

第一章

全球疾病负担研究（GBD）概述

摘要

本章系统概述了全球疾病负担（Global Burden of Disease, GBD）研究的发展背景与核心内容。

第一节介绍了 GBD 研究的总体目标、演进历程、核心指标体系构成及主要研究领域。

第二节总结了 GBD 研究的八项主要特征：①强调研究结果的全球性与区域间可比性；②作为健康政策制定与评估的重要工具；③构建以多重 Meta 分析为基础的整合研究平台，并发展相应的机制提升可信度；④建立统一的涵盖死因、疾病、伤害、后遗症及危险因素的多层级分类体系；⑤利用 DisMod-MR 模型，并结合小域估计策略，实现先验信息在地理层级间的传递；⑥数据覆盖面不断拓宽，信息来源持续增长；⑦建立系统化的数据质量评估与处理方法；⑧构建特色鲜明的全球健康数据交换平台，以提升数据共享与应用效率。

第三节简要介绍了 GBD 研究中所采用的核心统计方法，包括人口统计学中的间接估计技术、死因建模工具 CODEm，描述性流行病学的综合元回归框架，贝叶斯建模、高斯过程回归，多层级与小域估计模型，以及“举证责任研究（burden of proof studies）”等。这些方法共同构成了一个以多重 Meta 分析为基础的集成流行病学建模体系，用于系统估计全球疾病与风险因素负担。

第四节简要介绍了支持 GBD 研究的数据整合与共享平台，为数据获取、标准化处理和再利用提供了关键支撑。总体而言，本章的内容为后续各章对 GBD 研究方法的深入探讨奠定了理论与方法论基础。

第一节 GBD 研究的发展历程

20 世纪 80 ~ 90 年代, 随着全球生育率持续下降和儿童死亡率显著降低, 国际社会逐渐认识到, 人口老龄化不仅深刻影响工业化国家, 也对发展中国家产生了深远的影响。与此同时, 环境因素和个体行为的变化进一步加剧了慢性非传染性疾病 (NCD) 和伤害的风险。流行病学研究表明, 吸烟、高血压、高脂饮食、机动车辆使用的普及、职业暴露和贫穷等多种因素, 构成了全球范围内非传染性疾病与伤害的主要风险因素集群。

在此背景下, 发展中国家流行病学转型的政策与规划成为研究关注的焦点^①。著名学者 Omran 提出了“流行病学转型”的理论, 认为各类人群在死亡率下降的过程中, 其疾病谱与致病因素会发生系统性变化: 高死亡率人群以传染病、营养缺乏和生殖健康问题为主, 而低死亡率人群则面临更多慢性和退行性疾病的负担^{②③}。尽管人类社会刚刚经历了 COVID-19 大流行, 但从长期趋势来看, 该理论框架在阐释人群健康转变规律方面仍具有较强的适用性。

在疾病模式转变的背景下, 如何改善人群健康成为国际社会关注的核心问题。世界银行《1993 年世界发展报告: 投资于健康》^④对这一问题作出了系统性回应。由 Murray CJL 和 Lopez AD 牵头完成的 1990 年全球疾病负担研究 (GBD 1990), 首次系统量化并报告了全球 8 个地区 106 种疾病和 10 种风险因素所致的健康损失, 为《世界发展报告》的政策建议提供了关键的实证证据。这标志着 GBD 研究首次成为全球公共卫生研究与政策讨论的核心议题。

一、GBD 研究的目标

GBD 研究的初衷是构建一套综合性健康指标体系——伤残调整损寿年 (Disability-Adjusted Life Year, DALY) 和健康调整预期寿命 (Health-Adjusted Life Expectancy, HALE), 用于全面评估与预测全球各区域、各国家、不同年龄和性别人群的疾病与伤害负担。1990 年, GBD 研究团队首次提出 DALY 指标, 建立了一种能够同时衡量早逝与非致命健康损失的测量框架, 用以估算不同区域、国家、不同性别年龄人群在

① National Research Council (US) Committee on Population. The Epidemiological Transition: Policy and Planning Implications for Developing Countries: Workshop Proceedings [M]. Washington (DC): National Academies Press (US), 1993. PMID: 25144040.

② Omran, A R. The epidemiological transition: A theory of the epidemiology of population change [J]. Milbank Memorial Fund Quarterly, 1971,49: 509-538.

③ Omran, A R. Epidemiologic transition [M] // Ross JA (ed.). International Encyclopedia of Population. London: Free Press, 1982, 172-183.

④ World Bank. World Development Report 1993 [M]. New York: Oxford University Press, 1993.

全生命周期中的健康损失^①。

自 GBD 2010 起，研究目标进一步明确为：通过系统和科学的方法，按年龄、性别、地区和国家量化疾病、伤害及风险因素所致的健康损失^②。

二、GBD 发展历程：研究产出与代表性发表物

为系统呈现 GBD 研究的演进脉络与关键成果，以下按时间顺序梳理了其发展阶段、研究重点与代表性出版物，并据此其划分为以下三个阶段。

1. GBD 1990：首次全球疾病负担研究

GBD 研究成果首次公开发表于世界银行《1993 年发展报告：投资于健康》。该研究评估了全球 8 大地区人群的疾病负担及主要危险因素，系统比较 130 种疾病和 10 种危险因素在死亡率、发病率和患病率上的影响，并提出相应的干预优先顺序建议。这是首次在全球范围内系统应用一致性估算方法对健康负担进行量化，结果广泛用于卫生政策制定与资源配置。除 1994 年发表于 *WHO Bulletin* 的四篇方法与结果论文外，研究团队还于 1996 年出版了系统总结成果的书籍及多篇研究文章（详见附录 1-1）。GBD 1990 的关键发现显著提升了国际社会对精神疾病、道路交通伤害、结核病和疟疾等长期被忽视健康问题的关注。

2. GBD 2000—2010：方法拓展与范围扩大

2002 年《世界卫生报告》以“降低风险，促进健康生活”为主题，首次系统报告了 WHO 成员国在 2001 年不同性别人群的死因、DALY、HALE 及归因疾病负担，标志着 GBD 研究进入了第二阶段（GBD 2000）。此阶段研究把全球进一步分为 14 个地理区域，死因类别扩展至 159 类，危险因素增至 26 项。研究结果部分发表于 1999—2002 年的《世界卫生报告》，并由 WHO 发布《GBD 2004 报告》^③及其危险因素估计结果^④（详见本章附录 1-2）。

在方法学方面，本阶段还对若干关键问题进行了探索与改进，包括在数据缺失条

① Murray CJL, Lopez AD, Jamison DT. The global burden of disease in 1990: summary results, sensitivity analysis and future directions [J]. *Bull World Health Organ*, 1994,72 (3): 495-509.

② Murray CJL, Ezzati M, Flaxman AD, et al. The Global Burden of Disease Study 2010: design, definitions, and metrics [J]. *Lancet*, 2012, 380: 2063-2066.

③ World Health Organization. The Global Burden of Disease: 2004 update [R]. Geneva: World Health Organization, 2008.

④ World Health Organization. Global Health Risks: Mortality and burden of disease attributable to selected major risks [R]. Geneva: World Health Organization, 2009.

件下的死亡率估计^①，死因建模策略的优化^{②③}，以及自我报告的危险因素调整方法^{④⑤}等。同期的重要方法学文献详见本章附录 1-2。

3. GBD 2010 及之后：建立持续迭代的全球健康评估体系

由 Murray CJL 和 Lopez AD 开创的 GBD 研究，在世界银行和世界卫生组织的支持下逐步发展壮大。2007 年，在比尔及梅琳达·盖茨基金会等机构的支持下，华盛顿大学成立健康测量与评估研究所（Institute for Health Metrics and Evaluation, IHME），成为 GBD 研究的主要牵头机构。在 Murray CJL 的领导下，GBD 团队联合全球研究者，持续推进方法创新与研究范围拓展，整合流行病学、统计学及大数据建模策略，逐步建立起一套系统化、可持续迭代的全球健康负担评估体系^⑥。自 GBD 2010 以来，研究框架总体保持稳定，但覆盖范围显著扩展，至 GBD 2019 已历经六轮迭代，评估范围经扩展至 21 个地区、204 个国家，涵盖 369 种疾病与伤害，20 个不同的年龄组，以及 87 种危险因素。

GBD 各轮主要研究结果发布于《柳叶刀》（*The Lancet*）等国际权威期刊上，方法学细节详见论文附录。其核心研究论文大体分为五类：①总死亡率的估计；②死因别死亡数、标化死因别死亡率和早逝所致寿命损失年（Year Life Lost, YLL）的估计；③疾病与伤害流行的伤残所致寿命损失年（Years Lived with Disability, YLD）的估计；④估计死亡和疾病的综合负担的评估，包括 DALY 和 HALE；⑤对危险因素和归因疾病负担的估计。除核心论文外，GBD 研究团队还发布了针对特定疾病（肿瘤、心血管疾病、糖尿病等多种疾病）及特定危险因素，如环境风险因素（如空气污染、吸烟与二手烟暴露、饮食风险等）的归因疾病负担报告。关于 GBD 2010 后的具体方法学更新与估计策略，后续各章将作系统阐述，本节仅列出选取与 GBD 方法学相关的代表性论文列于框图 1-1。

后续章节将对 GBD 的核心方法学与应用进行深入分析。

① Murray CJL, Ferguson BD, Lopez AD, et al. Modified logit life table system: Principles, empirical validation, and application [J] . *Population Studies*, 2003, 57 (2): 165-182.

② Mathers C, Sadana R, Salomon J, et al. Estimates of DALE for 191 countries: methods and results [EB/OL] . Geneva: World Health Organization, 2000 [2023-12-01] . <https://apps.who.int/iris/handle/10665/66549>.

③ Salomon JA, Murray CJL. The Epidemiologic Transition Revisited: Compositional Models for Causes of Death by Age and Sex [J] . *Population and Development Review*, 2002, 28 (2): 205-228.

④ Salomon JA, Tandon A, Murray CJL. Comparability of self rated health: cross sectional multi-country survey using anchoring vignettes [J] . *BMJ*, 2004,328 (7434): 258.

⑤ Murray C, Evans D. Quantifying Individual Levels of Health: Definitions, Concepts, and Measurement IssuesM] // *Health Systems Performance Assessment: Debates, Methods and Empiricism*. Geneva: World Health Organization, 2003: 301-318.

⑥ Institute for Health Metrics and Evaluation. About IHME [EB/OL] . [2023-12-01] . <https://www.healthdata.org/about>.

框图 1-1 GBD 2010 后的重要文献

1. Murray CJL, Ezzati M, Flaxman AD, et al. GBD 2010: design, definitions, and metrics [J] . Lancet, 2012, 380 (9859) : 2063-2066.
2. Srebotnjak T, Mokdad AH, Murray CJ. A novel framework for validating and applying standardized small area measurement strategies [J] . Popul Heal Metrics, 2010, 8: 26.
3. Murray CJL, Ezzati M, Flaxman AD, et al. Comprehensive Systematic Analysis of Global Epidemiology: Definitions, Methods, Simplification of DALYs, and Comparative Results from the Global Burden of Disease Study [J] . The Lancet, 2012, 380: 1-140.
4. Flaxman AD, Vos T, Murray CJL. An Integrative Meta-regression Framework for Descriptive Epidemiology [M] . Seattle:Institute for Health Metrics and Evaluation, 2015.
5. Institute for Health Metrics and Evaluation (IHME) . Rethinking Development and Health: Findings from the Global Burden of Disease Study [R] . Seattle: IHME, 2016.
6. Zheng P, Afshin A, Biryukov S, et al. The Burden of Proof studies: assessing the evidence of risk [J] . Nat Med, 2022, 28, 2038-2044. <https://www.nature.com/articles/s41591-022-01973-2>.

三、GBD 研究测量人群健康的指标体系

1. 衡量人群健康的综合指标

如前所述，为实现不同疾病与伤害在死亡与患病层面的可比性，GBD 研究发展了一套人群健康综合指标（Summary Measures of Population Health, SMPH），主要包括健康调整预期寿命（HALE）与 DALY。DALY 由早逝所致寿命损失年（YLL）和伤残所致寿命损失年（YLD）组成。自 GBD 2010 起，DALY 被定义为 YLL 和 YLD 的简单相加（框图 1-1 文献 3）。该综合指标体系的构建体现了一系列明确的价值取向与方法选择，主要包括：设定统一的理论最小风险参考寿命表；采用公众对健康状态所致健康损失的主观判断所形成的共同价值作为伤残权重（Disability Weight, DW），并明确取消对时间贴现和年龄权重。这一理论立场为 DALY 的简化计算提供了伦理和方法学依据，但也带来了重要影响：5 岁以下儿童死亡的权重相对上升，而非致命性健康损失在总负担中所占比重相应下降。

需要特别指出的是，DALY 虽名为“伤残调整损寿年”，但并不直接反映人群中实际存在的残疾状态或功能障碍。其核心计算元素“残疾权重”实质上是公众对典型健康状态严重程度的共同价值判断，与个体实际生活质量或医学意义上的残疾评估并

无直接对应关系。

关于 DALY 和 HALE 的构建策略和方法，详见本书第二章。

2. 社会人口指数

为系统评估社会发展水平与人群健康之间的关联，GBD 研究从 2015 年起引入了社会人口指数（Socio-Demographic Index, SDI）^①。该指数借鉴联合国开发计划署（UNDP）于 1990 年提出的人类发展指数（Human Development Index, HDI），后者基于预期寿命、教育水平和人均收入三个维度的几何平均。

SDI 的构成包括以下三个变量：25 岁以下总和生育率（TFU25），反映生育水平；15 岁及以上人群的平均受教育年限（EDU15+），衡量教育水平；人均滞后分配收入（Lagged distributed income, LDI），体现经济发展程度。这三项变量的标准化得分（范围 0 ~ 1）按几何平均计算，得出 SDI 值。根据 SDI 水平，GBD 将全球国家和地区划分为五个发展等级：高 SDI、高 - 中 SDI、中 SDI、低 - 中 SDI 和低 SDI。在疾病负担估计中，GBD 进一步比较各国实际疾病负担与 SDI 水平对应的预期负担，据此判断某国实际健康负担是否高于或低于“社会人口发展应有水平”。

尽管 SDI 难以完全解释所有地区疾病模式的差异，但作为反映宏观社会经济发展程度的简明指标，SDI 与健康结果之间具有显著相关性，可为健康政策的制定提供重要参考，其具体应用详见后续章节。

四、GBD 研究的总体分析策略

1. GBD 研究的 18 个工作模块及其相互关系

框图 1-1 中文献 1 的图 1 系统展示了 GBD 2010 的总体分析框架，涵盖 18 个工作模块及其相互关系。总体而言，GBD 的估算体系可分为五个主要部分。

第一部分：估计不同年龄性别的总死亡率（模块 2）。该部分构成后续死因分布与疾病负担计算的总体约束。

第二部分：估计不同年龄性别的死因别死亡率（模块 3 ~ 5）。在总死亡率的约束下，进一步估计各主要疾病与伤害的死因别死亡率，为计算伤残调整寿命年（YLL）提供基础。

第三部分：估计不同年龄 - 性别人群中各种疾病的发病率、患病率及其病程参数（模块 6 ~ 9 和 11）。这部分是为计算伤残生存年（YLD）的关键输入。

第四部分：综合负担指标计算（模块 10、12、13）。这部分在前述死亡与非死亡负担估计的基础上，计算不同年龄 - 性别人群不同疾病的 YLL、YLD 和 DALY，以及 HALE（模块 10、12 和 13）。

^① Institute for Health Metrics and Evaluation. Socio-demographic Index (SDI) [EB/OL]. [2023-12-01]. <https://www.healthdata.org/taxonomy/glossary/socio-demographic-index-sdi>.

第五部分：危险因素暴露与归因负担估计（模块 14 ~ 18）。这部分估计不同危险因素的人群暴露水平，各风险 - 结果对的相对危险度、归因危险度及其对应的归因疾病负担（模块 14 ~ 18）。

此外，模块 1 负责构建和维护全球协变量数据库，为所有模块的建模和分析提供一致的基础数据与建模支持。上述各模块之间高度依赖并相互关联，形成一个统一建模框架下的迭代估算系统。例如，若总死亡率（模块 2）估计发生变化，则需同步更新死因别死亡率（模块 3 ~ 5），并进一步影响 YLL、YLD、DALY、HALE 及归因疾病负担的估计。因此，GBD 研究强调在统一建模框架下对不同估计模块进行迭代约束与修正，以确保各项估计结果之间的内在一致性。

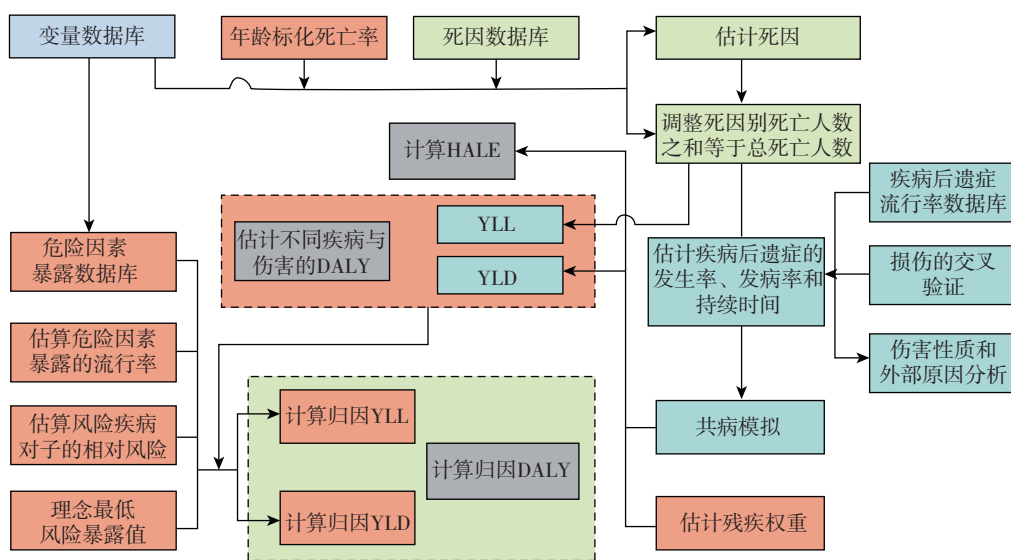


图 1-1 GBD 2010 的 18 个工作模块及其相互关系

注：本图参考框图 1-1 中文献 1^① 的图 1 重新绘制，经 *The Lancet* 授权转载

2. 研究方法的可重复性

自 GBD 2010 起，GBD 研究在分析框架上确立了系统一致的分析策略，旨在增强结果的可重复性与可验证性。研究团队在模型设计中明确识别各类估计中的不确定性来源，并通过量化不确定性区间来反映结果的统计稳定性和数据局限性。所有疾病、伤害和风险因素的估计均以 95% 不确定性区间（Uncertainty Interval, UI）报告，用以展示估计值的可信区间范围及其潜在误差边界。

此外，GBD 的建模流程普遍引入样本外预测效度（out-of-sample predictive validity）检验，以评估在新数据上的泛化能力。这种模型性能评估机制增强了方法的科学性和重复性，为跨时间、跨区域的比较提供了更加稳健的技术基础。这些方法的引入，标

① Murray CJL, Ezzati M, Flaxman AD, et al. GBD 2010: design, definitions, and metrics [J]. *The Lancet*, 2012, 380 (9859): 205-240.

志着 GBD 研究在确保估计结果的可重复性和透明性方面迈出了关键一步。

3. 流行病学转型理论作为分析框架

流行病学转型理论 (the Epidemiologic Transition Theory) 被广泛用于解释人口结构与社会经济发展对人群死亡模式和疾病谱变化的深远影响^①。该理论认为,随着死亡率水平的下降与生活条件改善,主要死因会由传染病、孕产妇和围生期及营养相关问题,逐步转向慢性非传染性疾病和退行性疾病。基于该理论框架,Salomon JA 和 Murray CJL 对 WHO 来自多个国家的、长达 50 年、并按年龄与性别分类的死因登记数据进行了重新分析,并结合人均收入水平和人口年龄结构等变量,探索各国死因模式演变的社会背景与驱动机制^②。

自 GBD 2010 起,为了深入理解不同时期死因别死亡数与 DALY 数的变动成因,GBD 研究团队引入“变化分解 (decomposition)”策略,将 1990 年到 2010 年估计值的变化划分为三类驱动因素:总人口增长、人口年龄结构变化以及流行病学因素的变化(表现为年龄别、死亡率或患病率的变化)。

至 GBD 2015,该分解方法在危险因素归因负担分析中进一步扩展,将归因于某类危险因素的 DALY 数的变化分解为总人口规模变化、人口年龄结构变化、危险暴露水平的变化,以及去除特定风险后的其他因素的贡献。GBD 在此处采用 Das Gupta 提出的分解技术来估计各因素对结果的贡献。

Das Gupta 在其著作《比率的标准化与分解:用户手册》(*Standardization and decomposition of rates: a user's manual*)^③中,系统阐述了如何将两个比率之间的差异拆解为可加性分量。该方法基于代数分解,而非回归建模或参数分析,因此特别适用于分析结构性变化对健康指标的影响。

有关该分解方法在 GBD 死因与归因负担估计中的具体应用,见本书第四章与第六章的相关论述。

第二节 GBD 研究的特点

前文已指出,GBD 研究在理念、方法与应用层面均显著区别于传统的描述性流行病学研究,为了全面理解其作为全球健康量化评估体系的独特性,本节概述 GBD 研究的八个主要特点。

① Omran, AR. The epidemiologic transition: A theory of the epidemiology of population change [J]. *Milbank Memorial Fund Quarterly*, 1971, 49 (4): 509-538.

② Salomon JA, Murray CJL. The Epidemiologic Transition Revisited: Compositional Models for Causes of Death by Age and Sex [J]. *Population and Development Review*, 2002, 28 (2): 205-228.

③ Das Gupta P. *Standardization and decomposition of rates: a user's manual* [M]. Washington, DC: US Bureau of the Census, 1993.

一、全局性与可比性

GBD 研究的核心目标，是在统一的分析框架下，对全球不同国家和地区的人群健康状况进行标准化与可比性评估。该研究通过构建一致的指标体系（如 DALY 和 HALE）、采用统一的数据处理与建模流程，确保不同疾病、伤害和危险因素在全球范围内得到一致的量度与比较^①。无论是高收入国家还是低收入国家，GBD 都力图在相同的方法学规则下对健康负担进行量化，从而支持跨区域、跨时间的人群健康比较，揭示全球和区域层面的主要健康挑战。

（1）在病因层面，GBD 对全球范围内的疾病、伤害和危险因素进行系统梳理与统一编码，将所有健康问题归入相互嵌套的多级分类体系，并使用统一的病因列表、后遗症划分及残疾权重估计方法，使传染病、慢性病、精神障碍、意外伤害、环境与行为危险因素等能够在同一指标体系中进行交叉比较。这种做法为疾病负担排序、卫生资源配置及优先领域的识别提供了统一基础，也使一些长期被忽视的健康问题，如精神健康问题伤害、及被忽视的热带病的非致命健康影响，以及更普遍的非传染性疾病得到应有的重视^②。

（2）在时间维度上，GBD 研究以年度为单位持续更新疾病负担估计，并系统追溯至 1990 年，形成跨越 30 余年的全球健康数据库。其时间序列的可比性依赖于统一的模型设定与参数规范，确保不同疾病或风险因素的趋势反映真实流行水平的变化，而非源于数据或方法差异的偏倚。这种一致性也使得 GBD 成为评估健康政策成效与应对突发公共卫生事件（如 COVID-19）的重要工具。

（3）在年龄和性别维度，GBD 采用标准化的人口分层结构，按 5 岁年龄组和性别进行细化估算，从而实现跨生命周期与性别人群的系统比较。这不仅有助于揭示不同年龄段人群主要疾病负担的异同，还能识别性别健康差异和易感人群特征，从而为制定精准和分层的健康干预策略提供科学依据。

综上所述，GBD 研究通过统一的指标体系、分析方法与时间分辨率，实现了在病因 - 时间 - 年龄 - 性别 - 地理空间五个维度的全球健康负担可比性。这种全局性与系统性，是 GBD 区别于传统单点式流行病学的根本特征，也是其在全球卫生治理与政策设计中发挥独特价值的基础。

① Murray CJL, Lopez AD. On the comparable quantification of health risks: lessons from the global burden of disease [J]. *Epidemiology*, 1999, 10: 594-605.

② Murray CJL, Vos T, Lozano R, et al. Disability-adjusted life years (DALYs) for 291 diseases and injuries in 21 regions, 1990-2010: a systematic analysis for the Global Burden of Disease Study 2010 [J]. *Lancet*, 2012, 380: 2197-2223.

二、政策评估工具

GBD 研究不仅是一项科学研究项目，也是一种面向政策制定的分析工具。其估计结果被广泛用于监测联合国千年发展目标（MDGs）和可持续发展目标（SDGs）的执行进展。例如，5 岁以下儿童死亡率（U5MR）的水平和变化趋势，常被用于评估各国实现千年发展目标 4（MDG4）的成效^①，并作为可持续发展目标（SDG）的核心指标^②。

GBD 研究通过引入社会人口指数（SDI），评估国家社会发展水平，并根据 SDI 水平设定相应的“预期健康状态”。各国人群的实际健康水平与 SDI 预期值之间的差距，为政策评估提供了可量化的依据。研究团队还据此识别在特定发展水平上表现优异或相对滞后的国家，为决策者提供可借鉴的经验与改进方向。

此外，GBD 研究借助流行病学转型理论，系统分析疾病与死亡模式的长期变化，并将死亡水平变化分解为三类驱动因素：人口规模增长、人口结构老化以及流行病学模式的演变。这一分析框架不仅揭示了健康水平变化的结构性原因，也提升了 GBD 在政策评估中的解释力与应用价值。

三、基于多重 Meta 分析的整合研究平台及其可信度提升机制

GBD 研究并不直接收集原始数据，而是依赖全球范围内的多种来源的二手资料，包括人口与健康调查报告、行政记录、科研文献、医院记录、死亡及疾病登记系统等。这些资料在数据质量、指标定义、样本代表性、分析方法等方面存在显著差异，传统 Meta 分析往往无法直接应用。

为应对这一挑战，GBD 团队开发了 DisMod-MR（disease modeling-meta-regression）这一核心建模工具。该工具突破了传统 Meta 分析的限制，能够在信息高度异质、数据来源多样且分布不均的情况下，构建稳健且可比较的分析模型，对传统 Meta 分析范式进行了突破性扩展。这类建模逻辑已远超传统 Meta 分析对同质性研究结果的汇总方式，成为系统性、多层次、面向估计推断的 Meta 建模平台，广泛用于非致命性疾病负担估计。该工具的理论基础来自 IHME 提出的《描述流行病学的综合元回归框架》（框图 1-1 文献 4^③）。

尽管 DisMod-MR 增强了数据整合能力，但其“包容性强”的建模逻辑也引发了

① Wang H, Liddell CA, Coates MM, et al. Global, regional, and national levels of neonatal, infant, and under 5 mortality during 1990-2013: a systematic analysis for the Global Burden of Disease Study 2013 [J]. Lancet, 2014, 384: 957-979.

② GBD 2015 DALYs and HALE Collaborators. Global, regional, and national disability-adjusted life-years (DALYs) for 315 diseases and injuries and healthy life expectancy (HALE), 1990-2015: a systematic analysis for the Global Burden of Disease Study 2015 [J]. Lancet, 2016, 388: 1603-1658.

③ Flaxman AD, Vos T, Murray CJL. An Integrative Meta-regression Framework for Descriptive Epidemiology [M]. Seattle: Institute for Health Metrics and Evaluation, 2015.

部分学界对于证据稳健性的担忧。为回应这一挑战，GBD 团队于 2022 年提出了“举证责任研究（Burden of Proof Studies, BoP）”方法框架（框图 1-1 文献 6^①），在后续研究中引入了更为严格的证据评估框架，从而形成了“系统整合 - 因果校准”的双重路径。BoP 研究可以被视为 GBD 团队对 Meta 分析方法合理性与科学性的一次回归式规范（normativeregression），在不削弱模型灵活性的前提下，加强了因果推断的可信度。需要指出的是，该框架目前还只针对归因负担研究，对非致死性健康研究还未涉及。

四、死因、疾病、后遗症与危险因素的分级结构

截至 GBD 2019，GBD 研究已纳入 369 种疾病和伤害、3000 余种后遗症及 87 种危险因素，并采用分级结构进行分类。死因、疾病和伤害的原因分类上，GBD 主要参照国际疾病和相关健康问题的统计分类（International Statistical Classification of Diseases, ICD）的分类原则进行，但第 1 级和第 5 ~ 6 级与 ICD 的分类框架则不完全对应。第 1 级把疾病或伤害的原因或死因分成三大类，这一层在 ICD 分类中并不存在；第 2 ~ 4 级与 ICD 疾病分类对应；第 5 ~ 6 级为基于问卷调查的症状和体征建立的“后遗症”体系，其健康状态的分类码与 ICD 的疾病原因编码并非一一对应。

在 GBD 的比较风险评估（the Comparative Risk Assessment, CRA）框架下，危险因素分为四级：第一级分为行为、环境职业和代谢危险因素三大类；第二级至第四级逐层细化。以 GBD 2016 为例，第 2 级包括 17 个危险因素，第 3 级包括 50 个危险因素，第 4 级包括 67 个危险因素。

此外，自 GBD 2015 以来，GBD 发展了新指标：每种危险因素的“总暴露值（Summary Exposure Value, SEVs）”，即相对风险加权的暴露流行率。该指标能够在总体水平上反映风险暴露的趋势，并用于估计死亡和 DALY 中可归因于特定危险因素的比例。

五、结合小域数据与地理层级先验的建模方法

从 GBD 2010 开始，GBD 研究团队在非致命疾病负担估计中引入了 DisMod-MR 模型序列，并发展出被称为“分析级联（analytical cascade）”的分层估计框架，对全球、区域和国家一级的非致命疾病负担进行估计。在 GBD 研究中，如何在缺乏高质量原始调查数据的背景下，对国家以下级别（如省、市、县区）的人群健康状况进行可靠估计，一直是重要挑战。为应对这一挑战，GBD 研究团队将小域估计（Small Area Estimation, SAE）原理与地理级联建模（Geographical Analytical Cascade）方法相结合，构建出一套多层次、可传递的建模流程。其核心工具即为 DisMod-MR 系列模型。

^① Zheng, P., Afshin, A., Biryukov, S. et al. The Burden of Proof studies: assessing the evidence of risk [J]. 2022, Nat Med, 28, 2038-2044.

使用 DisMod-MR 模型，通过沿地理级联传递的先验分布进行估计非致死性疾病负担，正是对小域估计理论的现实应用，也是 GBD 研究中使用的独特分析策略。其具体建模流程及方法细节将在第五章中作进一步阐述。

六、数据覆盖面持续扩大，数据增长迅速

GBD 研究自 2010 年以来，数据源数量和覆盖范围显著提升。主要动力包括：①多数国家科学研究与公共健康体系的完善带来更多高质量数据，包括行政记录、医学与人口监测数据；②通过与政府和机构谈判获取未公开或受限数据。由此 GBD 研究已成为全球最大规模的人群健康数据库之一。

七、系统化数据质量评估和处理方法

面对来源广泛、定义多样、质量参差的数据，GBD 研究开发了大量针对不同数据类型的数据清洗、再分配与校正技术。例如“垃圾代码”（garbage codes）重新归类、年龄性别拆分与不同标准之间的数据映射与转换（crosswalk）等。数据异质性处理机制保障了研究的系统性与准确性。

此外，GBD 从各种来源收集的数据的异质性显著，GBD 研究应用年龄 - 性别拆分（age-sexsplitting），数据集间的转换映射等方法，进行校正。

目前，GBD 研究正在评估暴露于危险因素后相关结果的相对风险（RR）的可靠性，并分析影响 RR 稳健性的相关因素^①。

八、有特色的全球健康数据交换平台

在国际公共卫生与医学研究领域，已有多个高影响力的文献数据库与平台，如由美国国立卫生研究院（NIH）的美国国家医学图书馆（NLM）的国家生物技术信息中心（NCBI）开发和维护的 PubMed 数据库；其次为 Elsevier 出版，以其强大的制药和药理学文献覆盖面而著称的 Embase 数据库，其中包括各种欧洲期刊和会议摘要；另外，如，Web of Science 是一个多学科书目数据库，提供高级搜索、引文分析和文献计量学工具的平台，使研究人员能够跟踪引文并分析学术著作的影响；作为 Cochrane 系统综述数据库（cochrane database of systemic reviews, CDSR）等，以及中国国内三大主要文献全文数据库：知网、维普网和万方数据。此外，还有一些人群调查和监测数据库，例如：SEER，即美国国立癌症研究所肿瘤监测、流行病学及结局项目数据库（Surveillance, Epidemiology and End Results Program, <https://seer.cancer.gov>）；CI5，世界卫生组织五大洲癌症发病率数据（WHO Cancer Incidence in Five Continents,

^① Zheng P, Afshin A, Murray CJL, et al. The Burden of Proof studies: assessing the evidence of risk [J]. Nat Med, 2022, 28 (10): 2038-2044.

<https://ci5.iarc.fr/Default.aspx>); the WHO Global Infobase (<https://library.mcmaster.ca/data/who-global-infobase-data-saving-lives>) 等。作为 GBD 研究的协调中心, 健康测量与评估研究所 (IHME) 建立的全球健康数据交换平台 (the Global Health Data Exchange, GHDx)。该平台通过系统汇集、编译 GBD 及其协作者提供的有关人群健康的数据, 直接支持 GBD 研究。根据 IHME 官方说明: “GHDx 的目标是通过编目数据信息 (包括涵盖的主题)、提供数据提供者的链接或获取指引, 以及在获得许可的情况下, 直接提供数据下载, 来帮助人们获取数据”^①。经过多年建设, GHDx 已经涵盖了与人口健康有关的多方面数据。①数据类型: 行政数据、人口普查数据、人口监测、疾病登记、环境监测、流行病学监测 (被动监测、主动监测和哨兵监测)、生命登记以及 GBD 研究的估计和建模数据; ②背景数据: 媒体报道的事件、财务记录、立法、地理空间数据、科学报告和科学文献信息等。GHDx 已经成为一个非常独特的平台, 任何对人口健康感兴趣的研究者都可以快速找到并分享来自不同地方和提供者的数据信息。因此, GHDx 也体现了 GBD 研究的特点。该平台的详情见本章第四节。

第三节 GBD 研究中的重要统计分析方法

GBD 研究在方法学上高度依赖统计建模与数据整合技术, 其核心估计结果均建立在复杂的量化分析框架之上。该研究创新性地运用了多种统计分析方法来估计全球不同地区、国家人群的疾病负担, 并逐步形成了一套系统而统一的统计方法体系。本书不对这些统计方法做深入阐述, 感兴趣的读者可参阅相关专业文献与方法学专著, 以进一步了解 GBD 研究中所采用的模型结构与算法细节。

一、对人口间接估计技术的运用和发展

1. 人口间接估计技术简介

人口间接估计技术从 20 世纪 60 年代发展至今, 已形成了一套完整的方法体系。该技术主要用于在缺乏人口普查和生命登记数据的地区, 估算生育率、儿童死亡率、成人死亡率及年龄别死亡率等关键人口学指标。联合国国际经济和社会事务部人口司于 1983 年出版的《人口间接估计技术》(Manual X)^② 及其中文版^③, 系统阐述了人口统计学的核心概念, 如稳定人口模型、死亡率模型 (模型生命表和生育率模型, 并介绍了适用于不完整或有缺陷的人口数据的分析技术。1990 年, 人口司又出版了《逐

① Institute for Health Metrics and Evaluation. GHDx help documentation [EB/OL]. Seattle: IHME, 2023 [2023-12-01].

② Department of International Economic and Social Affairs. Manual X Indirect Techniques for Demographic Estimation [M]. New York: United Nations, 1983.

③ 人口间接估计技术 [M]. 国家统计局人口统计司, 译. 北京: 中国统计出版社, 1992.

步估算儿童死亡率指南》(*Step-by-Step Guide to the Estimation of Child Mortality*)^① 进一步说明如何利用妇女生育子女资料, 通过 BRASS 方法和 Brass-Macrae 方法^② 估算儿童死亡率。

随着新的人口估计技术的出现, 以及《Manual X》中部分内容过时, 联合国人口司于 2009 年 3 月在海牙会议上决定启动修订项目, 并于 2013 年正式推出新版人口统计工具^③。新版工具书共 33 章, 涵盖数据质量评估、出生、生育、儿童死亡与成人死亡的间接估计方法, 并通过国际人口科学研究联盟(International Union for the Scientific Study of Population, IUSSP)建立的互动学习平台(<https://demographic. estimation. iussp. org/>)向公众免费开放。该平台提供清晰的演示实例与可更新的教学资源, 使用户能够方便地学习和应用最新的人口估计方法。

2. GBD 对儿童、成人和年龄别死亡水平的间接估计技术的运用和发展

GBD 在传统间接估计方法和技术的基础上进行了多方面的创新, 尤其在儿童、成人和特定年龄死亡率水平的不确定性量化方面取得了突破, 并在实际人口中进行了广泛验证和应用。

首先, GBD 提出了一种基于简明出生史的新型低成本方法^④, 用以推算 U5MR 及其不确定性区间宽度。该方法不仅提高了原标准方法的精度, 还扩展了时间覆盖范围。

死亡分布法(Death distribution methods, DDMs), 包括广义增长平衡法(Generalized growth balance, GGB)、综合迭代法(Synthetic Extinct Generations, SEG)及其组合方法(GGBSEG), 是评估生命登记系统、人口普查漏报情况、估算成人死亡率的常用方法, 并在实践中得到广泛运用, 但其方法选择缺乏科学文献指导。GBD 系统地评估了 234 种不同 DDM 方法在三种验证环境中的性能, 并基于此改进了这些方法在生命登记不完善国家中的应用^⑤。

同胞生存法(sibling survival)是一种间接估计成人死亡率的方法, 但由于存在回忆误差, 容易导致估计偏差。GBD 研究发展的“校正的同胞生存法”(the Corrected

① United Nations, Dept of International Economic and Social Affairs. *Step-by-step guide to the estimation of child mortality* [M] . New York: United Nations Department of International Economic & Social Affairs, 1990.

② Brass W, Macrae S. Childhood mortality estimated from reports on previous births given by mothers at the time of a maternity: I. Preceding-births technique [J] . Asian and Pacific Census Forum, 1984, 11 (2).

③ Moultrie TA, Dorrington RE, Hill AG, et al. *Tools for Demographic Estimation* [M] . Paris: International Union for the Scientific Study of Population, 2013.

④ Rajaratnam JK, Tran LN, Lopez AD, et al. Measuring Under-Five Mortality: Validation of New Low-Cost Methods [J] . PLoS Medicine, 2010, 7 (4), e1000253.

⑤ Murray CJL, Rajaratnam JK, Marcus J, et al. What Can We Conclude from Death Registration? Improved Methods for Evaluating Completeness [J] . PLoS Medicine, 2010, 7 (4).

Sibling Survival, CSS)^①用以修正传统同胞生存法中因回忆误差所致的选择偏差,对无存活家庭成员造成的死亡低估提出新解决方案。

此外,GBD研究在修订的对数模型寿命表(mortified logit life table, MLLT)^②基础上,进一步发展了具有灵活标准的模型生命表系统(model life table system with flexible standards, MLTFS)。该系统解决了两个关键局限:一是优化了在HIV高流行环境下的成人死亡率估计,二是改进了最大年龄组老年人口的年龄别死亡率估计。

通过整合上述方法,GBD实现了对全球几乎所有国家和地区儿童、成人及各年龄别死亡水平的系统估计^③。

二、死因估计的集成模型

1. 对国家死因模式的早期研究

20世纪70年代以来,学者们开始尝试通过统计建模方法系统地揭示各国死因结构与总体死亡率变化之间的关系。社会学家Samuel H. Preston基于165个人口数据,构建了国家死因模式与死因别死亡水平和变化趋势的关系模型,分析了各种死亡原因对总死亡率、年龄和性别死亡率的影响,并探讨了经济社会因素对20世纪死亡率下降的贡献^④。

基于这一研究框架,在1980年前后,不少学者对全球不同地区的主要死亡原因及其死亡率进行了估计。世界发达国家的死亡率数据直接来源于WHO的死亡率数据库,而对于缺乏死因数据的国家,则采用线性回归方法估计死因别死亡率。20世纪80~90年代关于全球部分国家死因模式估计的系列研究报告均采用了这种基于经验关系的建模途径^{⑤⑥}。

这些早期探索奠定了全球死因估计的理论与方法基础。其核心思想——利用跨国模式、社会经济指标及有限的可观测数据推断死因结构——正是后续全球疾病负担

① Obermeyer Z, Rajaratnam JK, Murray CJL, et al. Measuring Adult Mortality Using Sibling Survival: A New Analytical Method and New Results for 44 Countries, 1974-2006 [J]. PLoS Medicine, 2010,7(4).

② Murray CJL, Ferguson BD, Lopez AD, et al. Modified logit life table system: Principles, empirical validation, and application [J]. Popul Stud (Camb), 2003, 57: 165-182.

③ Wang HD, Dwyer-Lindgren L, Murray CJL, et al. Age-specific and sex-specific mortality in 187 countries, 1970-2010: a systematic analysis for the Global Burden of Disease Study 2010 [J]. Lancet, 2012,380(9859): 2071-2094.

④ Preston SH. Mortality Patterns in National Populations: With Special Reference to Recorded Causes of Death [M]. New York: Academic Pr, 1976.

⑤ Hakulinen T, Hansluwka H, Lopez AD, et al. Global and regional mortality patterns by cause of death in 1980 [J]. International Journal of Epidemiology, 1986,15(2): 226-233.

⑥ Bulatao R. Mortality by cause, 1970 to 2015 [C] //Gribble JN, Preston SH (eds.). The Epidemiological Transition: Policy and Planning Implications for Developing Countries. Washington, DC: National Academy Press.1993.

(GBD) 研究在构建死因估计模型 (如 CODEm) 时的重要理念源头。

2. 死因集成模型的发展

GBD 1990 也沿用类似思路, 对 1990 年全球不同国家人群的死因模式进行了系统评估, 并在死因分类和年龄组划分上进一步细化。另外, GBD 2000 提出了“组合模型” (the combinatorial models) 的思路来构建死因分布模型^①。GBD 2010 提出了死因模型开发、验证和报告的五项一般原则以规范模型的建立, 系统规范了模型的开发、验证与报告流程, 并开发了一种新的分析工具, 即死因集合模型 (the cause of death ensemble model, COMEm)^②。CODEm 的核心思想是通过大规模模型集成与性能评估, 从众多候选模型中筛选出预测性能最优的组合, 以提高估计的稳健性和透明度。在建模过程中, CODEm 首先基于不同算法可产生许多可能的协变量组合, 然后通过混合效应线性模型和时空高斯过程回归 (Gaussian process regression, GPR) 估计死因构成、死因别死亡率及其时空分布趋势。每一候选模型的样本外预测效率均被系统评估, 最终按预测性能加权整合为最优模型集。该方法不仅显著提升了模型预测力, 也为跨国、跨时间的死因估计建立了可复制的标准化框架。

三、描述流行病学的综合元回归框架

《描述流行病学的综合元回归框架》 (*An Integrative Metaregression Framework for Descriptive Epidemiology*), 是 GBD 2010 后发展形成的一套用于描述流行病学 Meta 分析的系统方法学框架。本框架旨在回应一个核心问题: 在高度异质且不完整的数据条件下, 如何对疾病的关键流行病学参数进行一致性估计。

DisMod-MR 是一种基于贝叶斯框架的多阶段元回归建模工具, 其设计初衷是整合来自不同来源、不同空间尺度、不同质量和统计方法的数据, 统一估计疾病的发病率、患病率、缓解率和死亡率等基本流行病学参数。该模型通过引入一组微分方程式, 系统描述疾病在发病、康复和死亡之间的动态平衡关系, 从而在数据不完整的情况下实现参数的“相互约束式补全”。

DisMod-MR 的发展经历了多个版本迭代, 从 DisMod-MR1.0 到 DisMod-MR2.1, 显著提升了计算速度、地理分辨精度和数据可视化能力, 逐步形成了 GBD 非致命性疾病负担估计的核心技术体系。

尽管 DisMod-MR 属于相对较新的建模工具, 但其基本研究思路, 即“当部分关键数据缺失, 但数据收集过程不存在系统性选择偏倚时, 利用描述疾病过程各变量之

^① Salomon JA, Murray CJL. The Epidemiologic Transition Revisited: Compositional Models for Causes of Death by Age and Sex [J]. *Population Development Review*, 2002, 28 (2): 205-228.

^② Foreman KJ, Lozano R, Lopez AD, et al. Modeling causes of death: an integrated approach using CODEm [J]. *Popul Health Metr*, 2012, 10: 1.

间的因果与结构关系来补充信息”^①，贯穿了 GBD 研究的整个发展历程。GBD 早期使用的 DisMod- I、DisMod- II 模型，已通过发病率、病死率、流行率和缓解率之间的数学关系，对疾病的完整流行病学图景进行重建。DisMod-MR 则在此基础上扩展了贝叶斯推断与层级建模结构，使这一方法能够在全球尺度上系统整合异质数据，实现对非致命性疾病负担的标准化估计。

1. 方法学挑战及其应对

GBD 研究中用于估算疾病和伤害流行率的数据来源极为多样，具有显著异质性，这与常规流行病学分析中按照统一方案获得的数据存在根本差异。为了应对这一挑战，GBD 研究采用 Meta 分析的总体思路，同时突破传统 Meta 分析在研究对象、结局指标和人群特征一致性方面的严格限制。

GBD 研究团队在《描述流行病学的综合元回归框架》一书中明确指出：“Cochrane 指南警告，在进行 Meta 分析时，不应合并效应指标显著不同或研究对象高度异质的研究^②。这一原则在效应量 Meta 分析中尤为重要，然而在描述流行病学数据的 Meta 分析中构成了现实挑战。事实上，对不同结局指标的研究进行系统比较，正是该框架试图解决的核心问题。该方法致力于整合和比较疾病流行率、发病率、缓解率以及死亡率等不同结果指标，这些研究往往涵盖不同年龄、性别、地区和不同时间段的人群（*The Cochrane guidelines caution not to compare studies with very different outcome measures of effect or very different patient populations when conducting a meta-analysis.*⁹ *This is a subtle point and is more clearly developed in the effect-size meta-analysis realm than in the meta-analysis of descriptive epidemiological data. In fact, comparing studies with different outcome measures is at the heart of this book, which develops a method for comparing the results of descriptive epidemiological studies of disease prevalence, incidence, remission, and mortality risk that are focused on subpopulations from varying age groups, sexes, regions, and time periods*”^③）。

简言之，虽然传统指南强调避免对高度异质研究进行合并，但 GBD 研究团队通过开发新的统计整合框架，实现了对不同指标与不同研究人群数据的系统比较与一致性估计。这一突破不仅构成了 GBD 研究的方法学核心，也标志着其有别于传统流行病学分析的根本特征。在技术实现层面，GBD 研究提出了集成系统建模（integrative systems modeling, ISM）的思路，即将描述疾病自然史的系统动力学模型与统计建模方法相结合，以整合所有可用的相关数据；并以贝叶斯元回归工具 DisMod-MR 为主

① Barendregt JJ, Oortmarssen GJ, Vos T, et al. A generic model for the assessment of disease epidemiology: the computational basis of DisMod II [J] . Population Health Metrics, 2003,1 (1): 4.

② The Cochrane Collaboration [EB] .2012. <http://www.cochrane.org/>

③ Page M J. Te PRISMA 2020 statement: an updated guideline for reporting systematic reviews [M] . Br Med J, 372, n71 (2021).

要实现平台，从而在高异质性数据条件下实现参数一致性估计。

2. 描述流行病学的综合元回归框架

《描述流行病学的综合元回归框架》（框图 1-1 文献 4^①），系统阐述了 GBD 研究在估计非致命健康损失时所使用的 DisMod-MR 工具的理论基础、模型构建原则及应用方法，是 GBD 研究方法体系中的最核心的技术文献之一。该书不仅揭示了 DisMod-MR 与传统元回归分析之间的根本差异，也展示了如何通过系列模型设计，应对描述性流行病学数据分析中的异质性与稀疏性难题。全书共 20 章，分为两大部分。第一部分（第 2 ~ 8 章）讨论了人群疾病综合系统建模的理论基础；第二部分（第 9 ~ 20 章）则通过十余种疾病的元分析实例，展示模型在实际研究中的应用路径。

在理论部分，作者首先概述了疾病率、比值和持续时间等基本统计模型，然后着重讨论了年龄模式建模，以解决原始数据中年龄分组不一致的问题。同时，该书强调了协变量建模的重要性——这是 GBD 研究能在数据稀疏地区实现疾病率和比估计的关键机制。具体而言，第 2 章论述率、比和持续时间模型；第 3 ~ 5 章集中探讨年龄模式建模及专家先验；第 6 章系统阐述协变量建模；第 7 章探讨不同类型数据的流行率估计。第 8 章则介绍了模型实现所需要的数值算法，并指出蒙特卡罗方法的发展使得使用贝叶斯建模方法在 GBD 研究中的应用成为可能。

应用部分通过对可卡因依赖、经前综合征、胰腺炎、心房颤动、丙型肝炎病毒感染、焦虑症、肝硬化、终末期肾病、膝关节骨关节炎、双相情感障碍，以及与危险因素相关的水果摄入和酒精依赖等多种疾病和危险因素的建模实例，系统展示了模型在不同疾病场景中的实施过程，同时也揭示了当模型假设受限或被违反时的应对策略。

综上所述，《描述流行病学的综合元回归框架》通过将描述疾病动态过程的系统动力学模型与进行 Meta 分析的统计学模型相结合，构建出创新性的综合元回归体系。该框架借助马尔可夫链蒙特卡罗（MCMC）等先进算法，实现了在高度异质数据条件下的贝叶斯估计，从而显著提升了 GBD 研究在疾病、伤害及危险因素流行率建模中的理论深度与计算效率。

3. DisMod-MR 模型的分析级联过程

在非致死性疾病负担估计中，DisMod-MR 模型采用一种基于地理层级的“级联分析（cascade analysis）”策略。该策略通过沿地理层次逐级传递先验分布，实现对不同空间尺度人群的系统性估计。这一方法是 GBD 研究区别于传统流行病学分析的核心特征之一。

在操作上，分析级联自全球层面启动。研究者首先使用包含所有可用国家数据的混合效应非线性回归模型，生成初始的全局估计结果。随后，这一全局估计结果作为先验信息（the prior），传递至下一个层级，以构建大区层面的 DisMod 模型。随后，

^① Flaxman AD, Vos T, Murray CJL. An Integrative Meta-regression Framework for Descriptive Epidemiology [M]. Seattle: Institute for Health Metrics and Evaluation, 2015.

采用大区特定的混合效应非线性回归结果作为先验来估计区域结果。大区模型在结合区域特有数据重新拟合后,其结果再作为先验分布传递至区域层级。上述过程在地理层次上逐步递进,直至国家及次国家级。

随着模型的更新迭代,DisMod-MR2.1 已实现了从全球到国家以下各级的完整级联过程,并可在统一的回归框架下对不同空间层次的数据进行一致性分析。整个分析过程在 GBD 定义的五个地理层次上依次展开:全球(global)、大区(superregion)、区域(region)、国家(country)和国家以下的行政区域(subnational)。每一较高级别的估计结果均为下一级的分析提供先验信息与参数约束,使建模过程在保持整体一致性的同时,兼顾局部地区数据差异。

此外,分析人员可根据数据密度与研究目的,在各层级上按时间和性别进行独立或联合的级联建模,以进一步提升模型的灵活性与估计精度。有关 DisMod-MR 及其在疾病、伤害与危险因素估计中的应用,可参阅本书第五章及第六章的相关内容。

四、小域估计模型的运用

小域统计方法(SAE)是 GBD 研究中常规使用方法之一。统计学中,“小域(small area)”与“大域(larger area)”是相对概念。如果该域内的样本量足以支撑满足一定精度要求的直接估计,则归为“大域”;反之,则视为“小域”,也就是说针对该域任何指标的直接估计均无法达到可靠精度。

按照 Rao 与 Molina(2015)在《小域估计》(*Small Area Estimation*)^①一书中的定义,SAE 是一类针对样本不足或无样本的区域的间接估计方法,可在不增加样本量的情况下,获得满意精度的估计结果。其核心思想是通过借助邻近区域、历史信息、协变量或层级结构建模,实现对数据稀疏地区目标参数的精细化估计。有效实施 SAE 的关键点包括选择高质量预测变量、构建稳健的预测模型、系统评估模型质量,并精确测量结果的不确定性。

在 GBD 研究中,小域统计方法被广泛用于预测数据稀疏或缺失地区的人群分年龄-性别死亡率、死因别死亡率、以及疾病和伤害的流行病学参数,形成系统的验证框架(框图 1-1 文献 2^②)。

在 GBD 的分析体系中,小域估计理念主要体现在以下几个方面:

(1) 多级建模与先验传播 DisMod-MR 利用全球级与地区级高质量估计结果作为下级区域(如国家、省级)模型的先验输入,实现“自上而下”的信息传递,并允许数据丰富区域的经验对数据稀缺区域进行“信息借力”(borrowing strength)。

(2) 空间相关结构的引入: DisMod-MR2.1 版本引入地理协变量与空间结构建模,

① Rao JNK, Molina I. *Small area estimation* [M]. 2nd ed. Hoboken: John Wiley & Sons, 2015.

② Srebotnjak T, Mokdad AH, Murray CJ. A novel framework for validating and applying standardized small area measurement strategies [J]. *Popul Heal Metrics*, 2010, 8: 26.

有助于提高小区域估计的空间平滑性与逻辑一致性。

(3) 不确定性传播机制：模型在生成平均估计值的同时，还提供 95% 不确定性区间，以反映数据稀疏地区估计的可信区间宽度，体现 SAE 方法中“以稳为主”的策略。

(4) 多源异质数据整合：SAE 强调多源数据融合，GBD 通过“数据标准化—交叉校准—地理级联”流程实现对行政记录、调查数据与登记资料的系统整合。

总体而言，GBD 通过 DisMod-MR 将小域估计的理念嵌入其全球健康数据建模框架，不仅有效弥补了传统疾病调查无法覆盖的数据空白区域，也显著提高了估计的地理分辨率与政策相关性。这一方法在支持国家以下层级健康资源配置、干预评估与政策响应方面展现出重要价值，标志着 GBD 在全球健康估计技术路径上的重要突破。

五、举证责任研究

传统 Meta 分析在推断健康风险与疾病结果的因果关系时，常因数据异质性较大、证据等级不一致或方法透明度不足而面临可信度争议。为应对这一挑战，GBD 研究团队自 2022 年起提出了“举证责任研究 (BoP)”方法，旨在构建一个更加系统、透明且稳健的风险评估框架^①。BoP 方法的核心思想是在面对复杂或尚存争议的风险因素时，通过系统评估所有相关研究的证据强度与一致性，构建“最保守可信的风险函数 (burden of proof risk function)”。它通过严格纳入标准、规范化质量评分、对研究偏倚的量化评估、以及敏感性分析和样本外预测等手段，识别因果联系的下限估计值，从而提升风险归因的稳健性。

BoP 研究进一步提出“证据等级评分体系”，将不同风险因素与健康结果之间的关联强度划分为“强”“中等”“有限”或“无确证”四级，为决策者提供基于证据质量的决策依据。这一机制的引入，不仅提高了 Meta 分析本身的效度，也增强了 GBD 对复杂公共卫生问题的应对能力和政策建议的可操作性。

六、整合多重 Meta 分析的集成流行病学建模体系

GBD 研究区别于传统流行病学研究的核心特征：它并不进行原始流行病学调查，而是建立在对全球现有健康数据和研究成果的系统性整合基础之上。其估算逻辑主要依赖于来自不同国家、不同时间、不同来源的二手数据的收集、整合与标准化处理，并通过基于 Meta 分析原理的建模方法，完成全球疾病与风险因素负担的系统估计。在分析过程中，GBD 研究采用了包容性极强的数据收集策略，大幅降低了纳入门槛，系统整合了来自全球不同国家不同人群的数据库资源。其后，GBD 对现有数据进行系统筛选、结构转化与质量调整，使数据以统一的格式作为分析输入。这种做法本质上等

^① Zheng P, Afshin A, Biryukov S. et al. The Burden of Proof studies: assessing the evidence of risk [J]. Nat Med, 2022, 28, 2038-2044.