

附件

《环境与健康统计分析技术指南（试行）
（征求意见稿）》编制说明

《环境与健康统计分析技术指南（试行）》指南编制组
二〇一七年八月

项目名称	环境与健康统计分析技术指南（试行）
项目编制单位	北京师范大学、环境保护部华南环境科学研究所、环境保护部环境与经济政策研究中心、中国医学科学院基础医学研究所（北京协和医学院基础学院）、华中科技大学公共卫生学院
编制组主要成员	程红光、孙海旭、赵欣怡、于云江、胡国成、王建生、许群、徐顺清
环境保护部科技标准司环境健康管理处项目负责人	宛悦

目 录

1 项目背景.....	1
2 必要性和可行性分析.....	1
3 国内外研究现状.....	1
4 主要技术内容说明.....	2
5 对实施本标准的建议.....	4

1 项目背景

1.1 任务由来

《环境与健康统计分析技术指南（试行）》（以下简称《指南》）为2017年度环境保护部环境与健康工作任务，编制任务由北京师范大学（以下简称北师大）承担。

1.2 工作过程

2016年10月：环境保护部科技标准司向北师大下达编制任务。

2016年11月-2017年3月：北师大成立《指南》编制组。编制组在广泛搜集、查阅相关标准、书籍、文献以及环境与健康调查项目资料基础上，拟定《指南》的基本框架，包括环境与健康统计分析方法及选取原则、污染源调查统计分析、环境暴露调查统计分析、人群健康调查统计分析、关联性分析等内容。

2017年4-6月：环境保护部科技标准司组织召开开题论证会，明确编制《指南》的指导思想、基本原则和技术路线。根据开题论证会上专家提出的意见，编制组在《指南》基本框架基础上加入“数据预处理及数据质量评价”，并且在污染源调查统计分析、环境暴露调查统计分析、人群健康调查统计分析中加入“数据特征”。在污染源调查统计分析、环境暴露调查统计分析、人群健康调查统计分析中按照“定指标、定水平、找差异”思路撰写，最后进行关联性分析。编制组经过多次研讨修改后，形成了《指南》（初稿）和编制说明。

2017年7月：环境保护部科技标准司组织召开技术审查会，对《指南》（初稿）和编制说明进行了技术审查。根据技术审查会意见，编制组形成了《指南》（征求意见稿）及编制说明。

2 必要性和可行性分析

2.1 必要性

2014年修订的《中华人民共和国环境保护法》首次设立专门条款，明确了环境与健康工作是环境保护事业的重要组成部分，同时新增第39条“国家建立健全环境监测、调查和风险评估制度”。为推动制度建设，《国家环境保护“十三五”环境与健康工作规划》提出制定系列环境与健康技术规范。2017年6月，环境保护部发布了《环境与健康现场调查技术规范 横断面调查》（HJ839-2017），配合标准的发布，考虑环境与健康调查数据具有非正态性、非线性、非独立性、因果关系复杂性、时空关系复杂性等特点，数据分析需要专门的统计分析技术文件指导并规范统计分析流程。《指南》的编制是环境与健康管理制度建设的客观需求，也是科学认识环境与健康关系的重要支撑。

2.2 可行性

《指南》编制有理论基础和实践经验。现有的《环境统计》《医学统计学》等教材的基础理论为《指南》编写提供了方法学依据。同时，编制组成员在污染源调查、暴露评价、流行病学调查、环境统计学、医学统计学等方面具有较好的理论知识和实践经验，所提出的统计分析技术路线和方法均在环境与健康现场调查中得到应用。

3 国内外研究现状

环境与健康统计分析方法多参考环境统计学、医学统计学以及数理统计相关的书籍进行，没有专门的统计教材，国内外也均未发布技术指导文件。无论环境统计学、医学统计学以及数理统计，在理论方法上存在一致性，但在具体应用时需要根据实际情况选择。本指南编制是建立在统计分析共性理论方法的基础上，根据环境与健康横断面调查目的及获得数据特点，有针对性地制定统计分析技术路线、筛选统计分析方法。

4 主要技术内容说明

4.1 层次框架

《指南》正文由 10 部分组成，包括：

- 1) 适用范围
- 2) 规范性引用文件
- 3) 术语和定义
- 4) 统计方法和选取原则
- 5) 数据预处理及数据质量评价
- 6) 污染源调查
- 7) 环境暴露调查
- 8) 人群健康调查
- 9) 关联性分析
- 10) 统计结果表达

4.2 技术要点

4.2.1 适用范围

本技术指南适用于环境与健康横断面调查所获得的数据的统计描述和统计推断，病例-对照研究、队列研究获得数据的统计分析另行制定。

4.2.2 术语和定义

指南共有 38 个术语定义，主要参考环境统计、医学统计和国家相关标准。

表 1 术语和定义及参考文献

术语和定义	参考文献
总体 样本 参数 统计量 计量资料 计数资料 等级资料 平均数 方差和标准差 分位数和百分位数 四分位数间距 相对数 率 构成比 相对比 假设检验 方差齐性检验 Z 检验 t 检验 t' 检验 方差分析 卡方检验 连续性校正 卡方检验 Fisher 确切概率检验 Wilcoxon 秩和检验 Kruskal-wallis H 检验 Pearson 相关系数 Spearman 相关系数 线性回归 Logistic 回归 死亡率 患病率 总率和专率 粗率和调整率	《统计学》(徐国祥主编, 上海人民出版社, 2007 年) 《医学统计学(第二版)》(刘桂芬主编, 中国协和医科大学出版社, 2007 年) 《中华医学统计百科全书: 描述性统计分册》(田考聪主编, 中国统计出版社, 2012 年) 《医学统计学》(贺佳, 尹平主编, 高等教育出版社, 2012 年) 《医学统计学(第六版)》(李康, 贺佳主编, 人民卫生出版社, 2013 年) 《中国公共卫生·方法卷》(王宇, 杨功焕主编, 中国协和医科大学出版社, 2013 年) 《医学统计学(第三版)》(颜虹, 徐勇勇主编, 人民卫生出版社, 2015 年)
横断面调查	《环境与健康现场调查技术规范 横断面调查》(HJ 839-2017)
等标污染负荷	《环境统计》(蔡宝森主编, 武汉理工大学出版社, 2004 年)
暴露量	《污染场地术语》(HJ 682-2014)
暴露参数	《环境化学污染物暴露评估技术指南(征求意见稿)》(环办科技函(2017)1211 号)

4.2.3 规范性引用文件

环境与健康调查数据的分布类型检验依据《GB/T 4882 数据的统计处理和解释正态性检验》进行，离群值处理依据《GB/T 4883 数据的统计处理和解释正态样本离群值的判断和处理》进行。为规范统计分析结果表达，需要依据《GB/T 8170 数值修约规则与极限数值的表示和判定》进行数值修约。

4.2.4 数据预处理及数据质量评价

对环境与健康横断面调查获取的数据进行数据预处理和数据质量评价是统计分析的前提。数据预处理包括缺失值、未检出数据及离群值的识别与处理。数据真实性和准确性直接影响统计分析结果，《指南》对应提出了数据质量评价指标，包括可溯源率和正确率两项指标。

4.2.5 统计分析方法选取原则

环境与健康横断面调查数据的统计分析主要涉及描述性统计分析、差异比较、相关与回归分析。需要依据数据的类型、样本量、数据分布特征以及方差是否齐性等几个方面来选取具体的描述性统计分析指标及推断性统计分析方法。此处列出了最基础、最常用的描述性统计分析指标及推断性统计分析方法的选取原则。

4.2.6 污染源调查统计分析

通过污染源调查（资料调研、现场监测），可以获得污染源的类型和位置、污染物的种类、排放量、排放浓度、污染物的排放方式和途径等数据。通过空间地图展示调查区内污染源的空间分布情况，可为高风险污染源的确定提供基础。

为了确定高环境健康风险的污染源和特征污染物，本指南借鉴了污染源评价中常用的方法——等标污染负荷法。等标污染负荷法可以将标准各异、量纲不同的污染源和污染物的排放量，通过一定的数学方法转变成一个统一的可比较值，从而确定出高环境健康风险的污染源和污染物。等标污染负荷法主要是基于污染物排放标准进行计算，考虑到相当一部分与人群健康关系密切的污染物尚无排放标准，本指南也提出了根据研究目的，依据国家发布的有毒有害污染物名录或者环境与健康调查研究结果，确定高健康风险污染源和特征污染物的方法。

污染源监测获取的数据主要为污染源排放的特征污染物的浓度数据，按照统计方法选取原则选取统计指标和方法，统计分析其检出情况、浓度水平及超标情况。

4.2.7 环境暴露调查统计分析

对横断面调查所获得的环境暴露数据进行规范化统计分析。统计各环境介质污染水平，获得调查区的污染程度。进行不同研究区之间污染水平的差异性检验，进一步明确污染影响范围及程度。结合人群暴露参数估算人群外暴露量，为环境暴露与人群健康的关联性分析做准备。

4.2.8 人群健康调查统计分析

人群健康调查统计分析强调了在按人口学特征（人群的年龄、性别、职业、民族等）以及疾病种类描述疾病率时，需要计算专率的要求；比较不同年龄结构人群的疾病率时，需要进行标化处理的要求。

4.2.9 关联性分析

污染源与环境暴露的关联分析通过定性方法描述，包括距离-浓度散点图和统计地图。此步骤即检验特征污染物是否呈现与源的距离衰减关系，以确定源的影响途径和影响范围。

环境暴露与人群健康的关联分析通过定性和定量方法描述，包括双变量散点图、相关分

析、回归分析和环境暴露组别间率的差异性检验。特别指出由于环境暴露导致健康结局的潜伏期可能长达数年,需要根据实际情况选取过去的环境暴露水平与当前的健康水平进行关联性分析。

4.2.10 附录

统计表和统计图是统计描述的重要工具。尤其广泛应用于统计结果表达及对比分析中。附录中针对统计表与统计图的制作原则和结构进行了详细说明,主要参考书目有《医学统计学(第六版)》(李康,贺佳主编,人民卫生出版社,2013年)及《医学统计学》(贺佳,尹平主编,高等教育出版社,2012年)。

5 对实施本标准的建议

本指南为现阶段指导性标准,建议指南发布实施后,根据实施情况适时对本指南进行修订,同步加强相关科学研究。