

附件 2

环境与健康统计分析技术指南（试行） （征求意见稿）

环境保护部科技标准司

二〇一七年八月

目次

前 言.....	i
1 适用范围.....	1
2 规范性引用文件.....	1
3 术语和定义.....	1
4 分析流程.....	4
5 数据预处理及数据质量评价.....	6
6 污染源调查.....	7
7 环境暴露调查.....	8
8 人群健康调查.....	9
9 关联性分析.....	10
10 统计结果表达.....	11
附录 A（资料性附录）统计表编制原则和结构.....	12
附录 B（资料性附录）统计图制作原则和结构.....	13

前言

为贯彻《中华人民共和国环境保护法》，科学分析环境污染与人群健康之间的关系，指导和规范环境与健康调查数据统计分析工作，推进环境与健康管理制度建设，制定本指南。

本指南规定了环境与健康统计分析方法的选取原则、分析流程、分析内容和分析方法。本指南由环境保护部科技标准司组织制定。

本指南为首次发布。

本指南主要起草单位：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

本指南由环境保护部 20 年月日批准。

本指南自发布之日起实施。

本指南由环境保护部解释。

1 适用范围

本指南规定了环境与健康统计分析方法的选取原则、分析流程、分析内容和技术要求。
本指南适用于针对环境与健康横断面调查所获得的数据进行的描述性统计分析和推断性统计分析。

2 规范性引用文件

本指南引用了下列文件中的条款。凡是不注明日期的引用文件，其有效版本（包括修改单）适用于本指南。

GB/T 4882	数据的统计处理和解释正态性检验
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
GB/T 8170	数值修约规则与极限数值的表示和判定

3 术语和定义

3.1 总体 population

根据研究目的确定的所有同质观察单位的集合。

3.2 样本 sample

根据研究目的从研究总体中抽取的部分观察单位。

3.3 参数 parameter

反映总体分布特征的统计指标。

3.4 统计量 statistic

反映样本分布特征的统计指标。

3.5 计量资料 measurement data

按观察单位观察值的大小表示其结果的资料类型。

3.6 计数资料 count data

又称分类资料，指按观察单位不同属性进行分组计数的资料类型。

3.7 等级资料 ordinal data

将观察单位按某种属性的不同程度或次序分成等级后分组计数观察的结果，具有半定量性质。

3.8 平均数 average

描述一组定量资料集中趋势的统计指标，常用以说明该组数据的平均水平。表示资料集中趋势的几个比较重要的指标有算术均数、几何均数、中位数。算术均数指所有观察值之和与观察值个数的商。几何均数指观察值对数的算术均数的反对数。中位数指将所有观察值从小到大排列后处于中间位置的值，如果观察值的个数为偶数，中位数则为中间两个数的算术均数。

3.9 方差和标准差 variance and standard deviation

方差和标准差是描述一个变量所有观察值与总体均数离散程度的指标。标准差是方差的算术平方根。

3.10 分位数和百分位数 quantile and percentile

分位数是指将样本按其值从小到大排列、将样本总量分为若干等分时每一个界点位置上的样本值。百分位数是分为 100 等分时的每个界值。

3.11 四分位数间距 quartile range

第三四分位数（第 75 百分位数）与第一四分位数（第 25 百分位数）之差。

3.12 相对数 relative number

两个相关联指标的数值之比，表示二者的相对大小。

3.13 率 rate

在某一时刻/时段内某现象或事件发生的频率或强度，可以用在某时刻/时段内某现象发生数与可能发生该现象的观察单位总数之比表示。

3.14 构成比 proportion

某事物内部各个组成部分所占的比重，是各个组成部分的观察单位数与总观察单位数之比。

3.15 相对比 relative ratio

两个相关指标的值之商。

3.16 假设检验 hypothesis testing

基于数理统计学原理根据一定假设条件由样本推断总体的一种方法。常用的假设检验方法有 Z 检验、t 检验、卡方检验、F 检验和秩和检验等。

3.17 方差齐性检验 homogeneity test of variances

判断两样本所代表的总体方差是否相等的方法。

3.18 Z 检验 Z test

又称 u 检验，指样本 n 大于等于 30 时的样本均数与总体均数间或样本均数间差异检验的方法，其基本前提是样本所在总体服从正态分布和已知总体方差。

3.19 t 检验 t-test

样本量小于 30 时对两个样本均数差异进行显著性检验时使用的方法，条件同 3.17。

3.20 t' 检验 t'-test

推论的两个样本所代表的总体方差不齐时使用的均数差异显著性检验方法。

3.21 方差分析 analysis of variance

又称变异数分析，用于两个及两个以上样本均数差别的显著性检验。其基本思路是将样本观测值的总变异分解为组间变异（处理）和组内变异（误差）后，通过检验二者的差异显

著性来分析影响总变异的各个因素的效应以判断各组样本间均数的差异是否存在显著性。

3.22 卡方 (χ^2) 检验 chi-square test

用于推断两个及多个总体率或总体构成比之间有无差别、两种属性或两个变量之间有无关联性以及进行频数分布的拟合优度的方法。

3.23 连续性校正 χ^2 检验 continuity-adjusted chi-square test

进行两个样本率差异显著性 χ^2 检验时, 如果发现总例数较少 (通常 <40) 或某个格子的理论频数较小 (通常 <5) 时, 为改善 χ^2 统计量分布的连续性而进行的校正。

3.24 Fisher 确切概率检验 Fisher's exact probability test

一种直接计算概率的假设检验方法。 χ^2 检验时, 当样本含量较少 (四格表资料总例数 $n < 40$ 或有期望频数 $E < 1$), 行列表资料 χ^2 检验结果可能会产生偏性, 此时常使用 Fisher 确切概率法直接计算概率。

3.25 Wilcoxon 秩和检验 Wilcoxon rank sum test

又称两样本秩和检验, 对两组不满足 t 检验条件的完全随机设计的计量资料或等级资料, 检验其分别代表的总体均数有无差别的方法。

3.26 Kruskal-wallis H 检验 Kruskal-wallis H test

利用多样本的秩和来检验多个独立样本是否来自同分布总体的非参数检验方法。

3.27 Pearson 相关系数 Pearson correlation coefficient

描述服从双变量正态分布的两个变量之间线性相关性的指标。

3.28 Spearman 相关系数 Spearman's rank correlation coefficient

描述不服从双变量正态分布、总体分布未知或原始数据用等级表示的两个变量之间的相关性的指标。

3.29 线性回归 linear regression

用直线方程来描述两个变量之间回归关系的一种方法。

3.30 Logistic 回归 Logistic regression

反应变量(因变量)为分类变量, 研究其与多个自变量之间关联性的一种回归分析方法。

3.31 等标污染负荷 equal-standard waste load

对污染源排放的各种污染物, 用统一的标化系数对排放量进行统一的标化评价, 得到的各种污染物的排放量都具有统一的量纲数值。

3.32 暴露量 exposure dose

人体经呼吸道、消化道和皮肤等途径接触环境污染物的量。

3.33 暴露参数 exposure parameter

用来描述人体暴露环境污染物的特征和行为的参数。

3.34 死亡率 mortality rate

某个人群在一定期间（一般指一年）内死亡发生的频率，通常用千分率或十万分之率表示。

3.35 患病率 prevalence rate

反映疾病在人群中流行规模与水平的一个指标，又称现患率或流行率，指一定期间内、一定人群中患某种疾病的病例数（新、旧病例）所占比例。根据观察时间的长短分为时期患病率和时点患病率。前者指在某观察时间内（一般长于1个月）特定人群中某病新旧病例数占同期该人群平均人口数的比例，后者指在某一特定时点（一般不长于1个月）特定人群中某病新旧病例数占该时点该人群总数的比例。

3.36 总率和专率 overall and specific rate

总率指不分人群、地区特征、不分疾病种类的疾病率。专率指按照观察对象的不同特征，如人群的年龄、性别、职业、民族、种族、婚姻状态等特征以及疾病种类分别计算的发生率（如发病率、死亡率、患病率、新生儿畸形发生率等）。

3.37 粗率和调整率 crude rate and adjusted rate

粗率指不分观察人群内部构成直接计算某事件/现象发生的频率或强度。调整率又称标化率，指为了比较不同时期/不同地区观察人群间事件/现象发生的频率或强度，采用标准人口构成以消除待比较人群间因构成不同可能带来的影响而进行标准化调整后的率。标化率并不代表某事件/现象实际的发生频度或强度。

3.38 横断面调查 cross-sectional study

指在特定时点或时期，对污染源、环境暴露水平和相应暴露人群的健康影响同时进行的调查。

4 分析流程

4.1 技术路线

环境与健康统计分析分为数据预处理、数据质量评价、统计描述和统计推断、关联分析、结果表达六个步骤（图1）。

4.2 统计方法选择原则

4.2.1 描述性统计

a) 计量数据。首先判断数据分布类型，根据不同分布类型选择统计方法。如果数据服从正态分布，采用算术均数和标准差描述平均水平和变异程度；如果数据呈偏态分布，经对数转换后呈对称分布，或数据之间成级数关系，采用几何均数和几何标准差描述平均水平和变异程度；如果数据呈偏态分布且不能通过数据变换转化为正态分布，采用中位数描述其平均水平，采用四分位数间距描述其变异程度。

b) 计数数据。采用率、构成比、相对比等进行统计描述。

4.2.2 推断性统计

a) 单变量计量数据的平均数比较。单变量计量数据的平均数比较方法主要根据样本量、

数据分布类型以及方差是否齐性等特征进行选取（图 2）。

b) 两个样本率或构成比的比较。列出四格表，当每一格子期望频数 $E \geq 5$ ，且总样本例数 $n \geq 40$ 时，采用卡方检验；当总样本例数 $n \geq 40$ ，但其中有一个格子期望频数 $1 \leq E < 5$ ，采用连续性校正卡方检验；当任何一个格子期望频数 $E < 1$ ，或总例数 $n < 40$ ，或检验所得 P 值接近于检验水准 α ，采用 Fisher 确切概率检验。

对于多个样本率或构成比等相对数的差异比较，列成行列表形式，当各格子的期望频数 E 均大于 1，且 $E < 5$ 的格子数不多于格子总数的 $1/5$ ，采用卡方检验；否则应采取增加观察例数等措施或进行 Fisher 确切概率检验。

c) 计量数据的相关分析。做两变量散点图，若发现呈直线相关关系，并符合双变量正态分布，采用 Pearson 相关系数；对非正态分布的变量以及等级资料的相关分析，则计算 Spearman 秩相关系数。

d) 回归分析。如果因变量是连续型变量且只包括一个自变量，因变量和自变量的散点图呈现直线相关关系，则采用一元线性回归。如果包括两个或两个以上的自变量，且因变量和自变量之间是线性关系，则采用多元线性回归。如果因变量是二分类变量，则采用 Logistic 回归分析。

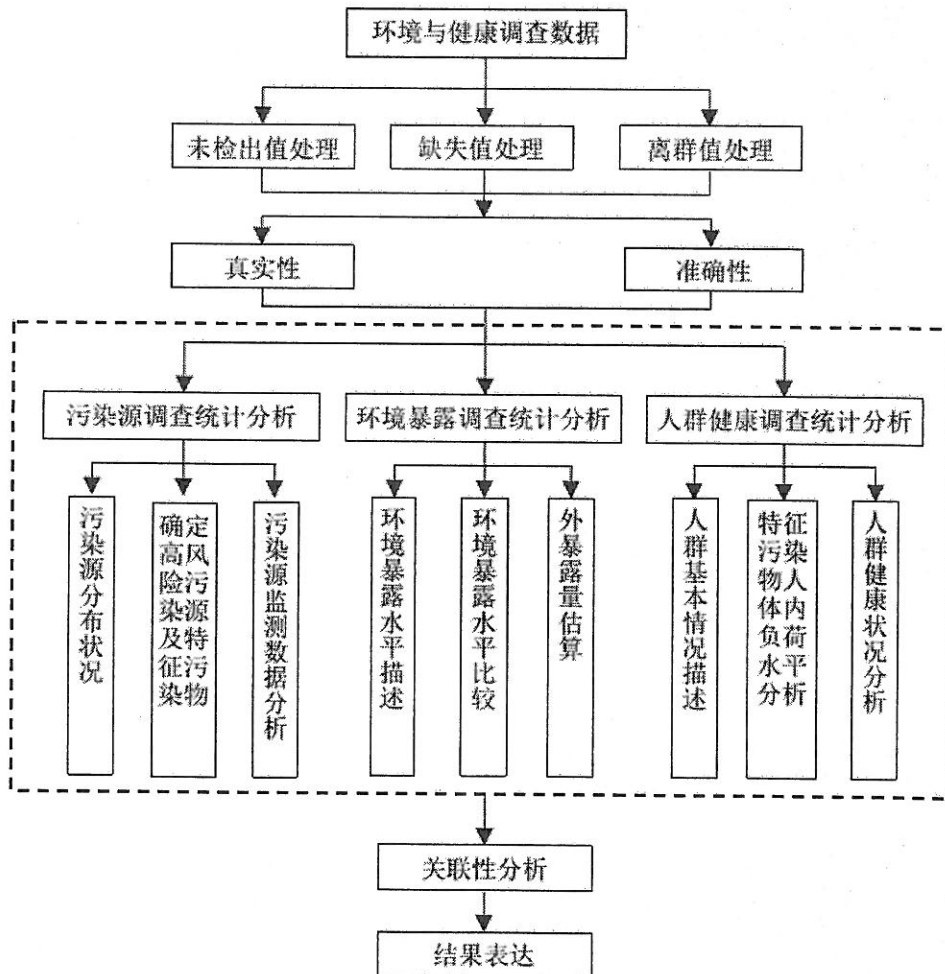


图 1 技术路线

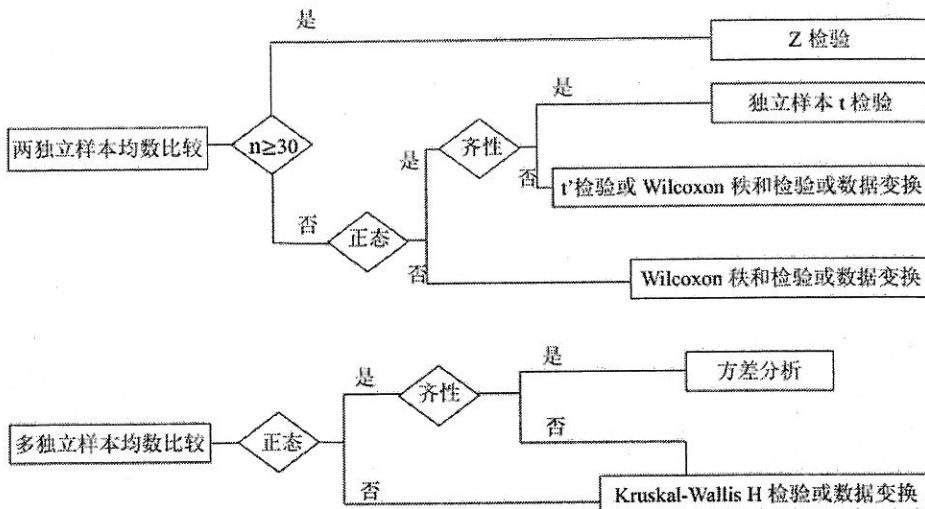


图 2 单变量计量数据均数比较方法选取程序

5 数据预处理及数据质量评价

5.1 数据预处理

5.1.1 缺失值处理

常见的缺失数据处理方法包括删除法、插补法等。对于污染源调查中现场监测数据、健康调查数据中的缺失值可采取直接剔除法，对于环境质量调查数据的缺失值采用插补法，通过空间插值的方法进行补充。

5.1.2 未检出数据处理

环境检测数据以方法检出限的二分之一替代未检出数据。医学检测数据以方法检出限的三分之一替代未检出数据。

5.1.3 离群值处理

应基于专业判断进行离群值处理。当样本数据分布服从正态分布时，依据 GB/T4883 剔除离群值；当样本数据分布不服从正态分布时，作样本数据箱线图（图 3）：

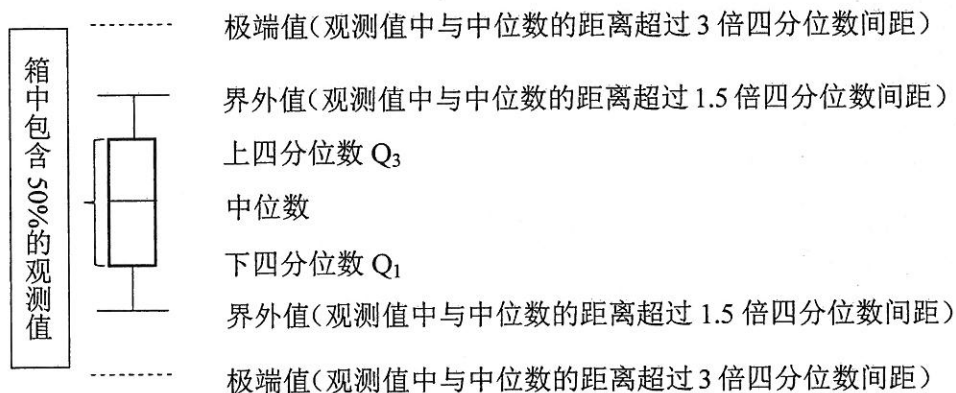


图 3 样本数据箱线图

箱线图中小于 $Q_1 - 1.5$ 倍四分位数间距或大于 $Q_3 + 1.5$ 倍四分位数间距的值为离群值，根据实际情况予以剔除。

5.2 数据质量评价

5.2.1 真实性

真实性的评价范围包括污染源现场调查数据、环境暴露调查数据、人群健康调查数据及实验室检测数据，数据真实性用可溯源率衡量评价。

5.2.2 准确性

数据准确性包括问卷调查和实验室检测两类数据的准确性。问卷调查数据的准确性用正确率衡量，问卷调查的正确率为问卷调查数据中无以下两类错误（非法值、逻辑错误）中任何一项的记录数占其上报总记录数的比例。实验室检测数据的准确性通过实验室质量控制结果衡量。

6 污染源调查

6.1 数据特征

污染源调查获取的污染物排放量、排放浓度及现场监测数据以计量资料为主，排放浓度和现场监测数据多呈对数正态分布。

6.2 统计内容及方法

6.2.1 污染源分布状况

根据污染源的空间位置信息，采用空间地图展示调查区内污染源的分布状况。

6.2.2 确定高风险污染源及特征污染物

6.2.2.1 分析内容

筛选调查区域内高环境健康风险的污染源及特征污染物。

6.2.2.2 分析指标

a) 等标污染负荷法

根据污染源类型、所属行业及污染物种类，选取污染物排放标准，计算污染源及污染物的等标污染负荷、等标污染负荷比（公式 1-6）。

1) 等标污染负荷

$$P_{ij} = \frac{G_{ij}}{S_{ij}} \quad (1)$$

式中：

P_{ij} —某污染源（j）中某种污染物（i）的等标污染负荷

G_{ij} —某污染源（j）中某种污染物（i）的年排放量

S_{ij} —某污染物（i）的评价标准，一般取排放标准

$$P_n = \sum_{i=1}^n P_{ij} \quad (2)$$

式中：

P_n —某污染源内各污染物的等标污染负荷之和

$$P_m = \sum_{j=1}^m P_{ij} \quad (3)$$

式中：

P_m —调查区内各污染源内某种污染物的等标污染负荷之和

$$P = \sum_{j=1}^m \sum_{i=1}^n P_{ij} = \sum_{i=1}^n \sum_{j=1}^m P_{ij} \quad (4)$$

式中：

P —调查区内所有污染源的等标污染负荷之和

2) 等标污染负荷比

$$K_i = \frac{P_m}{P} \quad (5)$$

式中：

K_i —调查区内某种污染物的等标污染负荷比

$$K_j = \frac{P_n}{P} \quad (6)$$

式中：

K_j —调查区内某污染源的等标污染负荷比

将污染物等标污染负荷比按大小排列，累计百分比大于 80% 的污染物为主要污染物。

将污染源等标污染负荷比按大小排列，累计百分比大于 80% 的污染源为主要污染源。

b) 其他方法

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，将其纳入统计分析。

6.2.3 污染源监测数据分析

6.2.3.1 分析内容

污染源排放的特征污染物的检出情况、浓度水平及超标情况。

6.2.3.2 分析方法

a) 检出情况。计算特征污染物的检出率（公式 7），描述其检出情况。

$$\text{检出率} = \frac{\text{测定结果高于方法检出限的样品数}}{\text{样品总数}} \times 100\% \quad (7)$$

b) 浓度水平。检验特征污染物浓度数据的分布类型，根据数据分布特征，参照 4.2.1 选取统计指标描述其平均水平。根据需要统计最大值，描述特征污染物浓度的极端水平。

c) 变异程度。根据特征污染物浓度数据分布特征，参照 4.2.1 选取统计指标描述其变异程度。

d) 超标情况。选取污染物排放标准，计算特征污染物排放浓度的超标率（公式 8），描述其超标情况。

$$\text{超标率} = \frac{\text{超过相应标准的样品数}}{\text{样品总数}} \times 100\% \quad (8)$$

7 环境暴露调查

7.1 数据特征

环境监测数据多呈非正态分布，具有严格时间或空间序列，时空相邻的数据具有更大可

能的相似度，人为或自然原因可能引起瞬间或局部环境监测数据变化。

7.2 统计分析内容及方法

7.2.1 环境暴露水平

对环境空气、环境水体、土壤、土壤种植的农产品、室内空气、饮用水、家庭购买的农畜水产品、室内积尘等样本中特征污染物的检出情况、平均水平、变异程度和超标情况进行统计描述。统计内容及指标包括：

- a) 检出情况。样本例数及特征污染物检出率。
- b) 平均水平。先做数据分布检验，再参考 4.2.1 选取描述性统计指标。
- c) 变异程度。用算数平均数描述暴露测量的平均水平时，须报告标准差、最小值、最大值和 95%置信区间。用中位数描述暴露测量的平均水平时，须报告最小值、最大值、四分位数和四分位数间距。
- d) 超标情况。有对应的环境标准或参考值时，报告样本超标率和平均超标倍数；无相关标准或参考值的，报告实测平均浓度相对于当地背景值或对照区污染物浓度值的比值；无当地背景值时，报告实测平均浓度相对于相关文献报告值的比值。

7.2.2 环境暴露水平比较

- a) 对污染区和对照区环境样本中特征污染物平均水平进行差异性比较，统计方法选取原则参照 4.2.2。
- b) 对多个区域的环境样本中特征污染物平均水平进行差异性比较，统计方法选取原则参照 4.2.2。

7.2.3 外暴露量估算

参见《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）。

8 人群健康调查

8.1 数据特征

人群健康状况往往与企业污染排放、调查区的环境质量状况在时空分布上存在一定的相关性。健康调查数据呈现组群间差异较大，群内差异较小的特点。

8.2 统计分析内容及方法

8.2.1 人群基本情况

8.2.1.1 分析内容

研究对象的人口学特征（如年龄、性别、民族、婚姻状况、文化程度等）、行为危险因素情况（吸烟、饮酒、户外活动习惯等）、职业暴露史、既往患病情况、家族史、就医行为等信息。

8.2.1.2 分析方法

参照 4.2.1 选取适当方法，对上述指标进行描述性分析；参照 4.2.2 选取适当方法，对上述指标进行差异比较分析。

8.2.2 特征污染物人体内负荷水平

8.2.2.1 分析内容

调查区域特征污染物的人体内负荷水平。

8.2.2.2 分析指标

根据人体内负荷水平数据的分布特征，参照 4.2.1 选取适当的统计指标描述其平均水平及变异程度；报告样本例数、体内负荷水平正常例数、异常例数，采用正常率、异常率、检出率等相对数进行统计描述。参照 4.2.2 选取适当方法，对研究区域人群内负荷水平及正常率、异常率、检出率等的差异进行比较。

8.2.3 人群健康状况

8.2.3.1 分析内容

调查人群健康状况（如症状、体征、疾病、死亡等）的分布及其影响因素。当对调查区域人群的疾病（死亡）率进行描述时，计算总率（如某病的患病率或全因死亡率）；当按人口学特征（人群的年龄、性别、职业、民族等）以及疾病种类描述疾病率时，则需要计算专率。在进行两组或几组人群的健康状况的比较时，必须对他们的年龄性别等可能影响健康状况的因素进行标准化。

8.2.3.2 分析方法

a) 症状和体征。统计各种症状体征的阳性例数、阳性率及其在不同人群中的分布，并参照 4.2.2 进行比较。

b) 生理生化。参照 4.2.1 选取统计指标描述不同人群生理生化指标的平均水平及变异程度；统计正常率、异常率并参照 4.2.2 进行人群间的比较。

c) 患病和死亡。统计不同疾病的患病（死亡）例数，估计相应的患病率（死亡率）等，并参照 4.2.2 进行不同人群间比较。

9 关联性分析

9.1 污染源与环境暴露

调查区内各环境介质中特征污染物浓度水平随污染源距离的变化情况可做散点图。若散点图呈现线性相关，则用相关系数和线性回归模型定量描述距离与污染浓度的相关程度，相关和回归模型选取原则参照 4.2.2。采用统计地图标出污染源地理位置，并采用空间插值等方法绘制特征污染物分布情况，定性观察污染源与环境污染水平在空间分布上的相关性。

9.2 环境暴露与人群健康

根据特征污染物和健康效应类型选取环境暴露与健康水平关联性分析指标。某一环境暴露因素导致某疾病的潜伏期可能长达数年，期间研究区域中人口迁入或迁出的发生概率很高。可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。

a) 健康数据呈现为计量数据时，采用散点图定量描述环境暴露水平与人群健康水平的关联性，若散点图呈现线性相关，采用相关系数和线性回归模型定量描述相关程度。相关和回归模型选取原则参照 4.2.2。

b) 健康数据呈现为计数数据时，采用 Logistic 回归模型定量描述环境暴露水平与人群健康水平，如疾病/症状发生率的相关性。也可以根据环境暴露的不同水平分组，描述每个组对应的人群健康指标（如阳性率/异常率/患病率等），采用 Spearman 相关系数描述其相关性，参照 4.2.2 选取卡方检验、连续性校正卡方检验或 Fisher 确切概率检验方法进行不同暴

露水平组间人群健康差异的显著性检验。

9.3 关联性判断

a) 散点图。散点的 Y 值随 X 值增加而上升, 正相关; 散点的 Y 值随 X 值增加而下降, 负相关; 散点的 X 与 Y 的数值增减无一定规律, 或 Y 值的变化不受 X 变化的影响, 零相关; 散点的 X 与 Y 的数值增减服从非直线规律, 非线性相关。

b) 相关系数。 $0 < r < 1$ 且假设检验统计量显著性 P 值 $< \alpha$ (一般取 0.05), 正相关; $-1 < r < 0$ 且假设检验统计量显著性 P 值 $< \alpha$, 负相关; $r = 0$ 且假设检验统计量显著性 P 值 $< \alpha$, 零相关。

c) 单变量直线回归、多变量直线回归、Logistic 回归。回归模型的假设检验统计量显著性 P 值 $< \alpha$, 单个自变量的回归系数 β_i 的假设检验统计量显著性 P 值 $< \alpha$, 则认为此自变量与因变量显著相关; 若回归系数的假设检验统计量显著性 P 值 $> \alpha$, 认为此自变量与因变量无显著相关性。一般情况下, 当两个变量之间的相关系数显著时, 单变量线性回归的回归系数也在同样的显著性水平下显著。

f) 卡方检验、连续性校正卡方检验、Fisher 确切概率检验。如果研究区的污染水平或人群特征污染物外暴露量显著高于对照区, 相应的人体内负荷水平、生理生化指标异常率、症状体征损害指标和疾病患病情况也显著高于对照区, 表现为假设检验统计量显著性 P 值 $< \alpha$, 则提示研究区内的环境污染与人群健康之间可能存在相关关系。

10 统计结果表达

通过统计表、统计图展示统计分析结果, 并给予必要的说明, 统计图表制作参见附录 A 和附录 B。数值修约依据 GB/T 8170 进行。

附录 A
(资料性附录)
统计表编制原则和结构

A.1 统计表编制原则

- a) 一张表一般只表达一个中心内容和一个主题。若内容过多，可分别制成若干张表。
- b) 主谓分明，层次清楚。统计表的内涵代表的是若干完整的文字语句，主谓语的位置要准确。定语部分放在标题内，主语放在表的左边作为横标目，谓语放在右边作为纵标目，横标目与纵标目交叉的格子放置数据，从左向右读，每一行便形成一个完整的句子。
- c) 数据表达规范，文字和线条尽量从简。

A.2 统计表结构

统计表可由标题、标目（包括横标目、纵标目）、线条、数字和备注 5 部分构成。

a) 标题。简明扼要地说明表的主要内容，包括时间、地点和研究内容，放在表的上方正中位置。如果有多张表格，标题前应加上标号，如“表 1-1”。如果表中所有数据指标的度量衡单位一致，可以将其放于括号内置于标题后面。

b) 标目。包括横标目和纵标目，有单位时需标明。横标目位于表的左侧，说明每一行数据的意义，纵标目位于表头右侧，说明每一列数据的意义。纵标目的总标目主要是对纵标目内容的概括，需要时设置。

c) 线条。通常采用“三线表”格式，即顶线、底线、纵标目下的横线。若某些标目或数据需要分层，可用短横线分隔。

d) 数字。用阿拉伯数字表示，同一指标小数点位数一致、位次对齐。表内不留空格，无数字用“-”表示，缺失数字用“.....”表示，以备注的形式说明。若数字是“0”，则填写“0”。

e) 备注。表中数据区需要插入文字或其他说明，可用“*”号标出，将说明文字写在表格的下面。

附录 B
(资料性附录)
统计图制作原则和结构

B.1 统计图制作原则

- a) 必须根据资料的性质、分析目的选用适当的统计图。
- b) 一个图通常只表达一个中心内容和一个主题，即一个统计指标。
- c) 编制图形应注意准确、美观，图线粗细适当，定点准确，不同事物用不同线条（实线、虚线、点线）或颜色表示，给人以清晰的印象。

B.2 统计图结构

统计图通常由标题、图域、标目、图例和刻度 5 个部分组成。

- a) 标题。简明扼要地说明资料的内容、时间和地点，位于图的下方正中位置并编号。
- b) 图域。即制图空间，除圆图外，一般用直角坐标系第一象限的位置表示图域，或者用长方形的框架表示。
- c) 标目。分为纵标目和横标目，表示纵轴和横轴数字刻度的意义，有度量衡单位时需标明。
- d) 图例。对图中不同颜色或图案代表的指标注释。图例通常放在图的右上角或图的正下方。
- e) 刻度。即纵轴与横轴上的坐标。刻度可在内侧或外侧，刻度数值按从小到大的顺序，纵轴由下向上，横轴由左向右。坐标原点必须从零开始。