学家然后应当在数据集中更正相关数据值。如果在原始数据源中也缺失数据值，流行病学家需要决定如何分析和报告缺失数据。通常而言，缺失值代码（例如 9、99、999 或“*”）被用来界定缺失值（6）。这一点可以从表 1 看到，在一个由 15 份病例组成的小数据集中，连续变量“年龄（周岁）”的一个缺失值被编码为“999”。即使查看原始表格后，流行病学家还是发现三份病例存在缺失数据。虽然存在缺失数据的病例数量看起来很少，但存在缺失数据的三份病例就意味着所有病例中有 20%的病例不具备年龄信息。

表 1
各年龄值的病例频次

年龄（周岁）	各年龄值的病例数量 （总数=15）
5	1
17	4
32	1
45	3
58	1
66	2
999	3

有时，流行病学家可能会给存在缺失数据的病例百分比设定一个限度（例如截断值），以此确定这类病例是要被排除还是要作为结果呈现。例如，如果一项变量缺失 >5%的数据，则比较适合将缺失数据作为结果呈现（例如在一个标题为“缺失”的类别中），有助于确保所有信息都被呈现。这确保在解释数据时对有问题的数据做出相应的记录（另请参见附录 2 中的表格）。相反，如果特定变量只有非常少的病例存在缺失数据，则比较适合在分析中排除存在缺失数据的病例。一定要谨记，在应用该项技巧时，如果许多病例在不同变量中缺失少量值，则被排除在分析之外的病例总数可能会非常大。已经开发了多重填补等技巧，利用数据集中其他病例的信息给存在缺失数据的变量填充或分配数据值（7）。本手册中并没有描述这些技巧，不过可以在《结核病现患率调查手册》的第 16 章中找到更多信息（8）。

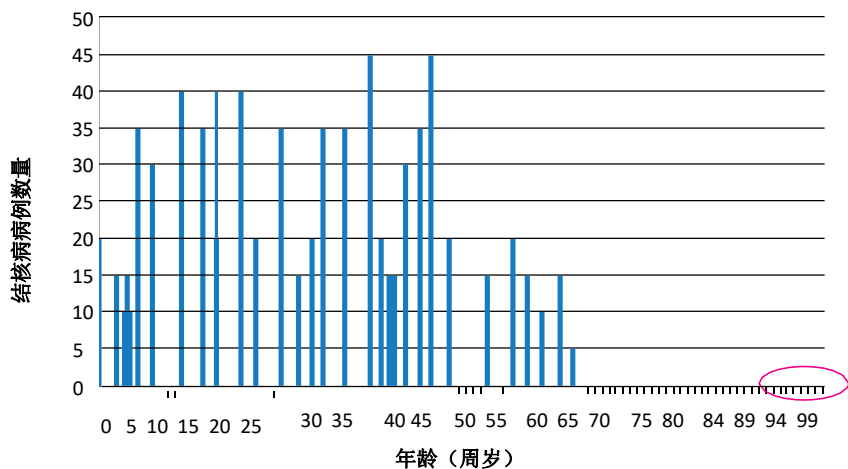
发现异常值

解决缺失数据后，流行病学家务必要探讨极值数据，极值也被称为异常值。异常值指落在特定变量被认为合理的上限和下限之外的数据值。按照定义，异常值是可以影响分析结果的极值，因此需要接受调查。

一些软件程序具备“发现异常”功能，可以发现个体数据错误。在调查异常值方面，一种更为简单的方法是查找与其他数据不兼容的可疑数据值。这方面可以考查最小值和最大值（例如 108 岁的年龄应当引起怀疑）。通常，采用柱状图或类似数据直观化方法绘制数据分布图可以最轻松地找出异常值。

例如，我们可以考虑下面一个示例，流行病学家希望确定 1500 份结核病病例的平均年龄。查看了数据集中的连续变量“年龄”后（图 3），流行病学家认为有两个数据值看起来像异常值：一份病例的年龄是 95 岁，另一份病例的年龄是 99 岁。接下来最高的记录年龄是 67 岁。

图 3
某个区县按年龄（周岁）划分的结核病病例数量



如果流行病学家已经发现了异常值，则建议流行病学家首先检查原始数据源，确定和更正数据值中的任何错误。在上述病例，流行病学家检查了结核病治疗卡（原始数据源），结果发现年龄为 99 岁的病例实际上是 9 岁。流行病学家据此在数据集中将 99 改成了 9。对于第二份病例（年龄 95 岁），无法找到治疗卡，因此流行病学家未能调整年龄数据。在这种情况下，流行病学家在分析中处理异常值的可能方法包括：将异常值编码为缺失、重新将年龄（周岁）（连续变量）编码为年龄段（类别变量）以最大限度地减少异常值的影响、将存在异常值的病例从分析中排除等。

剔除数据集中的重复项

对于病例相关数据，一个常见错误来源是存在重复的病例记录（例如监测系统对一名患者的相同信息捕获一次以上）。重复记录会造成问题，因为它们会人为地导致记录和报告的结核病病例数量的波动。这会在地方和国家层面上形成不准确的结核病负担数字。重复病例的出现可能会源于各种各样的原因。数据录入错误（例如拼写错误）是出现重复的最基本原因。另外，如果患者在不同机构间流动，则可能会被多次报告给国家结核病防治规划 – 文字框 1（9）。

为了确保数据的准确性，流行病学家务必要检查和解决病例相关监测系统重复病例。发现重复病例的一种简单方法是对系统内的唯一识别号运作频次检验；唯一编号通常是区县编号（例如区县代码/年份/区县序列号）或国内个人识别号（例如美国社会保障号和英国国民健康服务号码）。

确定所有重复后，流行病学家应当将重复的记录标记为重复，并在分析中排除重复病例（10）。在表 2 所示的示例中，区县编号为 546/2009/23 的第一份病例（Bobby Smith）被标记为重复，没有用于分析过程。

表 2
确定数据集中的重复编号

序列号	区县编号	姓名	出生日期	性别
1	546/2009/23	Bobby Smith	12/12/1973	M
2	546/2009/23	Bob Smith	12/12/1973	M
3	145/2009/10	Sue Mary	3/12/1999	F
4	897/2009/43	Tanya Jones	10/1/1956	F
5	234/2009/88	Melinda Frank	9/8/1955	F
6	546/2008/23	Bill Smith	11/9/1979	M
7	009/2009/01	Irwin Law	5/9/1983	M
8	115/2009/67	Burt King	4/10/1975	M
9	768/2009/99	Jason Kirk	5/6/1965	M

给变量重新编码

清理好数据后，流行病学家应当考虑给变量重新编码。监测系统收集的数据的格式结构可能会让流行病学家很难开展预期分析（例如年龄被作为连续变量捕获，而分析工作要年龄是类别变量）。在这类情况下，流行病学家可以操作现有变量或将现有变量重新编码为新的变量，以便更好地适应分析需求和所需统计计算。

文字框 1**个案研究：巴西预防病例重复**

在二十一世纪头十年的中期，巴西国家结核病防治规划致力于改善国家结核病通报系统。除了扩展监测系统的覆盖面之外，国家结核病防治规划还开发了流程，用于系统性地验证数据和链接数据库中已有的记录。这些流程非常重要，因为数据库中的记录由于各种原因出现了重复和分类不当的现象（例如复发病例，病例被一个以上的卫生保健机构通报，数据录入错误等）。在 2008 年，国家结核病防治规划采用了可能的记录链接，发现了数据库中的重复记录并更正了错误数据。这可以通过一系列连续步骤来完成。首先，国家结核病防治规划工作人员根据性别和姓名以电子方式对记录进行匹配，然后使用变量手动比较记录，例如患者的全名、患者母亲的全名、出生日期、通报日期和患者地址。最后，该流程发现检查的所有记录中几乎有四分之一的记录属于具有多重记录的患者（例如转出、复治和复发）；6.6%属于重复记录。清除这些记录极大地改变了国家结核病统计数据。从 2001 年到 2007 年，新病例的通报率从 2.3% 上升到了 6.3%，治愈病例的比例从 1.5% 上升到了 4.9%。这些结果展示了建立机制对于防止病例相关结核病监测系统出现重复记录的重要性。

资料来源： Bierrenbach AL et al. Duplicates and misclassification of tuberculosis notification records in Brazil, 2001–2007. *International Journal of Tuberculosis and lung Disease*. 2010;14(5):593–599.

在给变量重新编码时，流行病学家可以通过很多方式修改新变量（表 3）。例如，重新编码可能涉及到重新分配数据值，例如将数据值从字母改为数字（例如将男性和女性输入数据改为相应的数字 1 和 2）。它还涉及到将原始变量的范围进行组合或缩减，从而创建新变量。例如，创建“年龄段”变量时可以将 0–25 岁患者分配到类别 1，再将 26–50 岁患者分配到类别 2，最后将 50 岁以上的患者分配到类别 3。当原始类别中的答案数量太少无法进行统计分析时，创建这些广泛范围尤其有用。还可以给一个以上的变量重新编码，以此创建另一个变量。例如，身体质量指数（BMI）变量通过体重（公斤）/身高（米）计算²—可以源自两个独立的变量，即身高和体重。

表 3
在病例相关数据子集中给变量重新编码的示例

编号	性别 (原始, 文本变量)	性别 (重新编码过, 数字变量) 男性= 1 女性= 2	年龄 (岁) (原始, 连续变量)	年龄段 (重新编码过, 类别变量) 0-25 岁=1 26-50 岁=2 >50 岁=3	身高 (米) (原始, 连续变量)	体重 (公斤) (原始, 连续变量)	BMI (重新编码过: 身体质量指数= kg/m ²)
001	M	1	10	1	1.35	30	16.5
002	M	1	25	1	1.81	55	16.8
003	F	2	37	2	1.73	60	20.0
004	F	2	21	1	1.79	58	18.1
005	F	2	62	3	1.65	51	18.7
006	M	1	55	3	1.75	56	18.3
007	M	1	26	2	1.85	73	21.3

将数据集链接起来

将变量调整为正确格式后，流行病学家应当检查数据集，确保开展分析所需的全部数据要素都包括在内。在一些病例中，开展特定分析所需的变量可能并没有包括在常规结核病监测数据中，而是可能包括在其他数据源中（例如实验室和/或人口动态登记数据）。可以将来自病例相关监测数据的结核病病例与其他数据源链接起来，这样可以更为全面地了解疫情流行情况，从而极大地加强分析工作（11）。流行病学家可以使用记录链接技巧，帮助确定不同数据集中的同一个人，通常可以使用个人识别码来做这一点，例如社会保障号。

将多个不同数据库合并为一个中心数据库时，流行病学家使用记录链接准确地确定同一个人是否存在两项或多项记录。流行病学家还可以视觉比较各项记录，通过这种方式手动完成记录链接。不过，这需要很多劳动力，因此对于大数据集不太有效（9）。数据链接流程通常通过特殊软件（例如 Link Plus）来完成，如确定性记录链接或概率性记录链接。在确定性记录链接中，通过找出在一个或多个字段中包含相同数据的所有记录，可以发现重复病例，理想情况下是采用个人识别号。概率性链接是基于一个分数，该分数记录与相同个人有关的概率。

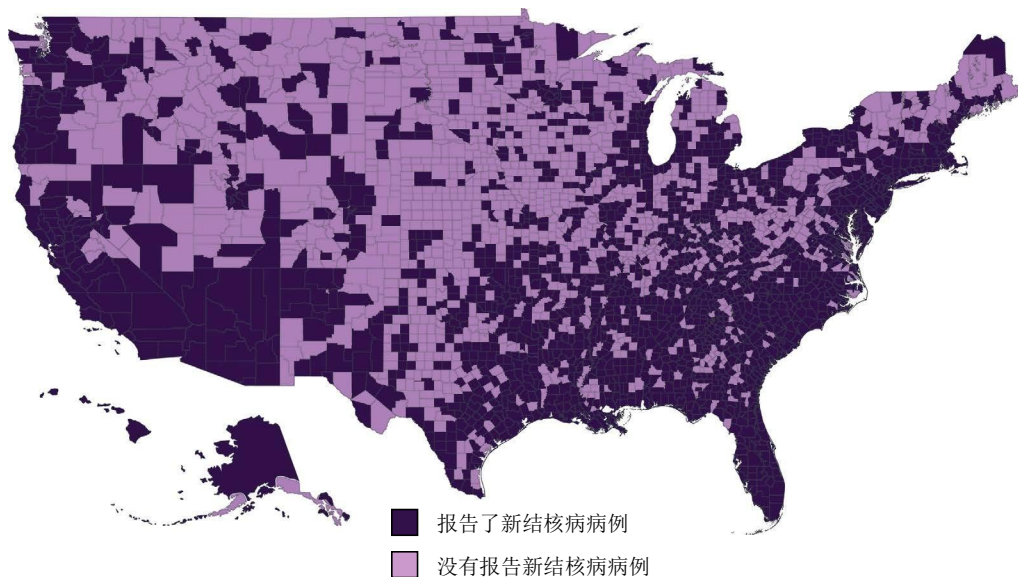
采取必要步骤清理和准备完病例相关数据后，流行病学家在开展分析之前必须采取的最后一个步骤是最终确定（或“冻结”）数据集。冻结最终分析数据集意味着流行病学家不应再对数据集做任何更改（例如更新或删除），因此所有分析应当源自一组相同的数据，并在病例数量和数据质量方面具有一致性。在被冻结数据集内对数据的任何操作都将作为分析流程的一部分，例如病例选择、数据子集设定、重新编码等。一旦手头上具备了被冻结的数据库，流行病学家可以进行下一个步骤，即数据分析。

2.4 数据分析：开展和解释描述性分析

描述性分析（通常也被称为描述性统计学）是以统计方式分析病例相关数据的第一步。开展常规描述性分析可以持续性地、及时地查看病例相关数据。流行病学家使用描述性分析来描述和呈现以人、地点和时间划分的数据。平均值、频次和发病率等计算都属于描述性分析的示例。描述性分析的结果通常采用表格、图形、图表和分布图等形式呈现（图 4）。

图 4

分布图是利用病例相关数据展示地域描述性分析结果的一种有用方式



（资料来源：原始发布网址 <http://www.cdc.gov/mmwr/pdf/wk/mm6211.pdf>）

一元和二元分析

描述性分析是开展分析过程中重要的第一步，这是因为描述性分析将数据转换为有意义的信息。病例相关监测系统生成的一系列原始数据很难解释，尤其是每份结核病病例捕获了很多信息时（例如许多变量）。流行病学家采用描述性分析总结一系列结核病病例，以此协助数据的解释。

为了了解描述性分析如何以及为什么对流行病学家有用，可以考虑 X 区县结核病病例数据的理论示例，该区 2013 年报告了 100 份结核病病例。表 4 显示了这 100 份结核病病例中 10 份病例的特征的系列清单。

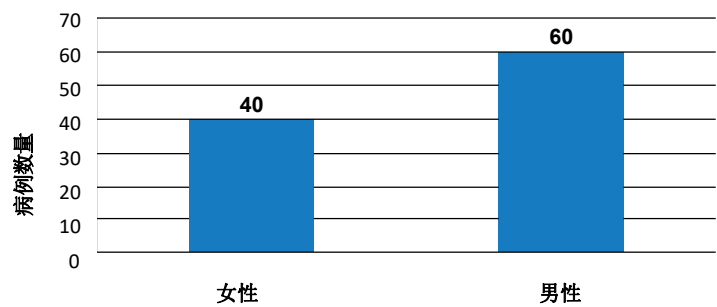
表 4
X 区县结核病病例数据的示例

年份	国家	性别	疾病部位	艾滋病病毒	职业
2013	X	男性	肺部	阳性	矿工
2013	X	男性	肺部	阴性	矿工
2013	X	女性	肺部	阴性	护士
2013	X	男性	肺外	阴性	农民
2013	X	女性	肺部	阳性	农民
2013	X	男性	肺外	阴性	矿工
2013	X	男性	肺部	阳性	矿工
2013	X	女性	肺外	阴性	店主
2013	X	女性	肺部	阴性	学生
2013	X	男性	肺部	阴性	矿工

如果表 4 包括全部 100 份结核病病例，则很难看出 100 份结核病病例的所有特征，也很难得出有关结核病流行情况的有意义的结论。通过这些原始数据很难回答相关的问题，例如该区哪种类型的结核病最常见、患有结核病的男性多还是女性多等。描述性分析有助于总结数据，并将数据转化为有用的信息。了解到这一点，X 区县的结核病官员采用了结核病监测数据的系列清单，并且利用描述性分析创建了多个图表，帮助回答有关当地结核病流行情况的基本问题。

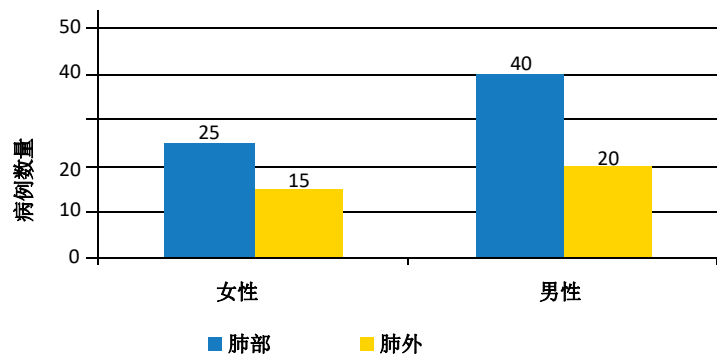
区县结核病官员总结了系列清单数据，结果发现在 100 份病例中，男性结核病病例多于女性（图 5）。这类分析只采用单一的变量，属于一元分析。一元分析是最简单的描述性分析方法，可以用来描述频次分布、比例和集中趋势衡量手段（例如均值、中位值和模式）。一元分析结果很容易呈现和解释，因此经常在报告中采用，例如年度结核病病例监测报告，这些报告被分发给广泛的目标受众。

图 5
X 区县按性别划分的结核病病例，2013 年



病例相关数据可供使用时，流行病学家可以更进一步地开展分析。在这种情况下，区县结核病官员查阅了病例相关数据。官员决定调查区县中男性和女性结核病病例之间结核病类型的区别。如果只有聚合数据应用，则无法做到这一点。

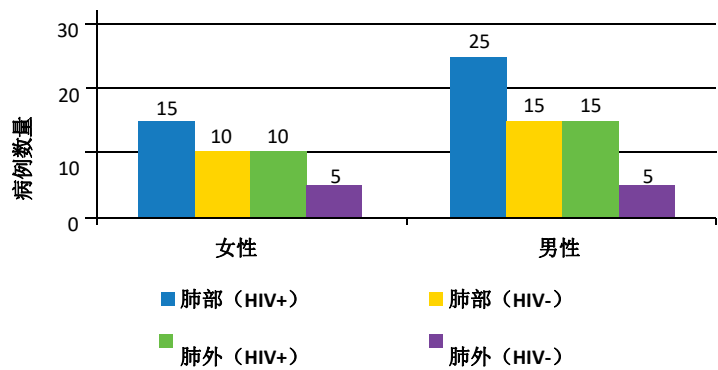
图 6
X 区县按性别和疾病部位划分的结核病病例，2013 年



与上一份示例相同，图 6 显示结核病病例中男性（n=60）多于女性（n=40）。它还显示肺部病例（n=65）多于肺外病例（n=35）。最为重要的是，它比较了相同性别之间和不同性别之间的肺部病例数量和肺外病例数量。通过这些信息，区县结核病官员了解到在男性和女性中肺结核都比肺外结核更为常见。患有肺结核的男性比例（ $40/60=67\%$ ）和女性比例（ $25/40=63\%$ ）接近。这是一种二元分析，因为它涉及到两个变量的分析（例如性别和结核病类型），并且评估了两个变量之间的关系。

利用病例相关数据，流行病学家还可以同时分析两个以上的变量。了解到这一点，区县结核病官员决定开展稍为更复杂一点的分析工作。官员想要了解艾滋病病毒感染状况对结核病类型的影响，同时会考虑到性别。因为官员具备病例相关数据，因此可以根据结核病类型（肺部或肺外）开展艾滋病病毒感染状况（阳性或阴性）的二元分析，并按性别给结果分层（图 7）。

图 7
X 区县不同类型结核病病例的艾滋病病毒感染状况，按性别分层，2013 年



区县结核病官员在结论中指出，不论性别如何，绝大多数结核病患者都是 HIV 阳性（女性：25/40=63%；男性：40/60=67%）。很显然，在男性和女性中，最大数量的结核病病例是肺结核和 HIV 阳性（女性：15/40=38%；男性：25/60=42%）。

发病率和趋势

流行病学家还利用描述性分析常规监测结核病发病率和病例数目在不同时间段的趋势，以及确定变化率，揭示数据和其他数据差异。这类分析通常按季度和/或年度开展。评估观察结果，了解它们是否反映了趋势的实际变化，或者是否属于数据输入或管理错误的结果。可以采用聚合数据开展这些类型的分析（参见第 1 章），不过它们对于病例相关数据的常规分析同样重要。

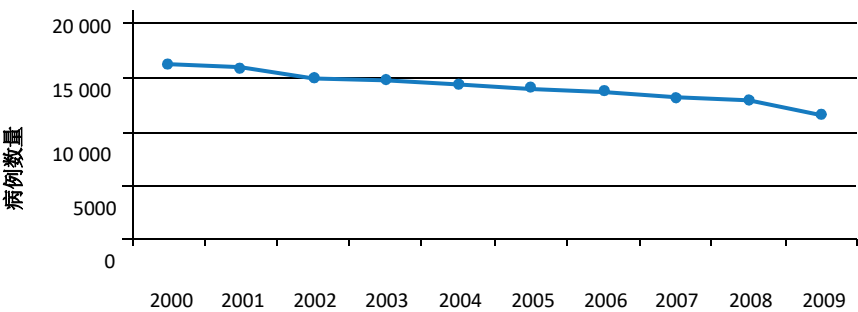
为了了解如何以及为什么将病例相关数据用于趋势分析，可以考虑美国的结核病监测工作，其中监测年度结核病趋势是结核病监测的一项重要内容。每年，公共卫生官员都会发布一份有关国内结核病病例的监测报告。下面是公共卫生官员在 2000 年到 2009 年看到的国内报告的结核病病例（表 5）。

表 5
美国报告的结核病病例数量，2000–2009 年

年份	报告的结核病病例数量
2000	16 209
2001	15 945
2002	15 056
2003	14 836
2004	14 499
2005	14 064
2006	13 734
2007	13 280
2008	12 906
2009	11 545

公共卫生官员利用这些数据回答有关不同时间段通报的结核病病例数量的趋势的各种问题。由于光是查看表格很难了解数据的更多内容，公共卫生官员按年度创建了结核病病例曲线图。这是视觉考查结核病病例趋势的一种简单方法，通常也是更为轻松的一种方法（图 8）。

图 8
美国按年度通报的结核病病例数量，2000–2009 年



表格和曲线图（表 5 和图 8）都显示美国的结核病通报病例数量自 2000 年以来一直在稳步下降。不过，通过曲线图考查了结核病趋势之后，公共卫生官员怀疑结核病通报病例从 2008 年到 2009 年出现了不寻常的急剧减少。他们决定开展进一步的分析，以便更好地考查这种减少，他们最终考查了发病率。

流行病学家通常在病例相关数据的描述性分析中计算和报告发病率。在医学领域内，发病率指某种疾病在特定时间段内在特定人群中出现的新病例数量。发病率可以比报告的病例数量提供更多的信息，因为发病率考虑到了人口规模的差异（参见[文字框 2](#)），该项因素会影响到报告病例的总数。计算发病率要求具备相应的人口或分母数据。例如，2012 年儿童结核病发病率是 2012 年生活在特定国家内的 15 岁及以下所有儿童中的结核病新病例数量。

美国卫生官员计算了年度结核病发病率，以此考查结核病病例从 2008 年到 2009 年的异常下降（[表 6](#)）。通过该项发病率分析，公共卫生官员有能力对不同时间段开展了基于人群的比较，结果发现从 2008 年到 2009 年的发病率下降较基于既往年度趋势得出的预期要大。

表 6
美国结核病通报病例发病率，2000–2009 年

年份	结核病病例发病率 *每 10 万人口
2000	5.8
2001	5.6
2002	5.2
2003	5.1
2004	4.9
2005	4.8
2006	4.6
2007	4.4
2008	4.2
2009	3.8

*采用每年的人口数据，计算了每 10 万人口中的发病率（一年的病例数量/当年的人口）。

文字框 2 利用发病率比较不同人群中的疾病负担

A 市

$$\frac{2012 \text{ 年 } 105 \text{ 份结核病病例} \times 100\,000}{2012 \text{ 年人口 } 1\,651\,894} = \frac{10\,500\,000}{1\,651\,894} = \text{每 } 10 \text{ 万人口中 } 6.4 \text{ 例}$$

B 市

$$\frac{2012 \text{ 年 } 48 \text{ 份结核病病例} \times 100\,000}{2012 \text{ 年人口 } 1\,680\,102} = \frac{4\,800\,000}{1\,680\,102} = \text{每 } 10 \text{ 万人口中 } 2.9 \text{ 例}$$

C 市

$$\frac{2012 \text{ 年 } 81 \text{ 份结核病病例} \times 100\,000}{2012 \text{ 年人口 } 553\,523} = \frac{81\,000\,000}{553\,523} = \text{每 } 10 \text{ 万人口中 } 14.6 \text{ 例}$$

根据同一个调整系数（每 10 万人口）计算发病率，即使三个城市有着不同的结核病病例数目和不同的人口数量，但还是可以比较三个城市的发病率。通过该项城市比较，我们可以得出下面的结论：C 市的结核病发病率较其他两个城市高。换句话说，即使 A 市有更多病例，但 C 市人口受到疾病的更多影响。

进一步考查数据所利用的其他类型的信息

对于某些类型的数据，例如人口统计学或行为学风险因素，可能无法获得准确的人口分母数据。例如，计算吸烟人口中的结核病发病率需要获得相关证据，了解目标地区有多少人在吸烟。不过，可能并不具备这类信息。在缺乏分母数据的情况下，可以采用其他方法分析数据趋势。像图 8 这样的曲线图显示了不同时间段的结核病病例数量，流行病学家可以利用这些数据直观地发现趋势是否有任何变化。

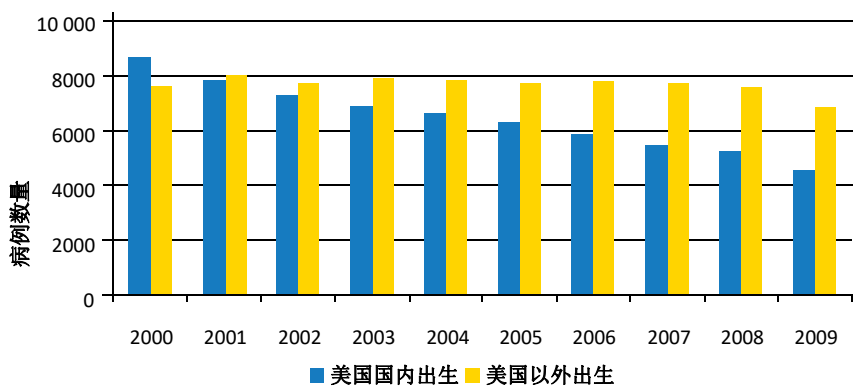
利用病例相关数据，可以给曲线图分层，显示一个以上变量的趋势。例如，美国公共卫生官员考查了各种人口统计学和风险因素，进一步了解到结核病负担在 2008 年到 2009 年间的下降，包括结核病患者的出生地（例如在美国国内出生或在美国以外出生）。

他们根据结核病患者的出生地点，制作了一份曲线图，用于考查结核病负担的趋势（图 9）。

考查了 2000 年到 2009 年美国按年龄划分的结核病病例数量，结果发现在美国国内出生的结核病病例数量下降速度要快于在美国以外出生的结核病病例数量（12）。而且，该曲线图还显示，从 2001 年到 2009 年，在美国以外出生的人口报告中报告的结核病病例数量要超过在美国国内出生的病例报告数量。鉴于 2009 年的结核病报告病例下降情况，很显然两类人群中都存在下降趋势（文字框 3）。

总之，描述性分析有助于描述地方人口和全国人口中的结核病流行病学，以及监测不同时间段的趋势。在 x 区县的理论示例中，结核病官员能够调查结核病患者的特征，并相应地调整区县结核病防治规划。在美国的示例中，开展了持续的病例相关数据常规分析，这有助于公共卫生官员尽早发现不断变化的趋势，了解趋势为什么有了变化。虽然国家结核病防治规划开展的绝大多数分析都类似于上文描述的分析，但利用病例相关数据还可以开展其他更深入的分析，下一节会对这方面作简要的介绍。

图 9
美国国内出生的结核病病例数量和美国以外出生的结核病病例数量



文字框 3 个案研究：美国

每年，各州结核病防治规划会将结核病确诊病例的信息发送给美国疾病预防控制中心（CDC），这些信息最终会被整合到国家结核病监测系统（NTSS）。在 2010 年年初，在分析了 2009 年初步病例数目后，美国疾病预防控制中心发现病例数量和计算出的发病率有了非预期下降。收到的初步数据指出总体发病率下降了 11.4%，从 2008 年的 4.2/10 万下降到 2009 年的 3.8/10 万。这是结核病监测自 1953 年覆盖全国范围以来记录的单年度最高下降，也是记录的最低发病率。对初步数据开展了广泛分析后，结果发现几乎所有亚群都出现了结核病发病率的下降。美国以外出生的结核病病例的发病率下降了 9.0%，美国国内出生的结核病病例的发病率下降了 15.8%。所有种族和民族人口中都出现了下降，下降幅度从非西班牙裔亚洲人中的 9.0% 到非西班牙裔白人中的 15.2% 不等。甚至，培阳肺结核病例发病率下降了 13.6%，而培阴肺结核病例发病率下降了 17.5%。发病率的非预期下降可能是因为结核病防治规划的更有效实施，从而使得更少的结核病病例被诊断出来，标志着国家在消除结核病方面迈出了一大步；不过，这种下降也可能是因为人口流动所导致，其他可能的原因还包括漏报或漏诊。

国家监测系统在前一年已经经历了显著的变化。许多州已经不再通过美国疾病预防控制中心开展的系统以电子方式报告结核病病例数据，而是开始采用他们自己的网络报告系统，并且已经改编了确诊结核病修订报告表，该表是美国全国范围内采用的标准化报表。美国在 2009 年引入了确诊结核病修订报告。一些新的网络报告系统缺乏美国疾病预防控制中心旧报告系统中已存在的所有验证规程。由于许多州的经济不景气，有些地方据报告裁减了结核病防治人员的人数。由于这些变化，美国疾病预防控制中心（以及州结核病防治规划）对州和地方层面上的病例发现和实践开展了专门的广泛调查。^{a,b} 调查人员将软件系统变化排除在结核病发病率下降的原因之外。采用不同类型的网络报告系统的各州都出现了下降。美国疾病预防控制中心在全国各地资助的临床分枝杆菌实验室报告从 2008 年到 2009 年少收了 5.9% 的结核病检测临床样本。

a. Centers for Disease Control and Prevention. Decrease in Reported Tuberculosis Cases — United States, 2009. *Morbidity and Mortality Weekly Report*. 2010;59:289–294.

b. Reported Tuberculosis in the United States, 2009. Atlanta: United States Department of Health and Human Services; 2010.

文字框 3 - 续

美国国家结核病控制人员协会调查了出现监测、诊断实验室规程和人员配置变化的全部 59 个报告地区（包括州和区县）。93% 的结核病防治规划管理人员接受了调查，没有人指出这些变化导致了病例的非预期下降。事实上，经历了人员裁减并报告了至少 100 例结核病病例的各州出现了较小的病例数目下降，而没有经历人员裁减的类似州则出现了较大的病例数目下降。在全国十个高负担区县中，没有发现任何漏报的病例。美国疾病预防控制中心调查人员在宾夕法尼亚州和格鲁吉亚州开展了详细的病例发现研究。

该项研究没有发现结核病病例的大量漏报或漏诊。^c 对过去十年的每月病例数目开展了统计分析，结果发现美国国内出生病例和美国以外出生病例按治疗开始日期划分的结核病观察病例数与结核病预期病例数之间有显著差异。美国以外出生病例从 2007 年 10 月开始经历了病例数目的大幅下降，这些下降与美国人口普查局报告的美国以外出生人口中的趋势保持了一致，也与美国国土安全局在非法移民中观察到的趋势相一致。^d 作者得出结论：观察到的 2009 年结核病发病率下降反映了美国结核病诊断病例数量的实际下降。

c. Centers for Disease Control and Prevention. Assessment of declines in reported tuberculosis cases — Georgia and Pennsylvania, 2009. *Morbidity and Mortality Weekly Report*. 2011;60:338–342.

d. Winston CA et al. Unexpected decline in tuberculosis cases coincident with economic recession — United States, 2009. *BMC Public Health*. 2011;11:846–857.

2.5 数据分析：开展和解释更为复杂的分析

虽然病例相关数据比聚合数据更有利于开展较为深入的描述性分析，不过病例相关数据的最大优势是能够开展更为复杂的分析，例如分层多元分析（例如回归）。流行病学家利用这些更深入的分析对两个或多个变量之间的关系进行统计考查。具体而言，这类分析可以确立各变量（例如风险因素和疾病转归）之间的关联，从而提供更强有力的证据，促进成功的结核病防治规划和政策制定。

分层分析可以在分析中增加第三个变量。流行病学家利用病例相关数据的分层分析，帮助确定第三个变量（例如混杂变量）是否会影响研究中其他两个变量之间的关系。混杂变

量与其他两个变量相关，会掩盖其他两个变量之间的真正关联。例如，某项研究旨在调查肺癌与饮酒之间的关系，在该项研究中，吸烟如果与饮酒相关则属于一个混杂变量，同时也是肺癌的一项独立风险因素。分层分析能够帮助梳理出真正的关系。[附录 3](#) 提供了分层分析的一个示例。如果考虑到多个变量的效果且这些变量可能会潜在地影响到被调查的两个主要变量之间的关系，则需要开展多元分析。流行病学家可以手动开展分层分析，而多元分析需要统计软件程序才能完成。

流行病学家在多元分析中最常使用的统计方法是逻辑回归、线性回归和生存分析。下文会提供这些分析的简要解释和示例。

流行病学家利用逻辑回归确定与分类回归相关联的独立变量，分类回归通常是二元转归（例如两个类别）或多元转归（例如超过两个类别）。二元转归的一个示例可能是治疗成功（1=是或 2=否），多元转归的一个示例可能是治疗转归（1 = 治愈或完成治疗，2 = 死亡、失访或失败，3 = 转出）。独立变量可能是分类变量（例如肺部、肺外或两者）、序级变量（例如不佳、一般或良好）或连续变量（例如年龄（周岁））。

独立变量的示例可能包括年龄、性别、酗酒或直接观察治疗类型。在逻辑回归中，优势比关联测量手段，用于量化独立变量与转归之间的作用。例如，多变量逻辑回归分析的结果有助于确定最容易出现不利转归的患者的特征，以及应当被优先考虑改善治疗转归的患者的特征。[附录 3](#) 提供了采用逻辑回归的多元分析的一个示例。

线性回归是一种类型的多元分析，被流行病学家用来界定独立变量与连续且正常分布的转归之间的关联（例如绘制所有数据值的分布图后，这些转归会以对称形式分布，就像钟形曲线，绝大多数数据值位于中位值/均值周围）。例如，线性回归分析中的转归可能是以厘米为单位的身高，模型中的独立变量可能是年龄、性别、民族或营养状况。在结核病流行病学中，线性回归通常被用来调查国家（而不是患者）之间的关联，例如国家层面指标（例如艾滋病病毒现患率或国内生产总值）与结核病发病率趋势之间的关联（13）。

流行病学家不仅利用生存分析来评估谁面临着某项转归（例如死亡）的更大风险，而且还评估这些转归的时机，而且在分析特定转归的风险因素时会考虑这一点。对于生存分析，Kaplan-Meier 曲线被用来评估事件时间或转归时间，Cox 比例危险率模型被用来确定转归的风险因素。一项重要的假设是转归的风险（危险比）在整个研究期间都不会变化。

总之，病例相关数据提供各种各样的监测数据深入分析机会。对开展这类分析感兴趣的人可以咨询经验更丰富的流行病学家或统计学家，同时参考其他文献，例如《基础流行病学》（3）。

2.6 传播分析结果

分享结核病监测数据分析结果可以让规划人员、政策制定者、公众和研究人员了解当地结核病流行情况、变化趋势和规划绩效。有效传播分析结果有助于国家结核病防治规划更好地认识到地方和国家层面上的结核病负担，分享在实现目标值方面取得的进展情况、激发对结核病服务的需求、加强对结核病政策变革的支持、提醒目标受众哪些行为会影响结核病预防和治疗（15）。国家结核病防治规划利用各种策略将结核病分析结果传播给不同的目标受众，下面提供了不同策略的示例。

规划人员：向结核病防治规划人员传播常规分析的结果时，可以分发总结病例相关结核病数据的年度报告。年度报告通常是在收集完数据后编制而成，可以给全国结核病流行情况提供全面的、最新的评估。国家结核病防治规划一般会根据国内重点工作，确定应被包括在报告中的关键指标，分析要素通常在各年度之间相同，以便国家结核病防治规划可以监测不同时间段实施结核病控制的进展情况和面临的挑战。美国^c和英国^d在线提供了每年利用监测数据开展和报告的常规分析的示例。

c. 参见：<http://www.cdc.gov/tb/statistics/reports/2012/default.htm>

d. 参见：http://www.hpa.org.uk/webc/HPAwebFile/HPAweb_C/1317139689583

为了更频繁地监测结核病控制工作，可以制作相应的简报，常规地与地方结核病防治规划人员进行有效的交流。例如，美国的国家结核病指标项目给地方层面提供相关报告，使得地方层面可以采用标准化指标追踪自身在实现国家目标方面取得的进展，这些标准化指标基于国家结核病控制重点工作开发而成。报告由地方层面提交的病例相关监测数据生成。数据每周更新一次，能够给地方层面提供最新数据，有利于开展规划的督导与评估工作。例如，国家结核病防治规划的目标之一是到 2015 年将新诊断结核病患者 12 个月内完成治疗的比例提高到 93.0%（图 10）。每个地区都可以利用国家结核病指标项目系统来监测当地进展情况，如果地方层面没有实现该项目目标，还可以帮助确定相关的障碍。

政策制定者和公众：向公众（包括政策制定者和媒体代表）传播监测分析结果时，情况说明书可以有效地描述国内结核病疫情的背景和现状。情况说明书是一种简报，篇幅通常不会超过两页，目的是面向一般目标受众。例如，情况说明书可能会描述结核病的原因、结核病的负担和分许（例如基于监测数据分析），针对要采取的行动提供必要建议。当具备更多数据时，轻松地进行更新和替换情况说明书。世界卫生组织的网站在线提供了结核病情况说明书的示例。^e

研究人员：给学术受众传播分析结果的最常用方法是在科学杂志上发表文章。绝大多数科学论文都遵循类似的格式，具体包括摘要、简介或背景、内容和方法、结果、讨论和结论。呈现监测数据的分析结果时，可以向从一般性内容开始，然后转到具体内容（例如先介绍国家结核病发病率，然后介绍不同人群的发病率）。还务必要采用简单的写作风格，使用简短的句子和过渡形式，从而在目标受众阅读过程中提供一个“路线图”。读者可以在《现场流行病学》（16）中找到有关撰写和呈现科学论文的更多信息。

e. <http://www.who.int/tb/publications/factsheets/en/>

图 10
美国国家结核病指标项目完成治疗示例报告

完成治疗

国家结核病指标项目

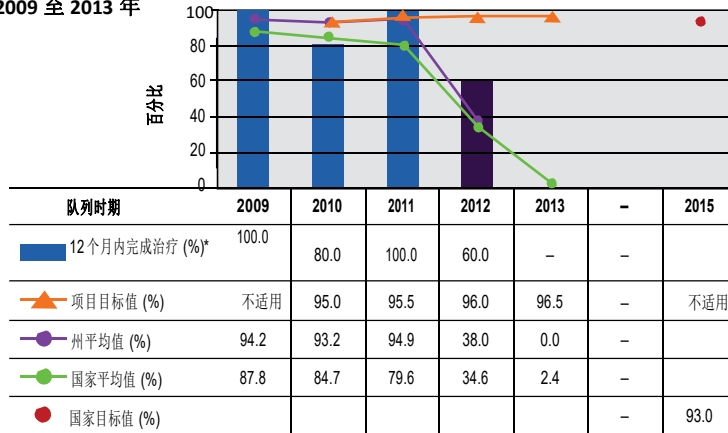
数据更新日期：03/18/2013

国家目标：

对于治疗疗程为 12 个月或不到 12 个月的新诊断结核病患者，到 2015 年将 12 个月内完成治疗的患者的比例提高到 93.0%。

某区县符合治疗条件的患者在 12 个月内完成治疗

2009 至 2013 年



* 图表上的特殊颜色代表着数据不完整。

队列时期	2009	2010	2011	2012	2013
符合治疗条件的患者总人数 (N)	4	5	1	5	0
完成治疗的患者人数 (累计) (n)	4	5	1	3	0
12 个月内完成治疗的患者人数 (n)	4	4	12	3	0
没有完成治疗的患者人数	0	0	0	2	0

目标：对于治疗疗程为 12 个月或不到 12 个月的新诊断结核病患者，提高 12 个月内完成治疗的患者的比例。

指标：在治疗疗程为 12 个月或不到 12 个月的新诊断结核病患者中，12 个月（366 天）内完成治疗的患者的百分比。

队列：推荐治疗疗程为 12 个月或不到 12 个月、诊断时存活、开始接受一种或多种药物的治疗、报告队列时期内计算在内的所有结核病患者。排除存在利福平耐药结核病或结核性脑膜炎结核病的患者以及患有播散性结核病的 14 岁或更年幼儿童。播散性结核病被定义为粟粒状疾病部位、胸片显示粟粒状结核病证据或血培养阳性。排除死亡的患者。对于 2009 年和以后的队列，还排除治疗期间搬离美国的患者。

数据源：确诊结核病修订报告字段：报告报告、统计日期、出生日期、结核病诊断状况、结核病患病部位、人体组织和其他体液的培养结果、初步胸片和其他胸部影像结果、治疗开始日期、初步药物治疗方案、初步药敏检验结果、转出、治疗停止日期、治疗停止原因。

计算公式：[建议在 12 个月内完成治疗的患者在 366 天或不到 366 天内完成治疗的人数 / 队列人数] x 100

2.7 结论

总而言之，高质量病例相关监测数据的分析对于国家和地方结核病防治规划都很重要，有助于了解结核病流行病学及其对规划的影响。本章提供的示例展示了病例相关监测数据如何被用来监测结核病疫情趋势、确定变化趋势的可能原因、确定和优先考虑特定亚群、更好地给结核病防治规划提供参考依据，从而更有效地将精力和资源引导到结核病减少或消除工作中。

1. Electronic recording and reporting for tuberculosis care and control. Geneva: World Health Organization; 2011 (WHO/HTM/TB/2011.22). 网址: http://whqlibdoc.who.int/publications/2012/9789241564465_eng.pdf
2. Antoine D and Abubakar I. Surveillance for tuberculosis. In Infectious Disease Surveillance, 2nd edition. Eds M'ikanatha NM, Lynfield R, Van Beneden CA, and de Valk H. John Wiley and Sons, Ltd; 2013 (pp234–247).
3. Basic Epidemiology. Second Edition. Geneva: World Health Organization; 2006. 网址: http://whqlibdoc.who.int/publications/2006/9241547073_eng.pdf
4. Definitions and reporting framework for tuberculosis (2013 revision). Geneva: World Health Organization; 2013 (WHO/HTM/TB/2013.2). 网址: http://apps.who.int/iris/bitstream/10665/79199/1/9789241505345_eng.pdf
5. Checklist of standards and benchmarks for case-based TB surveillance data. Geneva: World Health Organization; 2012. 网址: <http://www.who.int/tb/publications/standardsandbenchmarks/en/>
6. Altman DG. Practical statistics for medical research. London: Chapman and Hall; 1991 (pp122–151).
7. Aday L. Designing and Conducting Health Surveys: A Comprehensive Guide. San Francisco: Jossey-Bass Publishers; 1989 (pp305–321).
8. Tuberculosis prevalence surveys: a handbook. Geneva: World Health Organization; 2010 (WHO/HTM/TB/2010.17). 网址: http://www.who.int/tb/advisory_bodies/impact_measurement_taskforce/re-sources_documents/thelimebook/en/
9. Bierrenbach AL, de Oliveira GP, Codenotti S, Gomes AB, Stevens AP. Duplicates and misclassification of tuberculosis notification records in Brazil, 2001-2007. International Journal of Tuberculosis and Lung Disease. 2010;14(5):593–9.
10. Assessing tuberculosis under-reporting through inventory studies. Geneva: World Health Organization; 2012 (WHO/HTM/TB/2012.12). 网址: http://www.who.int/tb/publications/inventory_studies/en/
11. Olsen J, Saracci R, Trichopoulos D. Teaching Epidemiology; A Guide for Teachers in Epidemiology, Public Health and Clinical Medicine. Oxford: Oxford University Press; 2001 (pp. 455–466).
12. Reported Tuberculosis in the United States, 2011. Atlanta: U.S. Department of Health and Human Services (CDC); 2012.
13. Dye C. et al. Trends in tuberculosis incidence and their determinants in 134 countries. Bulletin of the World Health Organization. 2009;87(9):683–91.
14. Churchill R Elliott. Using Surveillance Information in Communications, Marketing and Advocacy. In Principles and Practice of Public Health Surveillance. 2nd edition. Eds Steven M Teutsch and R. Elliott Churchill. Oxford: Oxford University Press; 2001 (pp. 364–379).
15. Gregg MB. Communicating Epidemiologic Findings. In Field Epidemiology. 2nd edition. Ed. Michael B. Gregg. Oxford: Oxford University Press; 2002 (pp. 183–195).

目标：描述荷兰儿童结核病的趋势和特征，与荷兰成年人结核病的趋势进行比较。

具体目标：描述和比较儿童和成年人中结核病病例的特征。描述和比较儿童和成年人中结核病通报的趋势。

假设：儿童结核病发病率的变化与成年人结核病发病率的变化相同。

研究人群：2000 年至 2012 年荷兰通报的结核病患者。

定义：

- 儿童结核病：通报日期年龄为刚出生到 14 岁的患者。
- 成年人结核病：通报日期年龄至少满 15 岁的患者。

分析中要采用的变量：通报日期、出生日期、性别、出生国家，国外出生结核病患者自录入荷兰系统后的时间、卡介苗（BCG）疫苗接种状况（接种疤痕和/或疫苗接种本）、结核病患病部位（肺外和/或肺部）、与结核病患者的接触史、实验室诊断结果（涂片、培养）、艾滋病病毒感染状况、病例发现方法（被动、接触者调查、移民筛查）、治疗转归。

数据准备：

- 采用出生日期与通报日期之间的时间间隔计算的年龄。
- 通过链接到荷兰统计局的人口数据，将要包括在数据集中的儿童和成年人的规模。
- 特定变量出现缺失值的病例的百分比超过 5%时，将要包括在分析中和要呈现的缺失值。
- 分析软件：STATA、SAS、SPSS

保护数据保密性的规程：

从国家电子结核病登记簿登记委员会处获得相关许可，从而针对所述目标获得和分析国家电子结核病登记簿中的数据。国家电子结核病登记簿并不包含个人识别号，如姓名和地址，因此数据属于匿名性质，这意味着不需要获得患者的知情同意。

统计方法：

- 研究人群的描述（参见下文的表 A1）。利用卡方检验了解要检验的儿童和成年人结核病病例之间的比例差异。
- 计算儿童和成年人结核病病例每年的通报率。
- 通过线性回归描述 2000 年至 2012 年儿童和成年人结核病病例通报率的趋势。为了确定儿童和成年人之间的趋势是否存在差异，会检验年龄（儿童与成年人）与通报率之间的相互作用。如果存在显著的相互作用，儿童和成年人结核病病例的趋势线的斜率会有显著差异。
- 所有检验采用的显著性水平都是 0.05。

表 A1
研究人群的描述

	儿童结核病病例		成年人结核病病例		p 值
	n	%		%	
性别					
男性					
女性					
缺失					
出生国家					
荷兰					
外国					
缺失					
卡介苗疫苗接种的证据					
是					
否					
缺失					
结核病患病部位					
肺部					
仅肺外					
缺失					
与结核病患者的接触史					
是					
否					
缺失					
艾滋病病毒感染状况					
阳性					
阴性					
缺失/未知					
涂片结果					
阳性					
阴性					
缺失/为开展					

表 A1 - 续

培养结果					
阳性					
阴性					
缺失/为开展					
病例发现方法					
主诉					
接触者调查					
移民筛查					
缺失					
治疗转归					
治愈					
完成治疗					
死亡					
失败					
失访					
缺失					
总计					

图 A1
每年通报的儿童和成年人结核病病例数量

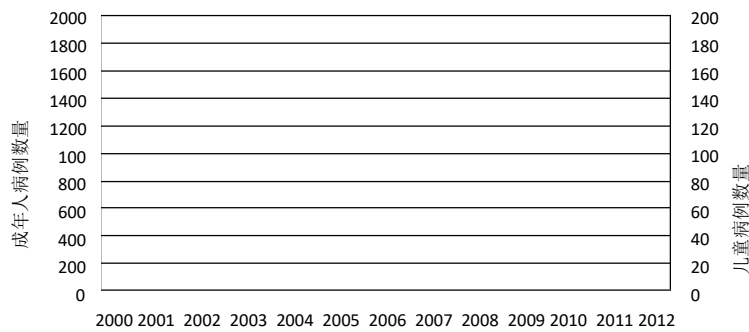
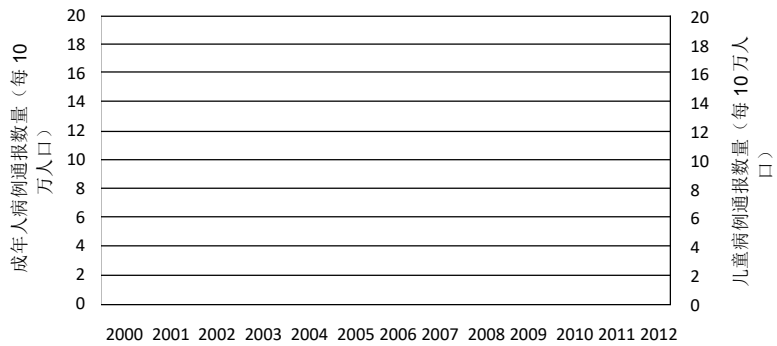


图 A2
每年儿童和成年人结核病病例通报率，线性回归预测趋势线



附录 3
用于评估失访因素的分层和多元分析的示例

总体而言，国家 A 在 2010 年的新涂阳肺结核患者治疗结果显示出现了 7% 的失访。让我们看看与治疗失访相关的一些患者特征。分析中包括了国家 A 在 2010 年开始接受治疗的 20 374 例新涂阳肺结核患者，其中治疗转归包括治疗成功和失访。

从表 A2 可以明显看出，成功治疗转归的绝大多数患者是男性（61.9%），不过治疗期间失访的最大比例的患者也是男性（74.3%）。成功治疗者中男性的调整前优势比是 $11\,832/7276=1.6$ ，失访者中男性的调整前优势比是 $948/328=2.9$ 。这意味着对于男性来说，失访的优势比要比女性失访的优势比高 1.8（ $2.9/1.6=1.8$ ）。在二元层面上（即与失访转归的关联不考虑任何其他变量）与失访显著相关的其他患者特征是无家可归、饮酒和居住在城镇地区。这意味着这些因素在统计学上与失访具有显著相关性，因为 95% 置信区间没有超过 1。

表 A2
国家 A 在 2010 年开始接受治疗的新涂阳肺结核患者按治疗转归划分的患者特征 (N=20 374)

变量		治疗成功 (n =19 108)		失访 (n =1266)		二元分析	
		n	%	n	%	调整前优势比 (OR)	95% 置信区间 (CI)
性别	男性	11 832	61.9	948	74.3	1.8	(1.6-2.0)
	女性	7276	38.1	328	25.7		
无家可归	是	309	1.7	107	2.1	5.6	(4.6-6.8)
	否	18 799	98.3	1168	97.2		
酗酒	是	825	4.4	168	13.2	3.4	(2.9-3.9)
	否	18 283	95.6	1107	86.8		
居住地	城镇	10 324	54.2	896	70.5	2.0	(1.8-2.5)
	农村	8716	45.8	375	29.5		

在分层分析中，考虑（或“控制”）第三个变量时可以计算一项患者特征的优势比。例如，无家可归患者所处的社会环境有时意味着他们还会经历其他共病或结核病风险因素。例如，相对于非无家可归患者来说，无家可归患者可能会更频繁地出现酗酒现象。在国家 A，监测数据提示实际情况就是如此，即更高比例的无家可归患者被登记为酗酒者。不过，酗酒者的失访优势比在无家可归患者（OR=1.8）和非无家可归患者（OR=1.9）中没有多大差别。这意味着按无家可归给结果分层时，酗酒优势比（1.8–1.9）较总体数据的酗酒优势比（OR=3.4，表 A2）要低。针对生活现状（无家可归或非无家可归）进行调整后，酗酒的优势比要比粗优势比或调整前优势比低。

表 A3
按患者生活现状和患者饮酒量划分的治疗转归分层

		无家可归				非无家可归			
		治疗成功		失访		治疗成功		失访	
		n	%	n	%	n	%	n	%
酗酒	是	144	46.7	65	60.8	2395	12.7	249	21.3
	否	165	53.3	42	39.2	16 405	87.3	919	78.7
		OR=1.8（（144/165）/（65/42））				OR=1.9（（2395/16405）/（240/919））			

在多元分析中，相同的规程可以用来同时考虑（控制）多个变量。这应当通过统计软件完成。表 A4 显示了一个多变量逻辑回归模型的结果，该模型包含上文描述的全部四个变量。相比仅针对无家可归对酗酒优势比进行调整（表 A3），表 A4 显示酗酒的优势比已经针对无家可归、性别和居住地进行了调整（控制）。结果显示，即使额外针对性别和城镇或农村居住地进行了调整，酗酒优势比还是 1.8，与仅针对无家可归进行调整后的优势比相近。

表 A4 中显示的所有优势比都超过 1（一致），并且具有不包括 1 的 95%置信区间，说明即使经过了调整，所有优势比仍然与失访显著相关。这显示男性性别、无家可归、酗酒和城镇居住地都是治疗期间失访的独立风险因素。这意味着非无家可归或非酗酒者且居住在农村地区的男性患者治疗期间的失访优势比要比具有类似其他特征的女性患者的优势比高出 1.5 倍。

针对“混杂”进行调整后，各种关联的程度下降。调整后优势比的值（表 A4）要比调整前优势比低（表 A2）。这意味着与失访相关的不同患者特征之间存在重叠。因此，相比不酗酒的患者来说，酗酒者更多的是男性、无家可归和居住在城镇地区。国家结核病防治规划可以使用这些结果给 DOT 提供方开展培训，以便给具有这些特征的患者提供额外（例如社会心理）支持，因为这些患者可能会失访，额外支持有助于他们完成抗结核病治疗。

表 A4
多元分析结果（治疗期间的失访）

变量		多元分析	
		调整后的优势比（OR）	95%置信区间（CI）
性别	男性 女性	1.5	（1.4-1.7）
无家可归	是 否	3.5	（2.8-4.3）
酗酒	是 否	1.8	（1.5-2.1）
居住地	城镇 农村	1.5	（1.3-1.7）

第 3 章

利用基因分型数据开展疫情暴发调查

目标受众：

- (1) 能够获取结核基因分型数据且想要利用这些数据开展结核病监测和疫情暴发调查的流行病学家和督导与评估官员。
- (2) 无法获取基因分型数据但有兴趣了解结核基因分型及其如何改善规划的国家结核病防治规划人员。

预期成果：

本章结束时，读者应当能够了解：

- 基因分型数据如何用于描述结核病流行病学和调查疫情暴发。
- 如何分析、报告和解释基因分型数据。
- 基因分型数据的分析如何促进实践和政策的变革。

作者：

Laura Anderson 、 Deanna Tollefson 、 Emily Bloss 、 Lori Armstrong、Rachel Yelk Woodruff、Susan van den Hof

3.1 基因分型数据：概述

简介

利用现代技术，科学家们能够从分子层面上了解结核病。基因分型是一种特殊的实验室方法，科学家们利用它^a分析结核分枝杆菌的遗传物质（例如 **DNA**），结核分枝杆菌是导致人类绝大多数结核病的细菌。通过对结核分枝杆菌的基因分型，科学家们已经了解结核菌具有许多不同的种类，这些种类被称为“菌株”。每种拥有唯一的“基因组”，即生物体内一组完整的遗传物质。除了全基因组序列测定，绝大多数基因分型方法都是考查结核基因组的选定部分，这些部分已知具有可变性，可以区分结核分枝杆菌的不同菌株。

了解结核基因组已经给结核病研究和项目规划提供了许多新的途径。因此，基因分型技术的可用性已经在全球范围内得到快速拓展。随着这些数据日益普及，公共卫生项目也获得了机会利用基因分型改善结核病监测和控制规划。

基因分型的目的和用途

基因分型让科学家们能够比较不同病例中的结核菌，从而确定是否出现了结核病传播以及如何传播。在这方面，科学家们主要是查看细菌分离株^b是否具有基因差异。如果结核分离株具有相匹配的（无区别的）基因型，则被认为属于同一菌株；如果结核分离株具有不匹配的（有区别的）基因型，则被认为属于不同的菌株。^c

流行病学家可以利用相匹配基因型的数据来追踪结核病是否已经在人和社区间传播以及如何传播。例如，相匹配结核分离株更有可能来自共同的源，被认为处于相同的传播链。如果两个或多个结核病病例具有相匹配的结核基因型，即已知为“群集”，流行病学家可以开展进一步的调查，确定病例之间是否有关联，如果有关联，说明可能存在传播。具体而言，流行病学家会调查病例之间的关联，具体包括时间、地域、耐药性和其他因素之间的关联。换句话说，流行病学家会询问是否具有共同性，例如共享的空间和时间，这种共同性能够解释传播。如果科学家们发现具有无区别基因型的结核病病例之间存在流行病学关联，则可能指示已经出现了传播。在另一方面，不相匹配的结核分离株更有可能来自不同的源。例如，具有唯一基因型的一个结核病病例可能源自潜伏结核病感染的激活。

-
- a. 可以在美国疾病预防控制中心的网站上找到有关基因分型的基本信息：
<http://www.cdc.gov/tb/publications/factsheets/statistics/genotyping.htm>
 - b. “分离株”指培养基中生长出来的结核菌，而不是直接从痰样中分离出来的结核菌。
 - c. 科学家们在什么时候可以判断基因型无区别方面具有不同的标准。这是因为始终存在着实验室污染或定型方法失败的可能性。这个主题不在本文讨论范围之内。

如果病例相关报告系统具有与常规监测数据相关联的实验室数据，基因分型数据可以发挥最大功效。一旦实验室数据和病例数据以这种方式关联起来，结核病防治规划可以结合使用基因分型结果和流行病学信息：

- 确定需要采取公共卫生行动的疫情暴发。
- 确定结核病病例之间近期是否可能出现了传播。
- 评估结核病发作是归于再激活还是再感染。
- 发现假阳性结核病病例。
- 确定和监测结核菌株不同时间段在不同人群中的传播。

基因分型数据的应用方式因具体目标而异。在公共卫生领域，结核基因分型数据通常被用来调查传播和疫情暴发。因此，本章首先介绍什么是基因分型以及如何开展基因分型。然后解释结核病防治规划如何使用这些数据调查传播和疫情暴发，以及如何了解更多地方结核病流行病学信息。为了更好地阐述这些内容，采用了虚构的英国基因分型数据，如果资源充足的话，流行病学家可以利用许多相同的流程和技术查看高负担地区的数据。在本章的最后部分，读者可以了解何时和如何利用基因分型分析，以及这些分析的结果如何证明传播情况。

预期目标受众

目前，低结核病负担国家的结核病防治规划要比其他国家和地区的规划更多地使用基因分型，因为低负担国家更容易获取基因分型技术，并且有着更少的结核病病例；这些使得可以在常规项目规划中实施基因分型。不过，高结核病负担国家的研究已经显示，结核基因分型不仅可以开展而且值得开展，但是应用的方式可能会稍微不同。例如，巴西高结核病负担城市的结核基因分型加强了对地方传播的了解，生成了有关如何改善结核病防治项目和在哪些地区改善项目的信息（1）。同样，埃塞俄比亚的基因分型阐明了地方传播（2,3），包括脆弱亚群中的传播，例如儿童（4），促进了结核病监测和控制的改善。

结核基因分型虽然具有很大的效用，但许多高结核病负担国家却没有能力开展结核基因分型，考虑到这一点，本章面向两类目标受众：能够获取结核基因分型数据且想要利用这些数据改善结核病防治规划的流行病学家和督导与评估官员等人员；有兴趣了解什么是基因分型及其对结核病防治规划有什么作用的人员。因此，本章描述基因分型现在和将来可以用于结核病防治规划的一些方式。本章不仅让读者有了基本的了解，还可以鼓励读者考虑如何更具体地将基因分型应用到高负担地区。

3.2 数据准备

虽然本章的重点并不是集中阐述实验室技术或实践，不过简单了解科学家们如何生成基因分型数据还是有助于我们更好地知道如何利用这些数据回答公共卫生问题。

区分结核菌株

科学家们使用四种主要的分子方法区分不同的结核菌株：限制性片段长度多态性；间隔区寡核苷酸分型；分枝杆菌散在分布重复单位-可变串联重复序列（MIRU-VNTR）；全基因组序列测定。科学家们可以单独使用这些方法，也可以将这些方法结合起来使用。这些方法在区分各基因型方面有着不同的能力，因为它们考查细菌基因组的不同部分。方法运作方式的差异可以确定被认为在同一个群集内的结核病病例数量（例如具有无区别基因型的结核病病例数量）。因此，科学家们使用一种以上方法来分析结核分离株，则可能会取得更好的效果，具体取决于所选择的方法（表 1）。

例如，如果科学家们只使用间隔区寡核苷酸分型方法，不相关的结核菌株也可能会看起来相关（例如结核分离株看起来具有相同的遗传模式），因为相比其他方法，间隔区寡核苷酸分型在区分结核分离株方面具有较小的特异性。这会使得结果看起来是少数几个大的结核病病例群集。这可能会误导公共卫生官员认为最近出现了高水平的结核病传播，导致他们错误地得出会有大规模结核病疫情暴发的结论。不过，如果科学家们将该方法与 MIRU-VNTR 分型结合使用，他们可能会发现这些大的结核病病例群集实际上是由许多较小的结核病群集构成。这样，公共卫生官员会认识到病例并不是通过共同源相互关联，因此没有理由相信正在发生传播。然后，规划可以避免花过多的时间和宝贵资源在群集调查上。

在这些方法中，每种方法都有独特的局限性，各国选择的方法取决于许多因素，例如结核病防治规划可用的技术、专业知识和预算（表 1）。

表 1
各基因分型方法的优势和劣势

方法	优势	劣势
限制性片段长度多态性	具有较高的菌株区分能力 高度稳定，半衰期为 3.2 年（生物标志物变化率）	要求具备活菌培养和大量 DNA 需要大量劳动力，并且非常耗时 要求技术专业知识和专业分析软件 不同实验室之间很难开展限制性片段长度多态性模式的比较分析 对于具有低拷贝数 IS6110 的菌株的区分能力差
间隔区寡核苷酸分型	要求少量 DNA，不要求活生物体 快速的和可复制的结果可以轻松地在不同实验室之间进行比较 精确区分具有低拷贝数 IS6110 的菌株	单独使用时区分能力低
分枝杆菌散在分布重复单位-可变串联重复序列	比间隔区寡核苷酸分型的区分能力要高 快速的和可复制的结果可以轻松地在不同实验室之间进行比较 比其他基因分型方法要便宜	比限制性片段长度多态性的区分能力要低 并不总是能够获得完整的 MIRU-VNTR 资料 生物标志物的稳定性/突变率不明确
全基因组序列测定	区分能力最高 稳定的生物标志物 可以确定传播的方向	成本高 要求技术专业知识和专业分析软件 结果的解释比较复杂

确定和命名群集

一旦科学家们基于基因型区分了不同的结核分离株，他们会使用特定的计算机软件（例如 BioNumerics）^d 将分离株分组到相应的群集（如果有的话）。国家结核病防治规划可以选择将这类软件整合到常规监测系统中。这样做的优势在于系统经过精心的设计后可以最方便和最有效地改善规划的现场工作和分析工作。

确定群集后，科学家们给每个群集分配唯一的号码或名称。务必要让该号码或名称保持一致性，这样才能轻松地分享信息。一致性的名称还可以确保流行病学家统计和调查群集中的所有病例。例如，如果一个国家的不同地区采用不同的群集编号或名称，则流行病学家可能会将单一群集作为两个单独的（而不是一个）事件来调查。同样，如果群集名称或编号随时变化，则公共卫生官员可能无法将最近的病例与既往的病例关联起来。许多国家已经根据菌株系或国家特定地理学确立了专门的术语。例如在英国，群集编号包含菌株系，而后是唯一编号（例如 B0001 = 北京菌株第一号）。此外，一些基因分型方法（例如 MIRU-VNTR）还具有标准化的国际术语，可以在各国之间开展群集比较（5）。

3.3 分析疫情暴发

基因分型数据可以用于有关结核病传播的特定问题，以及用于分析潜在疫情暴发。具体而言，我们可以提出下列问题：

- 是否有任何病例是假阳性？
- 我们是否应当将这种情况考虑为疫情暴发？
- 最近是否出现了传播？
- 我们看到的病例是归于再激活还是再感染？
- 结核病传播在不同人群之间、不同时期有什么不同？

开展疫情暴发调查的目的是快速诊断和治疗患有活动性疾病或潜伏感染的患者，从而预防进一步的结核病传播。通过利用基因分型数据回答上述问题，国家结核病防治规划可以最大限度地利用国内资源来找到、治疗和预防结核病。最终，这可以促使结核病防治规划推动更有效的政策，从而终止结核病传播和改善病例发现。

d. 参见：<http://www.applied-maths.com/bionumerics>

基因分型可以极大地提高疫情暴发调查的成功度。在低结核病负担国家，科学家们可以通过前瞻性方式确定所有结核分离株的基因型，从而快速发现和调查可能导致潜在疫情暴发的群集。在高结核病负担国家，国家结核病防治规划可以选择不同的方法，例如当怀疑选定人群中可能出现疫情暴发时，可以对这些人群的分离株开展基因分型，和/或当想知道某些地区结核病流行情况（包括传播）的更多信息时，也可以对这些地区的分离株开展基因分型。例如，可以对特定地点（例如医院）或特定亚群（例如儿童）中的病例开展基因分型（1-4）。

确定疫情暴发涉及到几个要素：排除假阳性病例；调查流行病学关联（例如各病例是否分享可能导致传播的共同时间或地点？）；考查耐药性；考虑每个病例的结核病史。所有这些要素都要考虑在内，示例被用来强调流行病学家应当如何在疫情暴发调查中考虑这些要素。虽然提供的各个示例都是基于英国（低结核病负担国家）的虚构疫情暴发调查，不过一般理论和流程可以应用到高结核病负担地区。

排除假阳性病例

一旦具备群集数据，流行病学家在调查潜在疫情暴发的过程中，应当采取的第一个步骤是确定观察到的结核病病例数量增加是否属于实际情况。我们必须考虑一种可能性，即当单个实验室连续处理的样本具有无区别的基因型时，一些病例可能属于假阳性。在这种情况下，应当向实验室查询结果，确定样本是否经过批处理或共用一个安全柜，这两种情况都有可能存在污染。其他有用的信息包括痰涂片状况、培养生长是否异常缓慢、采集的样本中是否仅有一个样本是培养阳性等。如果这些问题中有任何一个问题的答案是肯定的，则更有可能是发生了污染。最后一点，应当与主治临床医生讨论各个病例，找出患者是否具有结核病相关症状。如果除了实验室信息之外，临床表现看起来也不像结核病，则可能病例属于假阳性。

流行病学关联

一旦我们已经确认病例数量增加是实际情况，我们应当查阅人口统计学、临床学信息和风险因素信息，确定群集中各个病例之间是否存在流行病学关联。确认的流行病学关联要求具备下列证据：各病例发生在相同的地点和相同的时间，其中一个病例患有活动性和传染性结核病（例如痰涂阳肺结核）。可能的流行病学关联具备下列条件：各病例发生在相同的地点，但时间未知。流行病学家在调查流行病学关联时会考查群集内各结核病病例的特征。例如，他们可能会比较各病例的社会学风险因素或居住地。如果具有相似点，则可以合理地认为各病例具有关联性。

多种方法可以用来整理群集数据，以便考查是否存在流行病学关联。首先，最简单的方法是将群集看作病例系列清单，显示所有相关的人口统计学、临床学、实验室和社会学风险因素，例如吸毒史、酗酒、在押时间和无家可归。这样做的目的是查找各结核病病例的相似点和不同点。最有用的方法是按照临床表现或报告的时间顺序查看病例，确定可能的传播链和可能是源病例的个人。例如，源病例（即感染其他人的最初结核病患者）可能是群集中被诊断或被报告的一个病例，或者可能是最早出现症状的人。

表 2
群集评估报告（英国）

群集编号： B1006*
MIRU-VNTR: 424352332517333456443372**

唯一患者识别码	病例报告日期	年龄	性别	地区	邮政编码	出生地	进入英国后的年数	疾病部位	痰检状况	既往结核病	症状发作日期	风险因素	耐药性
1	01/05/2013	26	女	伦敦	NW11 5BT	英国	不适用	肺部	阳性	无	14/02/2013	吸毒	无
2	20/06/2013	33	女	伦敦	NW11 8EL	英国	不适用	肺部	阳性	无	12/05/2013	吸毒	无
3	28/08/2013	26	男	伦敦	NW11 8EE	牙买加	13	肺部	阴性	无	25/07/2013	在押	无
4	01/09/2013	31	女	伦敦	NW11 3AZ	印度	10	肺外	不适用	无	06/05/2013	吸毒	异烟肼耐药
5	12/12/2013	52	女	伦敦	NW11 9EQ	英国	不适用	肺外	不适用	有	04/11/2013	无	耐多药

* 数据是虚构的。
** 群集被定义为具有无区别的结核基因型的两个或多个病例。唯一菌株通常被认为源自再激活，而不是最近传播链。拥有一个不同位点的菌株的分离株不被认为是群集的组成部分。

文字框 1

英国 – 示例 1

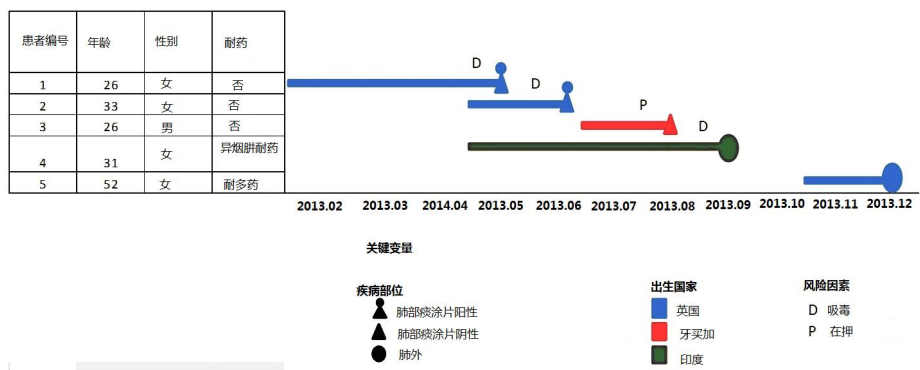
让我们考虑一下表 2 中呈现的数据，了解我们如何利用基因分型数据评估潜在疫情暴发。该表列出了英国被科学家们归类到群集 B1006 中的病例。这些病例之所以都在一个群集中，是因为 MIRU-VNTR（424352332517333456443372）发现每个病例都具备无区别的基因型。这意味着病例之间可能存在传播，因此我们现在需要确定是否存在可能导致传播的流行病学关联（例如分享的时间和地点）。请注意，如果所有基因型都不相同，它们就不可能是一个群集，所以没有证据证明存在关联，从而也没有必要调查疫情暴发。

我们查看群集时，我们考查的第一件事情是患者是何时生病的。在群集 B1006 中，我们看到在 7 个月时间内报告了 5 个结核病病例。我们利用病例报告日期来确定时间段，因为它是可验证的日期；它是患者成为实验室或临床确诊结核病病例的日期。此外，病例报告日期是一个强制变量，因此数据比较完整。我们通常不使用发病日期，因为这是患者自己报告的日期，更有可能不准确或缺失。不过，如果我们知道发病日期，则务必要使用它来评估获得结核病诊断或开始治疗之前的症状持续时间。通过这一点，我们可以确定病例在社区内有多长时间具有传染性，以及病例可能已经将结核菌传播给其他人的程度。基于这个认知，我们可以重新评估各病例在群集中的出现顺序。鉴于可能出现通报较晚的现象，所以病例可能在病例报告日期之前就已经生病了好几个月。

查看时间段（例如病例报告日期和发病日期）后，我们可以考查其他变量。在这种情况下，我们通过邮政编码可以看出所有病例看起来好像都是伦敦同一个小区的居民。绝大多数病例还分享着其他关键特征：她们是女性，年龄在 25–35 岁，主要出生在英国，患有肺部疾病（两个病例同时还是痰涂阳），具有吸毒史等。在其他国家出生的两个病例已经在英国居住了十多年。基于这些信息，看起来似乎各病例之间很可能存在流行病学关联（例如各病例可能处于相关的地点和相同的时间）。我们可以假设共同吸毒是最有可能的流行病学关联，不过需要开展更多调查来确认这些假设。

将数据看作系列清单后，通过图形或其他可视格式来查看数据会很有用处。可以在图形中展示表 2 中的数据，这样可以更轻松地看出感兴趣的关键变量，还能更轻松地确定流行病学关联（图 1）。在该图中，流行病学家已经根据病例报告日期将每个结核病例安排到时间轴上。个体的形状和颜色指明了疾病部位。每个病例后面的尾部显示了从症状发作日期到病例报告日期之间的时间长度。图上还指明了风险因素，具体指吸毒和在押史。最后，在该图左边的表格中显示了有关年龄和性别的额外信息。

图 1
集群图示例，英国（群集 B1006）



文字框 2
英国 – 示例 2

正如前面所讨论的一样，评估群集内各病例之间是否存在传播时，我们必须考虑病例报告日期、症状发作日期（如果可用的话）、疾病部位、病例是否具有传染性（例如肺部）和传染水平（例如痰涂阳病例被认为比涂阴病例更具有传染性）。虽然所有这些信息都可从表 2 中找到，但利用图 1 更便于解释。通过将病例按照疾病部位和痰检状况安排到时间轴上，我们可以观察到哪个病例可能已经传播了结核病。

在群集 B1006，第四名患者（来自印度）是一个肺外病例，即使她被报告的时间比第三名患者要晚，但她却在第三名患者（来自牙买加）之前开始显现症状。因此，第四名患者不可能通过第三名患者感染结核病，因为第四名患者的发病日期要早于第三名患者。而且，第三名患者可能也不是通过第四名患者感染结核病，因为第四名患者患有肺外结核，因此不具有传染性。

文字框 2 - 续

而且，第一名患者、第二名患者和第四名患者都是吸毒者，吸毒已知属于结核感染和结核病的风险因素（6）。将风险因素信息与症状发作日期结合起来，我们可以假设这些病例之间在与吸毒相关的社会场所可能已经发生了传播。

耐药模式

评估传播的可能性时，我们还必须考虑群集内各结核病病例的耐药模式。如果各病例的结核分离株具有完全不同的耐药模式，则这些病例不可能处于相同的传播链。如果一个群集同时拥有完全敏感或耐药的病例，这一点就尤其准确。虽然完全敏感结核病病例可能将结核病传播给某个人，然后这个人又发展出继发性耐药，但耐药结核病病例不可能将药敏结核病传播给另一个人。而且，分离株对多种药物耐药的结核病病例不可能将结核病传播给分离株对较少药物耐药的结核病病例。在卫生机构开展基因分型的国家，结核病防治规划通常会常规开展药敏试验，意味着群集中每个分离株都应当具有相应的药敏模式。

文字框 3 英国 – 示例 3

我们可以查看耐药模式，从而更好地了解群集 B1006 可能的传播链。在表 2 和图 1 中，我们可以看到第四名患者对异烟肼耐药，而第一名患者、第二名患者和第三名患者有着完全敏感的结核病。基于症状发作日期，我们已经确定第三名患者不可能感染第四名患者。耐药结果支持这种假设，因为在第三名患者感染完全药敏结核病之前，第四名患者已经感染了异烟肼耐药结核病。第一名患者可能感染了第四名患者，后者然后获得了异烟肼耐药。第五名患者出现耐药结核病。因此，也可能第一名患者、第二名患者或第三名患者将结核病传播给了第五名患者，后者然后发展出额外耐药。

既往结核病发作

调查传播链时，流行病学家还必须考虑一项额外的信息，即结核病史。结核病病例可能是新近感染结核病，也可能是潜伏结核病被激活，还可能是既往接受过结核病治疗后再次复发。基因分型数据可以帮助区分新感染与既往疾病复发，因此能够提供有关传播的重要信息。

在低结核病负担国家，具有结核病史的病例可能会具有历史结核分离株。如果具备这类分离株，科学家们应当检索它们（和基因型），将它们与当前分离株进行比较。如果各结核病发作之间的基因型存在差异，这提供证据表明患者经历了新感染（例如再感染），而不是复发（例如再激活）。因此，这名患者可以成为最近传播链的组成部分。在另一方面，如果基因型之间无区别，这提示更有可能是结核病再激活。

在高结核病负担国家，不太可能具有历史结核分离株，不过结核病史仍然可以提供重要的背景信息，帮助评估一名患者是否涉足传播和成为疫情暴发的组成部分。例如，应当查阅既往发作的治疗转归。如果该名患者既往没有完成治疗，则这提示当前发作属于复发或持续疾病，而不是最近传播，从而可以将患者安排到时间轴考前的位置。这名患者很有可能在较长一段时间内都具有感染性，可能导致了群集内病例的增加。如果该名患者既往完成了治疗，则更有可能处于最近的传播链中。即使具备历史分离株，但高负担地区经常出现常见菌株的传播，因此也可能再次感染相同的菌株。应当查阅既往发作的接触者追踪信息，因为也有可能既往未能发现源病例，而患者是被未诊断的源病例再次感染。

呈现病例之间的流行病学关联

一旦流行病学家通过时间、人和地点相关变量查看了群集，他们可以更好地调查流行病学关联。正如本章前面所讨论的一样，疫情暴发调查期间收集的基因分型数据和患者信息可以提供特定场所的传播证据，例如、监狱、学校、工作场所或居民小区，帮助确定病例之间的流行病学关联。流行病学家务必要通过适当的方式整理和呈现这些数据，以便明确地显示这类潜在关联。这有助于确定结核病防治规划应当在哪些群体或地区开展进一步筛查或发起其他公共卫生干预。图 2 和图 3 显示流行病学家如何以图示方式展现流行病学关联和传播地区。可以利用专业软件（图 2）或基本软件（例如 PowerPoint）（图 3）创建此类图示。流行病学家可以通过调整符号来改编这类图示，从而展示与疫情暴发调查最相关的信息。

文字框 4

英国 – 示例 4

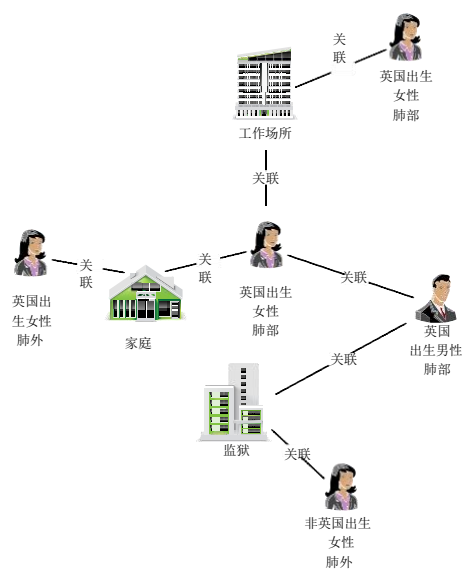
在群集 B1006 中，第五名患者是唯一有过既往结核病发作的病例。务必要了解到，如果第五名患者的当前结核病发作与第一名患者、第二名患者、第三名患者和第四名患者属于同一个传播链，我们就需要调查有关第五名患者既往结核病发作的详细信息。随访了既往治疗第五名患者的临床医生后，发现第五名患者的既往结核病发作是在 2012 年 6 月。当时该名患者被诊断为肺部痰涂阳结核病，但她没有已从医生嘱咐的治疗方案。虽然当时没有记录风险因素信息，但患者实际上是一名吸毒者。

基于有关第五名患者的这些新信息，我们需要考虑下列情景。第五名患者可能是因为治疗失败导致了复发，现在由于治疗依从性差发展出耐多药结核病（再激活）。这将意味着该名患者不属于最近传播链。如果这是实际情况，则该名患者不可能属于疫情暴发的组成部分，因此可能没有必要进一步调查该名患者与其他患者之间的流行病学关联。不过，如果第五名患者的既往结核病发作是由一种不同的菌株所导致，基因分型数据可能显示第五名患者是因为再次感染结核病而生病。在这情况下，她可能属于最近传播链的组成部分，因此我们应当调查该名患者，作为当前疫情暴发的一部分。

不论是上述哪种情景，我们应当追踪第五名患者的接触者，调查第五名患者是否属于疫情暴发的源病例，因为她是第一个被诊断的病例，具有传染性，并且与其他病例一样存在社会学风险因素（例如吸毒）。如果我们确定她是源病例，这可能提示接触者追踪未能发现接触者。因此，我们应当考查接触者，确保我们已经诊断了所有活动性病例。

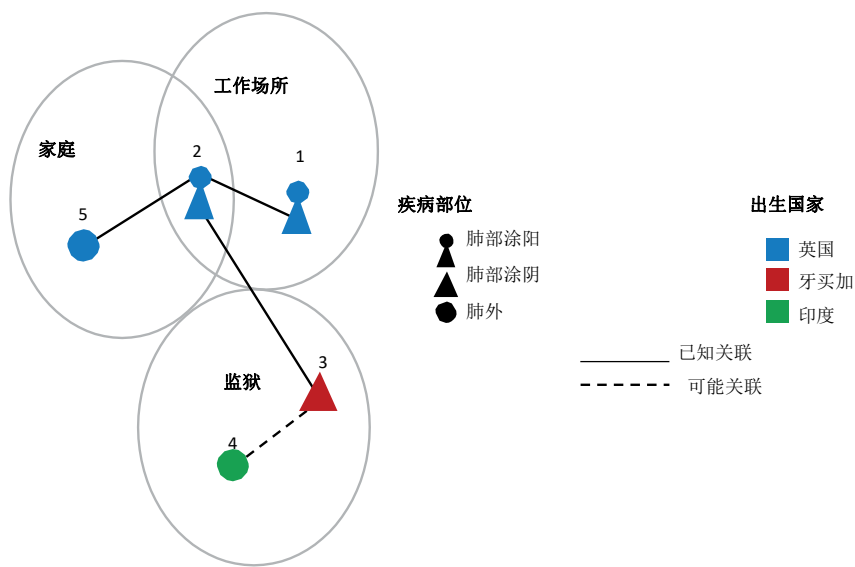
因为我们已经确定了可能的流行病学关联，我们应当进一步调查该群集，确认这些关联，确定要开展更多接触者追踪的传播场所。通过发现和治疗社区内尚未被国家结核病防治规划发现的任何其他活动性结核病病例，预防结核病的继续传播。在这种情况下，我们进一步的调查需要可以收集有关病例的更多信息，例如吸毒的场所、详细的在押史（例如在押机构和时间）、职业、定期社交场合、重要的旅游、常客和既往居住地。

图 2
群集内病例与传播场所之间的流行病学关联



该图是用 IBM® i2® Analyst’s Notebook® 中创建，网址：<http://www-03.ibm.com/software/products/us/en/analysts-notebook/>。每个病例下面显示了相关人口统计学或临床信息。每个病例上面用符号指示吸毒。各病例和场所之间的流行病学关联通过线条显示。

图 3
阐述群集内病例与传播场所之间流行病学关联的示意图



该图是用 PowerPoint 创建。每个病例通过符号显示，患者特征根据符号的形状或颜色来指示。显示了各病例之间的关联。各病例的报告顺序由每个病例上面的数字显示。传播地区在维恩图解风格中显示为圆点，病例被安排在相应的每个相关场所中。

文字框 5 英国 – 示例 5

我们可以使用图 2 和 3 以可视方式展现该群集中各病例之间的流行病学关联。在图 2 中，我们可以看出哪个病例可能有着最大数量的流行病学关联或处于疫情暴发的中心。在该图中，我们看到处于疫情暴发中心的病例与其他三个病例直接相关联；她与两个病例通过家庭相关联（一个是家庭成员，一个是在家庭中待过），与一个病例通过工作相关联（一个是工作同事）。此外，我们看到其他两个病例通过监狱相关联。在图 3 中，我们看到相同的信息以不同的格式呈现。该图将病例安排到相应的场所，有助于评估各病例分享哪些场所。它还采用符号更为明确地显示出生国家和疾病部位。

一旦流行病学家收集了所有相关信息并完成了必要的示意图，结核病防治规划应当考虑下列问题：

- 是否确定了患者在每个场所的所有接触者？是否所有接触者都接受了筛查？
- 是否可能有其他活动性结核病病例未被诊断出或未经培养进行确认？
- 针对存在既往结核病发作的病例开展了多广泛的接触者追踪？

一旦规划回答了这些问题，流行病学家应当与公共卫生官员和临床人员一起完成疫情暴发调查。这需要他们首先确定与传播场所相关的病例的接触者。接下来，他们会邀请接触者接受结核病筛查。最后，他们确保患有结核病或感染的人根据国家指南尽快开始接受结核病治疗。

3.4 分析大型群集

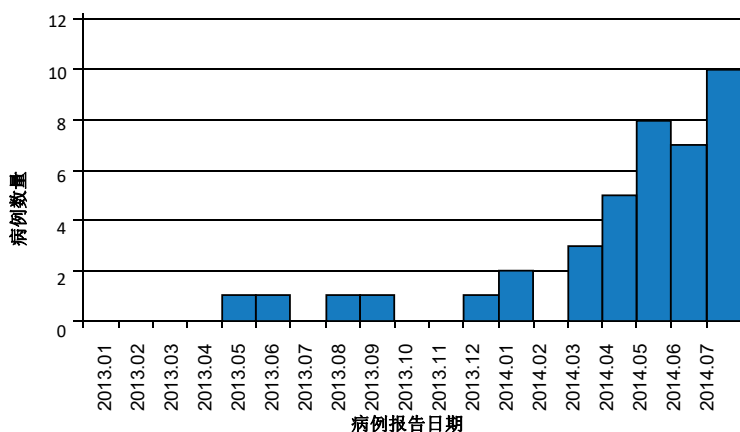
病例数量较少时，利用系列清单或群集图可以轻松地查看基因型群集，正如上面一节所述。不过，如果群集中有许多病例（例如>10），流行病学家必须使用不同的数据分析和展示方法。利用大量群集数据，流行病学家通常分别展示时间、人和地点等信息，然后将这些信息整合起来了解群集的全貌，然后他们能够以图示方式确定流行病学关联。

显示时间、人和地点

首先，流行病学家考查病例在不同时间段的分布。本章已经展示流行病学家如何使用时间轴来考查传播的真实性。不过，流行病学家还可以使用时间轴来考查群集的增长（例如不同时间段发现的新结核病病例数量）。这可以呈现为一种标准流行曲线，正如在图4中所见。

病例数量的突然增加可能指示着最近传播，因为一个人最有可能在感染结核分枝杆菌后很快发展为结核病。发展为活动性疾病的可能性会随着时间的流逝逐渐下降，许多人永远都不会发展为活动性疾病。因此，完全不可能出现多个病例自发由相同结核菌株再激活的情况，除非人群中通常传播的基因型数量有限。不过，流行病学家应当调查具有无区别结核基因型的结核病病例数量的任何增加，确定是否存在真正的疫情暴发，因为观察到增长并不总是实际情况，正如上文所讨论的一样（例如实验室诊断技术或病例批报告的错误）。

图4
群集内结核病病例的流行曲线



文字框 6

英国 – 示例 6

图 4 显示 2013 年 12 月至 2014 年 7 月在特定场所的单一群集中出现了病例数量的增长。在该群集中，每月平均结核病病例数量在 2013 年 12 月至 2014 年 6 月从 1 例增长到了 >3 例，并在 2014 年 7 月增长到了多达 10 例。很显然，这种模式不同于已知规范，可能指示着疫情暴发。不过，我们在得出已经出现结核病疫情暴发的结论之前，观察到的增长必须被确认属于实际情况。这要求排除实验室污染或报告变化，确定这些因素是否属于观察到的病例增长的可能原因。当单个实验室连续处理的样本具有无区别的基因型时，一些病例可能属于假阳性。在这种情况下，应当向实验室查询结果，确定样本是否经过批处理或共用一个安全柜，这两种情况都有可能存在污染。最后一点，应当与主治临床医生讨论各个病例，找出患者是否具有结核病相关症状、痰涂片状况是否属于阴性、采集的样本中是否仅有一个样本是培养阳性等。一旦我们已经确认病例数量增加是实际情况，我们应当查阅人口统计学、临床学信息和风险因素信息，评估可能的流行病学关联，正如上文所讨论的一样。

流行病学家确认完观察到的病例增长后，会将患者相关信息绘制到流行曲线上，包括人口统计学、临床学和风险因素特征（图 5）。这可以提供有关社区中如何发生传播的更多信息。例如，它可以凸显传播正在发生的人群，并且显示传播在不同时间段的变化趋势。

图 5
群集内按出生地划分的结核病病例的流行曲线

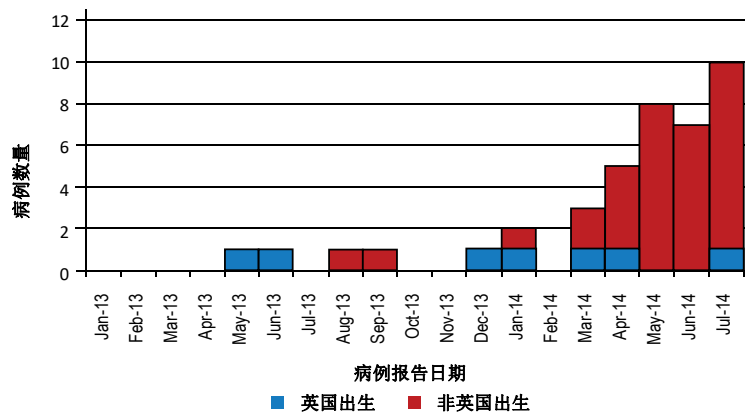


表 3
群集内病例数量和比例小结

特征	n	%
年龄（25-35 岁）	35	88
性别（女性）	35	88
英国出生	15	38
疾病部位	30	75
痰阳性*	24	80
既往结核病	2	5
社会学风险因素**	22	55
总计	40	100

* 仅肺部病例。
** 社会学风险因素包括吸毒、饮酒、在押和无家可归。

文字框 7
英国 – 示例 7

图 5 显示群集中最早诊断的病例来自英国出生患者，而非英国出生病例数量后来不断增加。这提示传播可能最开始是从英国出生人群向非英国出生人群蔓延。一段时间之后，传播在很大程度上继续在非英国出生人群中扩散。

表 3 显示群集中绝大多数病例的特征是年龄 25–35 岁（88%）、女性（88%）、非英国出生（62%）和肺部痰涂阳（80%）。我们还发现超过一半的病例（55%）具有结核病相关社会学风险因素（例如吸毒史、酗酒、在押时间和无家可归）。

利用流行曲线评估完基因分型数据后，流行病学家会评估特定群集内的结核病病例的特征，方法是总结群集中病例的人口统计学、临床学和社会学风险因素数据。流行病学家会计算具有每个特征的病例的数量和比例，然后将它们与群集的病例总数进行比较（表 3）。

如果结核病防治规划考查该群集，规划人员可以将这个地方群集的特征与国内具有相同结核基因型的所有病例进行比较，前提是具备这些数据（例如国家群集，包括这个地方群集的病例和国内具有相同基因型的其他病例）。这有助于规划评估地方群集中的病例是否不同于国内其他地方具有该基因型的其他病例。如果地方病例看起来不同于国家层面上的其他病例，这可能提示传播是在当地发生（表 4）。

表 4
病例数量和比例的比较（地方和国家群集）

特征	地方		国家	
	n	%	n	%
年龄（25-35 岁）	35	88	48	40
性别（女性）	35	88	42	35
英国出生	15	38	75	63
疾病部位	30	75	90	75
痰阳性*	24	80	72	80
既往结核病	2	5	7	6
风险因素**	22	55	40	33
总计	40	100	120	100

* 仅肺部病例。
** 社会学风险因素包括吸毒、饮酒、在押和无家可归。

除了查看结核病病例中人口统计学和社会学风险因素的相似点之外，流行病学家还务必要查看群集中各病例之间的地理位置距离，或者各病例与疑似传播来源地区之间的地理位置距离。在这方面，流行病学家可以利用邮政编码或地址推导出纬度/经度坐标。然后，流行病学家可以将数据导出到 GIS 软件包，^e用于绘制分布图。流行病学家可以在分布图上添加病例特征和地标或场所，计算病例之间的距离。如果不具备 GIS 或其他分布图绘制软件，流行病学家可以使用简单纸质地图。如果流行病学家拥有病例和/或疑似传播点的近似或完整地址，这个效果会特别好。通过确定潜在群集内结核病病例的居住地 and/或生活社区，两种方式都能协助展示可能已经发生结核病传播的地点。

e. 如 ArcGIS (<http://www.esri.com/software/arcgis/>)

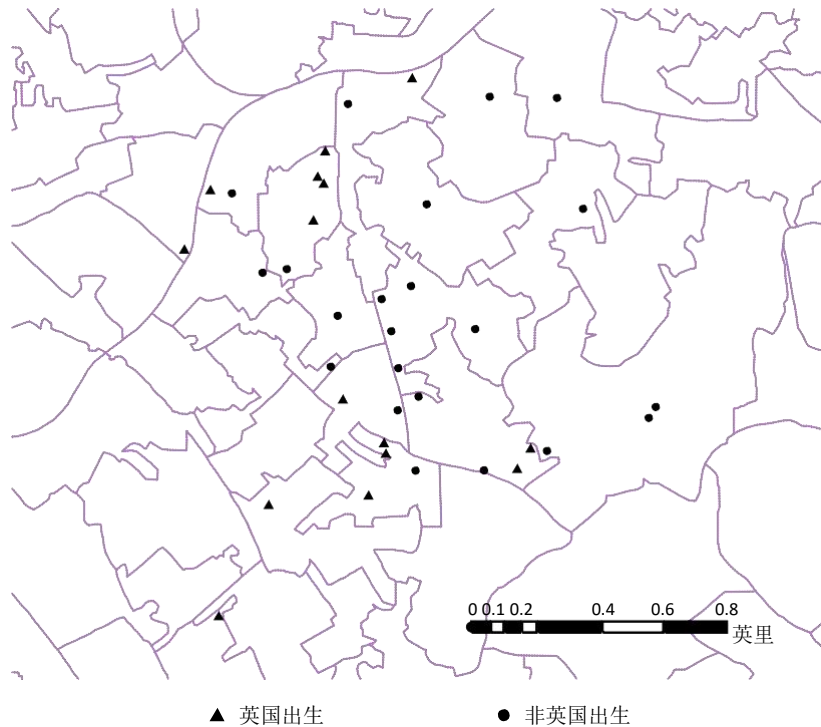
文字框 8 英国 – 示例 8

表 4 凸显地方群集与国家群集之间存在着明显的差异。地方群集的病例主要特征是非英国出生（62%）、女性（88%）、年龄 25-35 岁（88%）。国家群集的病例主要特征是英国出生（63%）、男性（65%）、仅 40% 年龄 25-35 岁。此外，地方群集比国家群集有更高比例的病例存在社会风险因素（分别为 55% 与 33%）。这可能提示地方传播发生在非英国出生高风险女性群体中。应当在地方层面上加强调查和监测，确定实际情况是否如此。

文字框 9 英国 – 示例 9

在图 6 中，我们看到按居住地绘制的群集结核病病例分布图（利用 ArcGIS 10）。各病例根据出生地（英国出生或非英国出生）采用不同的符号进行编码。该分布图揭示各结核病病例之间居住得是多么地靠近。此外，它还显示非英国出生病例聚集在分布图的中心，而英国出生病例零星聚集在分布图的外围地带。这有助于确定可能发生传播的地理区域。

图 6
按出生地划分的群集结核病病例的 GIS 分布图



3.5 基因分型数据的局限性

尽管基因分型数据具有良好的效果，但结核病防治规划在利用基因分型数据发现结核病疫情暴发之前必须了解这类数据的局限性。

首先，流行病学家必须谨记不同基因分型方法（即限制性片段长度多态性、分枝杆菌散在分布重复单位-可变串联重复序列、间隔区寡核苷酸分型和全基因组序列测定）有着不同的区别水平。也就是说，每种方法在区分各基因型方面有着不同的能力。如果一种方法不具备足够的区别能力，病例可能会被归类到错误的群集，无法代表实际的传播情况。

其次，流行病学家必须谨记即使利用具有高区别能力的方法来区分基因型，基于基因型的群集归类也并不总是代表传播情况。人群中传播的常见菌株或潜伏结核病的同时再激活都是很好的例子，说明各病例如何被归类到单一群集但可能并不属于最近传播的结

果。这在高结核病发病率地区尤其存在问题。例如在南非，北京结核株属于地方病，有着广泛的传播。这意味着绝大多数基因分型方法单独使用都不具备足够的区别能力，无法发现结核病的实际传播（9）。此外，有些低发病率国家的绝大部分结核病病例源自从高结核病负担国家传来的菌株，这些国家较大规模的分子群集可以代表国外常见菌株的传播。因此，不论是在高负担地区还是在低负担地区，流行病学家始终应当获得流行病学信息，例如发病时间、去过的地方、人口统计学和社会学风险因素等，这些信息有助于取得传播的可能原因。因此，流行病学数据的质量对于分析利用基因分型数据非常重要。缺失或不准确数据可能会导致各病例和场所之间的关联缺失或错误关联，进而导致更多有关联的结核病病例未被诊断或诊断延迟。因此，结核病防治规划应当尽量最大限度地减少关键变量的数据缺失，确保具备数据质量保障规程。^f

最后，抽样偏差通常不可避免。当抽样人群并不代表实际人群时，就会出现抽样偏差。基因分型依赖于培养技术，因此未通过培养确诊的病例被排除在基因型群集之外（8）。因此，一些结核病病例会被错误地归类为拥有唯一结核基因型，有些属于疫情暴发范围内的结核病病例也可能被错过。这会导致过低地估计传播程度。有些国家并不会对所有分离株开展普遍分型，或者有些国家很大比例的病例是可能在诊断前就离开的流动人口，抽样偏差在这些国家尤其具有局限性。在这种情况下，许多菌株可能会被错误地归类为唯一菌株，妨碍群集和传播事件被发现。还务必要在时间期限固定的研究中承认这种局限性，因为无法对研究前和研究后诊断的病例的分离株开展基因分型。如果科学家们在研究开始阶段发现受控群集中最后几个病例，或者在研究结束阶段发现某个群集中最初几个病例，可能会导致科学家们错误地将群集归类为增长数量小或增长速度慢。

3.6 在高结核病负担地区开展基因分型的特殊考虑事项

迄今为止，结核基因分型通常在低结核病负担地区常规开展，因此在高结核病负担地区利用基因分型数据的最佳方式仍然尚未确定。虽然基因分型的原则在不同地区之间相同，但是流行病学家在高结核病负担地区如何将基因分型应用于公共卫生目的还是有一些差别。结核病高传播水平地区应用基因分型数据尤其面临着特定的局限性。

^f 有关最大限度减少和处理缺失病例相关数据的更多信息，请参见第2章。

与低结核病发病率地区不同的是，国家结核病防治规划在高结核病发病率地区利用基因分型开展深入群集调查并不总是可行或值得。这通常是因为疫情暴发调查非常耗费资源，而资源和人员可能并不到位。具体而言，在结核病高传播水平国家，一般人群患结核病的风险较高，这意味着确定要开展筛查和公共卫生干预的具体传播地区尤其困难。如果人群中传播的菌株的数量有限，这一点会变得更困难（9）。

高结核病负担国家的结核病防治规划可以在疫情暴发调查期间使用基因分型确认或排除传播，而不是前瞻性地利用基因分型来发现和调查所有群集。例如，如果在学校内发现了结核病病例，然后筛查发现了进一步的活动性病例，则基因分型可能有助于确认儿童们是否已经处于传播链中，最终可能推动该场所的延展型接触者追踪。如果儿童们并不是感染相同的菌株，则尚未出现传播，可以花较少的资源开展调查。基因分型还可以用于特定的亚群和分场所，了解地方结核病流行病学和进一步改善结核病控制。例如，巴西和埃塞俄比亚的多项研究发现，亚群（例如儿童）或分场所（例如单一地理区域）中的基因分型可以更好地了解地方结核病流行情况和防治规划（2-5）。

如果国家结核病防治规划决定通过前瞻方式使用基因分型，则应当集中发现重点疫情暴发，这类疫情暴发中菌株与治疗转归差息息相关，例如耐多药结核病或广泛耐药结核菌株、导致高水平死亡率的菌株、导致毒性加强的菌株等。为了发现这些疫情暴发，最有用的要素是能将流行菌株与地域关联起来的基线数据库。然后，国家结核病防治规划可以监测具有特定菌株的病例数量的变化，作为常规监测的一部分。他们可以使用本章前面描述的工具以可视方式展现按空间和时间划分的趋势，凸显病例数量的非预期增长。基于结核病地域数据或“热点”，更为广泛的社区筛查可以更有效地发现更多活动性病例和减少继续传播，而不是将精力放到收集信息开展群集评估上，后者需要耗费大量资源，而且不太可能更多地发现继发性活动性病例。

3.7 结论：利用基因分型数据改善大众健康

尽管基因分型数据存在局限性，它们还是提供了革命性的机会，可以更好地了解低负担和高负担结核病地区的地方疫情。通过考查人群内不同时间段流行的结核基因型，可以确定最常见的或最流行的基因型。这些信息可以用于推导出每个基因型的基线传播水平，方法是计算每个月或每个季度的平均病例数量。只要发现特定基因型的结核病例数量相比基线数据有了增长，就可以确定传播的加剧。而且，通过将基因分型数据与流行病学数据结合起来，结核病防治规划可以对社区中的结核病传播有更详细的了解，从而使得公共卫生官员有针对性地面向高危人群开展干预。

还可以总结每个结核基因型中结核病病例的特征，利用这些特征制定有针对性的结核病控制策略。然后，可以监测基因型群集内病例数量的减少，以此评估这些策略的成功程度。最后，创建一个基因型库还有助于发现传入病例或再激活的最近传播在人群中出现的新基因型。考虑到在分析中利用基因分型数据可以获得宝贵的信息，这些信息如果利用得当，则可以成为公共卫生领域的巨大财富，能够帮助极大地改善结核病控制和预防。

1. Luiz, R, et al. Genotyping and drug resistance patterns of *Mycobacterium tuberculosis* strains observed in a tuberculosis high-burden municipality in Northeast, Brazil. *The Brazilian Journal of Infectious Disease*. 2013;17(3):338–345.
2. Esmael A, et al. Genotyping and Drug Resistance Patterns of *M. tuberculosis* in Eastern Amhara region, Ethiopia. *JBR Journal of Clinical Diagnosis and Research*. 2014. 2:102. doi: JCDR-1000102.
3. Mihret, A, et al. Modern lineages of *Mycobacterium tuberculosis* in Addis Ababa, Ethiopia: implications for the tuberculosis control programme. *African Health Sciences*. 2012; 12(3): 339–344
4. Workalemahu, B, et al. Genotype diversity of *Mycobacterium* isolates from children in Jimma, Ethiopia. *BMC Research Notes*. 2013, 6:352. doi:10.1186/1756-0500-6-352.
5. Smittipat N, Palittapongarnpim P. Identification of possible loci of variable number of tandem repeats in *Mycobacterium tuberculosis*. *Tuberculosis and Lung Disease*. 2000;80(2):69–74
6. Pedrazzoli D et al. Tuberculosis in the UK: 2013 report. London: Public Health England; 2013.
7. Murray M. Determinants of cluster distribution in the molecular epidemiology of tuberculosis. *Proceedings of the National Academy of Science USA*. 2002;99(3):1538–1543.
8. Cowley D, Govender D, February B, Wolfe M, Steyn L, Evans J et al. Recent and rapid emergence of W-Bei- jing strains of *Mycobacterium tuberculosis* in Cape Town, South Africa. *Clin Infect Dis* 2008; 47(10):1252– 1259.
9. Cowley D, Govender D, February B, Wolfe M, Steyn L, Evans J et al. Recent and rapid emergence of W-Bei- jing strains of *Mycobacterium tuberculosis* in Cape Town, South Africa. *Clin Infect Dis* 2008; 47(10):1252– 1259.

第 4 章

分析加剧结核病流行情况的因素

目标受众：

- (1) 流行病学家、统计学家和督导与评估官员。
- (2) 有兴趣了解加剧结核病流行的因素和如何利用所获信息改善规划实施的国家结核病防治规划人员。

预期成果：

本章结束时，读者应当能够：

- 了解生态分析的目的和流程。
- 了解利用结核病通报数趋势作为结核病发病率不完美代理变量的依据。
- 认真评估不同时间段的项目规划、卫生、宏观经济数据时间序列图与结核病病例通报系列的关系。
- 通过补充各种各样的数据，合理推断出病例通报数趋势在多大程度上反映基本结核病发病率的实际变化。

作者：

Charalambos Sismanidis、Ellen M.H. Mitchell、Emily Bloss、Philippe Glaziou

4.1 生态分析

生态分析指研究各人群或地区间的差异。生态分析是一项有用的工具，关于社会、经济、政治、文化、流行病学和卫生系统因素如何积极地或消极地影响结核病流行进程，生态分析可以提出宽泛的陈述和建立相关的假设。生态分析主要是一项探索型工具，而不是一项确认型工具。生态分析利用平均值或其他聚合度量指标，探索宏观层面促成因素与结核病流行动态之间的关系，而这类关系无法通过其他方式进行研究。

生态分析可以提供哪些解释？

生态分析被最常用来合理解释不同时间段和不同地域（国内或跨国）的结核病负担趋势。不过，它有助于探索细菌性确认病例的比例差异，查询肺结核比例与肺外结核比例之间出现差异的原因。

简单探索型生态分析可以由结核病防治规划人员开展，在高负担地区已经帮助显示了看似关系不大的社会 and 经济发展如何影响结核病病例通报（1）。本章中描述的分析原则和任务都是关于结核病病例通报（作为结核病发病率的代理变量），不过也可有效地解释结核病控制领域让人困惑的其他归因问题。

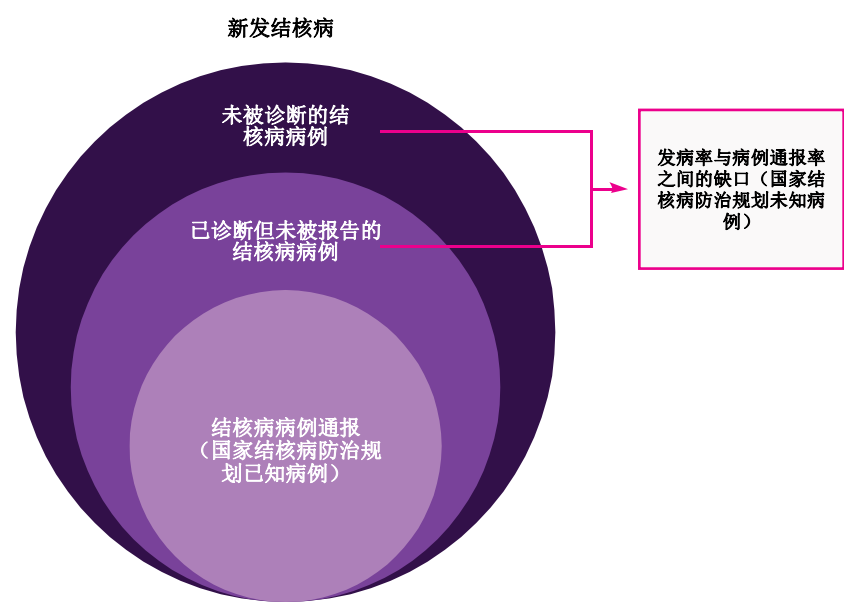
4.2 结核病发病率

发病率是特定时间段（结核病防治领域通常被界定为一个日历年）内出现的新结核病发作数量。鉴于当前结核病诊断工具的局限性、从感染到发病的长时间潜伏期和结核病临床表现的各种方式，直接衡量一般人群中的结核病发病率不仅不切实际，而且需要耗费大量资源。迄今为止，没有任何国家开展过具有全国代表性的发病率调查。发病率调查需要涵盖一个非常大规模的个人队列（例如超过 50 000 人），该队列在至少一年的时间内需要频繁地接受筛查，尽量避免出现失访现象，从而确保结果的有效性。^a

a. 痰涂阳肺结核的发病率无法从学校儿童的结核菌素调查结果中推导出来。虽然这是过去采用的一种方法，不过它最近受到了密切关注，而各项基本假设不再被认为令人满意。[van Leth F, van der Werf MJ, Borgdorff MW. Prevalence of tuberculous infection and incidence of tuberculosis: a re-assessment of the Styblo rule. Bull World Health Organ 2008; 86: 20-6.]

因此，发病率通常可由通报给国家结核病防治规划的常规结核病病例推导出来，尽管仍然存在着不确定性，包括已诊断但未被报告（例如通报给国家结核病防治规划）的病例数量以及完全未被诊断出来的病例数量（图 1）。

图 1
日历年内新发结核病病例与结核病病例通报之间的关系



在拥有高绩效结核病监测和卫生系统的高收入国家（例如英国或美国），病例通报系统可以捕获所有或几乎所有新发病例。

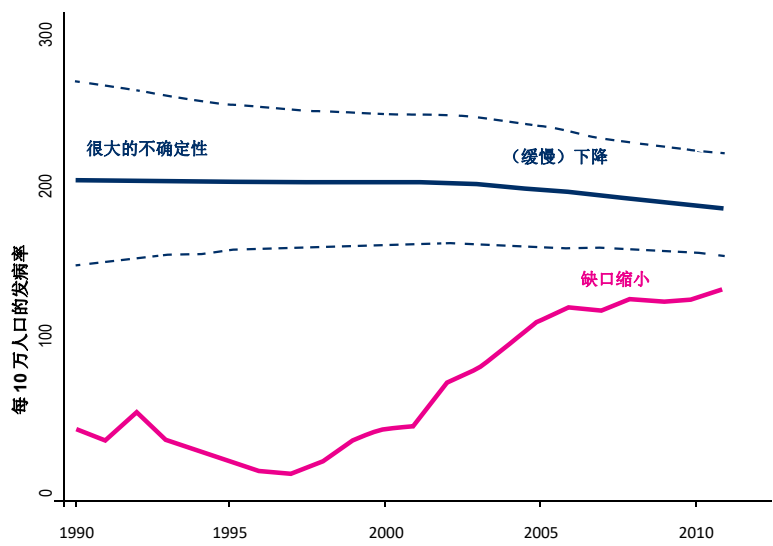
不过，在其他国家，常规病例通报数会提供有偏差的数据。

- 首先，私营部门更容易不报告新诊断病例，这是在亚洲和世界其他地区通常观察到的现象。
- 其次，并不是所有病例都会被诊断出来，尤其是在缺乏健康保险和社会保护或同等机制的国家，或者在卫生服务和实验室的覆盖面和绩效仍然较薄弱的国家。诊断水平低的常见原因包括：很大比例的现有结核病病例缺乏症状；患者自行判断症状不严重，因此没有求医；就诊者求医过程中没有被关怀服务提供方诊断出结核病。
- 第三，中低收入国家只有约一半的结核病报告病例最后被确诊为结核病病例——即通过细菌学确诊的病例。在许多国家，诊断规程主要局限于痰涂片镜检，痰涂片镜检只具备有限的敏感性。

因此，中低收入国家结核病发病率的水平和趋势是一项巨大的挑战。临时解决方案是利用现有数据（即病例通报）开展生态分析，提高对结核病流行决定因素的了解，以及更好地了解这些决定因素如何影响趋势变化。同时，我们还应当尽量增强结核病监测系统的绩效，以便覆盖所有卫生保健服务提供方，最大限度地降低漏报水平。虽然我们知道病例通报率并不接近于过去或现在的实际发病率，但我们期望结核病控制领域的各种改善应当在未来最终导向一个共同的目标。下面是印度尼西亚结核病发病率和病例通报率估计值的时间趋势示例（图 2），其中监测系统最近几年已经将规模扩展到包括大型公立和私立医院的病例通报数，这一举措已经减少了与发病率估计值之间的缺口。

图 2
印度尼西亚 1990 年至 2011 年结核病发病率（蓝线）与结核病病例通报率（红线）估计值的时间趋势。蓝色点线显示结核病发病率估计值的不确定区间。

资料来源：世界卫生组织结核病数据库



4.3 利用生态分析了解结核病流行情况

应当系统性地编制和评估有关推动结核病流行变化的已知因素的数据，这样有助于充分了解结核病负担的水平和趋势，尤其是在监测系统薄弱的国家，薄弱的监测系统会产生存在质量问题的数据。例如，为了评估结核病病例通报率能在多大程度上代表真实的结核病发病率，我们需要根据自己对推动结核病流行变化的各项因素的了解，确定观察到的病例通报趋势是否合理。我们会考查观察到的病例通报趋势与影响结核病流行情况的所有要素之间的逻辑一致性。在本章中，我们将这些影响要素称为推动力，因素或参数。表 1 总结了我们在生态分析范畴下开展的任务的类型以及各项任务的预期目的。

表 1
推荐的生态分析任务列表和各项分析的预期目的

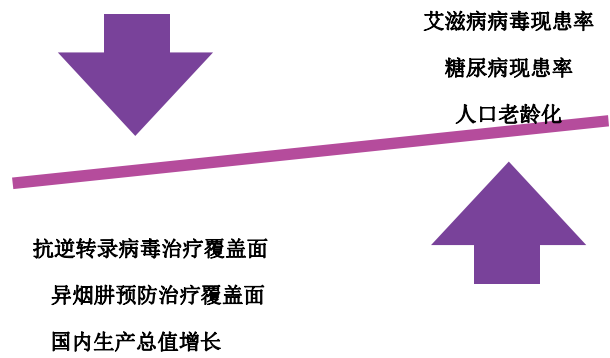
分析任务	分析目的
查阅发布的荟萃分析，如果没有，查阅可用的文献（证据），了解特定因素在结核病动态流行中的作用。	评估潜在关系的社会、经济、政治和生物 合理性 。 理解 关联的预期强度 。
观察病例通报的时间序列图（包括绝对数量和人均通报率）和已知与结核病相关的因素。	确定和解释观察到的趋势的变化。 评估各因素与病例通报之间的 临时关系 。 假设每个因素推动结核病负担的方向。
观察年龄、性别、区域和结核病病例类型之间的病例通报差异。	评估观察到的趋势的 一致性 。
观察推动力的变化程度，与病例通报的变化程度进行比较。	发现任何潜在的 剂量反应关系 。
与区域内具有类似特征的其他国家进行比较。	评估假设的趋势解释的一致性。

附录 4 详细列出了结核病生态分析中可以开展调查的各种因素的类型，以及相关的示例和推荐的解释。

4.4 生态分析的概念框架

为了帮助预期目标受众了解生态分析，我们建议提供一个概念框架，框架会指明针对主要变量之间的关系所做的基本假设。这些可以是非常简单的总结描述，了解专家对不同因素的作用的初步看法，例如图 3 中显示的概念框架。

图 3
关于结核病相关因素在哪个方向加剧疾病负担的概念框架



概念框架还可能会有许多更具细微差别的、多层次的假设，如附录 5 所示。

概念框架的设计越简单，就越容易描述您所探讨的积极关系和消极关系，并且有助于证明您选择的一系列因素的合理性。预期目标受众应当能够轻松地掌握生态分析的理论依据。虽然一些目标受众不需通过视觉方式就能很快认识到基本流行病学逻辑，但其他人可能需要清晰的概念框架才能更好地认识到消极影响和积极影响。

所有领域的特定关键信息都缺失时该怎么办？

鉴于并不是所有主题都具备全部所需数据（或以适当的单位或年份），有必要保持乐观态度 – 利用该时间段的已有数据和您掌握的分析单元，同时承认这些数据的局限性。

确定生态分析需要包括哪些要素无疑非常富有挑战性。流行病学家可能会就文献（例如荟萃分析）中发现的关联的相对强度提供建议。不过地方专家对领域（附录 4 中有相关解释）和指标的深入讨论应当可以指导优先次序的确定，因为结核病流行情况可能会与当地实际情况密切相关。一种方法是结核病防治专家集体讨论出可能会发挥作用的所有潜在因素的完整列表，然后商定适合探讨的重点潜在推动力的清单。下面是南亚一个国家的示例（表 2）。

表 2
影响结核病发病率的主要领域和因素的清单

领域	推动发病率下降的因素	推动发病率上升的因素
卫生部门层面保护和预防干预	改善的卫生系统绩效	有限的健康保险覆盖面
传播	过渡到门诊护理	快速城镇化加剧人口流动
易感性	经济不平等减少 吸烟减少	人口老龄化 艾滋病疫情蔓延 糖尿病发病率增加
筛查和诊断	DOTS 规模扩大（长期） 医防合作方法（长期）	DOTS 规模扩大（CNR） 医防合作方法（CNR）
结核病治疗	拓展型治疗中心	
记录和报告		

CNR - 病例通报率

请注意，虽然有些因素通过短期内提高病例通报率可以给结核病发病率带来明显的有益效果，但自相矛盾的是，这些因素会给人一种结核病负担再次加重的错觉。例如，DOT 规模扩大、医防合作、主动病例发现以及记录和报告改善会弥补实际发病率与病例通报率之间的差距，从而增加监测系统内接受关怀的结核病患者人数。

一旦界定了主要领域和推动力，应当尽可能多地编制相关的、及时的、有效的和以适当单位表达的支持数据。

如果生态分析是作为正式的规划和政策评估的一部分，我们建议重点考虑国家结核病防治规划和卫生利益相关方通过政策和实践变革可以积极影响的领域。确定超出预期目标受众控制范围之外的疫情推动力通常只能发挥有限的作用。

如果专家认为很重要的领域没有数据该怎么办？

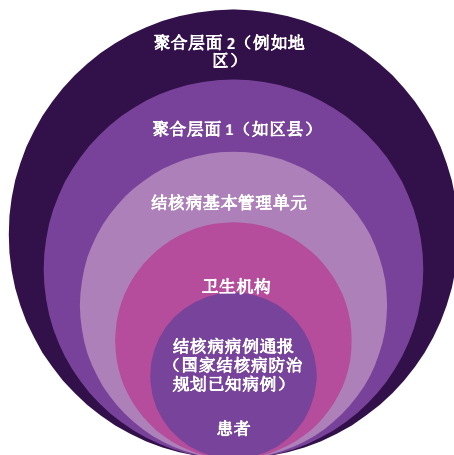
什么时候可以使用代理变量？如何确定代理变量是否合适？通常，我们想要衡量的变量无法从我们要求的聚合层面获得，或者无法在我们需要的时间段内获得。因此，我们经常依赖于代理变量提供帮助。合适的代理变量会与我们希望包括的衡量手段有着非常明确的紧密相关性。查询文献资料，了解其他场所采用的代理变量的优势。代理变量与想要衡量的推动力越不相关，就越不可能实际反映无法获得的数量。

4.5 准备分析所需数据

开始前，务必要冷静地评估选择要包括的数据的质量。避免太过老旧的数据（根据分析的目标需提前界定），还要避免与所选分析单位基本上没有任何对应关系的数据（例如预期分析单位是区县，但只具备聚合区域数据）。最重要的是，如果希望将因素趋势与病例通报率趋势关联起来，需要具备有关多个时间点的因素的信息。

在不相容的地域单位（图 4）之间整合数据非常耗费时间，但却是生态分析的一个必要步骤。如果希望使用的数据的呈现单位比已经选择的分析单位要低，则必须聚合数据。绝大多数统计软件包提供聚合指令。

图 4
不同层面的分析单位，从最低（个体患者）移到较高聚合地域层面

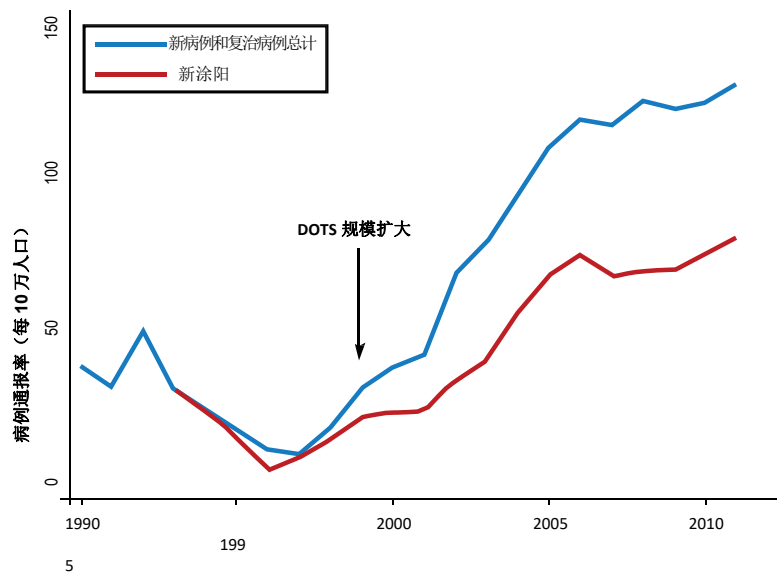


生态分析非常适用于以可视方式呈现结果，可以非常有效地将结果传递给各种各样的利益相关方。一旦数据被汇集起来，开始开展分析，以可视方式呈现与时间序列图和散布图的关系。很明显，分析的类型可能取决于可用的数据和每个国家的具体国情。

在本节中，我们提供一系列阐述性的示例，显示不同国家如何逐步地以图形方式呈现数据，从而对病例通报趋势进行了合理的解释。第一步可能是详细地查看不同时间段的病例通报情况，注意变化的方向、时机和斜率。

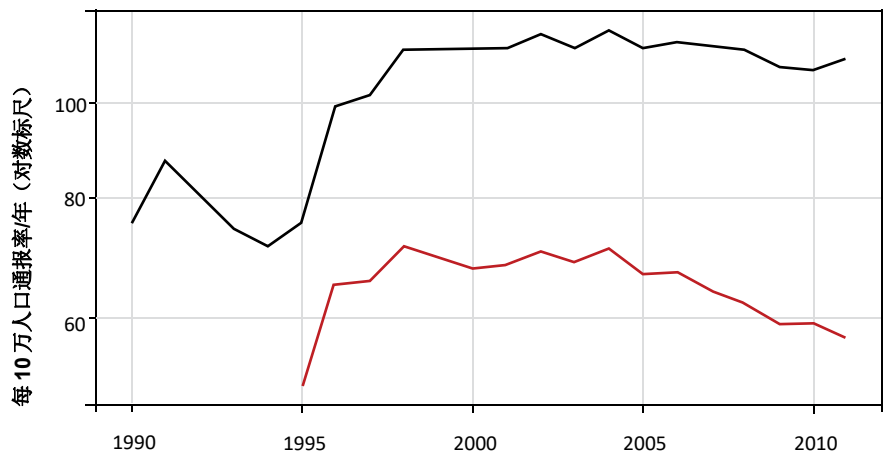
图 5 显示两个阐释性国家示例，强调三角测量和背景信息对于推导出有意义的结论的必要性。在过去十年里，印度尼西亚的病例通报率一直在稳定增长，一个箭头指示这种增长开始于 DOTS 策略被采纳之时。推动力与趋势之间的时机问题（也被称为时间关系）是一个很好的线索，有助于评估潜在的因果关联。我们期望在病例通报率发生变化之前先出现推动力变化。

图 5
印度尼西亚每 10 万人口病例通报率趋势（1990–2011 年）
资料来源：世界卫生组织结核病数据库



与此相反，越南的示例（图 6）揭示结核病报告病例通报率存在两个截然不同的轨迹。稳定的结核病通报率和下降的涂阳结核病病例比例在解释方面造成了困扰。

图 6
越南每 10 万人口病例通报率趋势，1990–2011 年（黑色：所有类型，红色：涂阳）
资料来源：世界卫生组织结核病数据库



查看所需期间的病例通报聚合趋势后，下一步是探讨年龄、性别和区域差异，如第 1 章所述。这既是一个数据质量保障步骤，也能弄清楚国家数据是否模糊了独特的推动力，这类推动力可能标志着国内存在区域动态或微观疫情。

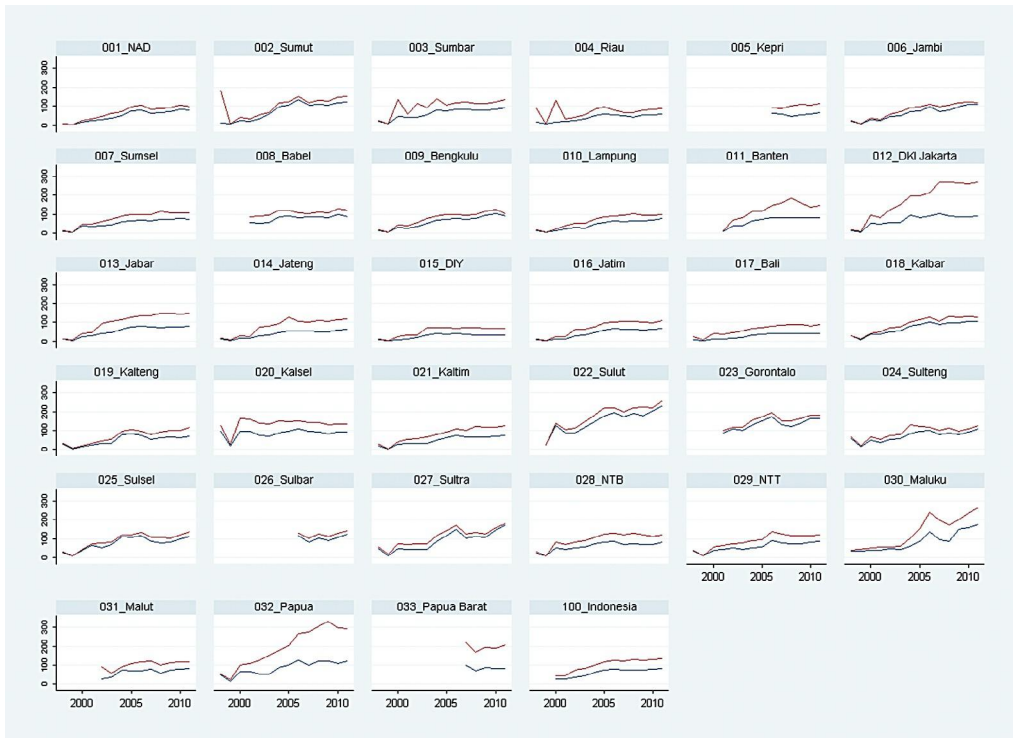
为了探讨时间序列数据中的地方差异，印度尼西亚的工作团队绘制了全部 34 个省的病例通报图（图 7）。这些时间序列数据的比较提示，虽然绝大多数省份出现了病例通报的普遍增长，但一些地区可能有着独特的疫情动态和疾病推动力。由于印度尼西亚是由群岛组成，这一点完全有可能存在。仍然需要调查省级层面上的病例通报率差异是否反映了疾病负担的差异或结核病发现和报告覆盖面的差异，或者是否同时反映了这两方面的差异。报告中的错误（包括错误的地域差异）也不能被排除在外。

在已经完成了全国结核病现患率调查的国家，为了假设成年人肺结核现患率水平和结核病病例通报中表达的地方疾病负担水平之间的关系，一种方法是将它们绘制成图（图 8）。

应当注意到，现患率调查不能足够有效地估计州级层面的现患率，不过它仍然是生成假设的一项有用练习，而且能进一步补充州级层面的现场调查。

图 7
印度尼西亚所有类型结核病（红色）通报率和涂阳（蓝色）结核病通报率省级和国家（100_印度尼西亚）层面时间序列（1998–2011 年）

资料来源：国家结核病防治规划 印度尼西亚



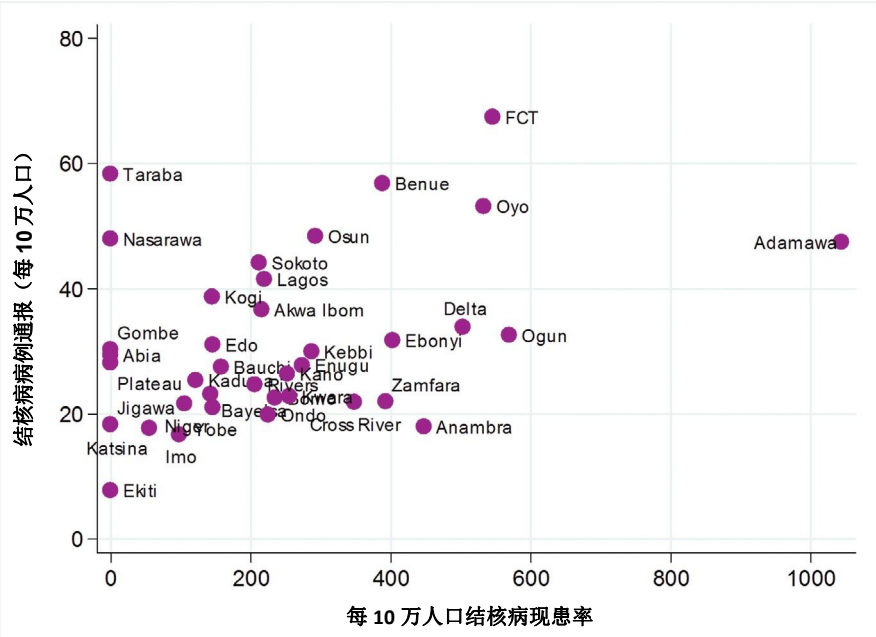
可以探讨多少影响要素？单个图形或解释模型中可以包括多少推动力？

只要具备高质量的数据，可以完成尽可能多的探索型图形和比较。不过，为了探讨单一解释模型内大量因素对病例通报的影响，您可能发现需要咨询一名统计学家，帮助找到适当的统计方法，以便克服众所周知的“大 p 小 n 问题”。该问题指面临着许多影响因素但却没有多少观察结果（例如年份或基本管理单元）。例如，一个小国有 34 个卫生区、13 个省和 4 个地区，则最大的样本量是 34，如果国家结核病防治规划想要了解几十个影响区县病例通报率的潜在要素，该样本量可能会太少。

生态分析的挑战和局限性

生态分析是一项宏观层面的工作，尝试解释复杂系统的形态，并且混杂和协方差通常也是一种风险。生态分析在一些圈子里一直有着不好的名声，因为它在很大程度上依赖于平均值和比例，这两类变量被认为无法完全反映特定地点或时间的可变性。这种过分简化的模式使得生态分析非常有吸引力和容易利用，但同时也容易出现错误。虽然结核病的主要推动力不太可能被忽略，但关键要素通常无法衡量或不可用。给所有相同分析单位和时间段找到足够多的数据几乎总是一项挑战。如果缺乏有关政策实施和项目规划扩展步伐的详细监测数据，则很难适当地将转归归因于工作。

图 8
尼日利亚州级层面成年人肺结核现患率和病例通报率散点图
(2012)
资料来源：尼日利亚国家结核病防治规划数据库和第一次全国结核病现患率调查



4.7 结论

可靠的生态分析可以帮助有效地确定不同时间段和不同地理区域影响结核病流行情况的大量要素。有时，分析过程不会得出单一的简单陈述，而是多个合理的和一致的解释。系统性地详细审查影响结核病流行情况的各类因素的主要要素，以及谨慎地应用真实性标准，这样可以帮助国家结核病防治规划弄清楚在结核病控制领域所做的努力是否产生了影响，以及如何扩大对结核病疫情的这种影响。

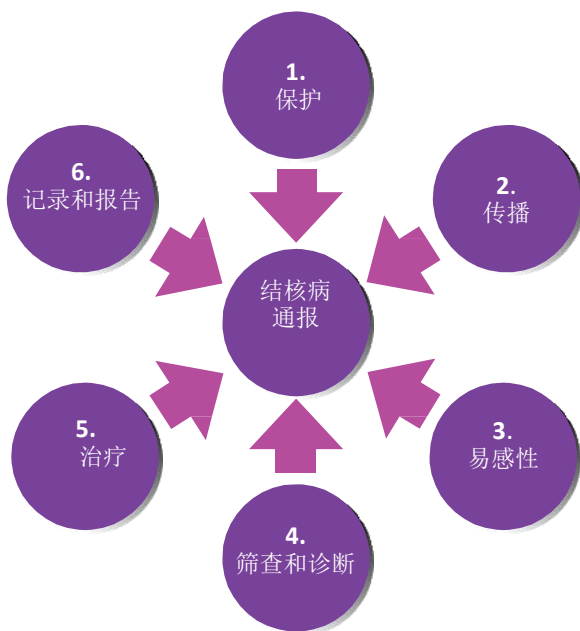
1. Chandra S et al. Resurrecting social infrastructure as a determinant of urban tuberculosis control in Delhi, India. *Health Research Policy and Systems*. 2014;12:3.
2. Basu S, Stuckler D, McKee M. Addressing Institutional Amplifiers in the Dynamics and Control of Tuberculosis Epidemics. *American Journal of Tropical Medicine and Hygiene*. 2011;84:30–37.
3. Stuckler D, Basu S, McKee M, Lurie M. Mining and Risk of Tuberculosis in Sub-Saharan Africa. *American Journal of Public Health*. 2011;101:524–530.
4. Stuckler D, King LP, Basu S. International Monetary Fund programs and tuberculosis outcomes in post-communist countries. *Plos Medicine*. 2008;5:1079–1090.

附录 4

结核病生态分析中应当调查哪些类型的数据？

有效的生态分析并不需要包括影响结核病疫情的所有因素，不过至少应当考虑使用下面六个领域的度量指标（图 A1）。

图 A1
影响结核病流行情况的各大领域的示意图



1. 卫生部门层面保护和预防干预

这些属于公共卫生干预，旨在预防结核病感染和保护感染者不进展到活动性疾病。这些因素被策略性地包括在生态分析中，因为它们都是在卫生部门可控范围内的因素。因此有关这些因素的洞察力可以直接被转化为改善的政策和实践。下面是可以被包括在生态分析中的各种因素的阐述性列表：

- 聚集区域的行政管理、环境和个人感染控制措施。
- 艾滋病病毒感染者中的抗逆转录病毒治疗覆盖面。
- 潜伏结核病感染治疗的覆盖面。
- 卡介苗接种覆盖面。
- 活产的五岁以下儿童死亡率（图 A2 和图 A3）。

图 A2

越南每 1000 活产中五岁以下儿童死亡率的时间趋势，1990–2010 年（对数标尺）。五岁以下儿童死亡率此处被用作卫生系统绩效的一项总体指标。在过去二十年间，该死亡率已经有了大幅下降，2011 年达到的水平相当于二十世纪六十年代中期在西欧观察到的水平。该项因素预期能推动结核病负担下降。

资料来源：世界银行

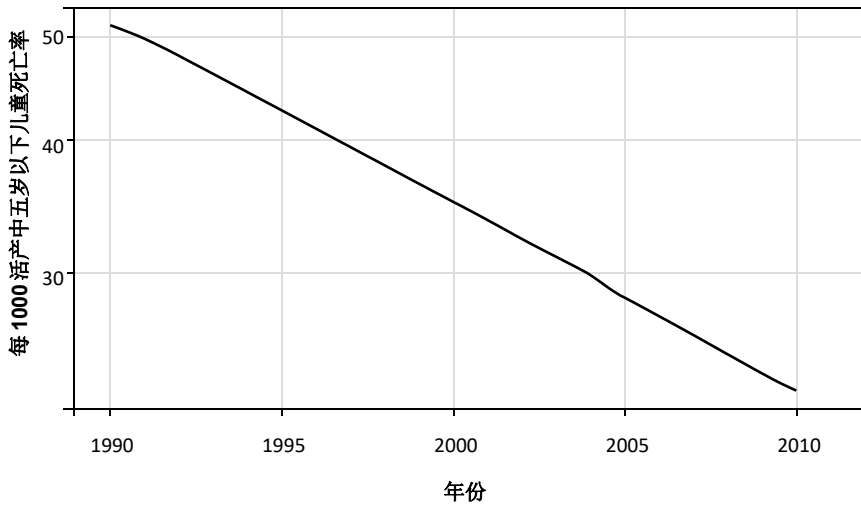
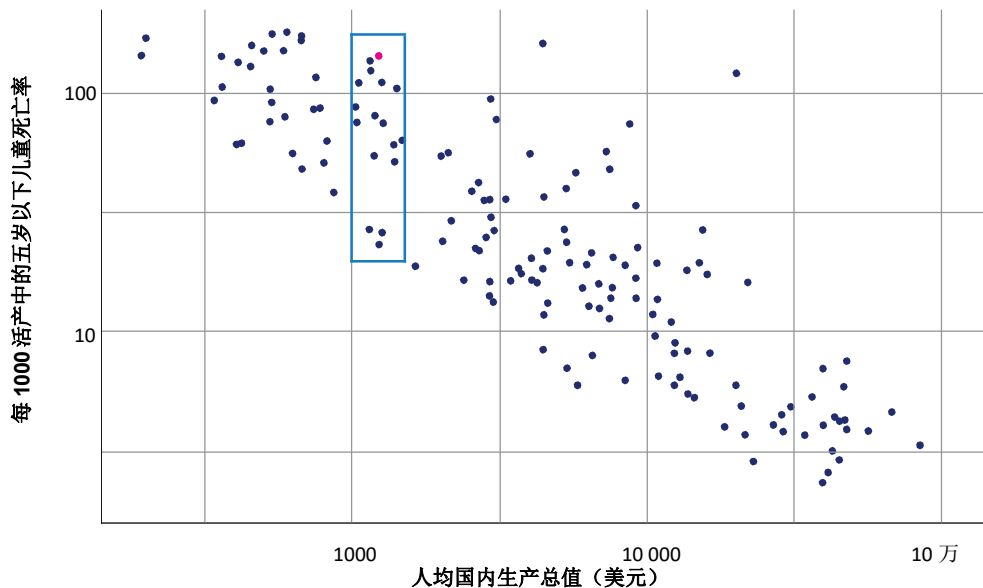


图 A3

每 1000 活产中五岁以下儿童死亡率相对人均国内生产总值的散点图（2010）。每个蓝点代表着一对国家数据点。尼日利亚被显示为红色。长方形内的国家的人均国内生产总值为 1000 到 1500 美元。这提示着尼日利亚卫生系统的绩效水平（通过五岁以下儿童死亡率指标衡量）比按照经济规模预期的水平要低。

资料来源：世界银行



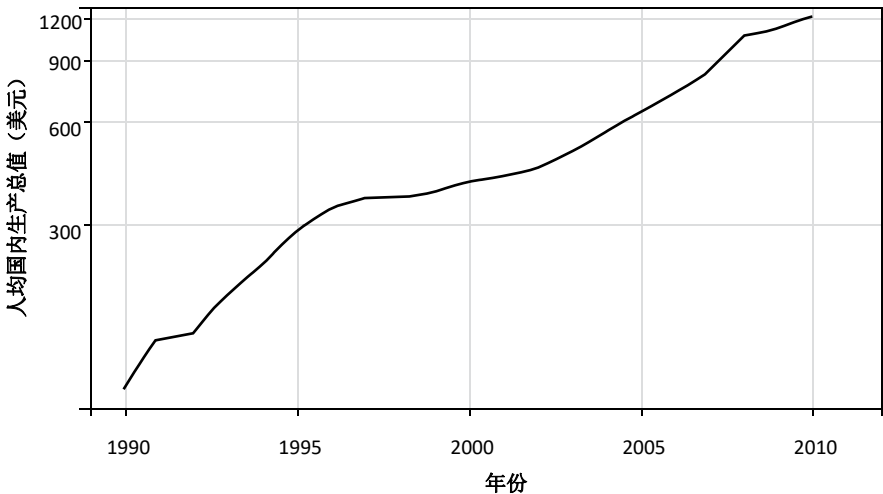
2. 传播

这些因素通常与城镇和工业设计、住房政策、劳教和拘留实践等有关。这些通常都不属于国家结核病防治规划的直接影响范围。不过，它们可能会在结核病流行情况发挥着强大的作用（2,3）。

结核病传播中发挥关键作用的因素包括：

- 监禁率。
- 监禁的平均时间（量刑标准）。
- 深坑硬岩采矿作业的覆盖面。
- 每间卧室或每个单元的平均人数。
- 人口密度（每平方米的人数）。
- 居住条件不佳（拥挤）的人口比例。
- 经济规模，例如通过人均国内生产总值衡量（图 A4）。

图 A4
越南人均国内生产总值的时间趋势（对数标尺），1990–2010 年。
过去二十年内，经济每年都以基本稳定的增长率在不断扩展。二十世纪九十年代后期的亚洲经济危机只产生了有限的影响。该因素预期能推动结核病负担下降。



3. 易感性

人们对结核病感染的易感性和罹患结核病是一个广泛的主题，我们对易感性的许多了解都来自生态分析（4）。这些指标可能是通过损害免疫系统与结核病易感性有着非常直接的关联（例如 HIV 和糖尿病），或者它们可能是源自不同形式社会排斥的现实易感性的上游指标，例如文盲、边缘化以及基于收入、性别、宗教、种族和国籍的不平等。反

- 健康保险和社会保护项目的覆盖面，个人自付额占卫生保健支出的百分比。
- 艾滋病病毒感染现患率（图 A5）。
- 糖尿病现患率。
- 吸烟、酗酒、注射吸毒和室内空气污染的发生率。
- 消瘦、生长迟缓和营养不良的发生率。
- 基尼系数。
- 识字率。
- 性别不平等指数。
- 失业率。
- 境内流离失所或无家可归的发生率。
- 矽尘暴露发生率。
- 既往结核病和纤维化病变现患率。
- 人口统计学变化，例如人口老龄化（图 A6）。

图 A5
尼日利亚 15–49 岁人群中艾滋病病毒现患率的时间趋势，1990–2011。预期 1990 年至 2000 年推动了结核病负担上升，自此后一直稳定。
资料来源：AIDSinfo 数据库

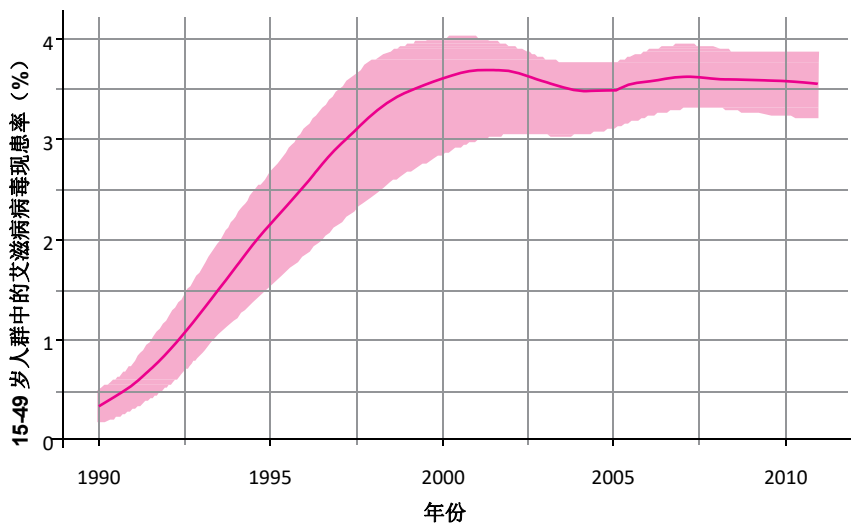
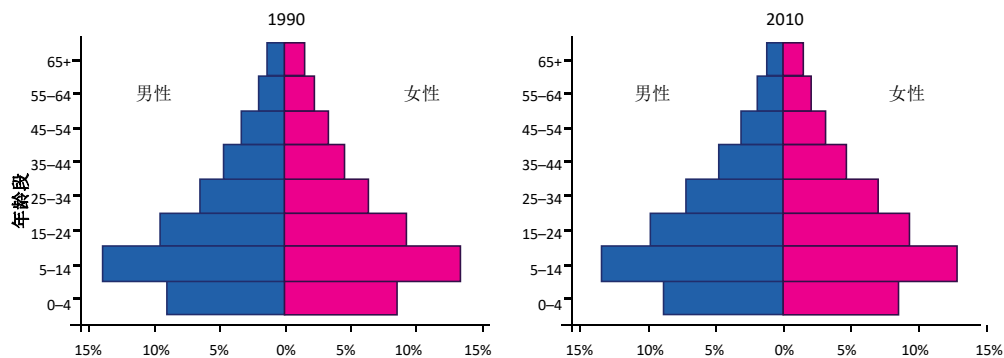


图 A6
1990 年和 2010 年代表人群按年龄和性别分布的一般人群人口金字塔。在过去二十年内，尼日利亚的人口统计学一直保持稳定。根据国内结核病负担在各年龄段的分布，人口统计情况可能会推动结核病疫情上升或下降。
资料来源：联合国人口司数据库



4. 筛查和诊断

筛查和诊断指标指卫生系统内影响求医行为以及结核病症状得到识别和疾病得到诊断概率的所有因素。这些指标相对简单，而且在国家结核病防治规划范畴内很容易得到改善，因此这些指标非常值得被包括在致力于考查结核病病例通报差异和结核病通报病例分布（肺部与肺外）的生态分析中。一些常用指标包括：

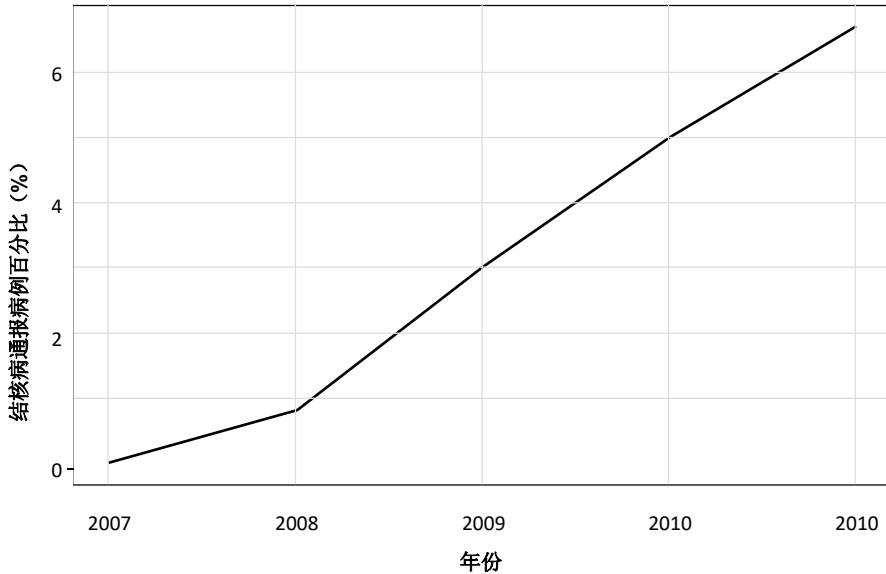
- 没有医疗保险的人口比例 – 卫生保健覆盖面不足 – 获得免费结核病诊断。
- 每 10 万人口中接受结核病筛查的平均人数。
- 高风险群体的结核病筛查率。
- 调查中报告的结核病症状知晓率（例如人口统计学健康调查结核病单元）。
- 调查中报告的长期咳嗽、发热和其他结核病症状的求医行为。
- 每 10 万人口中的结核病调查率（提交的样本数量/10 万人）。
- 每 10 万人口的胸部 X 线检查平均数量。
- 每 10 万人口的 GeneXpert 设施的平均数量。
- 每 10 万人口的诊断中心平均数量。
- 培养实验室和/或药敏试验实验室的平均数量。
- 每 10 万人口的护士、医生和/或放射科医师平均人数。
- 艾滋病病毒感染者污名指数，对求医行为的影响。

• 医防合作/社区/活动性病例发现倡议行动的覆盖面和绩效。下面是一类时间序列数据的示例，这些数据被编制来书面记录诊断能力和覆盖面的扩展（图 A7）。

图 A7

医防合作和主动病例发现对结核病病例通报的促进作用，表示为越南结核病病例通报总数的百分比。

资料来源：越南国家结核病防治规划



5. 结核病治疗

治疗措施也需要在生态分析中着重考虑，因为结核病治疗对疾病动态的良好效果已经得到证明。生态分析采用的一些结核病治疗措施包括：

- 治疗成功率（图 A8）。
- 平均痰转化率。
- 结核病治疗支持率。
- 结核病患者死亡率。
- 直接观察治疗的覆盖面。
- 药品管理指标。

图 A8 显示了不同时间段总体治疗成功百分比的示例。虚线代表全球 85% 的治疗成功目标值，2011 年首次实现该目标值。

图 A8
尼日利亚国家新涂阳病例通报队列的治疗成功（治愈或治疗完成）百分比，1996–2011 年
资料来源：世界卫生组织结核病数据库

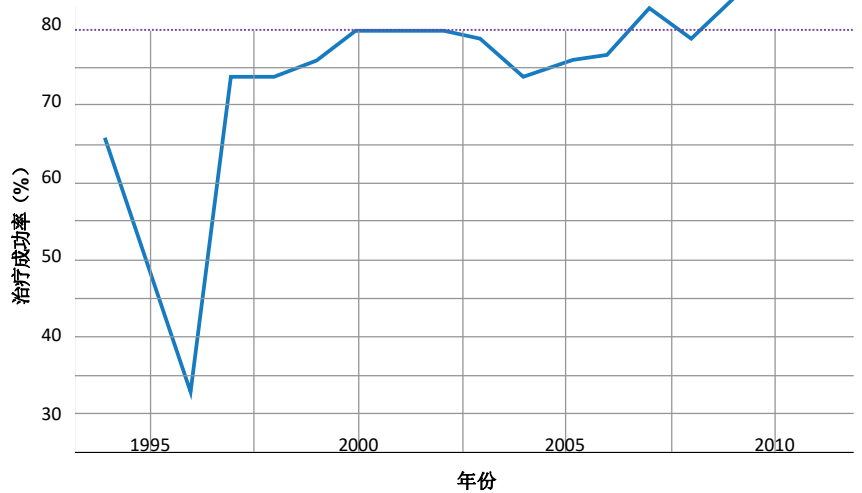
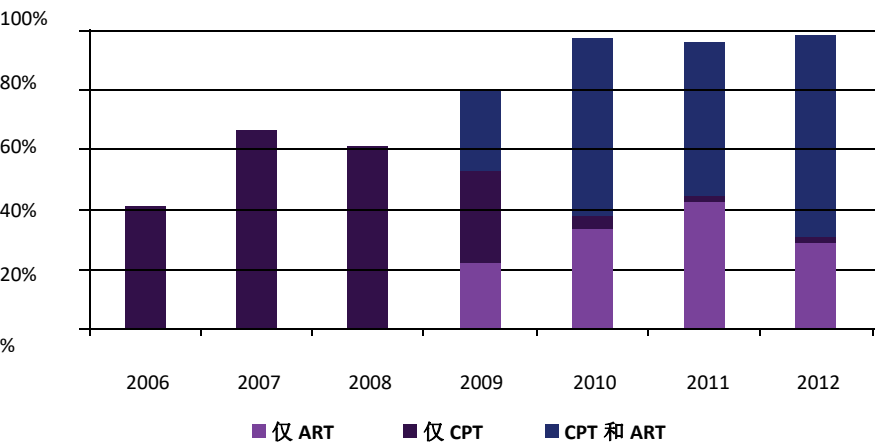


图 A9 显示了莫桑比克一个省份在不同时间段给合并感染结核病患者扩大辅助艾滋病关怀服务的示例。可以对这些数据进行叠加分析，以便探讨队列中的治疗转归。

图 A9
莫桑比克加扎省艾滋病病毒阳性结核病患者中复方新诺明预防（CPT）和抗逆转录病毒治疗（ART）覆盖面扩大的时间序列，2006–2012 年
资料来源：莫桑比克国家结核病防治规划

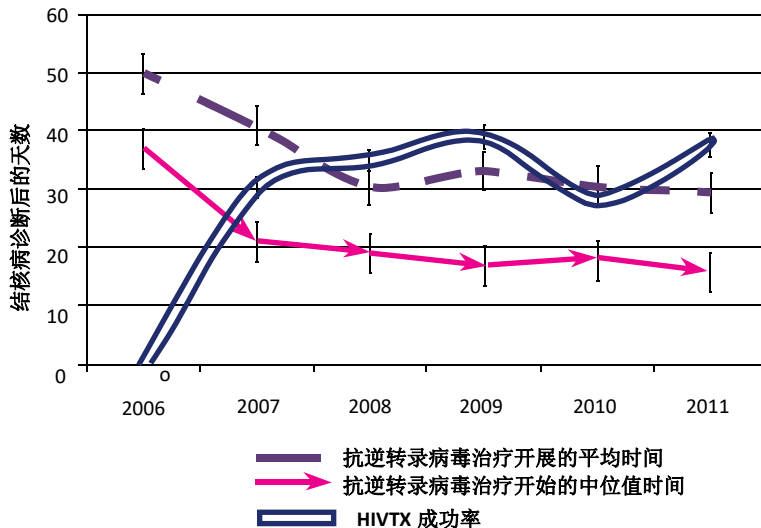


偶尔情况下，还值得考查诊断后治疗的时间延迟，以此了解延迟在关怀提供中可能扮演的角色。这些图合起来显示了覆盖面的上升和抗逆转录病毒治疗提供及时性的初步提高。这些趋势似乎与艾滋病病毒阳性结核病患者中的结核病治疗成功率上升相吻合（图 A10）。

图 A10

莫桑比克与结核病治疗成功相关的抗逆转录病毒治疗提供的时间轴，2006-2012 年

资料来源：莫桑比克国家结核病防治规划 [ART：抗逆转录病毒治疗；HIVTx：艾滋病病毒阳性结核病患者中的艾滋病治疗成功率]



6. 记录和报告

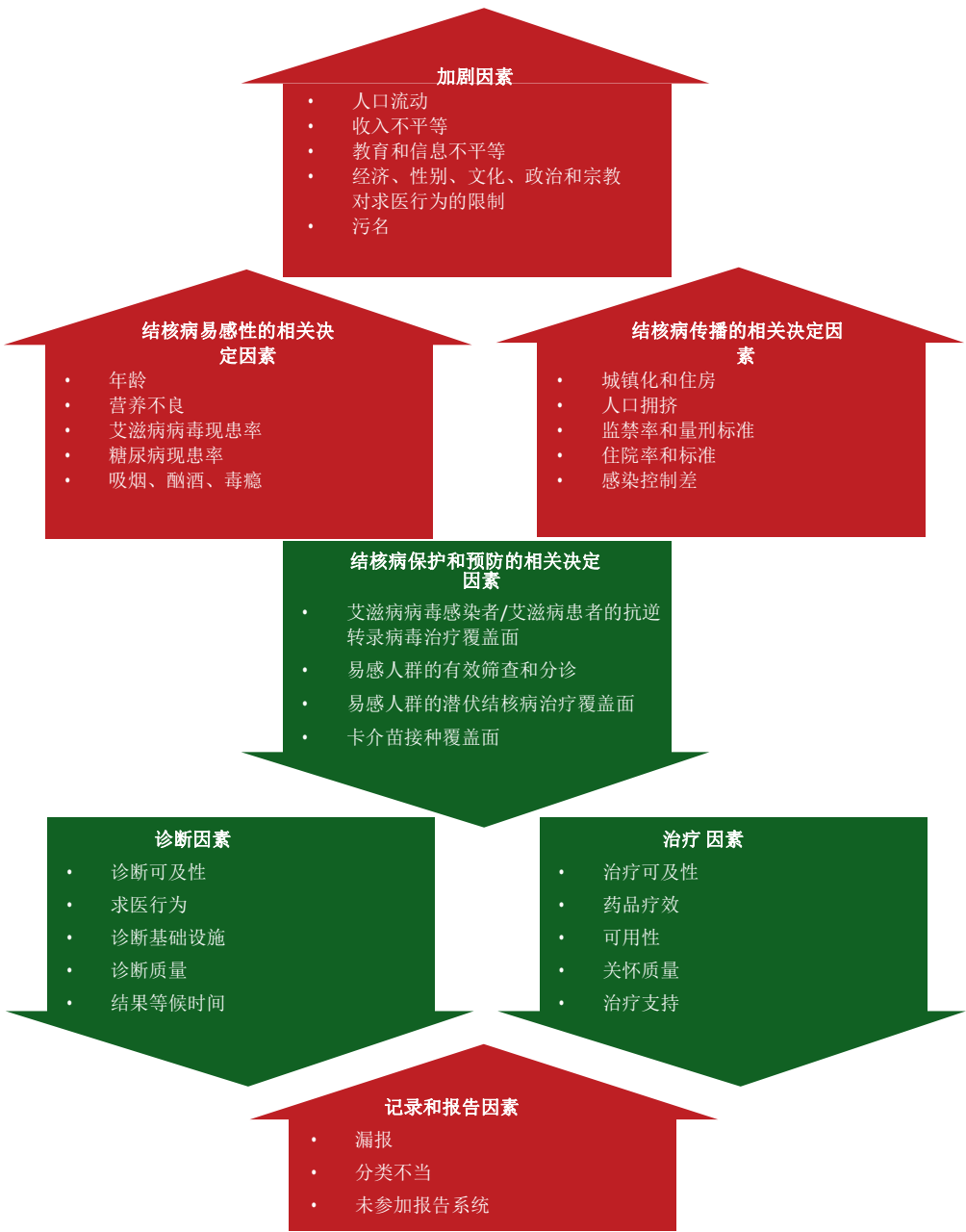
结核病病例记录和报告方式的变化通常对通报的数据的数量和质量有着深远的影响。这可能会使得生态分析的解释变得复杂。因此，始终应当密切关注记录和报告的任何变化，确保再次分类、从纸面过渡到电子系统和其他记录保持变革不会增加对基本结核病流行情况的理解难度。要考虑的一些潜在因素包括：

- 漏报水平的（例如病例存量研究的结果）。
- 基本管理单元的数量和/或地区的变化。
新记录和报表被引入的日期或病例定义变化生效的日期。
- 强制报告政策的变革。
- 私营服务提供方在报告系统中的参与率。
- 被利用世界卫生组织病例相关结核病监测数据标准和基准核查清单和人口动态登记系统开展结核病监测数据系统性评估的结果（第 1 章）。
- 电子系统被大规模采用的日期。

一些专家还包括第七个领域，即**政府和国际资助机构对结核病控制的经费支持**。这是一个重要的领域，无疑是结核病控制工作的一个强大的推动力，不过一些人认为它属于“上游”影响，许多其他领域都可受到影响。

附录 5
示例：关于各因素如何影响结核病负担的详细概念框架

（红色代表加重结核病负担的因素，绿色代表减轻结核病负担的因素。）



第 5 章

耐药结核病：负担和防治措施分析

目标受众：

本章旨在指导负责国家和地方层面结核病防治规划督导与评估的人员以及技术合作伙伴，以期改善耐药结核病的监测。本章还致力于协助相关人员设计应用性研究。

预期成果：

本章结束时，读者应当能够了解：

- 耐药结核病监测和调查以及疾病流行病学和相关疾病负担数据分析原则，包括疾病的地域分布（即按年龄和性别的分布和特定人群内的分布）。
- 耐药结核病防治规划绩效的主要指标，利用和解释规收集的病例发现、治疗开始和治疗转归数据。

作者：

Dennis Falzon、Anna Dean、Matteo Zignol

定义

通过专门调查或持续监测分别收集新（既往未经治疗）结核病病例和既往治疗结核病病例的耐药监测数据。理想情况下，专门调查每三到五年重复一次，衡量具有代表性的细菌学确诊结核病通报病例中的耐药情况。持续监测系统对所有细菌学确诊结核病患者开展常规诊断性药敏试验。（1）^a

用来评估规划绩效的主要定义和指标已在其他文献中阐述（2–4）。监测结核病防治规划中的耐药结核病病例时，重点放在感染了体外利福平耐药菌株的病例上（利福平耐药结核病）。这些病例可能是单耐药结核病、多耐药结核病、耐多药结核病或广泛耐药结核病。^b

数据源

在耐药调查中，主要数据源是设计用于捕获临床和实验室数据的调查数据库。

在持续耐药监测和耐药结核病项目规划管理（PMDT）绩效评估中，主要数据源包括如下：

- 基本管理单元（BMU）结核病登记簿（或区县结核病登记簿）。
- 二线结核病治疗登记簿（既往 IV 类登记簿）。
- 培养、Xpert MTB/RIF 和药敏试验实验室登记簿。
- （一线或二线）结核病治疗卡。

这些数据源只有纸面格式可供使用，不过它们正逐渐被电子系统取代（5）。这项发展成果极大地提高了管理数据、推导出必要指标和开展更频繁分析的效果。患者精确位置分布图绘制正获得越来越多的关注，有助于确定耐药结核病热点和改善服务（6），

-
- a. 文件 *Guidelines for surveillance of drug resistance in tuberculosis. 4th ed.*（《结核病耐药监测指南》第四版）（WHO/TB/2009.422）。详细描述本章中提到的有关专门调查设计、实施和数据分析的绝大多数概念（1）。
- b. 与其他结核病患者相比，利福平耐药结核病患者需要更长时间的治疗，因此监测时间也更长，而且有一组特定的诊断和转归指标。

而且越来越多的手机具备了全球定位系统坐标功能，因此地理信息系统（GIS）数据可能会在常规结核病关怀和预防中变得更加重要。在完成 3 期试验之前，新抗结核病药物就被投放用于耐多药结核病患者，这意味着药物警戒注定要成为人会常规监测系统的一个重要组成部分。这方面已经开发了专门的数据收集问卷和软件（7）。

描述和分析数据的方法

描述和分析结核病耐药数据的常用方法包括：

- 以标准指标呈现的简单描述性统计，如频次表，以及基于这些统计的图形（最常见的线形图和柱状图用于描述时间趋势和比例，分布图用于描述国内差异）。文字框 1 显示了一个国家示例。
- 检验与耐药决定因素和治疗转归风险因素相关的假设。这些情况下最常采用的技术在本章的众多国家示例中进行了阐述。这些示例主要源自报告给世界卫生组织的国家监测数据，旨在显示如何以及可以在多大程度上利用这些数据，应当注意一些数据集在准确性方面的局限性。相同模式和相关解释可以适用于没被包括的一些其他国家。

专门调查

分析专门调查数据时，应当考虑下面四项基本描述性分析。

患者招募。调查中每个诊断中心或群集包括的患者人数应当与预期通报病例数量进行比较，比较应当基于抽样方法，按治疗史分解。例如，如果新病例的目标群集大小是 40，则每个群集中招募的新病例的实际数量应当与目标大小 40 进行比较。如果招募的新病例数量比目标要少，应当采取措施改善患者参与，最终达到群集目标大小。

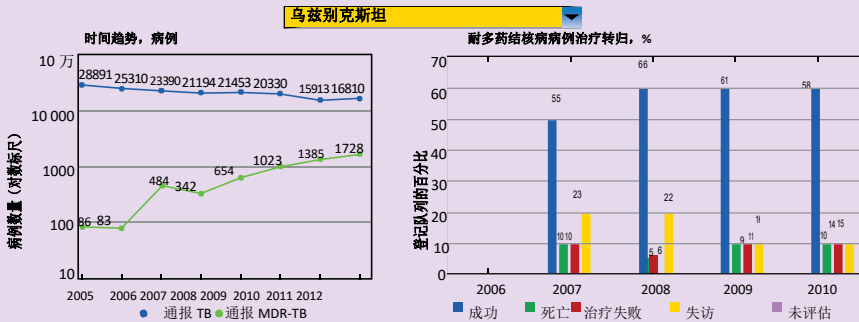
缺失数据。计算异烟肼和/或利福平耐药数据缺失的病例所占的百分比，按年龄段、性别、治疗史、群集和调查结果中任何其他重要变量进行汇总。

文字框 1
国家耐多药结核病数据汇总指标“一览表”

简单图形和表格等手段以可视方式呈现关键指标非常有帮助作用。电子记录和报告有着很多优势，包括灵活性和轻松地实现可视性呈现。最近几年，世界卫生组织全球结核病项目已经开发了一个一览表，直接从世界卫生组织全球结核病数据收集系统将国家或区域指标集合起来（参见图 1）。各国结核病防治规划可以按照国内重点自定义设计这类工具，同时根据需要添加其他特定信息。

图 1
耐多药结核病关键指标汇总
资料来源：世界卫生组织全球结核病数据库，
www.who.int/tb/data

耐多药结核病诊断、通报和治指标，按地区或国家和年份划分



	新病例				既往治疗病例			
2012 年肺结核通报病例中的耐多药结核病病例估计数量	2400 (1800–3000)				1600 (1400–1900)			
	2005	2006	2007	2008	2009	2010	2011	2012
结核病通报病例	28 891	25 310	23 390	21 194	21 453	20 330	15 913	16 810
通报耐多药结核病病例（绝对数）	86	83	484	342	654	1023	1385	1728
通报耐多药结核病病例（每 10 万人口通报率）	0	0	2	1	2	4	5	6
开始接受耐多药结核病治疗的患者					464	628	855	1491
具有治疗转归记录的耐多药结核病病例			330	294	464	628		
具有药敏试验结果的新结核病病例的比例	0	3	6	5	12	60	10	52
具有药敏试验结果的既往治疗结核病病例的比例	5	4	10	9	30	26	11	39

* 如果具有一线药敏试验结果的结核病病例的比例超过，则显示为 100%；如果结核病通报不完整，可能会出现这种情况，尤其是在结核病病例报告与药敏试验没有衔接的系统中。

* 招募结束后两年报告的耐多药结核病的转归。患者可能没有获得符合国际推荐的方案或规范的治疗。

本页上的指标旨在监测国家在实现“全球 2011-2015 年终止结核病计划”中的关键耐药目标值方面取得的进展情况。

指标	2015 年目标值
具有药敏试验结果的新细菌学阳性结核病报告病例的比例	≥ 20%
具有药敏试验结果的既往治疗结核病报告病例的比例	100%
成功完成治疗的确诊耐多药结核病病例的比例	≥ 75%
报告了转归的确诊耐多药结核病病例的比例	100%

耐药模式。应当制作一个表格，描述对单种药物/多种药物组合耐药的患者的比例。这些信息应当根据新患者和既往治疗患者进行划分。分析中可以包括/排除缺失数据的多重填补（表 1）。然后可以观察国内不同地区耐多药结核病患者比例的差异（文字框 2）。

耐药决定因素。耐药结核病病例通常按照年龄段、性别、艾滋病病毒感染状况和调查期间收集的其他患者人口统计学和临床数据进行分层。表 2 显示了白俄罗斯最近开展的一项调查的结果，其中 32% 的新病例和 76% 的既往治疗患者有耐多药结核病（8）。在这项分析，耐多药结核病被发现与既往治疗、监禁、残障导致的失业、吸烟、酗酒和艾滋病病毒感染独立相关。一些关联可能依赖于耐多药结核病的因果途径（例如既往治疗或监禁），而其他管理可能代表着慢性疾病对社会状况的影响（例如残障）。

表 1
一线抗结核病药物耐药模式示例

数量 (% [95% CI])			
耐药模式	新病例 (n=1 050)	既往治疗病例 (n=291)	总计 (n=1 344)
对所有药物敏感	87.7 (84.0-90.6)	56.8 (50.5-62.9)	81.3 (77.3-84.7)
任意耐药			
- 对 H 任何耐药	5.3 (4.0-7.0)	35.8 (29.9-42.0)	11.7 (9.1-14.8)
- 对 R 任何耐药	1.6 (0.9-2.9)	28.9 (23.9-34.4)	7.3 (5.2-10.0)
- 对 E 任何耐药	0.9 (0.4-2.1)	17.8 (13.5-23.3)	4.4 (3.2-6.0)
- 对 S 任何耐药	9.9 (7.4-13.0)	33.1 (27.1-40.0)	14.7 (11.8-18.0)
任何耐药总计	12.4 (9.4-16.1)	43.2 (37.1-49.5)	18.8 (15.3-22.8)
单耐药			
- 对 H 单耐药	1.5 (0.8-2.7)	2.5 (1.2-5.2)	1.7 (1.1-2.7)
- 对 R 单耐药	0.2 (0.0-1.0)	0.4 (0.0-2.7)	0.3 (0.0-1.1)
- 对 E 单耐药	0.2 (0.0-0.8)	0.0 (-)	0.2 (0.0-0.6)
- 对 S 单耐药	6.6 (4.4-9.8)	7.1 (4.8-10.3)	6.7 (4.8-9.3)
单耐药总计	8.5 (6.0-12.0)	10.0 (7.3-13.5)	8.8 (6.6-11.6)
耐多药			
- HR	0.4 (0.1-1.2)	4.3 (2.2-8.2)	1.2 (0.6-2.4)
- HRE	0.0 (0.0-0.7)	3.0 (1.8-5.0)	0.7 (0.4-1.3)
- HRS	0.4 (0.2-1.1)	7.0 (5.1-10.0)	1.8 (1.1-2.8)
- HRES	0.5 (0.2-1.3)	14.1 (9.9-19.7)	3.3 (2.2-4.9)
耐多药总计	1.4 (0.7-2.5)	28.5 (23.5-34.1)	7.0 (5.0-9.6)
多耐药			
- HE	0.1 (0.0-0.8)	0.0 (-)	0.0 (0.0-0.6)
- HS	2.4 (1.5-3.7)	4.0 (2.2-7.2)	2.7 (1.9-3.9)
- ES	0.0 (-)	0.0 (-)	0.0 (-)
- HES	0.0 (-)	0.7 (0.2-2.9)	0.2 (0.0-0.6)
- RE	0.0 (-)	0.0 (-)	0.0 (-)
- RS	0.0 (-)	0.0 (-)	0.0 (-)
- RES	0.0 (-)	0.0 (-)	0.0 (-)
多耐药总计	2.5 (1.6-3.9)	4.7 (2.6-8.5)	2.9 (2.1-4.2)

H=异烟肼；R=利福平；E=乙胺丁醇；S=链霉素；CI=置信区间

表 2
白俄罗斯结核病病例中耐多药结核病的风险因素，2010–2011 年

特征	病例		一元分析 ¹		多元分析 ¹	
	人数	耐多药结核病比例	优势比 (95%CI)	p	优势比 (95%CI)	p
性别						
男性	1075	47.6	对照组	-	对照组	
女性	269	37.2	0.7 (0.5-0.8)	0.000	1.1 (0.8-1.5)	0.657
年龄 (岁)						
15-34	328	49.7	对照组	-	对照组	
> 34	1016	44.2	0.8 (0.6-1.0) ²	0.046	0.7 (0.6-1.0)	0.021
出生国家						
白俄罗斯	1293	45.3	对照组	-	-	
其他	51	51.0	1.3 (0.8-2.0)	0.360	-	
治疗史						
新病例	934	32.3	对照组	-	对照组	
既往治疗病例	410	75.6	6.5 (5.2-8.2)	0.000	6.1 (4.8-7.7)	0.000
教育水平						
大学	43	34.9	对照组	-	-	
大专	345	47.3	1.7 (0.9-3.0)	0.080	-	
中学	887	46.6	1.6 (0.9-2.8)	0.087	-	
小学或更低	69	30.4	0.8 (0.4-1.7)	0.573	-	
生活条件						
自己家里	1104	45.7	对照组	-	-	
出租房里	185	43.8	0.9 (0.7-1.2)	0.587	-	
宿舍	29	37.9	0.7 (0.4-1.4)	0.345	-	
无家可归	26	61.5	1.9 (1.0-3.8)	0.069	-	
住户人数 (家庭成员数)						
1	329	49.9	对照组	-	对照组	
2	543	46.0	0.9 (0.7-1.1)	0.209	1.0 (0.7-1.3)	0.758
3	287	42.9	0.8 (0.6-1.0) ²	0.046	0.8 (0.6-1.1)	0.184
>3	185	40.5	0.7 (0.5-0.9)	0.020	0.8 (0.6-1.2)	0.266
就业状况						
用工作	401	42.4	对照组	-	对照组	
失业但体格健全	659	50.5	1.4 (1.1-1.7)	0.003	1.1 (0.8-1.4)	0.578
退休	175	27.4	0.5 (0.4-0.7)	0.000	0.7 (0.5-1.0)	0.052
残障导致的失业	87	59.8	2.0 (1.3-3.0)	0.001	1.9 (1.2-3.0)	0.010
学生	22	40.9	0.9 (0.4-2.0)	0.874	1.4 (0.6-3.2)	0.417
监禁史						
否	1145	42.2	对照组	-	对照组	
是	199	64.8	2.5 (1.9-3.3)	0.000	1.5 (1.1-2.0)	0.009

CI=置信区间；MDR-TB=耐多药结核病；OR=优势比。

¹所有决定因素都通过逻辑回归进行了研究。一元分析中 p 值显示统计显著性 (p<0.05) 的决定因素被包括在多元分析中。

²因为四舍五入，CI 包含数字 1.0。

表 2 - 续

饮酒情况（酗酒天数） ³						
0	573	40.0	对照组	-	对照组	
1	61	39.3	1.0 (0.6-1.6)	0.914	1.0 (0.6-1.7)	0.977
2-4	264	45.1	1.2 (1.0-1.6)	0.109	1.0 (0.8-1.4)	0.770
>4	446	53.8	1.8 (1.4-2.2)	0.000	1.3 (1.0-1.8)	0.038
吸烟史 ⁴						
否	288	31.6	对照组	-	对照组	
是	1056	49.3	2.1 (1.7-2.7)	0.000	1.5 (1.1-2.0)	0.021
艾滋病病毒感染状况						
阴性	1249	44.6	对照组	-	对照组	
阳性	72	68.1	2.6 (1.7-4.1)	0.000	2.2 (1.4-3.5)	0.001
未知	23	26.1	0.4 (0.2-1.0) ⁴	0.047	0.5 (0.2-1.3)	0.173

³患者至少饮过 5 单位的饮酒量时，上一个月报告的天数。

⁴“吸烟”在此处被界定为既往 5 年定期使用任何烟草产品。

析专门调查数据时，需要考虑两个主要问题：

- 为了解决缺失数据导致的潜在偏倚，应当填补缺失值。
- 在群集抽样调查中，应当解决同一个群集中不同个人的相关性。

世界卫生组织耐药调查指南中详细介绍了如何填补缺失值和描述群集（1）。

持续监测

分析持续监测系统收集的数据类似于专门调查数据的分析。主要差别在于前者并不需要填补缺失值和调整群集。

为了确保持续监测数据具有全国代表性，理想情况下较高比例（例如超过 75%）的新肺结核通报病例应当至少具有书面的利福平药敏试验结果，正如世界卫生组织结核病监测核查清单中标准 B2.1 所述。^c

不过，世界卫生组织当前要求国家持续监测系统通报的病例中至少 60% 的新肺结核病例至少具有书面记录的利福平药敏试验结果，这样才能具有代表性。

。

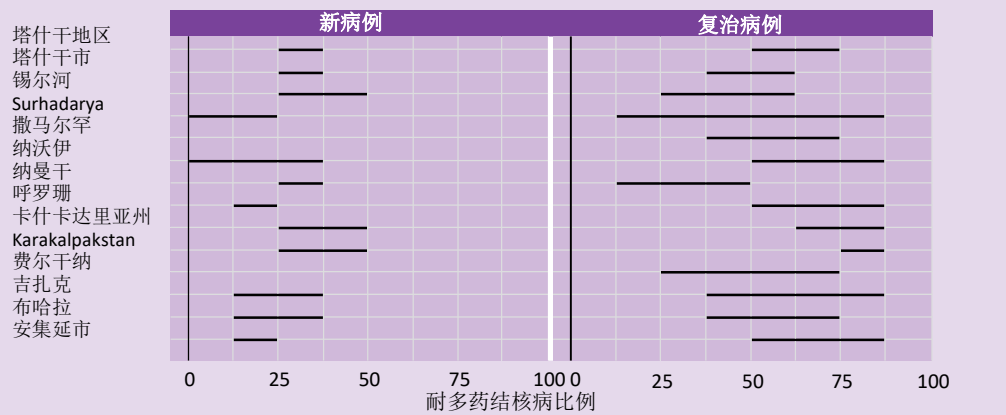
c. 核查清单和用户指南可以在世界卫生组织“全球结核病影响衡量”任务组的网站上找到：
http://www.who.int/tb/advisory_bodies/impact_measurement_taskforce/en/

有时，只具备国内一部分地区的数据（文字框 2）。在一些情况下，哨点监测系统可以作为中间措施，然后转移到全面的持续监测（1）。

文字框 2
地方统计和耐药监测

耐药调查通常可以有效地提供国家层面上结核病患者的耐药估计值。为了让地方层面的调查也具有代表性，需要大幅增加样本量，而绝大多数国家无法承担这方面增加的成本。因此，各省之间耐多药结核病现患率的差异通常只在具有极值的地区之间存在统计显著性（比较图 2 中的区域数值）。

图 2
新病例和既往治疗病例中耐多药结核病的百分比，95%置信区间，按地区/城市划分，乌兹别克斯坦（2010–2011 年）全国数据记录系统



相反，在综合持续监测系统只局限于特定行政区域的国家，区域数据可以用来推导出最有可能的耐药估计值。例如，俄罗斯联邦卫生部针对国内所有地区和城市（俄罗斯行政区划）发布有关新结核病病例和既往治疗结核病病例中耐多药结核病百分比的年度信息（10）。世界卫生组织查阅这些数据，根据上文描述的各项标准评估它们的代表性。

5.2 耐药结核病负担估计和时间趋势分析

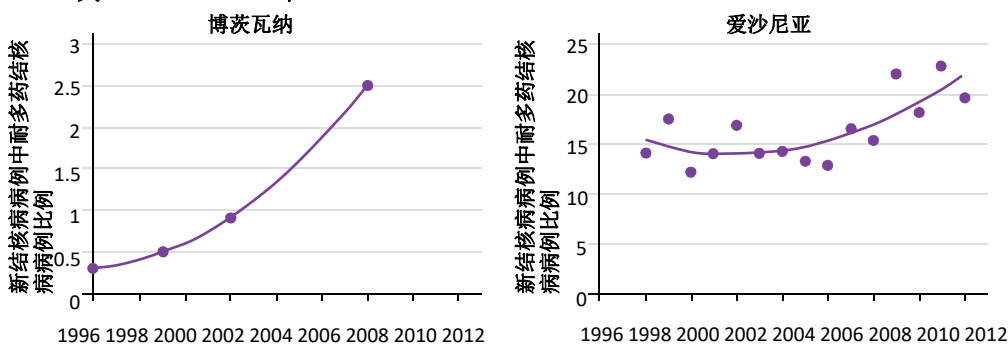
自 2008 年以来，世界卫生组织报告中就包括了各国耐多药结核病负担估计值（11）。这些估计值基于各国内接受治疗的结核病患者中耐多药结核病病例比例的最新和最佳可用证据。它们可以用于量化发现耐多药结核病患者并将他们纳入治疗的有效性。计算肺结核通报病例中估计的耐多药结核病病例数量时，将肺结核通报病例数量乘以最近专门调查或持续监测中发现的结核病病例中的耐多药结核病病例比例。新结核病病例和既往治疗结核病病例被分开计算。

监测耐多药结核病时间趋势的最合适指标是仅在新结核病病例（即并不是所有结核病病例）中发现的耐多药结核病病例比例，该比例可以通过持续监测或重复耐药调查的结果计算得出。调查通常是三到五年重复一次，以此评估耐药水平的变化。新结核病病例由一群同质患者组成，这些患者可以通过特定时间相关数据进行监测。这与具有既往结核病治疗史的患者形成鲜明的对比，后者通常是一群更为异质的患者，包括复发病例、失败后治疗和失访后治疗。

至少有两种方法可以用来分析新结核病病例中耐多药结核病的时间趋势：

- 绘制各年度新结核病病例中耐多药结核病病例比例的图形后，描述不同时间段新结核病病例中耐多药结核病病例比例的差异。
- 描述不同时间段通报新结核病病例中耐多药结核病病例估计数量的差异（图 3）。

图 3
博茨瓦纳和爱沙尼亚不同时间段新结核病病例中的耐多药结核病病例比例（1996–2012 年）



（请注意 Y 轴的尺度不同）

为了开展后一种计算（*estMDRT*），将新结核病病例中耐多药结核病病例比例乘以每年每 10 万人口中的新肺结核通报病例数量，并按年度绘制结果图形，其中：

pMDR = 特定年份新结核病病例中的耐多药结核病病例比例；

newTB = 同一年的新结核病通报病例数量；

pop = 同一年的国内人口。

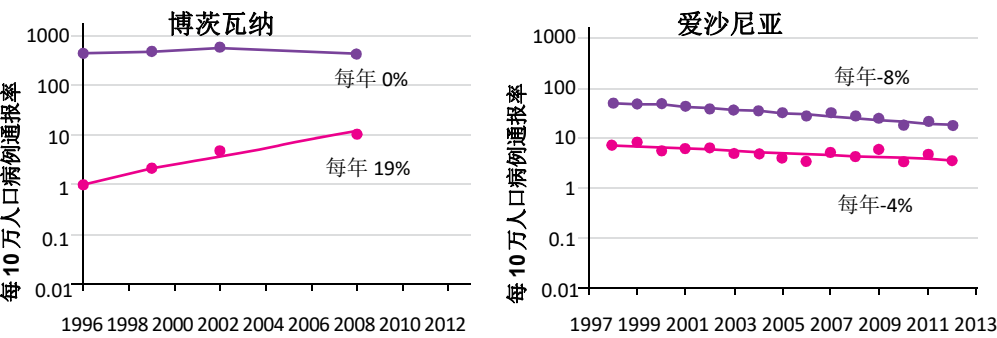
$$estMDRT = pMDR \times \frac{newTB}{pop} \times 100\,000$$

y 轴采用对数标尺。如果趋势在视觉上看起来呈线性，直线可能适合于数据。然后，可以计算新肺结核通报病例中预期可发现的耐多药结核病病例比例的平均年度百分比变化。

耐多药结核病病例估计数量的趋势可以与同一个图形中的新结核病病例通报数趋势进行比较。

给出每年的平均百分比变化。例如在博茨瓦纳，结核病通报率比较稳定，但估计的人均耐多药结核病率每年以平均 19% 的速度增长（图 4）。在爱沙尼亚，结核病通报率和估计的人均耐多药结核病率都呈下降趋势，不过耐多药结核病的下降速度较结核病要慢一些。

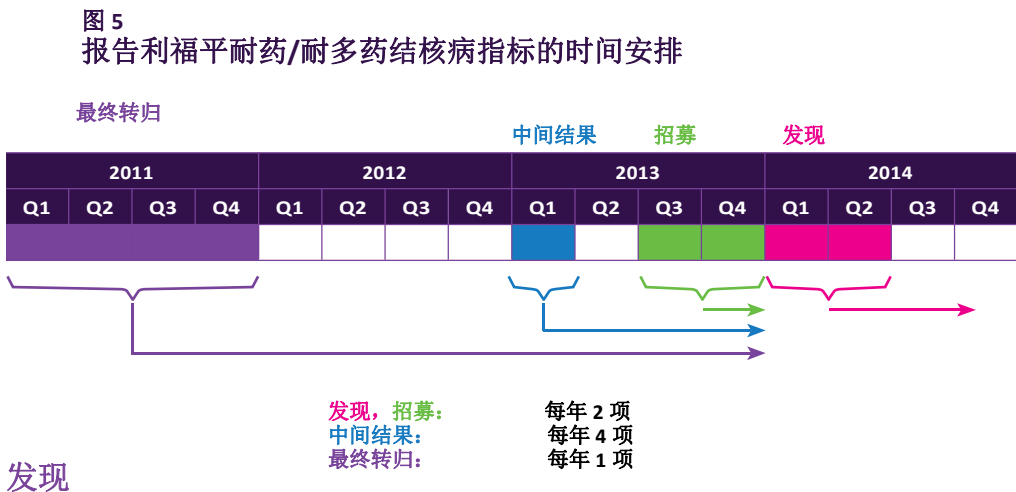
图 4
博茨瓦纳（1996–2008 年）和爱沙尼亚（1998–2012 年）不同时间段每 10 万人口新结核病病例通报率（紫色圆点）和新结核病通报病例中耐多药结核病病例估计数量（红色圆点）



5.3 监测规划有效性

本节根据四组指标进行划分：即发现、招募、中间结果和最终转归。这些指标由世界卫生组织推荐，用于评估规划在发现利福平耐药/耐多药结核病病例、给病例安排适当治疗和监测治疗转归等方面的有效性（2）。下面几节会概述数据源、如何计算指标（定义分子、分母和患者亚群）以及生成指标的时机和周期性。与季度和年度报告方法相比，指标可以帮助更多地关注监测活动和分析时间趋势。这些指标还通过一个在线表格进行了汇总（4）。

下面显示了报告四组指标的时间安排（图 5）。



早期发现利福平和异烟肼耐药可以确保从一开始就制定适当的药物治疗方案，从而提高治疗成功的可能性，减少获得额外耐药的概率。虽然世界卫生组织通常推荐给所有既往治疗患者和耐多药结核病病例接触者开展药敏试验，不过各国根据国家政策接受药敏试验的群体各不相同（12）。建议给所有耐多药结核病患者开展氟喹诺酮类药物和二线抗结核病注射药物的药敏试验。

五项指标可用来衡量在结核病患者中实现耐药结核病诊断普遍可及方面取得的进展情况。前面两项指标针对所有检测病例进行计算，同时针对国家政策中所列全部风险类别单独计算。还会评估不同风险类别中个体的检测延迟和耐多药结核病频次。这些参数对规划管理人员很重要，因为需要它们评估药敏试验的目标选择和时间表以及耐多药病例数量如何随着目标患者的风险类别有所不同。在单独开展 Xpert MTB/RIF 的场所，发现指标 1、2 和 5 可以修改为包括具有利福平检测结果的所有病例，主要目标可以调整为发现利福平耐药结核病病例（而不是耐多药结核病病例）。

发现指标应当每六个月计算一次，通常是从一月份到六月份和从七月份到十二月份，计算应当在每六个月周期结束后的三个月内完成，利用基本管理单元结核病登记簿、二线结核病治疗卡和实验室登记簿（或电子同等文件）的数据开展计算。文字框 3 提供了国家利用发现指标的示例。五项指标描述如下。

发现指标 1^d：具有异烟肼和利福平药敏试验结果的结核病患者

分子：评估期间国家政策指定的每个风险类别中同时具有异烟肼和利福平药敏试验结果的结核病病例数量。

分母：评估期间每个单独风险类别中发现的结核病病例数量。

数据源：分子数据源自实验室登记簿；分母数据源自基本结核病登记簿和治疗卡。对于一些风险类别（例如耐多药结核病接触者），信息可能并没有被包括在治疗卡中，因此需要从医疗记录的其他地方追踪。风险类别包括既往接受结核病治疗的所有结核病患者、确诊耐多药结核病病例的接触者或被认为面临其他耐药结核病风险的其他个人。

发现指标 2：检测结果对异烟肼和利福平敏感的结核病患者中发现的确诊耐多药结核病病例

分子：评估期间国家政策指定的每个风险类别中的确诊耐多药结核病病例数量。

分母：评估期间每个单独风险类别中同时具有异烟肼和利福平药敏试验结果的结核病病例数量。

d. 在单独开展 Xpert MTB/RIF 的场所，发现指标 1 的分子和发现指标 2 的分母可以修改为包括具有利福平检测结果的所有病例，而发现指标 2 的分子也可以调整为包括所有利福平耐药病例。同样，发现指标 5 可以改编为用于利福平耐药结核病检测。

数据源：分子数据源自实验室登记簿；分母与发现指标 1 的分子相同。风险类别包括既往接受结核病治疗的所有结核病患者、确诊耐多药结核病病例的接触者或被认为面临其他耐药结核病风险的其他个人。

发现指标 3：检测结果对任何氟喹诺酮类药物和任何二线注射药物敏感的确诊耐多药结核病病例

分子：评估期间检测结果对任何氟喹诺酮类药物和任何二线注射药物敏感的确诊耐多药结核病病例数量。

分母：评估期间的确诊耐多药结核病病例数量。

数据源：分子数据源自实验室登记簿；分母与发现指标 2 的（未分解的）分子相同。

发现指标 4：检测结果对任何氟喹诺酮类药物和任何二线注射药物敏感的确诊耐多药结核病病例中发现的确诊广泛耐药结核病病例

分子：评估期间发现的确诊广泛耐药结核病病例数量。

分母：评估期间检测结果对任何氟喹诺酮类药物和任何二线注射药物敏感的确诊耐多药结核病病例数量。

数据源：分子数据源自实验室登记簿；分母与发现指标 3 的分子相同。

发现指标 5：利福平耐药/耐多药结核病推定与药敏试验结果之间的间隔

定义：根据国家政策确定结核病患者处于某个风险类别的日期与异烟肼和利福平药敏试验结果被记录在实验室登记簿中的日期之间的持续时间（天数）。第一个日期由风险类别的类型确定。如果药敏试验被普及，该日期可能对应着结核病诊断日期，还可能对应着结核病治疗期间实验室结果指示治疗失败或持续痰涂片阳性的日期，或者对应着艾滋病相关结核病被发现的日期。对于已知与结核病病例有接触的可能病例，该日期指实验室确诊源病例患有耐多药结核病的日期。这可以出现在接触者被诊断为结核病之前或之后（基于实验室登记簿中的信息）。在单独开展 Xpert MTB/RIF 的场所，指标可以修改为包括具有利福平检测结果的所有病例，采用显示利福平耐药的第一次结果的日期，而不用管同一名患者随后是否被确诊患有耐多药结核病。

计算基于六个月评估期间被录入实验室登记簿中具有异烟肼和利福平药敏试验结果（敏感或耐药）的所有病例。计算所有患者在这两个日期之间的间隔（天数），然后除以具有检测结果的病例数量。指标被表述为计算中包括的最小和最大疾病发作范围的算术平均天数。应当指明计算中包括的疾病发作数量。

发现指标 1 没有将国内发现的利福平耐药/耐多药结核病病例数量与根据监测衡量手段估计发生的耐多药结核病病例数量联系起来（13,14）。为了评估病例发现对国内实际耐多药结核病负担的影响，可以将发现的利福平耐药/耐多药结核病病例数量与预期耐多药结核病现患率进行比例。只要国内已经给通报的所有肺病患者提供普及的药敏试验，各种比率就可以被用来将发现和招募数与可能发现的耐多药结核病病例估计数进行比较（15）（根据上文所示进行计算）。

文字框 3 国家示例

利福平耐药结核病的发现：相匹配临床和实验室数据源

发现的利福平耐药结核病和耐多药结核病病例的统计通常要求在不同数据源之间开展交叉检查。这样做主要出于下面三个原因：

- 为了确保所有实验室确诊病例都被计算在内。
- 将实验室数据与临床记录中的个人数据关联起来，从而尽量避免同一名患者的重复计数。
- 可以按照国家政策规定的不同风险类别对发现的病例进行分层。

文字框 3 - 续

下表是尼日尔在 2010 年 7 月至 12 月期间的一份药敏试验结果示例。根据异烟肼和利福平的药敏试验结果，33 个病例中有 24 个病例（73%）被确定面临着耐多药结核病风险。在确定的 24 个病例中，20 个病例（83%）患有耐多药结核病，所有这些病例都接受了二线药物检测。这意味着药敏试验有着较高的覆盖面，因此数据具有代表性。不过，尼日尔在 2010 年通报了 452 个结核病复发病例和 215 个其他结核病复治病例（14）。所以，表 3 中显示的统计数据只提供了面临较高水平耐多药结核病的一个选定患者群体的信息，该患者群可能来自专门负责结核病疑难病例管理的单位。

怀疑耐多药结核病与获得检测结果之间的时间相对较长（3 到 7 个月），这提示患者的转介会有一定的延迟，而且实验室检测是在与关怀点有一定距离的地方开展。使用中位值作为平均延迟的衡量标准，并采用相应的 25th 和 75th 百分位数（例如四分位数间距的限度），这样做会更有意义，因为中位值不太可能受到极端异常值的影响，不过算术平均数和范围得到了更广泛的理解，可以提供集中趋势和均分的另一个视角。

表 3
尼日尔按患者风险类别划分的药敏试验结果（2010 年 7 月至 12 月）

风险类别	结核病病例数量			
	总计	具有异烟肼和利福平检测结果	对异烟肼和利福平耐药（耐多药）	患有耐多药且接受一种氟喹诺酮类药物和一种二线注射药物检测
耐多药结核病接触者	1	1	0	0
结核病复治方案失败	20	15	15	15
结核病初步治疗方案失败	2	2	2	2
复发	10	6	3	3
失访后复治	0	0	0	0
总计	33	24	20	20

文字框 3 - 续

表 4
尼日尔耐多药结核病推定与药敏试验结果之间的延迟（2010 年 7 月至 12 月）

具有时间间隔信息的耐多药结核病病例的数量	怀疑耐多药与药敏试验结果之间的时间间隔（天数）		
	均值	最小	最大
24	139	89	215

资料来源：S Mamadou（尼日尔国家结核病实验室）、A Piubello（达米恩基金会）和 A van Deun（IMT Antwerp）。

不同时间段发现的耐多药结核病病例数量通常描述为简单频率图。

图 6
中国、印度和俄罗斯联邦通报的耐多药结核病病例（2009–2012 年）



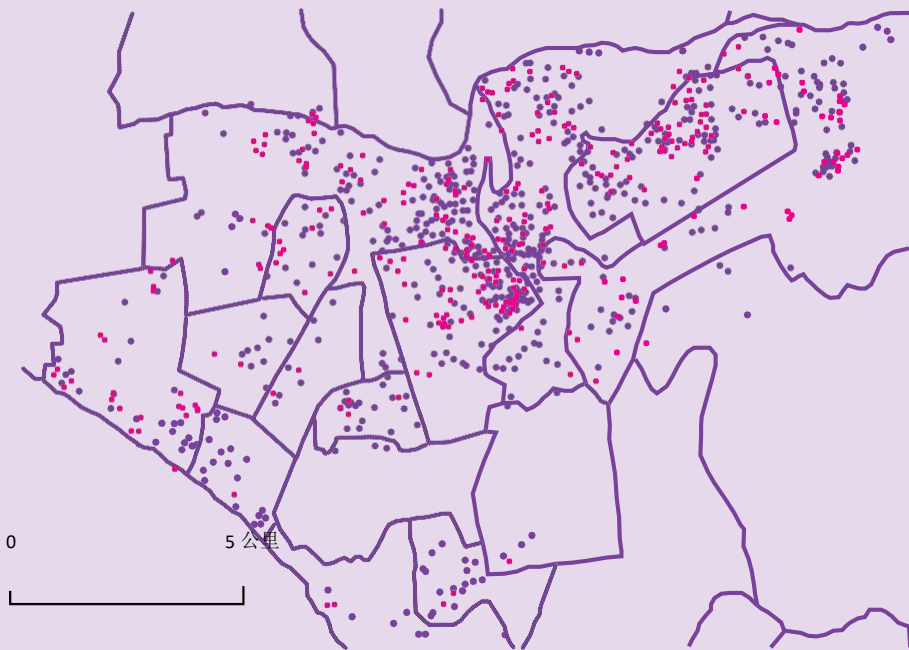
耐多药结核病病例的空间分布

规划数据可以用来确定面临更高耐多药结核病风险和应当重点开展药敏试验的患者。有关诊断医疗机构或患者居住地的空间信息可以改善风险分析的预测能

文字框 3 - 续

力。拥有 GPS 功能的手机和其他移动装置越来越多，它们可以记录患者的家庭地址或医疗机构的位置，现在甚至能够支持在资源匮乏地区开展这类研究。秘鲁首都利马两个卫生区发现了 1116 名具有药敏试验结果的结核病患者（其中 346 名患者被确诊为耐多药结核病）（2005–2007），我们重新分析了这些患者的数据，以期获得诊断时的耐多药结核病预测指标（图 7）。

图 7
接受药敏试验的患者中药敏（紫色圆点）结核病和耐药（红色圆点）结核病的空间分布（Lima Ciudad 和 Lima Este 地区，2005–2007 年）
(16)



（注：空间坐标中给每名患者添加了一个小的随机误差，以便保护保密性。）

这些结果提示如果卫生所的结核病患者进行其他风险因素调整后更有可能存在耐多药结核病，则扩大药敏试验入选资格范围可能会有帮助作用。这在绝大多数耐多药结核病病例预期出现于新结核病病例中的地区有着实际应用 – 许多病例并没有已知的耐多药结核病风险因素 – 资源被限制于给国内或大型行政区域的所有新结核病病例开展药敏试验。数据分布图绘制也有利于生成有关疾病疫源位置的假设，从而给不同地区的资源分配提供参考依据（6,16）。

招募

规划管理人员负责确保被诊断出利福平耐药结核病或耐多药结核病的所有患者都会在尽可能短的时间内开始接受适当的治疗。这还可以适用于面临利福平耐药结核病风险但尚未被确诊（即推定型诊断）的患者。如果一名患者随后开始接受充足的药物治疗方案，则尽早发现耐药情况会非常有益。

四项指标被推荐用来评估招募结核病患者接受二线结核病治疗的覆盖面，尤其是在儿童和女性中，这两个群体在特定地区可能无法获得平等的卫生保健服务。艾滋病病毒阳性利福平耐药/耐多药结核病患者患者的进一步分层可以评估接受抗逆转录病毒治疗的患者比例。通过比较新纳入的和已发现的利福平耐药/耐多药结核病病例，可以了解到卫生保健服务的可及性，尽管应当在评估期开始前发现那些开始接受治疗的患者。因此，前三个招募指标可以有用地表述为特定时间段内符合治疗条件的病例与已经开始接受治疗的病例之间的比率。

这些指标应当每六个月计算一次（通常是从一月份到六月份和从七月份到十二月份），计算应当在每六个月周期结束后的三个月内完成，利用基本管理单元结核病登记簿、二线结核病治疗卡以及实验室培养、Xpert MTB/RIF 和药敏试验登记簿（或电子同等文件）的数据开展计算。一些规划保留着尚未开始接受二线结核病治疗的耐多药结核病（推定或确诊）患者的清单：这样一份清单如果维护得当，可以成为规划的宝贵资源，有助于计算对药物和实验室资源的需求，以及相关卫生机构在可接受条件下具备给患者提供治疗时联系患者（[文字框 4](#)）。

计算

招募指标 1^e：被招募接受耐多药结核病治疗的利福平耐药/耐多药结核病病例（推定或确诊）。

定义：评估期间被登记且开始接受规定耐多药结核病治疗方案的利福平耐药/耐多药结核病病例数量（推定或确诊）。

对照数据：评估期间符合二线药物治疗条件的利福平耐药/耐多药结核病病例数量（推定或确诊）。

e. 在单独开展 Xpert MTB/RIF 的场所，指标可以修改为同样计算开始接受二线结核病治疗的利福平耐药结核病病例，并将它们与利福平耐药结核病病例（推定或确诊）进行比较。

数据源：耐多药结核病病例数量（推定或确诊）源自二线结核病治疗登记簿；对照数据源自本结核病登记簿和实验室培养、Xpert MTB/RIF 和药敏试验登记簿。确诊病例采用药敏试验结果的日期，其他病例采用被推定患有利福平耐药/耐多药结核病的日期（例如治疗失败的患者在痰涂片仍然为阳性时会据此进行分类）。

这指标的计算对象为：（i）所有病例；（ii）年龄超过 15 岁的病例；和（iii）女性。

招募指标 2：被招募接受耐多药结核病治疗的确诊利福平耐药/耐多药结核病病例

定义：评估期间被登记且开始接受规定耐多药结核病治疗方案的确诊耐多药结核病病例数量。

对照数据：评估期间发现的确诊耐多药结核病病例数量。

数据源：接受治疗的确耐多药结核病病例源自二线结核病治疗登记簿；对照数据源自实验室培养、Xpert MTB/RIF 和药敏试验登记簿（采用药敏试验结果的日期）。

这指标的计算对象为：（i）所有病例；（ii）感染艾滋病病毒且接受抗逆转录病毒治疗的病例；和（iii）感染艾滋病病毒但尚未接受抗逆转录病毒治疗的病例。

招募指标 3：被招募接受广泛耐药结核病治疗的确诊广泛耐药结核病病例

定义：评估期间被登记且开始接受规定广泛耐药结核病治疗方案的确诊广泛耐药结核病病例登记簿。

对照数据：评估期间发现的确诊广泛耐药结核病病例数量。

数据源：接受治疗的确广泛耐药结核病病例数量源自二线结核病治疗登记簿；对照数据源自实验室培养、Xpert MTB/RIF 和药敏试验登记簿（采用药敏试验结果的日期）。

招募指标 4：利福平耐药/耐多药结核病诊断与开始耐多药结核病治疗之间的时间间隔

定义：利福平耐药/耐多药结核病确诊日期（实验室登记簿中的药敏试验结果显示同时对异烟肼和利福平耐药）与二线结核病治疗登记簿中患者开始接受规定二线药物治疗方案的日期之间的持续时间（天数）。在单独开展 Xpert MTB/RIF 的场所，指标可以修改为包括所有确诊利福平耐药结核病病例，采用显示利福平耐药的第一次结果的日期，而不用管同一名患者随后是否被确诊患有耐多药结核病（即患者第一次被发现符合耐多药结核病治疗条件的日期）。

计算基于六个月评估期间被记录到二线结核病治疗登记簿中的所有确诊利福平耐药/耐多药结核病病例。所有患者在确诊日期与治疗开始日期之间的间隔（天数）被加起来，然后除以治疗事件的数量。指标被表述为计算中包括的最小和最大治疗事件范围的算术平均天数。如果是在报告确诊药敏试验结果之前开始治疗，则时间间隔标记为零天。应当指明计算中包括的治疗事件数量。

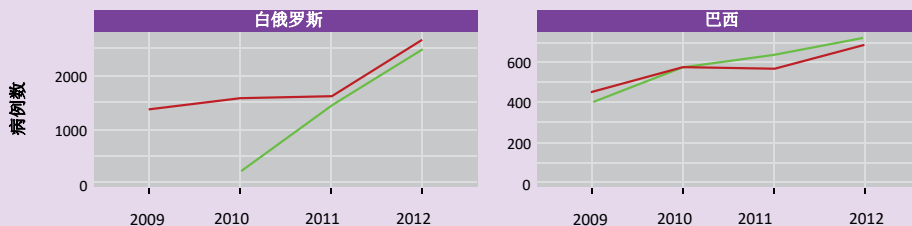
文字框 4 国家示例

i) 比较不同时间段的招募与发现

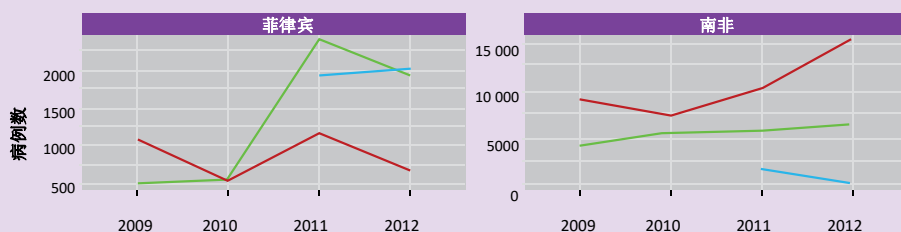
将不同时间段招募接受耐多药结核病治疗的病例数量绘制成图，同时绘制相应时间段内被诊断出耐多药结核病和其他利福平耐药结核病的病例（图 8）。这样有助于比较规划诊断患者的能力和治疗患者的能力。例如：

图 8A

四个选定国家内发现的耐多药结核病病例（红线）和额外利福平耐药结核病病例（蓝线）与接受耐多药结核病治疗（绿线）的结核病病例之间的比较（2009–2012 年）



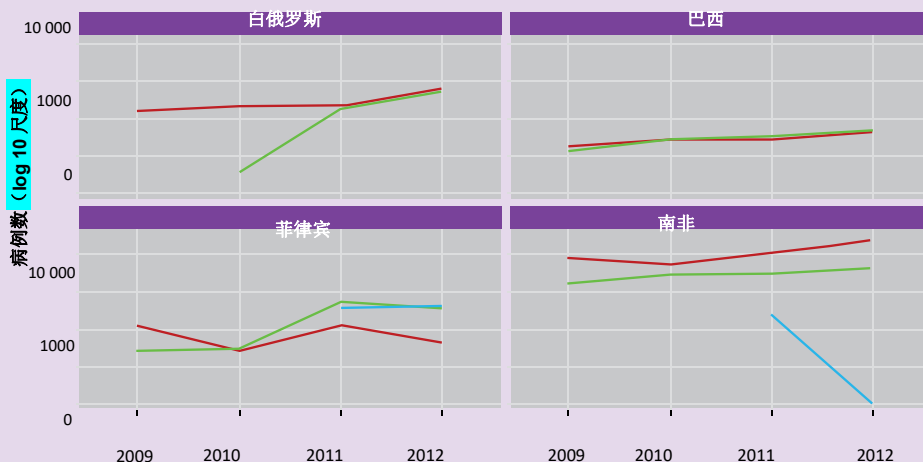
文字框 4 - 续



在更实用的层面上，如果同一张图上要显示的数字有着较大的范围（参见下文），可以采用对数尺度，通常是在纵轴上使用十进制。而且，对数尺度还有助于以视觉方式比较不同时间段的通报率，这种情况下趋势线的倾斜度代表着病例的平均年度百分比变化。如果数据沿着纵轴中的对数标尺呈一条直线，则意味着指标以指数方式发生着变化（例如从一个时间段到下一个时间段发生固定百分比的变化）。

图 8B

四个选定国家内发现的耐多药结核病病例（红线）和额外利福平耐药结核病病例（蓝线）与接受耐多药结核病治疗（绿线）的结核病病例之间的比较（2009–2012 年）。对数尺度上的纵轴。



ii) 将发现和招募与国家耐多药结核病病例估计值关联起来

为了评估规划活动对病例数量的影响，发现和招募指标还必须与估计会在国内发生的耐多药结核病病例数量进行比较。出于这个原因，可以确定病例估计数

文字框 4 - 续

中发现的病例所占的比例（发现比率）与发现病例数中估计的病例所占的比例（招募比率），这样有助于根据疾病负担的分担情况进行解释（15）。下面两组图形（图 9）比较了这两种方法，评估相同的三个国家的这些指标：第一组图形将发现和招募呈现为不同时间段两个线形图，病例的绝对数位于纵轴；第二组图形将发现的/估计的耐多药结核病病例（纵轴）的百分比与招募的/估计的耐多药结核病病例（横轴）的百分比进行了比较。三个坐标都标记了年份，添加了一个对角线，描述代表着发现和招募 1:1 关系的位置。虽然三个国家随着时间的发展都有了改善，但这些国家在估计病例中接受关怀的比例方面有着显著的差异。

图 9A

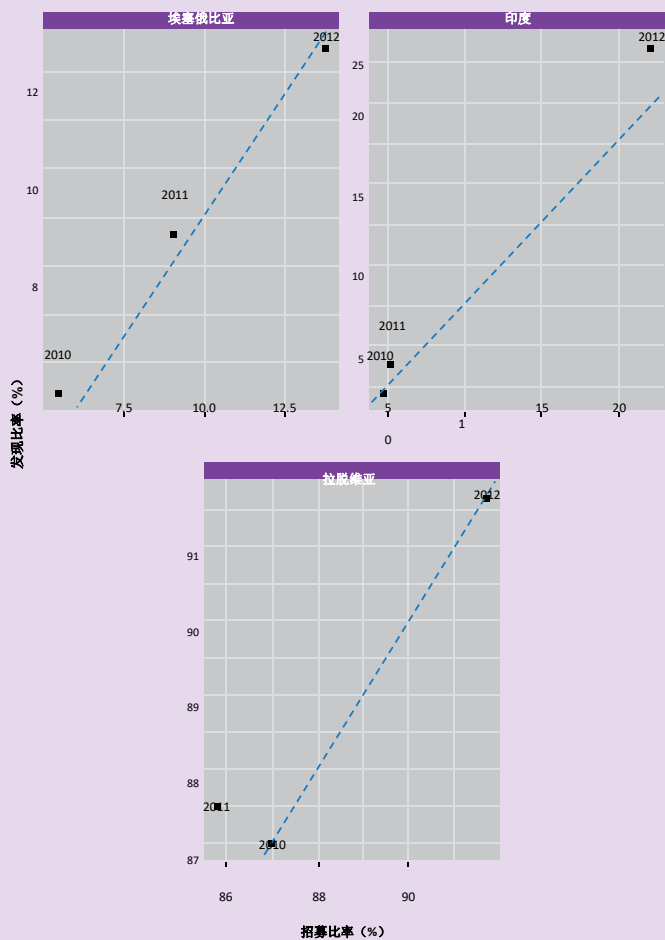
三个选定国家内发现的耐多药结核病病例（红线）与被招募接受耐多药结核病治疗的结核病病例（绿线）之间的比较（2010–2012 年）



文字框 4 - 续

图 9B

三个选定国家内发现病例数与耐多药结核病病例估计数之间的比率（y 轴）与招募病例数与耐多药结核病病例估计数之间的比率（x 轴）之间的比较（2010–2012 年）。对角线代表着两个比率的一致性。



中间治疗转归

耐多药结核病的治疗通常需要 20 个月或更长时间，因此最终转归只能在招募两三年后进行评估。不过，可能需要更早地了解患者如何对治疗做出反应，尤其是当耐药结核病治疗项目于近期启动的情况下。第六个月的培养阴转（针对确诊肺部病例）和六个月内的死亡都是经常使用的治疗响应指标。六个月内的失访信息也有帮助作用。还应

了解开始接受耐多药结核病二线药物治疗的患者中有多少人后来被发现并不患有耐多药结核病（同样适用于广泛耐药结核病）。这可以反映治疗流程在给需要二线治疗方案的患者提供治疗方面的有效性，还可以避免给不需要相关治疗的患者安排有潜在毒性的治疗方案。

评估期间是三个月（一个季度），通常为：（i）一月至三月；（ii）四月至六月；（iii）七月至九月；（iv）十月至十二月。评估期间被登记且开始接受治疗的所有患者都被包括在计算中。只有已经开始接受治疗的实验室确诊利福平耐药结核病、耐多药结核病和广泛耐药结核病病例被计算在中间转归的报告中。计算六个月内转为阴性培养的病例的比例时，开始接受治疗的所有患者都被包括在分母中，包括六个月内死亡或失访的患者。转归按照先满足的条件进行分配：例如，如果一名患者六个月内先失访后死亡，则病例被归类为“失访”。

指标在评估季度结束后九个月进行衡量。这让第六个月开展的培养有足够的时间出结果和供检索。所有数据都可以从二线结核病治疗登记簿中摘取。

计算

中间结果指标 1^f：接受耐多药结核病治疗方案六个月内出现阴性培养的利福平耐药/耐多药结核病病例

分子：被登记且开始接受规定耐多药结核病治疗的确诊肺部利福平耐药/耐多药结核病病例中治疗第六个月出现阴性培养结果的数量。

分母：评估期间被登记且开始接受耐多药结核病治疗的确诊利福平耐药/耐多药结核病病例数量。

中间结果指标 2：接受耐多药结核病治疗方案六个月内死亡的利福平耐药/耐多药结核病病例

分子：被登记且开始接受规定耐多药结核病治疗的确诊利福平耐药/耐多药结核病病例中治疗第六个月末因任何原因死亡的数量。

f. 在单独开展 Xpert MTB/RIF 的场所，指标还计算开始接受二线耐多药结核病治疗且被分配为中间结果的利福平耐药结核病病例，或者在中间结果指标 4 的情况下，被安排接受没有保证的二线耐多药结核病治疗方案的利福平耐药结核病病例。

分母：评估期间被登记且开始接受耐多药结核病治疗的确诊利福平耐药/耐多药结核病病例数量。

中间结果指标 3：接受耐多药结核病治疗方案六个月内失访的利福平耐药/耐多药结核病病例

分子：被登记且开始接受规定耐多药结核病治疗的确诊利福平耐药/耐多药结核病病例中治疗第六个月末失访的数量。

分母：评估期间被登记且开始接受耐多药结核病治疗的确诊利福平耐药/耐多药结核病病例数量。

第一个指标仅适用于肺部病例。所有指标的分母都是开始接受治疗的全部病例，包括开始接受规定二线药物治疗的广泛耐药结核病病例。

中间结果指标 4：接受耐多药结核病治疗方案但发现不是利福平耐药/耐多药结核病的患者

定义：评估期间开始接受规定耐多药结核病治疗方案但后来发现不是利福平耐药/耐多药结核病的患者的人数。

中间结果指标 5：接受广泛耐药结核病治疗方案但发现不是广泛耐药结核病的患者

定义：评估期间开始接受规定广泛耐药结核病治疗方案但后来发现不是广泛耐药结核病的患者的人数。

除了这些标准指标之外，结核病研究中还可以利用时间与事件统计，以便比较从开始治疗到出现特定终点（例如死亡或细菌学阴转）所需的时间。Kaplan-Meier 存活曲线描述结核病患者不同亚群中的趋势，不仅可用于汇总临床试验中的这些模式，而且还可用于汇总在规划条件下接受治疗的患者中的这些模式（17，18）。

最终治疗转归是衡量耐多药结核病防治规划有效性的最重要手段。治疗登记簿中的所有确诊耐多药结核病患者都应当在治疗结束时被分配六项相互排斥的转归中的一种（参见下文）。病例可能会因为被转出、最终评估时尚未完成治疗或缺失信息而没有接受评估，这些病例都被聚集到一起。“转入”患者在最开始给其提供治疗的治疗中心的转归队列中进行计算，而不被包括在接收其继续治疗的治疗中心的队列中。所有患者都应当被分配各自在治疗评估中经历的第一个转归。如果某个病例被分配“治疗失败”转归，然后该患者在队列同一年内再次开始接受修订治疗方案的治疗，则为了队列监测的目的，该病例不会被重新登记为另一个转归。

同样的规则也适用于患者失访，这类患者通常中断治疗两个月或更长时间后在同一年再次开始接受治疗。^g“治愈”转归局限于肺部病例。评估期间为 12 个月，通常是从一月至 12 月，被称为年度队列。评估期间被登记且开始接受治疗的所有患者都被包括在计算中。在单独开展 Xpert MTB/RIF 的场所，指标需要修改为也包括开始接受完整耐多药结核病治疗方案的利福平耐药结核病病例。仅实验室确诊利福平耐药结核病、耐多药结核病和广泛耐药结核病病例被包括在队列最终转归报告中。

指标在评估年度结束后 24 个月进行衡量。这让绝大多数患者有足够的时间完成治疗，也有足够的时间获得和记录最终培养结果。所有数据都可以从二线结核病治疗登记簿中摘取。[文字框 5](#) 描述了一系列国家示例，说明这些指标如何用于监测治疗项目绩效和有效地解释各种关联。

计算

转归指标 1–6：接受耐多药结核病治疗方案的利福平耐药/耐多药结核病病例的转归

- 1. 治愈
- 2. 完成治疗
- 3. 治疗失败
- 4. 死亡
- 5. 失访
- 6. 未接受转归评估

g. 因此，以这种方式评估的转归可能比较保守，因为有一定比例的患者在治疗失败或失访后会再次开始接受新治疗方案，并且最终可能会被治愈。如果维护一个患者相关登记簿（最好是电子格式）用来记录满足第一次转归后的事件（包括复发），则可以开展更具有实际意义的长期转归评估。

分子：评估期间登记到耐多药结核病治疗的确诊利福平耐药/耐多药结核病病例中被分配六种转归中一种转归的数量。

分母：评估期间被登记且开始接受耐多药结核病治疗方案的确证利福平耐药/耐多药结核病病例数量。

这指标被表示为每个相互排斥的转归中患者所占的百分比。如果规划有能力区分广泛耐药结核病与其他耐多药结核病病例，尤其是当广泛耐药结核病病例在耐多药结核病病例所占的比例超过 5% 时，应当分开报告非广泛耐药的耐多药结核病（包括其他利福平耐药结核病）和广泛耐药结核病病例的转归。如果在二线结核病治疗期间的任何时候发现耐多药结核病患者患有广泛耐药结核病，则应当将这些患者从非广泛耐药的耐多药结核病队列中移除，然后计算到广泛耐药结核病治疗队列中。

在孕产妇中艾滋病病毒现患率 1% 或结核病患者中艾滋病病毒现患率 5% 的国家，阳性艾滋病病毒感染状况的患者应当分开计算艾滋病病毒感染患者中的转归指标（2,19）。

文字框 5 国家示例

i) 比较不同单元之间的治疗转归

鉴于转归指标被界定为六个标准的、相互排斥的类别，它们可以被描述成患者队列的比例图表。描述转归指标常用的一种方法是层叠式条形图（带有 100% y 轴），层叠式条形图能比饼图更轻松地在不同单元和不同时间段进行比较。下面的示例（图 10）比较罗马尼亚两个不同患者群体的转归。通常而言，正如在本示例一样，治愈和完成治疗两个类别被整合到一个转归中（即治疗成功）。

ii) 改善治疗成功

随着耐药结核病项目人员在实施过程中积累越来越多的经验，耐药结核病项目有望得到不断改善。队列内患者转归的信息完整性非常重要，有助于得出有关患者中成功可能性或死亡风险的有意义的结论。有时，改善数据本身的恢复可以提高成功比例，正如本示例中显示了秘鲁向世界卫生组织常规报告转归数据（图 11）。

文字框 5 - 续

图 10

罗马尼亚绿灯委员会（GIC；左边）和非绿灯委员会（右边）场所 2007 年登记治疗的 840 名耐多药结核病患者的治疗转归

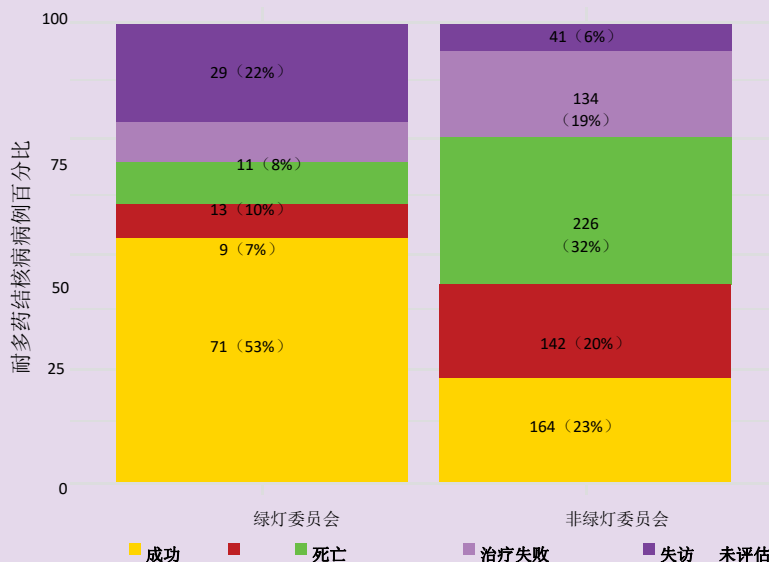
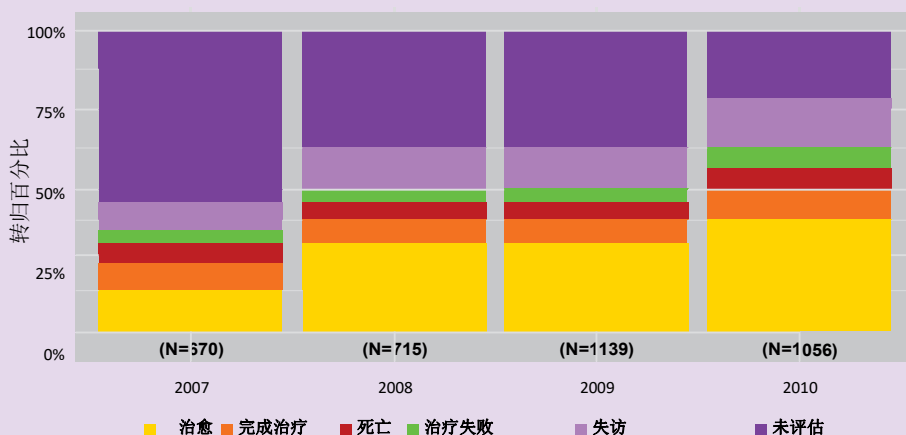


图 11

秘鲁 2007 年至 2010 年开始接受治疗的耐多药结核病队列的治疗转归（基于报告给世界卫生组织的数据）



除了改善数据完整性之外，治疗转归中的其他时间趋势也需要管理人员开展评估，例如：

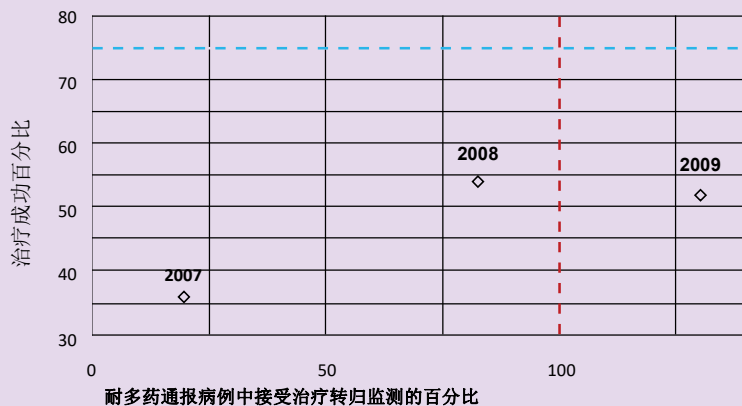
- 细菌学确定的治疗成功（例如治愈）的比例。如果这个比例很低，可能指示着培养监测比较罕见或结果数据没有从实验室获得。该比例不应随着时间的发展不断下降。

文字框 5 - 续

- 有时非常高的成功水平可能意味着监测过程中存在特定的问题。例如，由于有关细菌学状况的信息不充足，治疗失败的患者可能会被错误地归类到“治疗完成”。
- 在艾滋病病毒感染和其他共病比较常见的地区，死亡病例的比例可能会比较高。这还可能反映治疗开始的延迟和晚期疾病的存在。项目目标之一是尽可能地降低死亡的比例。
- 失访病例的比例是一个关键的管理指标，应当尽最大努力改善患者依从性和减少损失。
- 治疗失败病例的比例是另一个关键的管理指标。在具有复杂耐药模式的地区或患者中，预计治疗失败会更加普遍，指示治疗方案可能需要修订或应当解决容易引发失败的其他患者相关因素。

iii) 耐多药结核病治疗成功对耐多药结核病病例数量的影响

图 12
格鲁吉亚耐多药通报病例中接受转归监测的百分比（排除未经评估的病例）与耐多药结核病队列中成功百分比之间的关系（2007 至 2009 年）



蓝线（水平虚线）= 到 2015 年耐多药结核病队列 75%成功率阈值（“全球 2011-2015 年终止结核病计划”）；红线（垂直虚线）= 100%治疗转归报告（排除队列中未经评估的病例）。耐多药结核病的不完整或延迟报告可能会导致接受转归监测的病例百分比超过 100%（基于报告给世界卫生组织的数据）。

如果耐多药结核病治疗规划要对国内耐多药结核病负担产生有效的影响，就必须结合高治疗成功率和 high 患者覆盖面。如果两者中有任何一个较低，公共卫生效果都会大打折扣。

文字框 5 - 续

在上文格鲁吉亚的示例中（图 12），成功率和覆盖面在 2007 年到 2008 年都有所增长。在 2009 年，成功率维持不变，但覆盖面看起来好像超过了 100%。出现这种现象可能是因为被招募接受治疗的患者是在前一年通报，或者病例被实验室漏报，计算通报病例的系统与汇总患者转归的系统之间没有关联。这种情况下，应当查看国内估计会出现的耐多药结核病病例数量。例如在格鲁吉亚，2009 年包括在治疗队列中的病例占该国当年肺结核通报病例中估计会出现的耐多药结核病病例的约三分之二。^h

iv) 比较不同结核病患者亚群之间的死亡可能性：离散数据和连续数据

某个风险因素与某个事件（例如死亡）之间关系的强度可以利用简单的统计检验来量化。“2x2”表可以用于开展卡方检验和生成粗略相对风险估计值。例如，巴西的示例（表 5）反映了一个事实，即耐药模式越广泛，治疗选择方案就越少，因此广泛耐药结核病患者通常比其他耐多药结核病病例有着更差的转归（20）。

表 5
巴西（2010 年队列）广泛耐药结核病相比耐多药结核病队列的死亡优势比

	死亡		总计
	是	否	
广泛耐药结核病	7	18	25
耐多药结核病	50	473	523
总计	57	491	548

优势比：3.68（95%置信限度，1.47-9.24）， $P < 0.05$

如果与上述示例相反，风险因素数据属于连续类型（例如年龄、体重、病例发现和治疗开始之间的间隔），可以采用其他统计数据。在下面的示例中，白俄罗斯（2009–2010 年）一个地区有 436 名耐多药结核病患者开始接受治疗（表 6），开始治疗时年龄方面的细微差别看起来并没有以统计显著方式与死亡独立相关。

h. 注：计算这类覆盖面，需要从分子中减去队列内未经评估的病例。

文字框 5 - 续

表 6

白俄罗斯 2009 年到 2010 年开始接受治疗的耐多药结核病患者
中年龄和死亡之间的关联。

死亡	人数	平均年龄	标准差	中位值年龄	25%百分位数	75%百分位数
否	284	44.1	12.7	43.8	34.9	52.6
是	152	45.0	10.7	46.0	37.3	52.0

(韦尔奇两样本 t 检验: $t = -0.7869$, $df = 357.584$, p 值 = 0.4319)

通过在回归方程中考虑一个以上的风险因素, 可以改善这些分析 (参见下文的 [vii](#) 小节)。

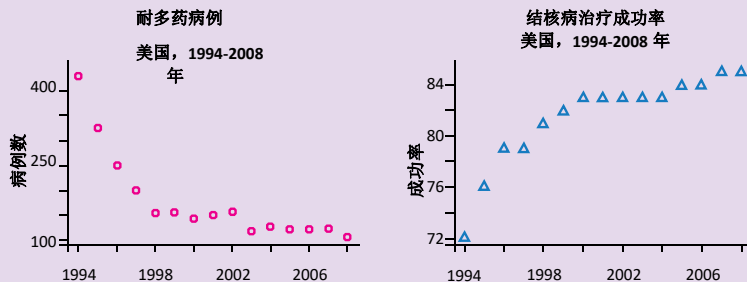
v) 增强一线治疗项目的绩效, 减少耐多药结核病负担

除了耐多药结核病治疗项目之外 (上文 [iii](#) 小节), 还需要有效开展面向药敏结核病的基本项目, 从而减少耐药的产生。下面美国的示例 (图 13) 阐述了不断上升的成功率与二十世纪九十年代中期后大幅减少的耐多药结核病病例数量之间的反向关系。

基于总成功率的耐多药结核病病例数量线性回归 (图 14) 显示了强大的反向关系, 这还可以通过视觉方式呈现, 即绘制图形显示耐多药结核病病例数量与治疗成功率的关系。这表明存在一种线性关系, 以及耐多药结核病数量的变化与治疗成功率有着强烈的关联。

图 13

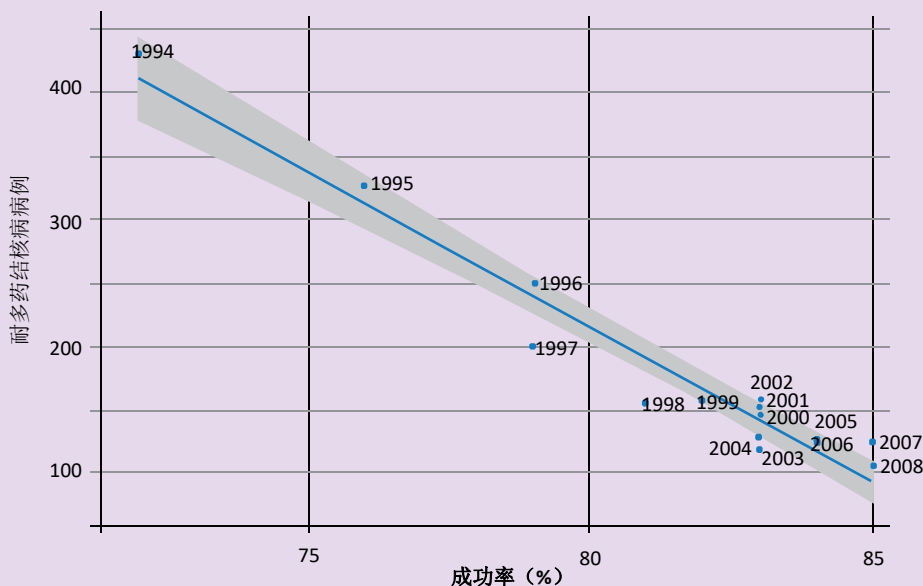
美国耐多药结核病病例报告 (红色圆圈, 左边) 与结核病总治疗成功率 (蓝色三角形, 右边) (1994 年至 2008 年)



文字框 5 - 续

图 14

美国所有结核病病例治疗成功率(%)与耐多药结核病病例数量之间的关系(1994年至2008年)



vi) 耐多药结核病的不良事件：将队列分析应用于药物警戒

与药敏结核病标准六个月疗程相比，耐多药结核病治疗方案一般包含四种二线药物和持续 20 个月或更长时间。这些治疗方案通常会导致药物相互作用和不良药物反应，从轻度到致命不等，严重时可能需要中断治疗。这可能表明患者出现选择性耐药株，因此增加治疗失败和死亡风险。全球范围内治疗项目的持续拓展加大了这些不利因素的机会，提出有必要适当地监测治疗规模扩大带来的潜在危害。

耐多药结核病治疗中引入的新药物尚未完成三期临床试验，最先引入的药物是 2012 年的贝达喹啉，这种情况使得不良事件监测显得更重要。世界卫生组织在 2012 年已经开发了有关自发和主动药物警戒的指南（7）。该指南还提供了指示说明如何解释警戒信号和因果关联。虽然危害的报告不属于耐多药结核病防治规划最低指标的一部分，不过收集有关不良事件、严重性、发作日期和后遗症等额外患者数据也很重要，可以结合人口统计学、既往治疗史、疾病部位、细菌学终点和治疗转归等更标准的数据要素，有效地量化药物引起的不良事件和确定风险因素。

文字框 5 - 续

表 7 显示了如何开展这类分析的一个良好示例，这类分析有助于确定不良事件的预测指标（这种情况下量化为采用名义回归的数字）（21）。同一文献提供了其他描述性方法，可以用来汇总队列的不良经历。这类分析可以帮助集中解决与不良事件相关的患者或项目规划决定因素，从而改善关怀的质量。

表 7
人口统计学和临床学特征与耐多药结核病治疗期间每个病例经历的不同不良事件的数量之间的调整后关联（ $n=1015$ ）*

特征	1-3 个事件与 0 个事件 aOR (95% CI)†	≥4 个事件与 0 个事件 aOR (95% CI)†
诊断时的年龄（岁）	1.02 (1.0-1.04)	1.06 (1.04-1.07)
性别		
男	-	-
女	1.94 (1.24-3.06)	3.61 (2.26-5.76)
诊断时的结核病症状		
<3 种症状	-	-
≥3 种症状	1.11 (0.79-1.55)	1.54 (1.07-2.20)
双侧肺空洞形成		
否	-	-
是	1.19 (0.77-1.85)	1.84 (1.18-2.89)
接受治疗的天数		
≤366	-	-
367-550	1.64 (1.05-2.58)	1.63 (1.00-2.68)
551-642	1.64 (1.04-2.58)	1.80 (1.10-2.96)
≥643	1.99 (1.22-3.23)	2.88 (1.73-4.80)
* 本项分析中的总数 n 比病例总数稍少，因为存在缺失信息的观察对象被排除在多元分析之外。 † aOR 和 CI 在第一列中针对患者特征与经历 1-3 次不良事件（相比没有不良事件）之间的关联，在第二列中针对患者特征与经历 4 次或更多次不良事件（相比没有不良事件）之间的关联。 MDR-TB = 耐多药结核病；aOR = 调整后优势比；CI = 置信区间；SD = 标准差		

文字框 5 - 续

vii) 接受治疗的耐多药结核患者的转归决定因素

可以利用监测数据开展更高级的分析 – 检验这些假设和量化风险或干预对特定效果的影响。这些分析要求个体病例层面具备可用的数据，数据可在治疗登记簿中找到。首先应当评估数据是否完整、没有错误和属于最新状态。

分析接受治疗的所有结核患者的数据以了解转归的决定因素时，可以发现感染耐药菌株通常与失败或死亡相关联，而不论是在什么场所。在一些国家，耐药结核患者的病例数量非常大，完全可以开展耐多药结核病病例的单独分析。在下面秘鲁的示例中（表 8）（22），分析帮助开展了一项研究，其中药物方案构成与死亡风险之间的关联具有统计显著性。在这种情况下，危险比被用来表述患者观察期间死亡的瞬时风险（即与上一个示例中使用的优势比相对，后者捕获与研究中各项变量相关联的耐多药结核病累积风险）。

表 8
秘鲁积极治疗方案与死亡时间的多元分析（1999 年至 2002 年）
（n=699）

变量	危险比，多元分析	95% CI, 多元分析
每月接受积极治疗方案的时间	0.63	0.43, 0.93
既往接受过≤2 种治疗方案，没有 CER	0.43	0.25, 0.74
女性	1.45	1.02, 2.07
年龄	1.01	1.00, 1.03
低 BMI 或营养不良	2.45	1.63, 3.68
心动过速	2.19	1.50, 3.19
肺外结核病	1.68	1.05, 2.68
至少一种共病，不同于艾滋病病毒感染	1.71	1.21, 2.43
艾滋病病毒感染	2.72	1.03, 7.24
耐药因子的数量	1.03	0.92, 1.15

doi: 10.1371/journal.pone.0058664.t004

积极治疗方案 = 至少五种可能有效的药物，包括一种注射药物（阴转后至少持续六个月）和一种氟喹诺酮类药物；CER= 秘鲁最初用于治疗耐多药结核病的标准化治疗方案（包含卡那霉素-环丙沙星-吡嗪酰胺-乙胺丁醇-乙硫异烟胺）；BMI = 身体质量指数；CI = 置信区间。

5.4 结论

本章应当让读者更好地了解相关方法，这些方法被用来描述耐药结核病数据、分析耐药结核病调查、生成耐多药结核病关怀相关规划绩效的主要指标、开展一些额外分析来检验假设（例如耐药或特定转归的风险因素）。

为了更好地掌握调查的设计和分析，或者了解哪些统计检验最适合回答特定问题，建议参考更多专业文献以及获得专家在调查方法和生物统计方面的支持。

还鼓励读者更多地熟悉数据分析软件，一些软件可以免费下载和使用，例如其中一种软件可以用来制作本章中的绝大多数图形（23）。

1. Guidelines for surveillance of drug resistance in tuberculosis. 4th ed. (WHO/TB/2009.422). Geneva: World Health Organization; 2009.
2. Multidrug-resistant tuberculosis (MDR-TB) indicators. A minimum set of indicators for the programmatic management of MDR-TB in national tuberculosis control programmes (WHO/HTM/TB/2010.11). Geneva: World Health Organization; 2010. 网址: whqlibdoc.who.int/hq/2010/WHO_HTM_TB_2010.11_eng.pdf
3. Definitions and reporting framework for tuberculosis—2013 revision (WHO/HTM/TB/2013.2). Geneva: World Health Organization; 2013. 网址: www.who.int/iris/bitstream/10665/79199/1/9789241505345_eng.pdf
4. Companion handbook to the WHO guidelines for the programmatic management of drug-resistant tuberculosis. Geneva, World Health Organization. *In production, 2014.*
5. Electronic recording and reporting for tuberculosis care and control (WHO/HTM/TB/2011.22). Geneva: World Health Organization; 2012.
6. Jenkins HE, Plesca V, Ciobanu A, Crudu V, Galusca I, Soltan V, et al. Assessing spatial heterogeneity of MDR-TB in a high burden country. *Eur Respir J.* 2013 Nov;42(5):1291–301.
7. A practical handbook on the pharmacovigilance of medicines used in the treatment of tuberculosis: enhancing the safety of the TB patient. Geneva: World Health Organization; 2012. 网址: www.who.int/medicines/publications/Pharmaco_TB_web_v3.pdf
8. Skrahina A, Hurevich H, Zalutskaya A, Sahalchik E, Astrauko A, Hoffner S, et al. Multidrug-resistant tuberculosis in Belarus: the size of the problem and associated risk factors. *Bull World Health Organ.* 2013;91:36–45.
9. Ulmasova D, Uzakova G, Tillyashayhov M, Turaev L, van Gemert W, Hoffmann H, et al. Multidrug-resistant tuberculosis in Uzbekistan: results of a nationwide survey, 2010 to 2011. *Euro Surveill.* 2013 Oct 17;18(42). 网址: <http://www.eurosurveillance.org/images/dynamic/EE/V18N42/art20609.pdf>
10. Ministry of Health of the Russian Federation, et al. Tuberculosis in the Russian Federation 2011: an analytical review of statistical indicators used in the Russian Federation and in the world. Moscow, Russian Fed; 2013.
11. Anti-tuberculosis drug resistance in the world. Report No. 4. (WHO/HTM/TB/2008.394) Geneva: World Health Organization; 2008.
12. Guidelines for the programmatic management of drug-resistant tuberculosis, Emergency update 2008. (WHO/HTM/TB/2008.402). Geneva, World Health Organization. 2008.
13. Global tuberculosis report 2012 (WHO/HTM/TB/2012.6). Geneva: World Health Organization; 2012.
14. Global tuberculosis report 2013 (WHO/HTM/TB/2013.11). Geneva: World Health Organization; 2013.
15. Falzon D, Jaramillo E, Wares F, Zignol M, Floyd K, Raviglione MC. Universal access to care for multidrug-resistant tuberculosis: an analysis of surveillance data. *The Lancet Infectious Diseases.* 2013 Aug;13(8):690–7. 网址: [http://www.thelancet.com/journals/laninf/article/PIIS1473-3099\(13\)70130-0/abstract](http://www.thelancet.com/journals/laninf/article/PIIS1473-3099(13)70130-0/abstract)
16. Lin H-H, Shin SS, Contreras C, Asencios L, Paciorek CJ, Cohen T. Use of Spatial Information to Predict Multidrug Resistance in Tuberculosis Patients, Peru. *Emerg Infect Dis.* 2012 May;18(5):811–3.
17. Briefing package: Division of Anti-Infective Products Office of Antimicrobial Products CDER, FDA. *Sirturo™ (bedaquiline 100 mg tablets) For the treatment of adults (≥ 18 years) as part of combination therapy of pulmonary multi-drug resistant tuberculosis (MDR-TB).* Applicant: Janssen Research and Development, L.L.C. FDA Anti-Infective Drugs Advisory Committee Meeting Silver Spring, MD Date: November 28, 2012. 网址: <http://www.fda.gov/downloads/AdvisoryCommittees/CommitteesMeetingMaterials/Drugs/Anti-InfectiveDrugsAdvisoryCommittee/UCM329258.pdf>
18. Holtz TH, Sternberg M, Kammerer S, Laserson KF, Riekstina V, Zarovska E, et al. Time to sputum culture conversion in multidrug-resistant tuberculosis: predictors and relationship to treatment outcome. *Ann Intern Med.* 2006 May 2;144(9):650–9.

19. Improving the diagnosis and treatment of smear-negative pulmonary and extrapulmonary tuberculosis among adults and adolescents. Recommendations for HIV-prevalent and resource-constrained settings (WHO/HTM/TB/2007.379). Geneva, World Health Organization; 2007. 网址: whqlibdoc.who.int/hq/2007/WHO_HTM_TB_2007.379_eng.pdf
20. Jacobson KR, Tierney DB, Jeon CY, Mitnick CD, Murray MB. Treatment outcomes among patients with extensively drug-resistant tuberculosis: systematic review and meta-analysis. *Clin Infect Dis*. 2010 Jul 1;51(1):6–14.
21. Bloss E, Kuksa L, Holtz TH, Riekstina V, Skripconoka V, Kammerer S, et al. Adverse events related to multidrug-resistant tuberculosis treatment, Latvia, 2000–2004. *Int J Tuberc Lung Dis*. 2010 Mar;14(3):275–81.
22. Mitnick CD, Franke MF, Rich ML, Alcantara Viru FA, Appleton SC, Atwood SS, et al. Aggressive Regimens for Multidrug-Resistant Tuberculosis Decrease All-Cause Mortality. *PLoS ONE*. 2013 Mar 13;8(3):e58664.
23. R Project. R: A language and environment for statistical computing. Vienna: R Foundation for Statistical Computing; 2013. 网址: www.R-project.org/

第 6 章

艾滋病相关结核病：负担和防治措施分析

目标受众：

普通读者，尤其是国家结核病防治规划督导与评估官员和国家规划管理人员。

预期成果：

进一步了解如何实施下列各项工作：

- 开展艾滋病相关结核病监测数据的各种分析；
- 评估艾滋病相关结核病的负担；
- 考查数据质量；
- 了解“代表性”的含义。

作者：

Philippe Glaziou

6.1 艾滋病相关结核病简介

据世界卫生组织估计，在 2012 年 870 万例新发结核病病例中，110 万例（13%）是人类免疫缺陷病毒（艾滋病病毒、HIV）感染者。感染 HIV 会极大地加强活动性结核病的风险。平均而言，HIV 感染者的结核病发病率大约为类似地区未感染 HIV 者的结核病发病率的 30 倍。

结核病是 HIV 感染者最常发生的疾病，包括接受抗逆转录病毒治疗（ART）的感染者。大约 75% 的双重感染者都在撒哈拉以南非洲。结核病是 HIV 感染者的头号杀手，占艾滋病相关死亡的五分之一。据估计 2012 年有 32 万人死于艾滋病相关结核病。

监测新发现的结核病病例（特别是在高艾滋病病毒现患率地区）有多种必要性。首先，它可以为估测抗逆转录病毒治疗的需求量提供基准。所有发现的结核与 HIV 合并感染病例都应当在结核病治疗启动后的 8 周内开始接受抗逆转录病毒治疗（1）。其次，它可以揭示结核病报告病例当中的 HIV 阳性病例的发现人数和预期人数之间的差额。第三，通过在不同的时间点收集信息，连同其他信息（例如新发现的结核病患者的 HIV 检测率）一起可以揭示结核与 HIV 合并感染负担的变化趋势。

6.2 规划数据的分析

结核与 HIV 合并感染数据的常规分析应当提供两方面的信息：结核与 HIV 合并感染联合防治工作的实施力度（2）以及结核与 HIV 合并感染的负担及其时间趋势（3）。本节探讨如下重要主题：确定患有艾滋病相关结核病的患者人数，并与预期人数进行比较；发现同时接受结核病治疗和抗逆转录病毒治疗的患者（抗逆转录病毒治疗是能够挽救生命的十分重要的治疗，所有 HIV 感染者都应当在结核病治疗启动后的 8 周内开始接受抗逆转录病毒治疗）；结核病关怀的转归；数据质量。

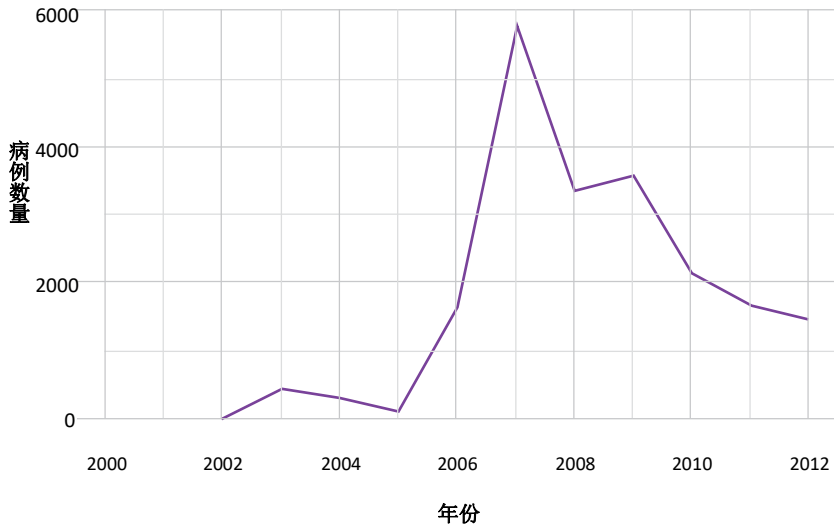
艾滋病与结核病病例发现

通常而言，国家艾滋病防治项目维护着在卫生机构接受医疗的 HIV 感染者的登记簿，其中包含关于他们的结核病患病情况的信息。国家结核病防治规划也维护着结核病患者登记簿，其中包含他们的艾滋病病毒感染状况信息。当 HIV 感染者被确诊患有结核病时，可以通过艾滋病关怀服务体系发现 HIV 感染与结核病之间的关联。当结核病病例被诊断或在结核病治疗期间接受 HIV 检测时，同样可以通过艾滋病关怀服务体系发现这种关联。

理想情况下，通过艾滋病监测系统报告的 HIV 阳性结核病病例数应当与通过结核病监测系统报告的 HIV 阳性病例数相等。而现实中，数字很少匹配，原因是负责管理各个疾病信息系统的团队之间的沟通不顺畅、对于指标的理解有偏差，以及两种信息系统的局限性。不过，如果 HIV 和结核病信息系统所登记的病例是相关的（第二章）并且使用全国唯一识别号，那么通过记录的关联可以改善结核病和 HIV 感染患者相关记录的匹配程度。此外，结核病和艾滋病联合防治服务可以维护相互关联的综合信息系统，从而有效地共享两种疾病的信息。

图 1 显示了柬埔寨发现的和报告的 HIV 阳性结核病病例数据。

图 1
柬埔寨发现和报告的 HIV 阳性结核病病例（2002 至 2012 年）
资料来源：世界卫生组织全球结核病数据库，2013 年



在 2002 年之前，新发现的结核病患者的 HIV 检测结果没有得到系统地报告。

发现和报告的 HIV 阳性结核病病例数目在 2002 至 2005 年间一直比较低，在 2007 年则达到 6000 例，而后逐步降低。

2005 至 2007 年间的合并感染病例数上升可能是由于结核病的流行病学变化，或者由于病例发现工作的改善或报告和发现的病例数上升。例如，下列方面可能得到改善：结核病通报病例的 HIV 检测结果报告工作得到加强；HIV 检测覆盖面加大；HIV 感染者当中的结核病筛查和检测覆盖面加大；结核病患者的 HIV 负担加重。

同样，2007 年以后的病例数降低可能是由于：HIV 检测结果的报告减少；HIV 检测覆盖缩小；HIV 感染者当中的结核病筛查和检测覆盖面缩小；结核病患者的 HIV 负担减轻。

还有必要了解接受了 HIV 检测的结核病病例总数量。例如，如果某一年内接受了 HIV 检测的结核病病例数量很低，那么其中检测结果为 HIV 阳性的病例的百分比在各年间可能会因几率的波动而有差别。

此外，如果大量结核病患者当中只有相对较少的患者接受了 HIV 检测，则可能当初他们不是随机被抽取到的。一些患者可能被视为重点检测对象，例如当他们的医生感觉他们比其他患者健康状况更差时。当解释接受检测者 HIV 现患率的时间趋势时，务必要考虑接受了检测的人群是否与没有接受检测者有系统性的差异。图 2 显示了在结核病诊断时或在结核病治疗期间有明确 HIV 检测结果的新结核病病例的时间序列。

2006 至 2010 年间，接受 HIV 检测的结核病病例飞速增长，而后则进入平台期。

如果我们用第一个指标除以接受了 HIV 检测的结核病病例数量，则可以得出这些病例的 HIV 现患率，如下面的图 3 所示。

图 2

柬埔寨接受 HIV 检测的结核病病例数量（2003 至 2012 年）

资料来源：世界卫生组织全球结核病数据库，2013 年

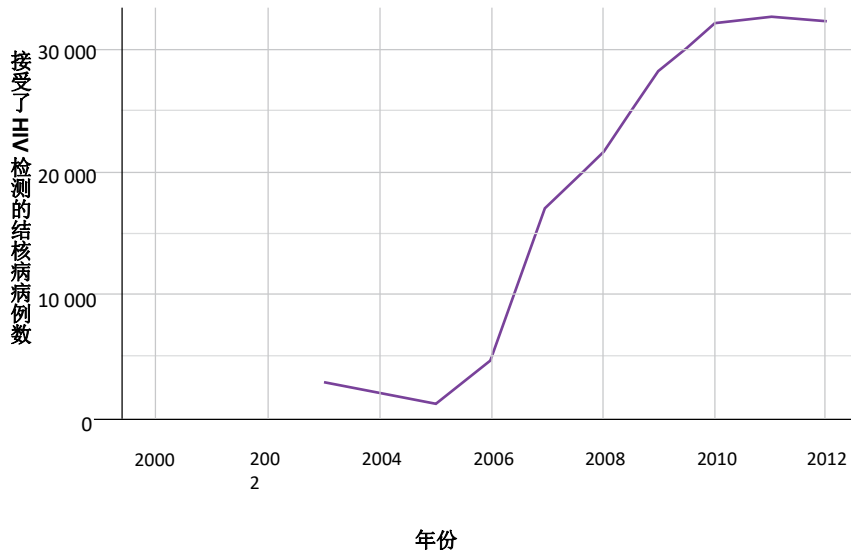
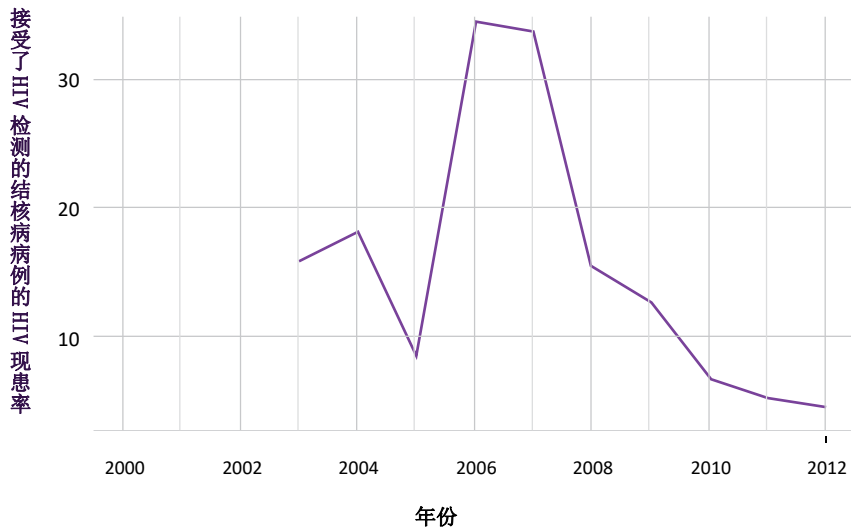


图 3

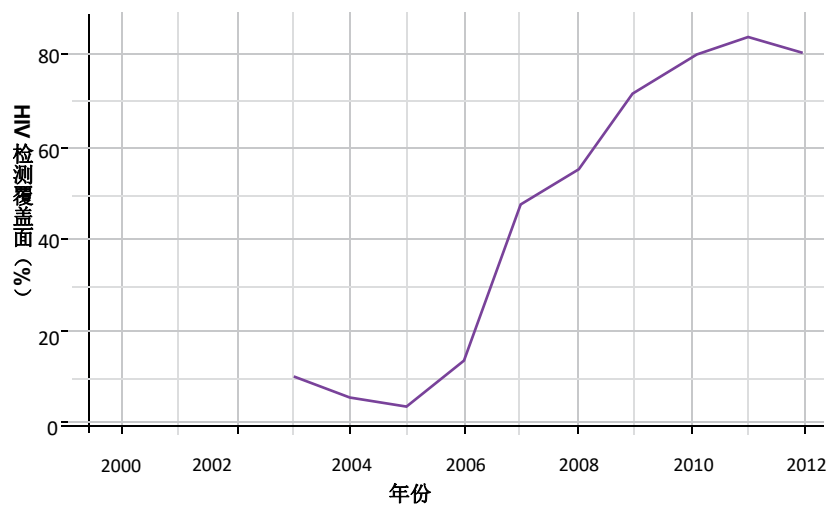
柬埔寨接受了 HIV 检测的结核病病例的艾滋病病毒现患率（2003 年至 2012 年）

资料来源：世界卫生组织全球结核病数据库，2013 年



在解释上述数据时，HIV 检测覆盖面是一个重要的考虑因素。如果覆盖面低于结核病病例的 100%，并且接受了检测的患者不是随机抽选的，则接受了检测的人的 HIV 现患率可能会与未接受检测的人的 HIV 现患率不同。图 4 显示了结核病患者的 HIV 检测覆盖面，以报告的结核病病例总数的百分比表示。

图 4
柬埔寨结核病病例的 HIV 检测覆盖面（2003 年至 2012 年）
资料来源：世界卫生组织全球结核病数据库，2013 年



2006 年以前，HIV 检测覆盖面低于 20%，而后于 2010 至 2012 年间逐渐升至 80%。2010 年以后，检测覆盖面比较高，因此我们信任通过常规检测得出的结核病病例艾滋病病毒现患率。这是由于受检人群很可能在统计学上可以代表整个结核病患者人群。但是 2010 年以前的信息质量较差。当检测覆盖面低于 80% 的时候，受检人群和未受检人群的艾滋病病毒现患率的差异性对于结果可能会产生较大影响，对于发现的结核病病例的估测 HIV 现患率和实际现患率的差异性也可能有较大影响。

为了优化资源利用、给患者提供最佳关怀，有必要确定结核病患者 HIV 现患率。在覆盖面不足的情况下，根据常规检测结果做出的任何估测都可能因患者选择偏倚而不准确。

不过，仍有可能获得已知结核病病例 HIV 现患率的不偏倚的估测值，方法是检测每一例新发现的结核病病例。然而，如果缺乏资源并且难以实地实施，最好开展一次调查，随机抽选新结核病病例（例如国内 1000 个病例）并对他们进行检测。如果所选的样本的确是随机抽选的，那么理论上讲调查结果是不偏倚的。通过调查得出的艾滋病病

a. 简单随机抽样要求一个国家要有所有结核病患者完整清单，这种方法是最基本的无偏倚结核病患者抽样方法。简单随机抽样可以是复杂随机选择方案的组成内容。例如，如果一个国家没有结核病患者完整清单，但是有结核病门诊清单，那么可以随机抽选结核病门诊。各门诊可以制定随机抽样患者清单（Smith PG and Morrow RH. Field trials of health interventions in developing countries: a toolbox. Macmillan Education 2005. ISBN 0 333 64058 6）。

毒现患率估测值仍有可能不同于所有结核病病例的实际现患率值，但是如果在类似条件下反复多次开展调查，则调查的平均估测值会与真实值匹配。柬埔寨于 2003 和 2005 年分别开展了一次调查（4）。根据调查结果制定的表 1 显示了结核病患者的艾滋病病毒现患率不像上述常规检测数据所提示的那样高。

这两次调查还允许人们对结核病病例以及不同亚人群的艾滋病病毒现患率进行比较。这两次调查都表明 25 至 44 岁年龄段的患者的艾滋病病毒现患率最高。这项发现允许项目确定常规 HIV 检测的重点目标人群，从而更有效地运用宝贵资源。

在 2008 和 2009 年，开展了另外两次有全国代表性的结核病病例艾滋病病毒现患率调查，测得的现患率分别为 6.4% 和 6.3%。调查结果比根据两轮常规检测得出的结果（分别为 15% 和 13%）要低许多，这提示常规检测的结核病病例的艾滋病病毒现患率比实际值要高。在免费 HIV 检测资源不足的情况下，医生往往会优先检测那些病重的患者或者可能更容易感染 HIV 的患者。因此，在受检人群中 HIV 阳性病例的比例可能会比全部结核病病例中 HIV 阳性病例的比例要高。图 5 表明了常规检测的结核病患者的 HIV 现患率时间趋势，以及上述四次全国调查估测的结核病患者艾滋病病毒现患率。

图 5
柬埔寨受检结核病病例的艾滋病病毒现患率（实线）以及四次全国调查的结果（红色 x 符号）（2002 至 2012 年）。资料来源：世界卫生组织全球结核病数据库，2013 年

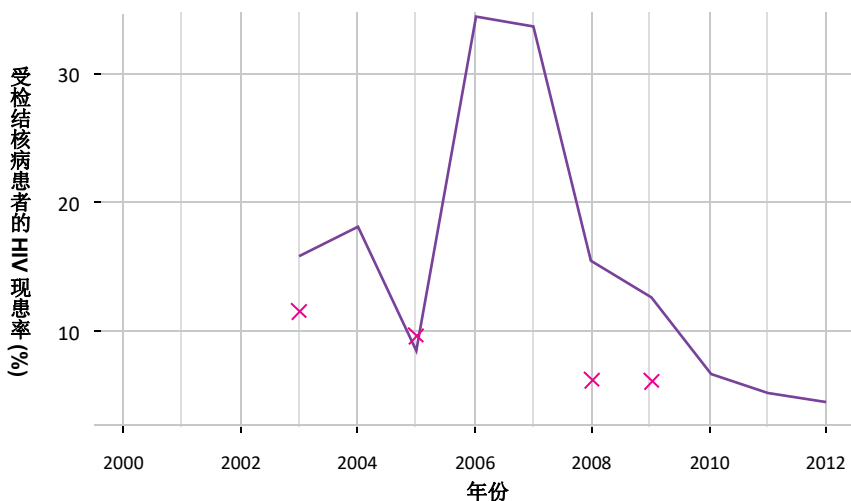


表 1
柬埔寨结核病患者 HIV 阳性率（根据 2003 和 2005 年调查）（4）

	2003		2005	
	HIV 阳性 (%)	95% CI*	HIV 阳性 (%)	95% CI*
所有受检结核病	11.8	10.5 – 13.2	9.9	8.8 – 11.1
男性	13.5	11.7 – 15.6	10	8.4 – 11.6
女性	9.9	8.2 – 11.9	9.8	8.2 – 11.5
年龄段				
0-14	15	7.1 – 26.6	7.7	0.4 – 14.9
15-24	9.7	6.1 – 14.5	6.2	3.1 – 9.3
25-34	27.2	22.6 – 32.2	24.8	20.6 – 28.9
35-44	16.9	13.8 – 20.5	17.5	14.5 – 20.6
45-54	5.7	3.8 – 8.4	5.2	3.3 – 7.2
55-64	4.7	2.9 – 7.4	1.7	0.5 – 2.9
65+	4.3	2.2 – 7.4	0.5	0 – 1.3

* 95%置信区间（CI）。95%置信区间的显著性是指：如果多次开展一项调查并且每次都计算 95%置信区间，则在置信区间的 95%范围内可以得出现患率真值。由于每项调查只开展一次，所以无法得知真值是否确实位于该区间内。

关怀流程

艾滋病治疗流程是一套系统，用于监测有多少 HIV 感染者实际获得他们需要的医疗关怀和治疗。它旨在确立一系列步骤，以便需要艾滋病关怀的每个人都能够一直获得关怀 – 从接受 HIV 检测一直到通过治疗抑制病毒。

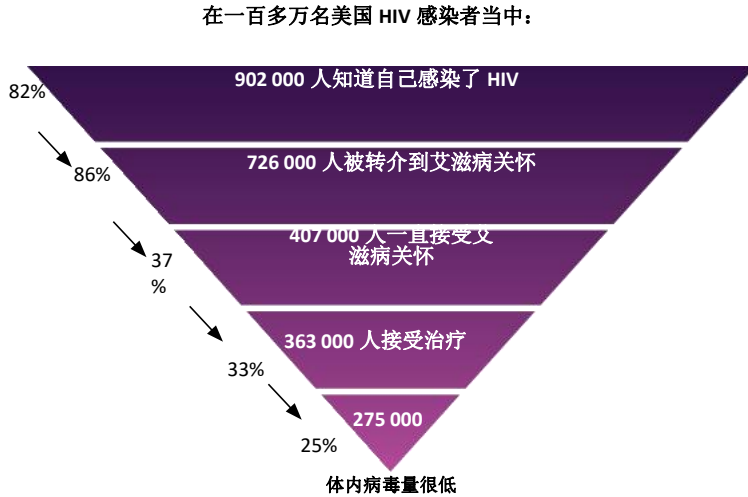
艾滋病治疗流程有助于探讨一些关键问题，包括：

- 有多少 HIV 感染者获得检测和诊断？
- 其中有多少人获得医疗关怀转介？
- 其中有多少人一直获得关怀？
- 其中有多少人获得抗病毒治疗？
- 其中有多少人能够依从治疗方案并且实现了病毒抑制？

通过密切考查这些不同的步骤，政策制定者和服务提供方能够发现在转介 HIV 感染者、使其获得持续的优质关怀方面存在哪些缺口。知道主要的缺口及其影响的人群，

这有助于国家、州和地方政策制定者和服务提供方改善系统和服务，从而支持患者在关怀流程中从一个步骤前进到下一个步骤。图 6 举例说明了美国的艾滋病关怀流程（5）以及容易发生患者中途退出治疗的几个关键点。

图 6
美国的艾滋病关怀流程（2010）



世界卫生组织推荐的结核与 HIV 合并感染管理策略（2）包括（以结核病诊断作为切入点）：

- 及早对新诊断的结核病患者进行 HIV 检测；
- 对 HIV 阳性结核病患者进行复方新诺明治疗；
- 在启动结核病治疗后的八周（对于严重艾滋病病例则为 2 周）内对 HIV 阳性结核病患者进行抗逆转录病毒治疗。

以 HIV 诊断作为切入点：

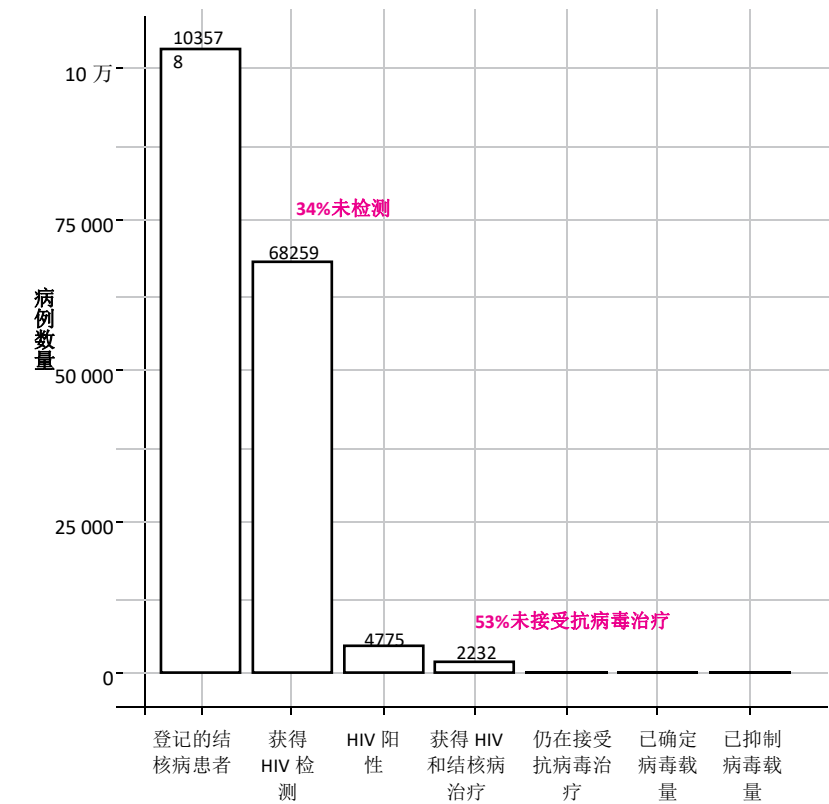
- 系统地筛查结核病症状，采用 Xpert MTB/RIF 进行诊断，并且根据需要提供相应治疗；
- 对于已排除活动性结核的患者提供异烟肼预防治疗。

从卫生机构管理的角度：

- 实施感染控制以便预防结核传播。

图 7 显示了越南在 2012 年以前采用的结核与 HIV 合并感染关怀流程，其中存在的问题用红色字体标示出来：大约四分之一的新登记的结核病患者没有接受 HIV 检测，并且超过一半的 HIV 阳性病例没有获得抗病毒治疗。

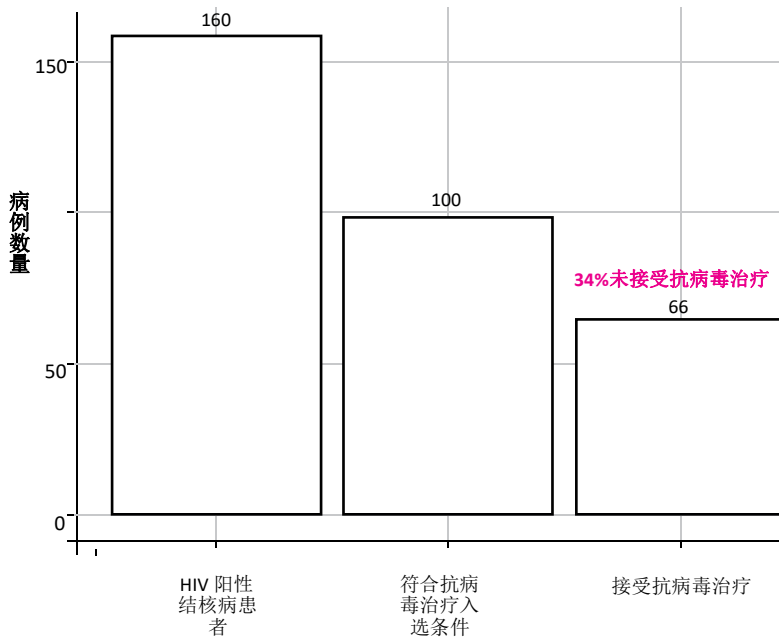
图 7
越南结核与 HIV 合并感染关怀流程（2012 年）



也可以从 HIV 诊断切入点来审视结核与 HIV 合并感染关怀流程，因为所有新诊断的 HIV 阳性患者都应当接受结核病症状筛查，然后在确诊结核病病例后会接受结核病和 HIV 治疗。

下图（图 8）显示了南非 2011 年采用的结核与 HIV 合并感染关怀流程（6），它表明三分之一的符合抗病毒治疗条件的结核病患者没有接受抗病毒治疗（在 2011 年，南非结核患者的抗逆转录病毒治疗的入选条件是 CD4 细胞计数大于 200 个细胞/ μ l，但是该标准目前在国际上已不适用：国际标准是不论 CD4 计数如何，所有结核病患者都应当在启动结核病治疗（1,2）后 8 周内开始接受抗逆转录病毒治疗）。通过考查结核病与 HIV 合并感染关怀流程，应当可以确定需要采取哪些改进措施来确保所有患者获得关怀。

图 8
南非某门诊的结核与 HIV 合并感染关怀流程（2011 年）



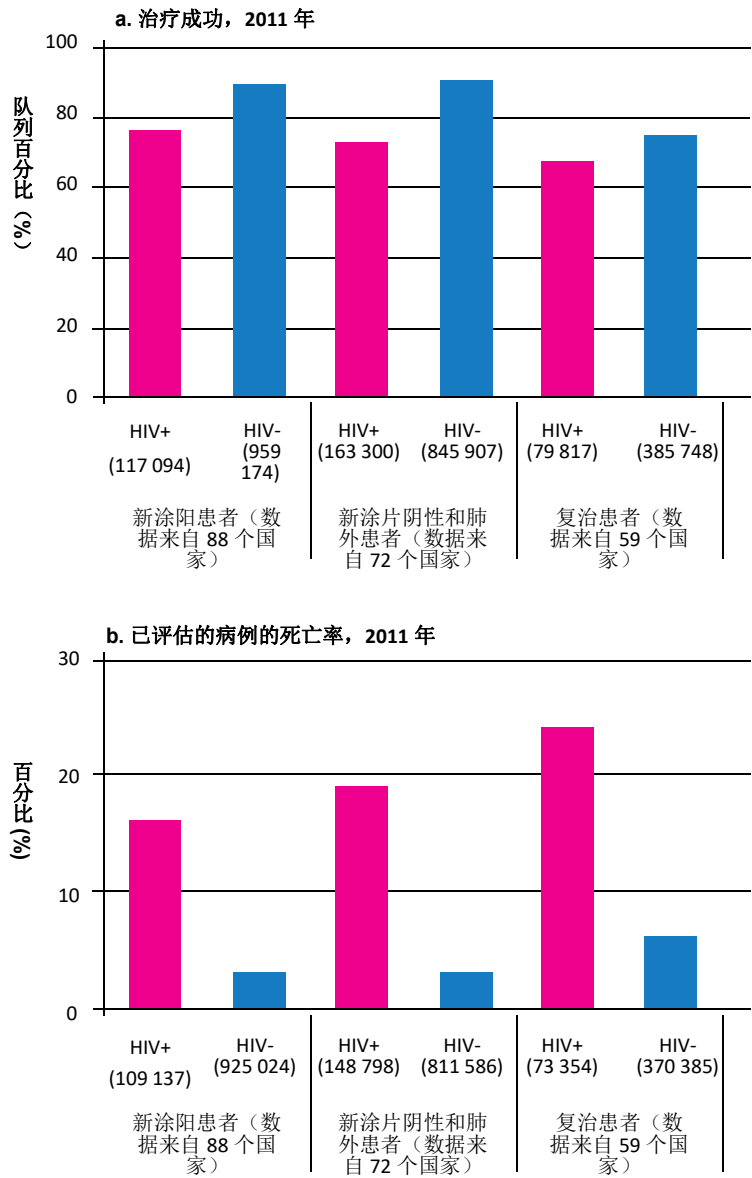
治疗转归的队列分析

评估结核病治疗和关怀（包括患者支持）的绩效的传统方法是分析 HIV 阳性结核病患者的治疗转归。世界卫生组织建议单独报告 HIV 阳性结核病患者队列的治疗转归（2,7,8）。

摘自《2013 年全球结核病报告》（3）的图 9 显示了全球 HIV 阳性和 HIV 阴性结核病患者的治疗转归分布情况（9）。

为了更全面地分析治疗转归，需要考查未接受治疗的确诊结核病患者的比例，以及治疗失访者当中的死亡病例的实际比例。在有病例相关信息系统的地方（参见第二章），为了确定与不良转归相关的因素，需要按年龄段、性别和关键人群类别（例如在押犯人、吸毒者等）来细分转归。了解这些因素有助于确定改进措施。

图 9
全球结核病治疗成功率和死亡率（死亡率涵盖治疗失访的患者）（按艾滋病病毒感染状况划分）（2011 年队列）



HIV 阳性患者的治疗成功率要低得多（图 9“a”部分顶端），这是由于较高的病例死亡率（“b”部分底部）。在结核病治疗期间及早启动抗逆转录病毒治疗可大幅降低死亡风险（1, 10–13）。

1. Consolidated guidelines on the use of antiretroviral drugs for treating and preventing HIV infection. Geneva: World Health Organization; 2013.
2. WHO policy on collaborative TB/HIV activities. Geneva: World Health Organization; 2012.
3. Global TB Report 2013. Geneva: World Health Organization; 2013.
4. Tamura M, Eam KK, Kimura K, et al. National HIV prevalence surveillance among TB patients through periodic surveys: experience in Cambodia. *International Journal of Tuberculosis and Lung Diseases* 2008; 12(3): 520-5.
5. Today's HIV/AIDS epidemic, 2013. Atlanta: US Centers for Disease Control and Prevention; 2013.
6. Pepper DJ, Marais S, Wilkinson RJ, Bhaijee F, De Azevedo V, Meintjes G. Barriers to initiation of antiretrovirals during antituberculosis therapy in Africa. *PLoS One* 2011; 6(5): e19484.
7. Treatment of Tuberculosis Guidelines. Geneva: World Health Organization; 2010. p. 1-160.
8. Definitions and reporting framework for tuberculosis - 2013 revision. Geneva: World Health Organization; 2013.
9. Global Tuberculosis Report 2012. Geneva: World Health Organization; 2012.
10. Holtz TH, Kabera G, Mthiyane T, et al. Use of a WHO-recommended algorithm to reduce mortality in seriously ill patients with HIV infection and smear-negative pulmonary tuberculosis in South Africa: an observational cohort study. *The Lancet Infectious Diseases* 2011; 11(7): 533-40.
11. Blanc F-X, Sok T, Laureillard D, et al. Earlier versus later start of antiretroviral therapy in HIV-infected adults with tuberculosis. *New England Journal of Medicine* 2011; 365(16): 1471-81.
12. Abimbola TO, Marston BJ, Date AA, Blandford JM, Sangrujee N, Wiktor SZ. Cost-Effectiveness of Tuberculosis Diagnostic Strategies to Reduce Early Mortality Among Persons With Advanced HIV Infection Initiating Antiretroviral Therapy. *JAIDS Journal of Acquired Immune Deficiency Syndromes* 2012; 60(1): e1-e7.
13. Abdool Karim SS, Naidoo K, Grobler A, et al. Timing of initiation of antiretroviral drugs during tuberculosis therapy. *N Engl J Med* 2010; 362(8): 697-706.

第 7 章

利用人口动态登记和死亡率调查数据 估算结核病死亡率

目标受众：

1. 流行病学家、统计学家和国家结核病防治规划督导与评估工作人员。
2. 希望了解如何估测结核病死亡率以及其趋势分析对于项目的指导意义的国家结核病防治规划工作人员。

预期成果：

本章结束时，读者应当：

- 熟知死亡率的主要数据源，包括死因归类所依据的“国际疾病分类”（ICD）系统；
- 了解如何得出 HIV 阴性和 HIV 阳性人群的结核病死亡率；
- 能够解释结核病死亡率的时间趋势和地域分布；
- 了解死亡数与通报数（M/N）比率指标；
- 了解如何估测耐多药结核病导致的死亡例数。

作者：

Charalambos Sismanidis, Philippe Glaziou, Dennis Falzon

7.1 死亡率数据的来源

国际疾病分类

“国际疾病分类”（ICD）系统是国际标准工具，用于系统地界定死亡率和发病率数据代码（1）。现行的 ICD 第十版（ICD 第 11 版将于 2015 年发布）界定了全面标准化的临床关怀和研究活动，例如疾病定义；管理卫生保健；分配资源和监测转归。此外，它还界定了死因（COD）标准代码。

表 1 列出了 ICD 最近两个版本中界定的结核病相关的主要死因代码以及不具体的或所谓“不明确”的死因的代码（例如结核病后遗症，即结核病导致的慢性阻塞性肺病等后遗症）。应当尽量避免使用“不明原因”。

表 1
结核病相关死因的标准代码

死因	ICD-9 代码 (1979-1998 年)	ICD-10 代码 (1999 至今)
结核病	010-018	A15-A19
呼吸系统结核病	010-012	A15
其他结核病	013-018	A17-A19
不明原因	137.0-137.4	R00-R99, B90.9, P37.0 以及 J65

户籍登记和人口动态统计数据系统

“户籍登记”是指重要事件的发生和特征的记录，例如出生、死亡（带有死因信息）和居民状态（例如婚姻记录）。

“人口动态统计数据”系统记录取自户籍登记系统的重要事件信息，目的是产生统计数据（2）。

有效的户籍登记和人口动态统计数据（CRVS）系统可以连续不断地产生与全人群数据一致且可比的数据。数据应当由与事件最接近的训练有素的专业人员来记录（3）。这种 CRVS 系统可以持续地产生死因统计数据。对于发生在卫生机构内的死亡以及在卫生

工作者在场情况下发生的死亡，可以根据 ICD 进行医学鉴定和编码（图 1）。如果死亡发生在卫生机构之外，可以采用死因推断（VA）方法来确定可能的死因。由于能够监测死因，CRVS 系统成了一个重要资源，例如针对特定疾病的项目（例如结核病项目）可以利用该系统确定工作重点、制定政策、分配资源和评估进展情况。

图 1
死亡医学证明示例（4）

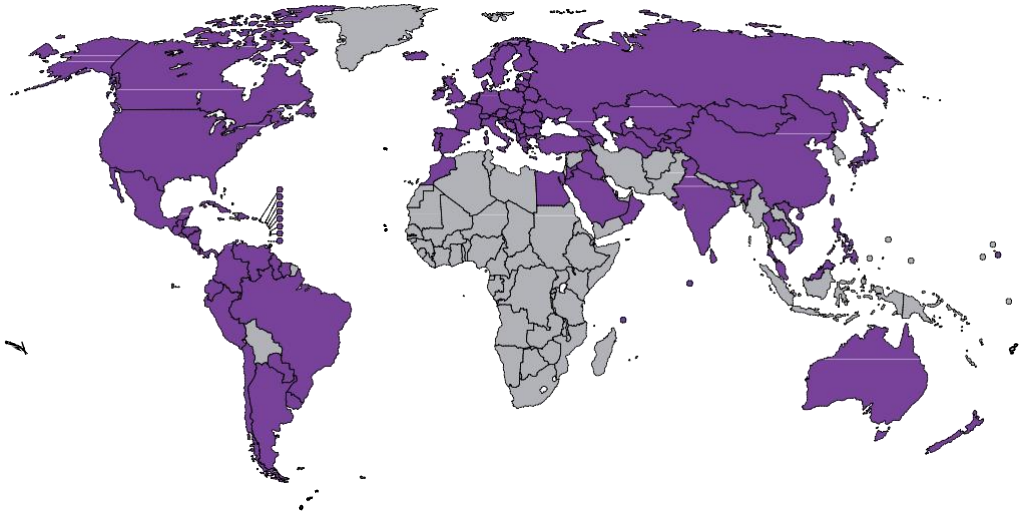
死亡原因	从发作到死亡的时间
I 直接导致死亡的疾病或健康状况* (a) <u>呼吸衰竭...</u> 起因（或成因）： 前因 (b) <u>肺纤维化...</u> 导致上述死因的病情（如果有的话），说明最近的潜伏问题 起因（或成因）： (c) <u>肺结核...</u> 起因（或成因）： (d)	<2 周.. 2 年.. 6 个月..
II 促成了死亡但与死因不相关的其他严重问题 <u>慢性支气管炎...</u> *这不是指死亡方式，例如心脏衰竭、呼吸衰竭，而是指导致死亡的疾病、损伤或并发症。	10 年..

医学证明从病原学角度描述了导致死亡的一系列事件。在这个示例中：
 （a）是指死亡的直接原因，（b）是指死亡的连带原因，（c）是（b）的连带原因，指的是触发一系列事件的潜在死亡促成因素。在本示例中，（c）是在 CRVS 系统中报告的原因。

虽然 CRVS 系统有这些显而易见的好处，但是许多国家没有在这方面进行投入，导致缺乏高质量的产出（5）。因此，大部分结核病高流行的国家目前都没有产生关于潜在死因的可靠信息（图 2）。世界卫生组织评估了全球各地人口动态登记数据的质量（6）。统计方法可以用于发现不完整的覆盖面或错误的编码。国际社会的政治承诺目前很强，“全球 CRVS 合作伙伴行动倡议”强调指出在亚洲和非洲开发 CRVS 系统的必要性，因为这两大洲的结核病疾病负担很高，导致需要量很大（7）。

图 2

2012 年以来有国家人口动态登记系统（人口动态登记系统示例）并且向世界卫生组织报告死亡率数据和死因分布情况的国家（用紫色表示）



通过人口动态登记系统报告的结核病死亡局限于 HIV 阴性个体的结核病死亡。在 ICD-10 的艾滋病相关章节中界定了 HIV 阳性结核病患者的死亡促成因素代码。不过，由于错误编码率较高，HIV 阳性个体的死亡促成因素往往难以分析。

死因推断

如果缺乏能够连续不断地产生高质量数据的 CRVS，或者死亡发生在卫生机构以外，则可以采用死因推断方法来确定死亡的可能原因。该方法涉及访谈家庭成员和/或在患者死亡前照料患者的人，以便确定死因（8）。有时候在人口统计学和健康问卷调查过程中采用该方法。死因推断有一些局限性，例如该方法不适用于确定某些死因。事实上，根据死因推断来判断结核病患者的死因往往会造成挑战，特别是当现有患者记录没有给出结核病（例如细菌学、胸部 X 线）证据时。

7.2 监测 HIV 阴性个体的结核病死亡率

在可以获得人口动态登记数据的国家（参见图 2），有可能监测结核病死亡率，方法是根据 CRVS 捕获的数据和每年报告给世界卫生组织的数据来绘制出结核病死亡分布情

况（9）。质量不可靠的数据，例如从覆盖面很低的系统获得的异常值和数据点，被排除在分析之外或者在经过调整后用于分析。报告的结核病死亡率数据会被调高，以便体现不充分的覆盖面（即没有注明死因的死亡例数的估计值）和不明原因的死亡（[文字框 1](#)）。

一个国家在某一年的估计死亡总例数是从根据联合国人口估计数据的生命年表获得的死亡率估计值得来（10，11）。通常而言，假定没有记录在人口动态登记系统中的结核病死亡比例与人口动态登记系统记录的死亡例数中结核病死亡的比例相同。对于人口动态登记系统记录的不明原因死亡，假定归因于结核病的死亡的比例与登记的死亡中观察到的比例相同。

经过调整的结核病死亡例数 d_a 根据人口动态登记系统报告的结核病死亡例数（ d ）计算得出，公式如下：

$$d_a = \frac{d}{c(1-g)}$$

其中： c 是覆盖面（即注明了死因的死亡例数除以估计死亡总例数）， g 是不明原因的死亡所占的比例（[文字框 1](#)）。

根据标准差来估算这样的调整所涉及的不确定度：

$$SD = \frac{d}{4} \left[\frac{1}{c(1-g)} - 1 \right]$$

不确定度的计算没有考虑到错误编码，例如艾滋病相关死亡被错误地归类为结核病相关死亡。

文字框 1
针对 CRVS 记录的不明原因的死亡来调整结核病死亡率：
泰国的示例

在 2013 年泰国全国结核病防治规划评估过程中，全国结核流行病学评估人员详细审视了常规收集的监测数据。这些数据中包含国家 CRVS 系统死因信息记录（表 2）。

表 2
国家 CRVS 系统中的死因数据

年份	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
1990	399134	252512	109184	0.63	0.43	3937	10964	6.96	19.38
1991	414130	264347	113536	0.64	0.43	3663	10041	6.41	17.59
1992	434683	275313	117130	0.63	0.43	3595	10011	6.23	17.13
1994	450482	305526	111137	0.68	0.36	3473	7980	5.94	13.76
1995	459558	324842	114598	0.71	0.35	4144	8979	7.03	15.36
1996	490440	342643	121648	0.70	0.36	4622	10317	7.76	17.22
1997	436101	298084	103739	0.68	0.35	3697	8364	6.14	13.78
1998	447274	310512	116981	0.69	0.38	4252	9939	6.98	16.14
1999	524710	362607	151140	0.69	0.42	5265	13156	8.54	21.20
2000	514438	365583	149567	0.71	0.41	6246	14910	10.02	23.86
2002	514526	380364	143060	0.74	0.38	6751	14714	10.58	22.94
2003	533220	384131	127487	0.72	0.33	6906	14316	10.71	22.25
2004	565150	393592	149212	0.70	0.38	6076	14000	9.34	21.59
2005	528387	395374	151096	0.75	0.38	5534	11901	8.44	18.26
2006	526734	391126	149918	0.74	0.38	5214	11364	7.91	17.28

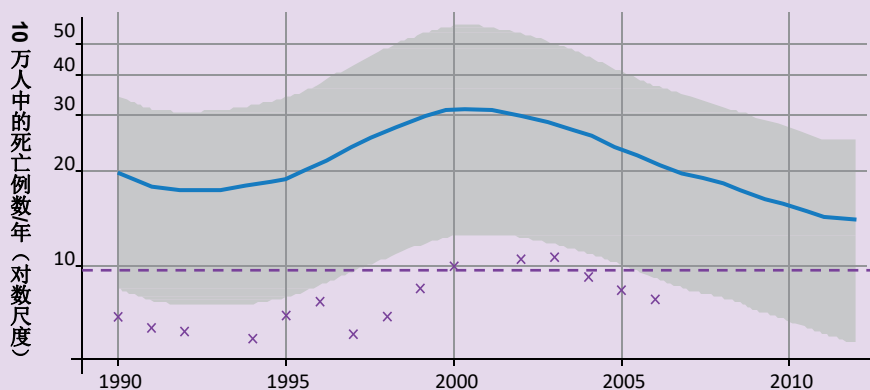
- (1) 国内估计总死亡例数
- (2) 记录了死因的死亡例数
- (3) 死因不明的死亡例数
- (4) 覆盖面比例 (c) : (2) / (1)
- (5) 死因不明的死亡占记录了死因的死亡的比例: (3) / (2)
- (6) 死因为结核病的死亡例数 (d)
- (7) 经过调整的 (针对覆盖面和不明原因) 结核病死亡例数 (da)
- (8) 每 10 万人口中的“原始”结核病死亡率: (6) / 国家人口估计数量
- (9) 每 10 万人口中的“经过调整的”结核病死亡率: (7) / 国家人口估计数量

评估人员绘制了 HIV 阴性患者的结核病死亡率时间趋势图（图 3）。x 符号代表根据 CRVS 记录测算的原始结核病死亡率（表 2 的（8）列中的数据）。原始数字被调整，以便体现不充分的 CRVS 覆盖面以及不明确的 ICD-10 死因代码，用蓝色线表示（表 2 的（9）列中的数据）。根据 2004 至 2006 年的降低

文字框 1 - 续

幅度测算了缺少 CRVS 数据的值（2007–2010 年）。蓝色条带显示了估计的死亡率的不确定度范围，虚线表示到 2015 年结核病死亡率在 1990 年基础上减半的国际目标值。1997 年以后的结核病死亡率增幅可能是由于泰国在实施全民卫生保健项目以前经济危机对该国的影响。

图 3
泰国结核病死亡率（排除 HIV）的时间趋势

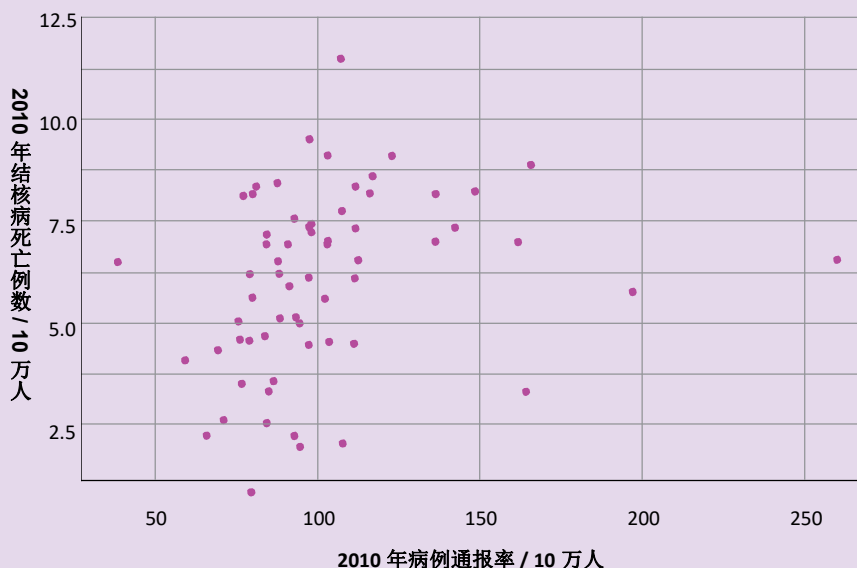


评估人员还考查了地方层面的结核病死亡率水平，目的是更好地了解国内的结核病死亡分布情况以及它们是否与结核病通报病例数量相关。计算了国内各个省份每 10 万人口中经过调整的死亡率，并与每 10 万人口中各类人群、各年龄段的结核病病例通报率进行了比较（图 4）。没有观察到任何关联性，因而难以做出确定性的解释（另请参见第 7.4 节）。

文字框 1 - 续

图 4

泰国不同省份结核病死亡率与病例通报率散点图



CRVS 系统数据的另外一个问题是死亡的错误分类。有明确记录的一个例子是南非和津巴布韦的关于结核病死亡的人口动态登记数据，其中大量艾滋病死亡病例被归类为结核病死亡，目的是避免艾滋病相关污名。最近发布了经过改进的以经验为依据的调整规程（12）；有必要研究确定如何针对错误分类进行具体的补救性调整以便更好地衡量结核病死亡率。

可以对现有 CRVS 数据点之间的缺失数据进行插值计算，同时可以利用指数平滑法模型测算某个时间序列的后续缺失值（13）。在缺少 CRVS 数据的情况下，还可以通过反复开展死因推断调查来估算结核病死亡率的时间序列（文字框 2）。

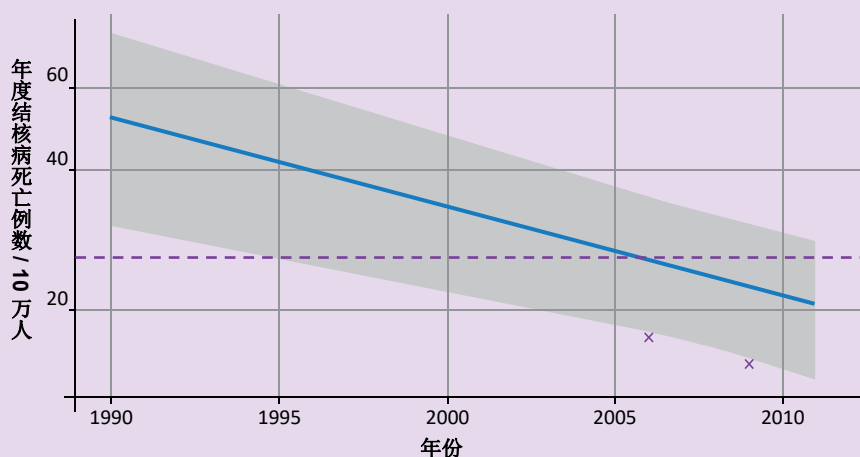
7.3 监测艾滋病病毒感染者的结核病死亡率

HIV 被记录为结核病患者的潜在死因，而结核病则被记录为促成因素。由于三分之一的有人口动态登记系统的国家都只向世界卫生组织报告潜在死因而不报告促成因素，人口动态登记数据通常不能用于估算 HIV 阳性人群中的结核病死亡例数。

文字框 2 个案研究：越南

利用从 2006（14）和 2009 年（15）分别开展的全国代表性死因推断调查得出的数据，越南确定了国内结核病死亡率水平（排除艾滋病相关结核病死亡）和时间趋势。从调查得出的死因数据被调整，以便体现不明死因和不充分的覆盖面。既往结核病死亡率基于经过调整的 2006 至 2009 年趋势得以估测，假定每年的结核病发病率指数递减幅度是相同的。图 5 显示了结核病死亡率的趋势（排除 HIV）。1990 至 2005 年缺少死亡率的直接估测，这影响了死亡率时间趋势估测的准确性。根据“遏止结核病合作伙伴关系”设立的国际结核病死亡率目标（即：到 2015 年结核病死亡率在 1990 年的基础上减半），1990 年被当作对照年份。

图 5
结核病死亡率（排除艾滋病相关结核病死亡），对数尺度



蓝色条带显示了不确定度范围，虚线表示到 2015 年结核病死亡率在 1990 年基础上减半的国际目标值。x 符号代表根据死因推断研究得出的原始数据。原始数据被调整，以便体现不明原因的死亡和不充分的 CRVS 覆盖面（蓝色线）。水平虚线（紫色）表示目标值：到 2015 年结核病死亡率在 1990 年的基础上减半。

而后，必须间接估测 HIV 阳性人群的结核病死亡率，例如采用 Spectrum 软件的方法。^a 结核病死亡率是 HIV 阳性人群结核病发病率和病死率的比值：

$$M = (I - N)F_u + NF_n$$

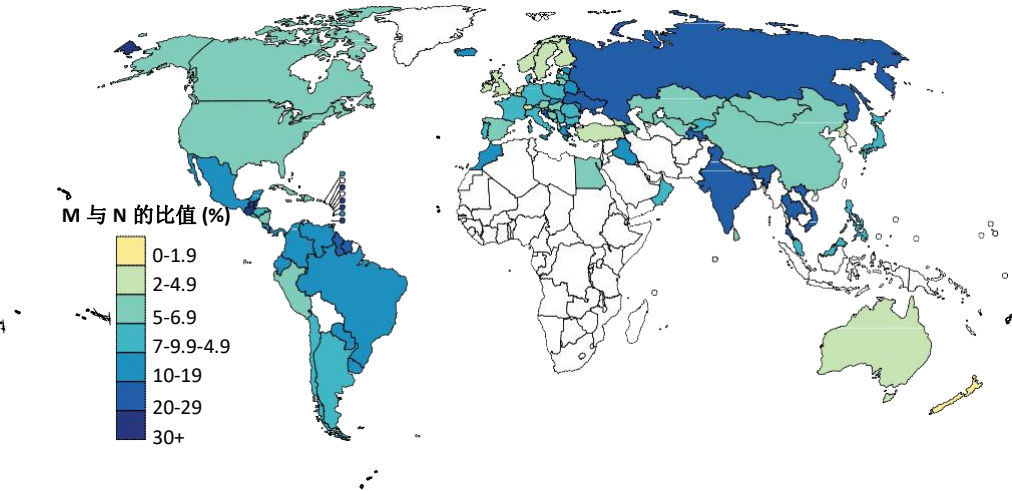
a. <http://www.futuresinstitute.org/spectrum.aspx>

其中， I 代表 HIV 感染者当中的结核病病例， N 代表通报的 HIV 阳性病例， $(I-N)$ 代表未通报的 HIV 阳性结核病病例， M 代表 HIV 阳性人群的结核病死亡率。 F_n 和 F_u 分别是通报的结核病病例和未通报的结核病病例的病死率。

7.4 死亡数与通报数的比值

根据 CRVS 系统得出的结核病死亡率（HIV 阴性人群中）与结核病病例通报数的比值可以成为有用的指标，用于帮助确定哪些人群有较高的死亡率，从而开展有针对性的行动。图 6 展示了有高质量死因数据的国家的该指标的值。东欧的结核病死亡率较高，可能与高耐药结核病水平相关，而在东南亚地区，较高的结核病死亡率可能是由于大量结核病病例没有得到治疗。

图 6
各国结核病死亡数与通报数的比值（2012 年）



还可以根据亚人群及其相关风险因素来分析结核病死亡率，例如种族、民族、社会经济状况、年龄、性别和地理位置（16）。因此，将死亡率与通报数据联系起来有助于确定国内面临较大死亡风险的亚人群。

7.5 耐多药结核病死亡率

目前，ICD 中没有区分因药敏和耐药结核病导致的死亡。用户需要自行界定类别代码（例如南非将耐多药结核病和广泛耐药结核病导致的死亡分别编码为 U51 和 U52，如表 3

所示)。不过,即便在南非,所有肺结核病例的药敏试验覆盖面也很低,这可能导致耐药结核病死亡被错误地归类。因此,为了直接根据 CRVS 系统估测耐药结核病死亡率(尽管这样已经是可行的),有必要进一步做出投入,以便产出高质量数据。

表 3
南非用户界定的耐药结核病死亡原因代码 (17)

	2006	2007	2008	2009	2010
耐多药结核病 U51	604	597	712	1184	856
广泛耐药结核病 U52	3	84	135	151	171
肺结核病例的药敏试验覆盖面	无	无	27%	10%	12%

通过间接估测耐多药结核病死亡率,根据结核病死亡总例数、耐多药结核病病例所占的比例以及相对死亡风险(RR)(耐多药结核病患者与非耐多药结核病的患者之比),可以推导出耐多药结核病导致的死亡例数。虽然根据人口动态登记数据(即结核病总死亡例数)和有代表性的监测或调查数据,可以估测出多年以来的结核病总死亡率和耐多药结核病现患率,但是相对风险则很难估测。相关人员对已发表的文献进行了系统性回顾和荟萃分析(18),以便估测耐多药结核病病例与非耐多药结核病病例的相对死亡风险。已开展的 25 项研究涉及有关接受耐多药结核病和结核病(非耐多药结核病)治疗的患者在治疗期间和治疗后的死亡率数据。在此基础上,可以计算死于耐多药结核病的相对风险的全球估计值。这些研究覆盖很广的地理范围,包括较高和较低耐多药结核病和艾滋病负担的国家,但是却仍不足以估测地区性或其他亚人群面临的相对风险。

全球耐多药结核病死亡例数根据下列公式计算:

$$m = M * p * r$$

m = 全球耐多药结核病死亡例数;

M = 全球结核病死亡例数;

p = 结核病病例中耐多药结核病的比例(根据耐多药结核病新病例和复治病例的加权平均值计算) = 0.06, 范围(0.04-0.07);

r = 死于耐多药结核病与死于非耐多药结核病的相对风险 = 2.36, 范围(1.67-3.05)。

对于 2012 年, $M=1\,300\,000$, $p=0.057$ (19)。计算得出 170 000, 范围(10 万至 24 万)。

1. International Classification of Diseases (ICD) Information Sheet. Geneva: World Health Organization; 2014 (update). 网址: <http://www.who.int/hiv/pub/guidelines/arv2013/en/>
2. Strengthening CRVS Systems through innovative approaches in the health sector. Geneva: World Health Organization; 2013. 网址: http://www.who.int/tb/publications/2012/tb_hiv_policy_9789241503006/en/
3. Civil registration and vital statistics 2013: challenges, best practice and design principles for modern systems. Geneva: World Health Organization; 2013. 网址: http://www.who.int/healthinfo/civil_registration/crvs_report_2013.pdf
4. Implementation of ICD. Geneva: World Health Organization; 2010. 网址: <http://www.who.int/classifications/icd/implementation/en/index.html>
5. Setel et al. Who Counts? A Scandal of Invisibility: making everyone count by counting everyone. The Lancet. 2007;370(9598):1569–1577.
6. Mathers CD, Fat DM, Inoue M, Rao C, Lopez AD. Counting the dead and what they died from: an assessment of the global status of cause of death data. Bull World Health Organ. 2005;83(3):171–7
7. Global Summit on Civil Registration and Vital Statistics (2013) (Website). 网址: <http://www.who.int/tb/publications/2010/9789241547833/en/>
8. Verbal autopsy standards: ascertaining and attributing causes of death. Geneva: World Health Organization; 2012. 网址: <http://www.who.int/tb/publications/definitions/en/>
9. www.who.int/healthinfo/global_burden_disease/GlobalCOD_method.pdf
10. Global Health Observatory: Mortality and global health estimates (Website). Geneva: World Health Organization; 2014. 网址: http://www.who.int/gho/mortality_burden_disease/en/index.html
11. <http://esa.un.org/wpp/sources/country.aspx>
12. Birnbaum JK, Murray CJL, Lozano R. Exposing misclassified HIV/AIDS deaths in South Africa. Bulletin of the World Health Organization. 2011;89:278–285.
13. Hyndman R et al. Forecasting with exponential smoothing: the state space approach. Berlin: Springer-Verlag; 2008.
14. Ngo AD et al. Mortality patterns in Vietnam, 2006: Findings from a national verbal autopsy survey. Biomed Central Research Notes. 2010;3:78.
15. Nguyen PH et al. Mortality measures from sample-based surveillance: evidence of the epidemiological transition in Viet Nam. Bulletin of the World Health Organization. 2012;90:764.
16. Lonnroth K et al. Drivers of tuberculosis epidemics: The role of risk factors and social determinants. Social Science and Medicine, 2009, 68:2240–2246.
17. Mortality and causes of death in South Africa, 2010: Findings from death notification. Pretoria: Statistics South Africa; 2010. 网址: <http://www.statssa.gov.za/publications/p03093/p030932010.pdf>
18. Nair H, Brondi L, Campbell H. A meta-analysis to estimate the risk of dying from MDR-TB compared to non MDR-TB in the world (publication in preparation).
19. Global tuberculosis report 2013. Geneva: World Health Organization; 2013. 网址: http://apps.who.int/iris/bitstream/10665/91355/1/9789241564656_eng.pdf

第 8 章

整合监测数据和调查数据来估测结核病负担

目标受众：

1. 流行病学家、统计学家和国家结核病防治规划督导与评估工作人员。
2. 希望了解如何估测结核病负担以及其趋势分析对于项目的指导意义的国家结核病防治规划工作人员。

预期成果：

本章结束时，读者应当能够：

- 了解如何采用不同来源的数据来解释观察到的结核病病例通报数和结核病死亡率趋势。
- 解释不同人群之间的差异性。
- 了解如何根据可用的数据来源得出结核病发病率、现患率和死亡率的估计值。
- 了解在间接估测某人群的结核病负担时要注意的主要问题。

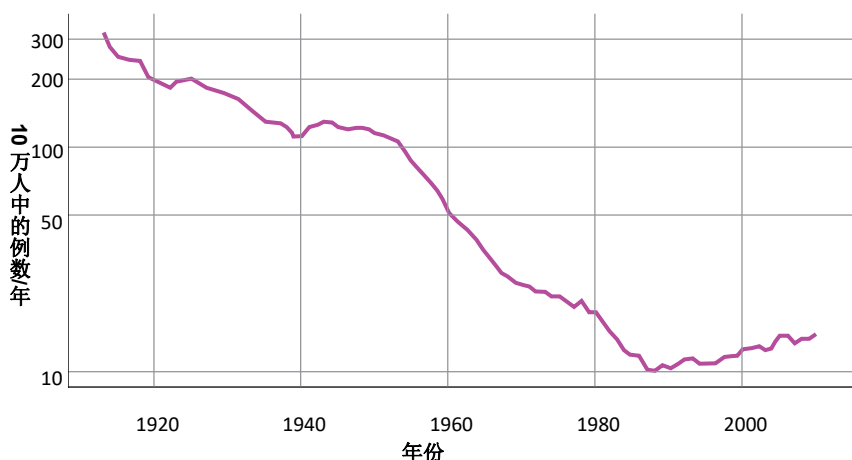
作者：

Philippe Glaziou, Charalambos Sismanidis, Irwin Law, Dennis Falzon, Ikushi Onozaki

8.1 结核病发病率

结核病发病率是“千年发展目标”（MDG）目标 6C 的核心：到 2015 年减半并开始逆转疟疾和其他重大疾病的发病率（1）。对于艾滋病疫情严重并且在二十世纪九十年代结核病发病率快速上升的国家，逆转结核病发病率的目标具有很大的相关性。不过，在世界上的绝大多数地方，结核病发病率自二十世纪九十年代以来都在缓慢下降（2）。英格兰和威尔士的病例通报率（被视为结核病发病率的替代指标）早在广泛开展有效的结核病干预之前就已经开始下降，这表明结核病发病率与多种因素相关，包括社会经济境况和生活条件的改善（图 1，另请参见第 1 章）。

图 1
英格兰和威尔士结核病病例通报率（对数尺度）的趋势（1913-2011 年）
资料来源：英格兰公共卫生署（结核病防治部）



在大多数国家，结核病发病率难以测量或估计。旨在测量结核病发病率的调查要求开展大规模人群队列研究。这种调查耗费资源，并且实施难度大，在大部分地区都难以落实，而发病率调查需要跟踪考查规模很大的队列（例如超过 50 000 人），时间跨度达一年以上。在调查期间，需要密切跟踪考查队列，以便减少失访和确保结果有效性。因此，迄今为止，尚无国家开展全国结核病发病率调查。

实际上，发病率通常是根据常规病例通报数推导出。不过，这类数据有两个方面的不确定性：1）并非所有被诊断的病例都得到通报；2）并非所有结核病病例都被发现。在有着完善的结核病监测和卫生系统的高收入国家，病例通报系统被要求收集所有（或者几

乎所有）新发结核病病例的信息。然而，在其他地区，常规病例通报数不能提供准确的发病率估计值。这有几个原因。首先，私营部门并不总是报告新发现的病例。例如，最近在印度开展的一项研究发现大约 46% 的新发现的病例没有得到报告（3）。其次，并非所有病例都会被发现，特别是在缺乏医疗保险和社会保障（或类似机制）的国家，以及在卫生服务和实验室的覆盖面和绩效薄弱的地方。在高结核病负担国家开展的全国人群现患率调查的结果一致表明大多数病例在调查前都没有得到结核病诊断。没有求医的常见原因包括：没有症状；认为症状不严重，无需求医；难以获得卫生服务。第三，中低收入国家仅有大约一半的结核病报告病例获得微生物学确诊，因为许多国家的诊断局限于痰涂片镜检，其灵敏度有限。对于临床诊断病例（没有获得细菌学确诊的结核病病例）而言，诊断技术的质量和准确性难以量化，在不同地区有差别。并非所有报告的临床确诊病例都是真正的结核病病例。通过引入灵敏度和特异性更高的诊断技术，应当能够提高所有结核病报告病例中细菌学确诊病例的比例，并且降低结核病误诊率。

无法从在校儿童的结核菌素调查结果可靠地获得痰涂阳肺结核的发病率。虽然既往采用了这种方法，近期该方法受到质疑，其所基于的前提假定已经不再获得认同（4）。

测量中低收入国家的结核病发病率是一项挑战。未来的主要解决方案是将结核病监测范围扩展到所有卫生保健服务机构并且减少漏报（即没有报告给公共卫生当局的病例在发现的病例总数当中所占的比例）和过诊（例如假阳性结核病病例）现象。需要采取行动减少严重的漏报情形，以便确保常规病例通报所提供的信息能真实地体现疾病负担和趋势。

病例存量研究

为了更好地评估和解决结核病漏报问题，世界卫生组织“全球结核病影响衡量”任务组于 2011 至 2012 年重点开发病例存量研究设计、实施和分析指南，用于指导结核病

漏报问题评估（5）。^a 病例存量研究正在得到越来越多的应用，有助于规划和实施医防合作活动。病例存量研究旨在确定所有或部分公立和私营卫生机构的符合标准病例定义的结核病病例数量，并且将这些数据与报告给地方和国家当局的结核病病例记录进行比较。比较的过程被称为“记录关联”，在过程中会确定重复的和唯一的记录。可以利用现有数据库进行记录关联，或者需要做出特别的安排才能进行关联（涵盖有限的时间段），具体取决于现有的数据管理系统，以便收集国内所有卫生服务机构诊断或在特定地区随机抽选的卫生服务机构所登记的病例相关数据。

捕获-再捕获分析

在某些情况下，病例存量研究的结果可以与“捕获-再捕获”分析模型进行组合，以便估算结核病发病率。

伊拉克的一项病例存量研究显示结核病漏报占发现的病例总数的 16%（图 2）（6）。国家结核病防治规划未涵盖的私营服务提供方和公立服务提供方诊断了这些病例。根据捕获-再捕获模型，估计额外有 473 个病例（95% 置信区间：394–565）未被任何卫生服务机构（共三类）发现。因此，利用捕获-再捕获分析可以更新国家结核病发病率估计值。漏报率低是完善的结核病监测系统的重要特征。

有必要指出，捕获-再捕获方法有天然的局限性，在许多国家并不可靠，这是由于其关键的基本模型前提假定并不一定符合现实情况（5）。

根据现患率估测发病率

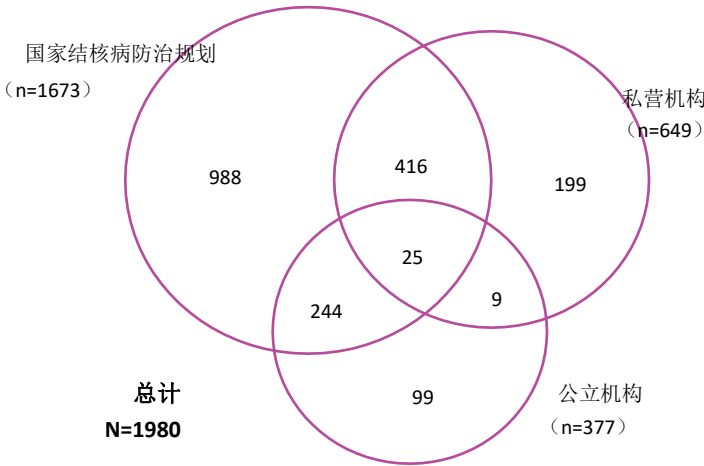
理论上有可能根据结核病现患率数据估算结核病发病率。结核病发病率大致是结核病现患率与平均疾病持续时间的比值。根据现患率估算发病率时有以下难点：

- 根据人群调查估算的结核病现患率通常而言不够精确（标准差通常为最佳估计值的 20–25%）。
- 关于疾病持续时间的数据有限。疾病持续时间受多种因素的影响，包括病例发现工作的绩效、HIV 现患率和结核病患者的耐多药率。尽管关于症状持续时间的数

a. 病例存量研究指南旨在描述和解释如何设计、实施和分析病例存量研究，以便评估结核病漏报水平。该指南还解释了如何运用捕获-再捕获方法来估算结核病发病率，同时强调了采用该方法所必须满足的条件。

据可以通过人群现患率调查来收集，但是有三项考虑因素制约了其在估算全国平均疾病持续时间方面的效用。首先，疾病的自然进程会因调查而缩短（现患率调查期间会进行主动病例发现），而没有参加调查的患者的疾病平均持续时间会较长。其次，现患率调查通常会发现较大比例的无症状或症状很少的细菌学确诊的肺结核病例。第三，结核病的一些最常发生的症状（例如慢性咳嗽）在一般人群（特别是吸烟人群）中也常见，并不是结核病特异性症状。

图 2
病例存量研究显示伊拉克的结核病患者在哪里得到诊断（2011 年）

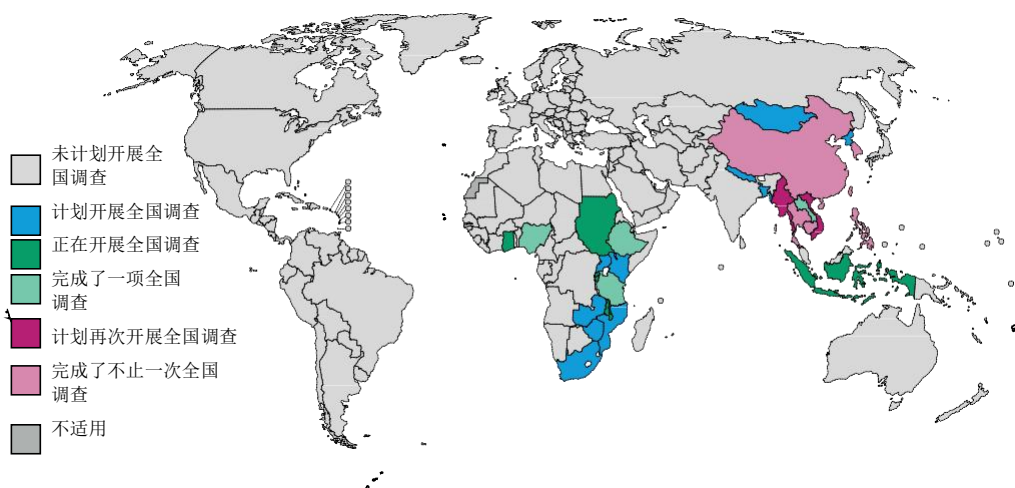


8.2 结核病现患率

结核病现患率是结核病负担的十分有用的指标，它指明人群中可能存在什么水平的传播。可以通过人群调查间接衡量结核病现患率（7）。在亚洲和非洲的许多国家，近年来在利用世界卫生组织推荐的方法规划和实施现患率调查方面有了显著进展（图 3）。现患率调查已被证明对于设计和调整国家结核病防治规划策略极为有用。

图 3

1990 年以来或不久的将来采用当前推荐的筛查和诊断方法开展全国结核病现患率人群调查的国家（2013 年 7 月）（2）

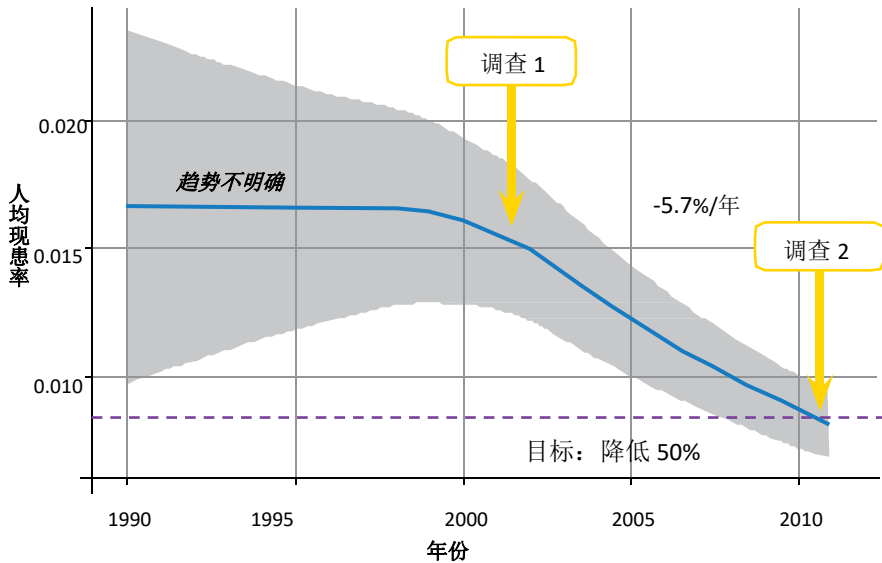


在过去的二十年间，柬埔寨的结核病现患率（占总人口的百分比）一直处于全球最高水平。在经历了数十年的内战和经济困境后，柬埔寨于 1994 年恢复了结核病防治工作。结核病服务最初仅在省级和区县医院开展。在 2002 年 DOTS 扩展到卫生所的初期，国家结核病防治规划决定通过开展全国现患率调查直接评估结核病负担。各类结核病的现患率估计为人群的 1.5%（范围：1.2%至 1.8%），这是近代史上记录的最高的现患率之一。2011 年开展了第二次全国调查。各类结核病的现患率估计为人群的 0.8%（范围：0.69%至 0.96%），这与第一次调查的结果有显著统计学差异（95%置信区间没有重叠）。

第二次调查提供的充分证据表明自 2002 年 DOTS 扩大工作之后，柬埔寨的结核病负担大幅下降。结果表明在短短九年内细菌学确诊的病例的现患率下降了 45%（图 4）。

随着结核病负担的减轻和监测系统的加强，现患率调查将不再是评估结核病负担的高效手段，将来很可能会停止采用该方法。例如，韩国从 1965 年到 1995 年每五年开展一次调查，共开展了七次。而后，该调查不再继续，这是由于在低结核病现患率的情况下所需的样本量过大，而且易于获取卫生保健服务的高收入人群逐渐不愿接受调查（8）。在 2010 年完成第五次调查后，中国可能也会基于相同的原因终止此类调查。在缺少人群调查数据的情况下，难以准确估测结核病现患率（2）。

图 4
柬埔寨人均结核病现患率的趋势



虚线表示“到 2015 年结核病死亡率在 1990 年基础上减半”这一国际目标值。

8.3 结核病死亡病例和病死率

第 7 章描述了结核病死亡率的数据来源。结核病死亡例数与估计的结核病发病例数的比值被称为结核病病死率（CFR）。计算结核病病死率时可以排除 HIV 阳性结核病死亡病例，也可以包括 HIV 阳性结核病死亡病例。

20 世纪英格兰和威尔士的数据可以用于考察不同时期的上述比值（图 5）。虚线表示结核病死亡率，它与表示病例通报率的实线基本平行，直到大约 1945 年，这表明死亡率与病例率的逐年下降幅度大致相同。到了第二次世界大战，大约 40%至 50%的结核病患者死于结核病。不过，之后结核病死亡率的平均下降速度加快。随着有效的化学疗法的采用，结核病病死率大幅下降。在二十世纪九十年代和二十一世纪的头十年，尽管结核病通报率有了提高，但是结核病病死率进一步缓慢下降（图 6），表明如今占大部分病例的移民所面临的死亡风险有所减小。

图 5
英格兰和威尔士 1913 至 2010 年的结核病病例通报率（实线）与结核病死亡率（虚线）（对数尺度）

资料来源：英格兰公共卫生署（结核病防治部）

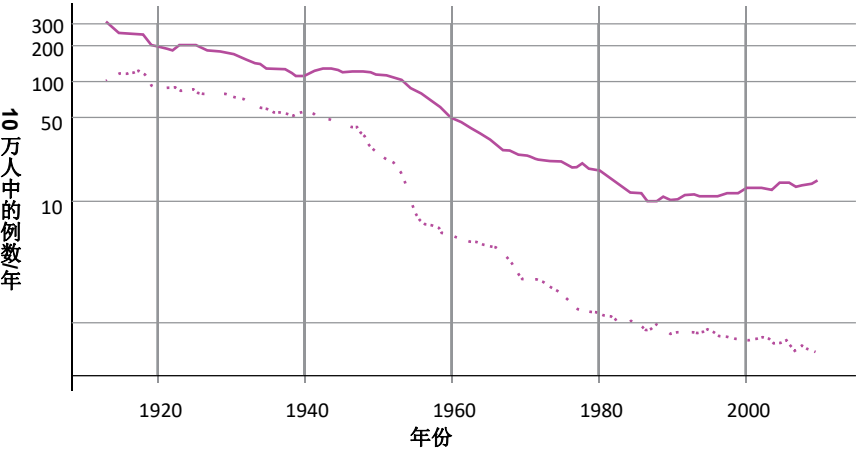
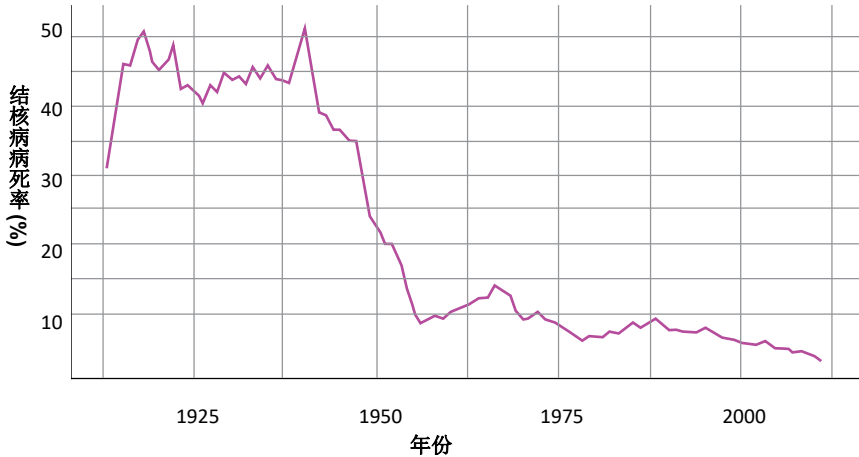


图 6
英格兰和威尔士 1913 至 2010 年的结核病病死率（结核病死亡例数与结核病病例的比值）

资料来源：英格兰公共卫生署（结核病防治部）



病死率是结核病发现工作和结核病防治规划的绩效的理想指标。利用国家人口动态登记系统中的结核病死亡率数据，可以分析地方层面（例如省级）的病死率（排除 HIV 阳性结核病死亡），从而可以确定哪些地区的项目绩效不佳。此外，可以按年龄和性别比较病死率，从而提高干预工作的针对性。

在全球范围内，病死率（包括 HIV 阳性患者的估计结核病死亡例）从 2000 年的 20%（范围：18–23%）降低到 2010 年的 16%（14–18%），这解释了为什么死亡率每年平均降低 3.5%，比同期的发病率的降低幅度（每年 1%）要大。随着覆盖面的扩大，预期抗逆转录病毒治疗将在预防结核病死亡方面发挥越来越大的作用（2010 年，全球大约 16% 的结核病通报病例为 HIV 阳性并且接受抗逆转录病毒治疗）。抗病毒治疗和异烟肼预防也将越来越有助于 HIV 感染和潜伏结核病感染个体预防结核病。

病死率还受到耐多药结核现患率和耐多药结核病项目管理绩效的影响。如果耐多药结核病项目管理薄弱，则观察到的病死率会比较高。

1. MDG 6: combat HIV/AIDS, malaria and other diseases. Geneva: World Health Organization; 2014. Available at: http://www.who.int/topics/millennium_development_goals/diseases/en/
2. Global tuberculosis report 2013. Geneva: World Health Organization; 2013. 网址: http://www.who.int/tb/publications/global_report/en/index.html
3. Satyanarayana S et al. From where are tuberculosis patients accessing treatment in India? Results from a cross-sectional community based survey of 30 districts. PLoS One. 2011;6(9):e24160.
4. van Leth F, Van der Werf MJ, Borgdorff MW. Prevalence of tuberculosis infection and incidence of tuberculosis: a re-assessment of the Styblo rule. Bull World Health Organ 2008;86:20–6.
5. Assessing TB under-reporting through inventory studies. Geneva: World Health Organization; 2012. Available at: http://www.who.int/tb/publications/inventory_studies/en/index.html
6. Huseynova S et al. Estimating tuberculosis burden and reporting in resource-limited countries: a capture-recapture study in Iraq. Int J Tuberc Lung Dis. 2013;17(4):462–7.
7. Tuberculosis prevalence surveys: a handbook. Geneva: World Health Organization; 2011. 网址: http://www.who.int/tb/advisory_bodies/impact_measurement_taskforce/resources_documents/thelime-book/en/index.html
8. Hong YP et al. The seventh nationwide tuberculosis prevalence survey in Korea, 1995. Int J Tuberc Lung Dis. 1998;2(1):27–36.

结语

第 1、2 和 3 章给出的分析解释了根据信息系统（产生聚合数据和个体数据）得出的病例通报率的趋势。第 4 章以及第 5 章的部分内容探讨了结核病的主要决定因素，揭示了病例随时间变化的趋势，以及结核病病例通报的趋势的影响因素。进而，根据病例通报数的趋势、结核病疫情的主要驱动因素（例如艾滋病，参见第 6 章）的变化、死亡率趋势（第 7 章）和结核病现患率，可以推测结核病发病率的趋势。

世界卫生组织“结核病影响衡量”任务组制定了一套标准和基准指标，用于评估结核病监测系统的绩效、确定数据缺口和数据质量问题，以及描述未满足的督导与评估需求。^a 对于缺少发病率数据、现患率或死亡率的国家而言，仍有可能估算结核病负担指标（参见世界卫生组织 2013 年全球结核病报告的附录 2）^b，但是估算结果会有很大的不确定性。

理想的情形是：根据结核病病例通报数据得出的数据可以提供可靠的结核病发病率替代指标，并且国家人口动态登记系统记录的死因信息能够用于准确地测算结核病死亡率（包括结核病促成的艾滋病死亡）。

-
- a. TB surveillance expert group. Geneva: World Health Organization; 2011. 网址: http://www.who.int/tb/advisory_bodies/impact_measurement_taskforce/meetings/surveillance_expert_group_meeting_sept2011/en/index.html
 - b. Global tuberculosis report 2013. Geneva: World Health Organization; 2013. 网址: http://www.who.int/tb/publications/global_report/en/index.html

人口动态登记系统和结核病通报系统收集的监测数据提供了有关国家和地方层面结核病流行情况和疾病防治项目的基本信息。分析这些数据有助于项目管理人员和其他人员跟踪考查结核病负担的水平和趋势，发现疫情暴发，并且确定改善现有结核病预防、诊断和治疗服务的途径。本手册提供了有关分析和使用这类监测数据的实用指南，适合从事结核病防治事业的各类人员参考使用。本手册是世界卫生组织“全球结核病影响衡量”任务组的合作项目工作成果之一。



**World Health
Organization**

ISBN 978 92 4 154878 6



9 789241 548786