

卫 生 学

第一单元 绪论

考试要点

一、卫生学及其发展简史

研究对象和目的；

②预防医学研究包括的内容。

二、我国社会主义初级阶段的卫生保健

我国新时期的卫生工作方针；

三级预防措施；

卫生保健策略。

重点、难点、疑点解析

一、卫生学及其发展简史

1. 卫生学的研究对象和目的

卫生学是在“预防为主”的卫生工作方针指导下，以辩证唯物主义的立场、观点和方法，研究外界环境因素与人体健康的关系，阐明环境因素影响人体健康的规律，提出改善和利用环境因素的卫生要求的理论依据和措施的原则，以达到预防疾病、促进健康、提高劳动能力的目的。

2. 预防医学研究包括的内容

预防医学针对人群中疾病发生发展的规律，运用基础医学、临床医学和环境卫生科学的理论、知识和技能、研究社会和自然环境中影响健康和造成疾病的主要因素；应用卫生统计学的原理和方法，探求病因和分析这些致病因素的作用规律，给予定量评价；并通过公共卫生措施进行预防，以达到保护健康和促进健康的目的。卫生学是预防医学的一门

学科，其主要研究内容包括：卫生学的形成和发展，特别是我国卫生保健事业的方针、政策和主要目标与措施；自然环境对人类健康影响的基本规律（生活环境、空气、水、地球物理因素与土壤、食物、生产环境、职业性有害因素）对人群健康的影响及预防疾病，增进健康的卫生要求、措施和原则；此外还包括人群健康研究（医学统计学方法、流行病学研究方法）的原理与方法。

二、我国社会主义初级阶段的卫生保健

1. 我国新时期的卫生工作方针

我国新时期的卫生工作的指导方针是：以农村为重点，预防为主，中西医并重，依靠科技和教育，动员全社会参与，为人民健康服务，为社会主义现代化建设服务。

这个指导方针的核心是为人民健康服务，为社会主义现代化建设服务。这是党和政府对卫生事业改革和发展的基本要求，也是卫生工作必须坚持的正确方向。

农村卫生、预防保健、发展中医药是我国卫生工作的战略重点。

2. 三级预防措施

公共卫生措施通过不同级别的预防在全体居民中实施，统称为三级预防。

第一级预防，又称病因预防，主要针对无病期，消除或控制各种危害健康的有害因素，并采取促进健康的各种措施，以防止健康人群发病。加强对病因的研究，减少对危险因素的接触，是一级预防的根本。它包括健康促进和健康保护两个方面。开展一级预防时常采取双向策略，即把对整个人群的普遍预

防和对高危人群的重点预防结合起来，二者相互补充可以提高效率。对病原或致病因素及条件明确的疾病，以第一级预防为主，如地方病、传染病和职业病等。

第二级预防，也称临床前期预防，即在疾病的临床前期（又称亚临床期）做好早期发现，早期诊断、早期治疗的“三早”预防工作。慢性病多是致病因素经过长期作用引起的，而且疾病的发展过程较长。诊期时病期愈早，预后愈好。因此早期发现和及早治疗各种癌前期病变也是第二级预防的重要一环。早期发现的具体办法有普查、筛检、定期健康检查、高危人群重点项目检查以及设立专科门诊等。对病原或致病因素及条件尚不完全明确的疾病，以第二级预防为重点，如慢性病、大部分肿瘤、脑血管疾病等。

第三级预防即临床预防。对已患某些病者，及时治疗，防止恶化。对慢性病患者则通过医学监护，减少疾病的不良作用，预防并发症和伤残；对已丧失劳动能力者或残废者通过康复医疗，使其能参加社会活动并延长寿命。

3 卫生保健策略

(1)全球卫生保健目标

世界卫生组织（WHO）提出的卫生保健目标是到2000年世界上所有的人在社会上、经济上都达到富有成效的健康水平卫生保健目标。这个目标的口号是：“2000年人人享有卫生保健”其含义是到2000年时，人们在社会、家庭、学校、工作场所从小到老一生中都可以方便地享有各种应该享有的卫生保健服务，以达到预防疾病、减少残疾的目的；特别是儿童、妇女应该得到适宜的预防、医疗保健服务；不同的国家、地区间应当均匀地分配卫生资源，发达国家的卫生经费应有适当的比例流入发展中国家；一个国家的卫生经费应有适当的比例用于初级卫生保健；人们懂得可以用自己的力量来摆脱疾病的束缚。

(2)初级卫生保健

初级卫生保健是实现“2000年人人享有卫生保健”目标的基本策略和途径。它是一种基本的卫生保健，依靠切实可行、学术上可靠而又能为社会接受的方式和技术，是社区的个人与家庭通过积极参与能普遍享受的，其费用也是国家和社区依靠自力更生和自觉精神能够负担得起的。它既是国家卫生系统的一个组成部分、功能中心和活动的焦点，也是社区及整个社会经济发展的一个组成部分。它是个人、家庭、群众与国家卫生系统接触的第一环节，能使卫生保健尽可能接近于人民居住地及工作场所；它还是卫生保健持续发展的起始一级。

初级卫生保健主要包括四个方面和九项内容。初级卫生保健的四个方面是：

促进健康：包括健康教育、保护环境、合理营养、饮用安全卫生水、改善卫生设施、开展体育锻炼、促进心理卫生、建立良好生活方式等；

预防保健：在研究社会人群健康和疾病的客观规律及它们和人群所处的内外环境，人类社会活动的相互关系的基础上，采取积极有效的措施，预防各种疾病的发生、发展和流行；

合理治疗及早发现疾病及时提供医疗服务和有效药品，避免疾病发展、恶化，促使早日好转痊愈，防止向慢性发展；

社区康复：对丧失正常功能或功能上有缺陷的残疾者，通过医学的、教育的、职业的和社会的措施，尽量恢复其功能，使他们重新获得生活、学习和参加社会活动的 ability。

初级卫生保健具体包括以下九项内容：

对当前主要卫生问题及其预防控制方法的健康教育；②改善食品供应和合理营养；提供足够的安全饮用水和基本的环境卫生设施；④开展妇幼保健和计划生育服务；⑤对主要传染病的免疫接种；⑥对地方病的预防控制；⑦对常见病的妥善处理；⑧提供基本药物；⑨预防控制非传染疾病和促进精神卫生。

我国初级卫生保健的规划目标及其具体

内容：我国农村初级卫生保健实施两个步骤、三个阶段。第一步，在 1995 年以前有半数县达到最低标准；第二步，到 2000 年再有半数县达到规定的标准。具体分为三个实施阶段：第一阶段为规划试点阶段；第二阶段为全面普及阶段；第三阶段为加速发展、全面达标阶段。初级卫生保健实施的关键要素有：组织领导、动员社区参与、做好组织保障和资源保障。

练习题及答案精讲

一、练习题

[A₁ 型题]

- 我国新时期的卫生工作方针是
 - 预防为主，依靠科技进步，卫生工作与群众运动相结合，面向工农兵，为人民健康服务
 - 以农村为重点，预防为主，中西医并重，依靠科技和教育，动员全社会参与，为人民健康服务，为社会主义现代化建设服务
 - 预防为主，依靠科技进步，动员全社会参与中西医并重，为人民健康服务，把卫生工作的重点放在农村
 - 预防为主，康复防残，动员全社会参与，中西医并重，发挥中草药优势，把卫生工作的重点放在农村
 - 预防为主，康复防残，促进健康，中西医并重，为人民健康服务，把卫生工作的重点放在农村
- 下列哪些是实现“2000 年人人享有卫生保健”目标的最关键措施
 - 发动整个社会广泛参与
 - 实施初级卫生保健
 - 增加政府对初级卫生保健的投入
 - 普及全民健康教育
 - 提高全民族素质
- 初级卫生保健又称
 - 农村卫生保健
 - 基层卫生保健
 - 低级卫生保健
 - 综合卫生保健
 - 二级卫生保健
- 世界卫生组织为达到“2000 年人人享有卫生保健”采取的基本措施是
 - 将个体防治转变为群体防治
 - 加强精神卫生
 - 消灭烈性传染病
 - 预防为主
 - 实施初级卫生保健
- 初级卫生保健的主要内容之一是
 - 延长人的寿命到 76.5 岁
 - 提供治疗的全部药物
 - 建立人类基因库
 - 地方病的预防和控制
 - 使人们脱离职业危害
- 有关“2000 年人人享有卫生保健”的含义下列选项哪个不对
 - 儿童、妇女应该得到适宜的预防、医疗保健服务
 - 人们懂得自己有能力来摆脱疾病的限制
 - 卫生资源能够在不同的国家和地区之间被均匀地分配
 - 人们从婴幼儿到老年都能健康地度过一生
 - 人们患的疾病都可以被治愈
- “适宜的技术”是指
 - 复杂的方法和技能、先进的设备及使用它们的人的总称
 - 科学合乎实际，方法易于掌握，群众乐于接受并负担得起的技术
 - 在实践中较易操作，但只适合个别情况
 - 费用低廉 群众接受 但效果一般
 - 技术先进，设备精良，疗效较好的技术，但价格并不昂贵
- 当今我国预防医学研究的对象是
 - 致病因素的高危人群

- B.对疾病抵抗力弱的人群
 C.有职业致病因素接触的人群
 D.低收入、无医疗支付能力的人群
 E.全社会的人群
- 9.初级卫生保健任务主要内容包括下列哪四方面内容
 A.合理治疗、综合环境保护、健康促进和慢性病控制
 B.健康促进、合理营养、环境保护和预防保健
 C.预防保健、慢性病控制、社区康复防残和中西药合理使用
 D.健康促进、合理营养、环境保护和预防保健
 E.资源合理配置、预防保健、计划生育和社区康复
- 10.下列关于一级预防措施中不正确的是
 A.免疫接种
 B.控制和消除环境致病危险因素
 C.加强体育锻炼,增强对疾病的抵抗力
 D.定期体格检查,对疾病早发现、早诊断、早治疗
 E.怀孕期和儿童期的卫生保健
- 11.初级卫生保健的内容不包括
 A.预防疾病
 B.增进健康
 C.康复防残
 D.增加对医疗设备的投资
 E.及时治疗
- 12.一级预防的重点疾病是
 A.心、脑血管病
 B.呼吸道疾病
 C.传染病、职业病
 D.传染病、运动系统疾病
 E.职业病、地方病、消化道疾病
- 13.下列有关亚临床期预防的选项中错误的是
 A.针对发病早期,早发现、早诊断、早治疗
 B.消除和控制各种危害健康的因素
 C.控制疾病,防止残疾,防止疾病的复发
 D.防止疾病发展演变成慢性病
 E.对接触有害因素的人群进行定期体格检查
- 14.三级预防的内容不包括
 A.对已患某些疾病者及时治疗,防止恶化
 B.对慢性病患者通过医学监护,减少疾病的不良作用,预防并发症和伤残
 C.对已丧失劳动能力者通过康复医疗,使之能参加社会活动
 D.消除或控制各种有害因素
 E.对残疾者通过康复医疗,使之能参加社会活动并延长寿命
- [A₂型题]
- 15.初级卫生保健是实现“2000年人人享有卫生保健”目标的基本策略和途径。我国的初级卫生保健实施工作是分两个步骤、三个阶段进行的。其中初级卫生保健规划试点阶段的目标已经基本达到,各个省市也总结了许多成功的经验。但是,由于我国现实条件的制约,初级卫生保健本身又需要全社会的长期参与。因此,我国全面实施初级卫生保健仍面临着挑战。下列哪个选项不是顺利实施初级卫生保健的关键
 A.各级组织参与
 B.资源有效保障
 C.组织领导重视
 D.世界卫生组织指导
 E.社区积极参与
- [B₁型题]
- A.对慢性病患者则通过医学监护,减少疾病的不良作用,预防并发症和伤残
 B.合作医疗、医疗保险、单项医疗保险
 C.促进行为、环境向有利于人们心身健康的方向转变,并做到早发现、早诊断、早预防和残疾康复
 D.提供足够的安全饮用水和基本的环境卫生设施
 E.开展健康教育,使人人讲究卫生,锻炼

身体，增强体质，提高机体的抗病能力；有系统有组织的进行预防接种，提高人群免疫水平；做好婚前卫生工作

- 16. 属于初级卫生保健内容的是
- 17. 以上属于三级预防的是
- 18. 属于一级预防的是
- 19. 属于农村医疗保健制度内容的是

二、答案精讲

- 1. B 2. B 3. B 4. E 5. D
- 6. E 7. B 8. E 9. D 10. D
- 11. D 12. C 13. C 14. D 15. D
- 16. D 17. A 18. E 19. B

2 初级卫生保健是“人人享有卫生保健”目标得以实现的最关键的措施，也是其基本策略和途径。

5. 初级卫生保健共有 9 项要点，对地方

病的预防和控制是其中第 5 点。

8. 卫生学的主要研究内容包括：卫生学的形成和发展；自然环境、生活环境、生产环境对人类健康影响的基本规律；人群健康研究的原理与方法，其对象为全社会人群。

10. 一级预防，又称病因预防，主要针对无病期，消除或控制各种危害健康的有害因素。早发现、早诊断、早治疗属于二级预防。

13. 第二级预防，也称临床前期预防，因临床前期又称为亚临床期，所以也称亚临床期预防。控制疾病，防止残疾，防止疾病的复发属于三级预防。

14. 第三级预防即临床预防，是对已患某些病者，及时治疗，防止恶化。消除或控制各种有害因素属于一级预防。

第二单元 人群健康的研究方法

考试要点

一、医学统计方法

医学统计在医学工作中的地位和作用；

统计工作的步骤及资料分类；
统计分析中常用的几个基本概念；
统计研究设计的分类及特点；
定量资料的统计描述；

⑥正态分布及其应用；

⑦定量资料的统计推断；

⑧分类资料常用的相对数指标及其意义；

⑨分类资料假设检验(X^2 检验)；

⑩死亡统计常用指标的意义；

⑪疾病统计常用指标的意义；

⑫列统计表的基本原则和要求；

⑬绘制统计图的基本原则和要求。

二、流行病学研究方法

描述性研究的概念、人群调查(现患调查)的含义、特点及临床医学上的应用；

诊断试验和筛检试验的概念与应用；

病例对照调查的概念、病例对照设计的要求及其在临床医学上的应用；

定群研究的概念、种类及其在临床医学上的应用。

三、实验性研究

实验性研究的概念；

②临床实验设计原则；

在临床医学上的应用。

重点、难点、疑点解析

一、医学统计方法

1. 医学统计学的概念和其在医学工作中的地位

统计学是研究数据的收集、整理、分析与推断的一门科学。医学统计学是应用概率论和数理统计的基本原理和方法,结合医学实际,阐述统计设计的基本原理和步骤,研究资料或信息的收集、整理与分析的一门学科。随着医学科学的迅速发展以及医学模式的转化,医学科学不仅限于生物领域,还涉及到社会、心理、行为等诸多方面的因素,这些对人体健康有潜在影响的因素都需要人们去研究,并认识其内在的联系,从而克服和消除一些不利因素对人们健康的影响。

医学统计学的研究内容包括医学统计的基本理论和方法;医学科研的设计及其数据处理的基本理论和方法,例如医学科研中的现场调查设计和实验设计;科学、准确、及时地收集整理统计资料的方法,以及科学地进行统计描述、估计、比较和预测等的方法;居民健康统计,如出生、疾病及死亡统计等;卫生资源分布和卫生业务统计,如反映卫生机构、人员与床位等卫生事业的基本统计等。

2. 统计工作的基本步骤及资料分类

统计工作包括四个步骤。分别为设计、收集资料、整理资料和分析资料。

设计:设计是后续步骤的依据,是最关键的一环。首先要明确研究目的,要对被研究的事物有一定的了解。为此,需要广泛查阅文献,了解实际情况,作预试验、预调查等。

设计的内容包括资料收集、整理和分析全过程总的设想和安排。例如，什么是研究目的和假说？什么是观察对象和观察单位？需要收集哪些原始资料？用什么方式和方法取得这些资料？怎样对取得的原始资料做进一步的整理汇总和计算统计指标？如何控制误差？预期会取得什么效果？需要多少经费等。凡此种种，都要结合实际，周密安排，做出科学详尽的计划。

收集资料：按照研究设计的规定，及时取得准确完整的原始数据。只有原始数据准确可靠，才能得出可信的结论。医学统计资料主要有三个方面的来源：统计报表；经常性工作记录，例如医疗卫生工作记录和报告卡片等；专题调查和试验。

整理资料：按照研究设计的规定，使原始数据系统化、条理化的过程，以便进一步计算指标，作出分析结论。数据整理的过程一般可分为三步：核对校正原始数据；根据研究数据的特征分组；根据分组情况，将原始数据分别归入相应的组，并做简单的总结。

分析资料：按照研究设计的有关规定，计算有关指标，反映数据的综合特征，阐明事物的内在联系和规律。分析资料包括：统计描述和统计推断。分析资料的过程也可分成三步：计算相关指标；绘制适当的统计图；进行统计处理，并做出恰当的结论。

统计资料的分类。统计资料按其特点，可分为不同的类型。不同的统计资料，其处理方法也不同。

计量资料：通过度量衡的方法，测定每一个观察单位的某项研究指标的量的大小，而得到的一系列数据资料，即观察值，就称为计量资料。计量资料通常有单位，如身高，单位为 cm；体重，单位为 kg；血压，单位为 Kpa 等。对于大多数计量资料来说，当观察单位足够多时，这些观察值可以充满某一个数值范围，故属于连续性资料。

计数资料：将全体观察单位，按照某种性质或特征分组，再分别清点各组中观察单位

的个数，这样得到的数据资料，就是计数资料。计数资料没有度量衡单位，另外计数资料都是整数，没有小数点，是一种间断性资料。计数资料又分为二项分类资料（如治愈和未治愈）和多项分类资料（如人的血型可分为 A 型、B 型、AB 型和 O 型）。

等级资料：等级资料是介于计数资料和计量资料之间的一种资料，通过半定量的方法测得。具体来说，将全体观察单位按照某种性质的不同程度分成若干组，再分别清点各组中观察单位的个数，这样得到的数据资料就称为等级资料。如人的镜下血尿可分为 1—4 个“+”。与计量资料相比，等级资料的每一个观察单位没有确切值；与计数资料相比，等级资料的分组有程度差别，各组按性质的程度大小顺序排列。因此等级资料被称为半定量资料。

3. 统计分析中常用的几个基本概念

(1) 变量

无论用何种方式收集资料，都要首先确定观察单位，也叫做个体，它可以是一个人、一个家庭、一个社区、一个样品等；然后对每个观察单位的某项特征进行测量和观察。观察单位的特征称为变量。

(2) 变异

性质相同的事物，称为同质的事物。即使同质的事物，其中各个体（即观察单位）也不是完全相同的。各个体之间的差异，称为变异。如同年龄的女大学生的身高有高有低，各不相同，这就是个体变异。同质观察单位的个体变异，尽管是有各种内外环境作用表现的偶然性，但是揭示了事物的重要特征。统计的任务就是要在同质的基础上，通过对个体变异的研究透过偶然现象，揭示同质事物的本质特征和必然规律。因此变异是统计学的基础，统计学是研究变异的科学。

(3) 总体

理论上讲，总体是根据研究目的所决定的同质研究对象的全体。实际上，当研究有具体而明确的研究指标时，总体是指性质相

同的符合研究要求的所有观察单位的该项变量值的全体集合。当研究没有明确而具体的研究指标时，其研究总体就只能是性质相同的符合研究要求的所有观察单位了。研究总体可分为有限总体和无限总体。只包括有限数量观察单位的总体，称为有限总体。数量巨大无比，或数量是无法搞清楚的，其总体就属于无限总体。

(4) 样本和随机抽样

样本是指从总体中随机抽取的部分观察单位其实测值的集合。我们通常对样本进行研究，用样本的信息来推断总体的特征。为了保证样本研究的准确性，对于总体，样本必须具有一定的代表性。样本要有代表性，挑选样本时必须遵循随机化原则，就是使总体中每一个个体都有相等的被选机会作为样本。样本研究通常就是抽样研究。常用的选择样本的方法有概率抽样和非概率抽样。概率抽样就是使总体中每一个个体都有一个不为零的被选机会进入样本。概率抽样分为两大类：等概率抽样和不等概率抽样。前者每一个个体被选机会相等，后者每一个个体的被选机会不等。等概率抽样也称随机抽样，它能避免抽样过程中人为因素的影响，保证样本的代表性。对于不等概率抽样，可以采取某些加权的方法对不相等的概率做调整。

(5) 系统误差

系统误差是指在数据收集过程中，由于仪器不准确、标准不规范等原因，使得数据倾向性偏高或偏低。这种误差具有累加性，会因测量次数的增加而增加。

(6) 随机误差

随机误差是指在数据收集过程中，由于一些非人为的偶然因素，使得测量结果或大或小。这种误差，会因为测量次数的增加，偏大的测量结果和偏小的测量结果相中和，而变得较小甚至消除。

(7) 抽样误差

抽样误差是指样本指标和总体指标的差异。只要是抽样研究，就必然存在抽样误差，

其中有两方面的原因：①研究对象之间存在个体差异即变异。这是抽样误差存在的根本原因。②抽样研究只研究一部分研究对象，部分不可能等于全部，也即抽样是抽样误差存在的直接原因。抽样误差属于随机误差，但是随机误差不可能被消除，只能被控制。方法有：改进抽样方法，使样本更具有代表性；增加样本量；选择变异程度小的指标。

(8) 频率与概率

数学上，频率是发生某现象的观察单位数占可能发生该现象的观察单位总数的比值；概率是频率的极限，说明某事件 A 发生的可能性的概率大小。一般用 P 表示。概率在 0 到 1 之间。肯定不发生的事件称为不可能事件 记为 $P(A)=0$ 。肯定发生的事件称为必然事件 记为 $P(A)=1$ 。发生概率介于 0 到 1 之间的事件既可能发生，也可能不发生，这类事件称为随机事件或偶然事件。习惯上常常将概率值小于 0.05 或 0.01 的事件称为小概率事件；概率值大于 0.05 或 0.01 的事件，称为大概率事件。

(9) 参数和统计量

统计学中总体的指标统称为参数，样本的指标称为统计量。例如某地所有成年女子的脉搏数是该总体的参数；从该地抽取 2000 名成年女子进行测量，所得的样本平均数是该样本的统计量。

4. 统计研究设计的分类及特点

(1) 调查设计

调查是客观地反映事物的实际情况，未加任何干预措施，为今后决策和进一步深入研究提供依据。

调查研究的特点 在调查研究中 研究的因素是客观存在的；研究者通常不能用随机化分组的方法来平衡研究中的混杂因素。以上两点决定了在调查研究中，研究者只能对研究对象做被动观察。

调查方法 按调查的范围 可分为普查和非全面调查，后者又以抽样调查和典型调查最为常用，分别介绍如下：

普查，又称全面调查，是将组成总体的所有观察单位全部加以调查，如我国的四次人口普查。

典型调查，亦称案例调查。即在对事物做全面分析的基础上，有目的地选定典型的人、典型的单位进行调查。典型调查可对总体特征做经验性推论，但不属于统计推断的范畴。

抽样调查，是从总体中随机抽取一定数量的观察单位组成样本，然后用样本信息来推测总体特征。抽样调查在实际工作中应用最多，应用中应针对观察对象的不同特点采用不同的抽样方法。

基本抽样方法。

单纯随机抽样：即先将调查总体的全部观察单位编号，再用随机数字表或抽签等方法随机抽取部分观察单位组成样本。

系统抽样：又称等距抽样或机械抽样。即先将总体的观察单位按某一顺序号等分成 n 个部分，再从第一部分随机抽第 k 号观察单位，依次用相等间隔，机械地从每一部分各抽一个观察单位组成样本。

整群抽样：先将总体划分为 K 个“群”组（如 K 个地区），每个群包括若干观察单位；再随机抽取 k 个“群”并将抽取的各个群的全部观察单位组成样本。

分层抽样：又称分类抽样。即先按影响观察值变异较大的某种特征，将总体分为若干个类型或组别（统计上称为“层”）再从每一层内随机抽取一定数量的观察单位，合起来组成样本。

调查设计基本步骤。

明确调查目的和指标：首先应根据防治工作的需要明确调查目的。从统计学上说，即了解参数和研究变量间的相关联系。指标要精选，要重点突出。

确定观察对象和观察单位：即划清调查总体的同质范围。

确定调查方法和收集原始资料的方式：调查方法如上所述。收集资料的方法有直接

观察法和采访法。

确定调查项目和调查表：根据调查指标确定对每个观察单位的调查项目，包括分析项目和备查项目。分析项目是直接用于计算的调查指标，以及分析时排除混杂因素影响所必须的内容。备查项目是为了保证分析项目填写的完整、正确，便于对其核查、补填和更正而设置的，通常不直接用于分析。

制定调查的组织计划：包括组织领导，宣传动员，时间进度，调查员培训，任务分工与联系，经费预算，调查表格和宣传资料的准备以及调查资料的检查（完整性、正确性）制度等内容。在大范围的正式调查之前，应先作小范围的试调查，以便检验调查表设计，并作必要的修改。

(2) 实验设计

实验研究是指将一组随机抽取的实验对象随机分配到两种或多种处理组中，观察比较不同处理因素的效应或结果的研究。如将病人随机分配到实验组和对照组中，以比较一种新疗法的疗效。

实验研究的特点：研究者能人为设置处理因素；受试对象接受何种处理因素或水平是由随机分配而定的。

由于以上特点，实验研究和调查研究相比，能使比较组间具有较好的均衡性，即研究者使非处理因素对不同处理组的影响相同，组间具有可比性，在此条件下来探讨和评价处理因素的作用。

实验研究的分类：动物实验：许多药物、营养保健品实验首先在动物身上进行，取得肯定的结果后才能逐步过渡到人体。动物实验持续的时间一般在一年以内。

临床试验：临床试验通常局限在患者人群中，持续的时间较长，目的是了解某种治疗措施的疗效。

社区干预试验：社区干预试验往往在某个地区的所有人群中，持续的时间一般较长，目的是通过干扰某些危险因素或施加某些保护性措施，然后了解它们在人群中产

生的预防效果。由于社区干预试验中难以贯彻随机分配，故又被称为半试验研究。

实验设计的基本要素。

处理因素：实验研究的特点是研究者能人为地设置处理因素。确定处理因素应当找出实验中的主要因素；找出非处理因素；标准化。

受试对象：受试对象可分为人和动物。对受试对象的选择非常重要，它对实验结果有重大影响。

实验效应：包括实验结果指标的选择和指标的观察两部分。实验效应的观察指标应当客观性强、灵敏度高和精确性好。实验效应的观察应避免带有偏性。

实验设计的基本原则。

对照的原则：对照是比较的基础，要确定处理因素与实验效应的关系，没有对照不能说明任何问题。对照的形式有很多，常用的有空白对照、安慰剂对照、实验对照、标准对照、历史对照、自身对照等。

重复的原则：重复是指各处理组及对照组的例数或实验次数要有一定的数量。如果例数太少，有可能把个别情况错误地认为普遍情况，把偶然性或巧合的现象当做必然的规律现象，甚至于将实验结果错误地推论到群体。但例数太多又会增加严格控制实验条件的困难造成不必要的浪费。为此，应该在保证实验结果具有一定的可靠性的条件下，确定最少的样本例数以节约人力和经费。

随机化的原则：在实验研究中，不仅要求有对照，还要求各组间除了处理因素外，其他可能产生混杂效应的非处理因素尽可能保持一致，即均衡性好。

常用的几种实验设计方法。

完全随机设计：完全随机设计仅涉及一个处理因素（但可以为多水平），故又称单因素设计。它是将受试对象按随机化的原则分配到各个处理组中，观察实验效应。

配对设计：是将受试对象按一定的条件配成对子，再随机分配每对中的两个受试对

象到不同的处理组中。配对因素是影响实验效应的主要非处理因素。

配伍组设计：亦称随机区组设计。它是配对设计的扩大，是将几个受试对象按一定条件划分成配伍组或区组，再将每一配伍组的各受试者随机分配到各处理组中去。每个配伍组的例数等于处理组个数。配伍的条件同配对设计。

5. 计量资料的统计描述

(1) 计量资料的频数表

为了解计量资料的分布规律，当观察值个数较多时，可将观察值及其相应的频数列编制成频数分布表，简称频数表。频数表的编制步骤如下：

求全距：找出观察值中的最大值和最小值 求出二者的差值（即全距或极差）

②确定组距：根据观察单位个数 n 大小酌情定组数，一般把 n 个变量分为 10—15 组，以揭示变量的分布规律。组距 = 全距 / 组数 为计算方便 近似取整）

划分组段：每个组段的起点称为组下限，终点称为组上限。第一组段应包括变量最小值，最后组段应包括变量最大值。

划记：将各变量值划记（正或 + + +）归入相应组段 合计变量个数 称频数 构成频数表。某地于 1997 年抽取 100 名女大学生进行身高测量，编制频数表并计算如下：

某地 1997 年 100 名 19 岁女大学生
身高均数的计算表（加权法）

身高组段	频数 f	组中值 X	fX
(1)	(2)	(3)	(4) = (2) × (3)
154—	2	155	310
156—	4	157	628
158—	11	159	1749
160—	13	16	2093
162—	22	163	3586
164—	19	165	3135
166—	15	167	2505
168—	9	169	1521
170—	4	171	684
172—174	1	173	173
合计	100 ($\sum f$)		16384 ($\sum fX$)

从频数表可以粗略地看出频数分布的两个重要特征：集中趋势和离散趋势。如上表，100名 19 岁女大学生身高向中央部分（即 162—164cm 组段）集中，且身高在 162—164cm 及接近 162—164cm 者居多，这反映的是集中趋势；从中央到两侧（即由中等身高到较矮或较高）频数分布逐渐减少，这反映的是离散趋势。平均指标应根据变量的不同分布类型选用。

(2) 集中趋势的统计描述

频数的分布有集中趋势和离散趋势。变量值分布的集中趋势用平均指标来描述，常称为平均数。平均数是反映同质的一群变量值的平均水平或集中位置，常用的平均数有算术均数、几何均数和中位数。

算术均数：算术均数简称均数，总体均数用 μ 表示，样本均数用 \bar{X} 表示。计算方法有：

直接法由 n 个变量 X_1, X_2, \dots, X_n 求均数 \bar{X} 的公式为： $\bar{X} = (1/n) \sum X$ 。式中 \sum 是求和符号， n 是样本含量。

加权法：当观察值个数较多时，用上述公式计算均数比较麻烦，为计算方便，可先将各观察值分组列成频数表，用加权法求均数。组中值 $X_i = (\text{本组段下限} + \text{上组段下限}) / 2$ ， f_i 为 X_i 的权数，公式为：

$$\bar{X} = (\sum fX) / (\sum f)$$

本例中：

$$\bar{X} = 16384 / 100 = 163.84 \text{cm}$$

均数适用于对称分布资料，这时均数位于分布的中心，最能反映分布的集中趋势。对于偏态分布资料，均数则不能很好地反映分布的集中趋势，可用几何均数或中位数等描述。

②几何均数：某些资料如抗体的滴度、细菌计数等，其频数分布呈明显偏态，变量值变化呈倍数关系，取对数后服从正态分布，称为对数正态分布资料。几何均数用 G 来表示。

直接法公式为：

$$G = \lg^{-1} [(\sum \lg X) / n]$$

加权法公式为：

$$G = \lg^{-1} [(\sum f \cdot \lg X) / \sum f]$$

几何均数常用于呈等比数列的资料，即变量呈倍数关系，尤其适合呈对数正态分布的资料。计算几何均数时应注意：观察值中不能有 0；观察值不能同时有正值和负值，若全是负值，计算时可把符号去掉，得出结果后再加上负号。

③中位数和百分位数：把 n 变量从小到大排列，位于中间位置的变量值称为中位数，用 M 表示。第 X 百分位次对应的变量值称为第 X 百分位数，用 P_x 表示。中位数为 P_{50} 。计算方法有：

直接法：将观察值按大小排列，

$$n \text{ 为奇数时, } M = X_{(n+1)/2}$$

$$n \text{ 为偶数时, } M = [X_{(n/2)}, X_{(n/2+1)}]$$

频数表法：列出频数表，求百分位数 P_x 和中位数 M 。

$$P_x = L + (I/f_x) \cdot (n \cdot X\% - \sum f_L)$$

式中 P_x 第 X 百分位数；

L : P_x 所在组段的下限；

f_x : P_x 所在组段的频数；

I : 该组段的组距；

$\sum f_L$ 小于 L 各组段的累计频数。

对于中位数 $M = f_{50}$ 正态分布的资料中，总体均数等于中位数；对数正态分布资料中，及格均数等于中位数。

中位数适宜情况：不宜用于几何均数的偏态分布资料，分布的一端或两端无确定数值的资料及资料分布不清楚时。

(3) 离散趋势的统计描述

变量值分布的离散趋势用变异指标来描述，常用的变异指标有：全距、四分位数间距、方差、标准差和变异系数等。

全距亦称极差，用 R 表示。全距是一组观察值中最大值与最小值之差。它反映了个体差异的范围，全距大，说明变异大。

四分位数间距：用 Q 表示，为上四分

位数 Q_0 (即 P_{75}) 与下四分位数 Q_L (即 P_{25}) 之差。这样, P_{25} 、 M 与 P_{75} 将数据四等分, P_{25} 与 P_{75} 之间包括了全部观察值的一半。常用于描述偏态分布以及分布的一端或两端无确切数值资料的离散程度。

方差: 方差既考虑了离均差 $\sum (X - \mu)$, 又考虑了观察值个数 N 。

总体方差:

$$\sigma^2 = \frac{\sum (X - \mu)^2}{N}$$

样本方差:

$$S^2 = \frac{\sum (X - \bar{X})^2}{n - 1}$$

虽然方差比较全面地考虑了观察值的变异情况, 但它使原来的度量单位都变成了平方单位而使用不方便, 因此, 常用标准差。

标准差: 标准差是方差的平方根, 公式为:

总体标准差:

$$\sigma = \frac{\sqrt{\sum (X - \mu)^2}}{\sqrt{N}}$$

样本标准差:

$$S = \frac{\sqrt{\sum (X - \bar{X})^2}}{\sqrt{n - 1}}$$

当相同变量值个数较多时, 可以按照加权法计算均数法计算:

$$S = \frac{\sqrt{\sum fX^2 - (\sum fX)^2/n}}{\sqrt{n - 1}}$$

标准差全面地反映了一组观察值的变异程度, 两组观察值在单位相同, 均数相近的条件下, 标准差大说明这组观察值的变异大, 即观察值围绕均数分布较离散, 均数的代表性较差; 反之, 说明各观察值多集中在均数周围, 均数的代表性较好。

标准差可用于: 表示观察值的变异程度; 计算变异系数; 计算标准误; 描述正态分布特征和估计正常值范围。

变异系数: 常用于比较度量单位不同, 或均数相差悬殊的两组或多组资料的变异度。变异系数记为 CV 亦称离散系数, 计算公式为:

$$CV = (S/X) \times 100\%$$

CV 大, 变异大 CV 小, 变异小。

6. 正态分布及其应用

(1) 正态分布的概念和特征

正态分布又称为 Gauss 分布, 其条件为频数分布以均数为中心, 左右两侧基本对称, 均数两侧频数较多, 距均数越远, 频数越小, 即以钟型曲线下的面积表示频数分布, 是医学、生物学中常用的总体分布。均数为 μ 标准差为 σ 记作 $N(\mu, \sigma)$ 为了使用方便, 常对正态分布进行变量变换。正态分布 $\mu = (X - \mu)/\sigma$ 如果 $\mu = 0, \sigma = 1$ 时, 称为标准正态分布, 亦称为 u 分布。

(2) 正态分布曲线下面积的分布规律

正态分布曲线下横轴上的总面积为 100% 或 1。因此, 对于正态或近似正态分布的资料, 只要得出均数和标准差, 就可以对其频数分布作出概率估计。常用的两个区间: $\mu \pm 1.96\sigma$ 和 $\mu \pm 2.58\sigma$ 的区间面积分别占总面积或总观察例数的 95% 和 99%。

(3) 正态分布的应用

某些医学现象, 如同质群体的身高、体重、红细胞数、血红蛋白量等, 以及实验中的随机误差, 均呈现为正态或近似正态分布; 有些资料虽然为偏态分布, 但经数据变换后可成为正态或近似正态分布, 故可按正态分布处理。

制定医学参考值范围: 医学正常值又称参考值指正常人(或动物)的形态、功能和代谢产物等各种生理、生化常数。由于个体指标的变异使其有一定的波动范围, 因此, 也称为正常值或参考值范围, 受时间、地点条件影响而异。制定正常值范围时: 首先要确定一批样本含量足够大的“正常人”; 需根据实际用途确定单侧或双侧界值; 根据研究目的和使用要求选定适当的百分界值, 习惯的界值常取正常人变量值的 95% 根据资料的分布特点, 选用恰当的计算方法。

正态分布是许多统计方法的理论基础。

7. 计量资料的统计推断

(1) 均数的抽样误差与标准误

均数的抽样误差：由于总体中各个体存在着差异，即使严格遵守随机抽样的原则，从同一总体中抽取相同样本量的不同样本，所得到的样本均数可能也会不同。这种由于抽样引起的样本统计量与总体参数之间的差异称为抽样误差。由抽样引起的样本均数与总体均数之差 $(\bar{x} - \mu)$ 就称为均数的抽样误差。

标准误：样本均数的标准差叫做标准误。标准误是表示抽样误差大小的指标。总体的标准差为：

$$\sigma_{\bar{x}} = \sigma / \sqrt{n}$$

σ 为总体标准差； n 为样本例数； $\sigma_{\bar{x}}$ 为总体标准误。由于总体标准差的大小通常是无法知道的，可用样本的标准差 S 代替 σ 则：

$$S_{\bar{x}} = S / \sqrt{n}$$

S 为样本标准差； n 为样本例数； $S_{\bar{x}}$ 为样本标准误。

标准误越小，说明抽样误差越小，用样本均数估计总体均数的可靠性越大。标准差 S 一定时， n 越大，标准误越小。所以可以通过增加样本量来减小抽样误差。

(2) 总体均数可信区间的估计

由样本统计量推算总体均数有两个重要方面：区间估计和假设检验。仅由样本均数估计总体均数称为点估计；总体均数的区间估计是按一定概率（可信度）估计总体均数在那个区间（范围），称为总体均数的可信区间。可信度用 $1 - \alpha$ 表示，常用的可信度为 95% 和 99%。95% 的可信区间相当于每 100 个由含量相同的样本总体计算得到的可信区间，平均有 95 个可信区间包括总体均数，5 个不包括。表示为 $\bar{x} \pm t_{0.05} S_{\bar{x}}$

区间估计的准确度指区间包含总体均数的可能性大小，用 $1 - \alpha$ 来衡量。99% 的可信区间的准确度比 95% 的可信区间的准确度要高。精确度是指区间范围的宽窄，范围越宽精确度越差。所以，99% 的可信区间其

精确度差于 95% 的可信区间。准确度和精确度是互相矛盾的。在 α 确定的情况下，增加样本量可以在不影响准确度的情况下提高精确度，因为当 α 确定后， n 越大， $S_{\bar{x}} (S / \sqrt{n})$ 越小，在同一 μ_0 或 t_0 （即同一准确度）的情况下，则区间范围越小，即精确度越高。

(3) t 分布与 t 分布的特征

以 0 为中心，左右对称的单峰分布；
t 分布曲线是一簇曲线，其形态变化与自由度 ν 的大小有关。自由度 ν 越小，则 t 值越分散，曲线越低平；自由度 ν 逐渐增大时，t 分布逐渐逼近 u 分布（标准正态分布）；当 $\nu = \infty$ 时，t 分布即为 u 分布。

由于 t 分布是一簇曲线，故 t 分布曲线下面积为 95% 或 99% 的界值不是一个常量，它随自由度的大小变化。为了便于应用，编制了 t 界值表。

(4) 假设检验的步骤

建立假设：建立两种假设，一种是“检验假设”或称“无效假设”用 H_0 表示， H_0 假设是假设两样本所代表的总体均数相同；另一种是“备择假设”用 H_1 表示， H_1 假设两样本的均数不同，即两样本来自不同的总体。 H_0 与 H_1 是相反的，当计算出的结果使我们拒绝 H_0 时就接受 H_1 ；不能拒绝 H_0 时就不接受 H_1 。

确定显著性水平：显著性水平是我们用来区分大概率和小概率事件的标准，是人为规定的。当某事件发生的概率不大于 α 时，则认为该事件为小概率事件，是不太可能发生的事件。通常 α 取 0.05 (5%) 或 0.01 (1%)。

计算统计量：根据资料类型与比较的目的，选择适当的公式计算统计量，比如计算 t 值或 u 值。

确定概率值（概率用 P 表示）：查 t 值或 u 值表，将计算得到的 t 值或 u 值与它们进行比较，得到 P 值与 α 大小的关系。根据 t 分布或 u 分布的规律，如果 $|t| > t_{\alpha, \nu}$ ($|u| >$

u_α 测 $P < \alpha$ 如果 $|t| < t_{\alpha, \nu}$ ($|u| < u_\alpha$ 测 $P > \alpha$ 。

作出推断结论：如果 $P > \alpha$ 我们认为在检验假设 H_0 成立的条件下，得到等于或大于现有统计量 u 值或 t 值的可能性大于 α ，不属于小概率事件，则不拒绝 H_0 。差别无统计学意义，结果是不认为两总体均数不相等。如果 $P < \alpha$ ，我们认为在检验假设 H_0 成立的条件下，得到等于或大于现有统计量 u 值或 t 值的可能性小于或等于 α ，可判断为小概率事件，则拒绝 H_0 ，接受 H_1 ，差别有统计学意义，结论是两总体均数不相等。

(5) u 检验和 t 检验

u 检验和 t 检验可用于样本均数与总体均数的比较。理论上要求样本来自正态分布总体。但在实际应用时，只要样本例数 n 较大或 n 小但总体标准差 σ 已知时，就可以应用 u 检验； n 小且总体标准差 σ 未知时，可应用 t 检验，但要求样本来自正态分布总体。两样本均数比较时还要求两总体方差相等。

样本均数与总体均数的比较：比较目的：比较一个样本的均数所代表的总体均数和已知的总体均数 μ_0 是否不同。

u 检验：用于 σ 已知或 σ 未知但 n 足够大。

$$\text{公式：} u = \frac{\bar{x} - \mu}{\sigma_x}$$

t 检验：用于 σ 未知且 n 较小时。

$$\text{公式：} t = \frac{\bar{x} - \mu_0}{S_x}$$

自由度 $\nu = n - 1$

② 配对资料的比较：配对设计主要有四种情况：同一受试对象处理前后的数据；同一受试对象两个部位的数据；同一样品用两种方法检验的结果；配对的两个受试对象分别接受两种处理后的数据。第一种情况的目的是推断其处理有无作用；后三种情况的目的是推断两种处理或方法的结果有无差别。

$$\text{公式：} t = \frac{|\bar{d} - 0|}{S_d} = \frac{|\bar{d}|}{S_d}$$

自由度 $\nu =$ 对子数 $- 1$ 。解决配对资料的检验问题时，首先要计算两组配对观察值

的差值 d 以及其均数 \bar{d} 。如果两组资料无差别理论上差值 d 的总体均数应为 0。所以可将这类问题看成样本均数 \bar{d} 与总体均数 0 的比较。

完全随机设计的两样本均数的比较。目的是推断两样本各自代表的总体均数 μ_1 和 μ_2 是否相等。根据样本含量 n 的大小和已知条件，分别进行 u 检验和 t 检验。

u 检验：用于两样本含量 n_1 和 n_2 均足够大时，一般大于 50。

公式：

$$u = \frac{|\bar{x}_1 - \bar{x}_2|}{\sqrt{S_1^2/n_1 + S_2^2/n_2}}$$

t 检验：可用于两样本含量 n_1 和 n_2 均较小，且要求两总体的方差相等，即方差齐。若被检验的两总体方差相差显著则需用 t' 检验。

$$\text{公式：} t = \frac{|\bar{x}_1 - \bar{x}_2|}{S_{\bar{x}_1 - \bar{x}_2}}$$

式中两样本均数之差的标准误：

$$S_{\bar{x}_1 - \bar{x}_2} = \sqrt{S_c^2(1/n_1 + 1/n_2)}$$

其中合并标准差的平方：

$$S_c^2 = \frac{S_1^2(n_1 - 1) + S_2^2(n_2 - 1)}{n_1 + n_2 - 2}$$

自由度为 $n_1 + n_2 - 2$ 。

(6) 均数假设检验中可能出现的两种错误

拒绝了实际成立的检验假设 H_0 这叫做第 I 类错误。如果显著性水平 α 定为 0.05 则犯 I 类错误的概率小于 0.05。

接受了实际上不成立的检验假设 H_0 ，这叫做 II 类错误。犯 II 类错误的概率用 β 表示， β 值一般不能确切地知道。当样本含量一定时 α 愈小，则 β 愈大；反之 α 愈大， β 愈小。当 α 一定时，增加样本量，可以减少 β 。

8. 分类指标常用的相对数指标及其意义

(1) 常用相对数指标

相对数是两个有联系的指标之比，医学中的相对数包括比和率两大类，其中比又分

为构成比和相对比。

构成比：又称为构成指标，它是指一事物内部某一组成部分的观察单位数与该事物各组成部分的观察单位总数之比，用来说明一事物内部各组成部分所占的比重或分布。

常用百分数表示，计算公式为：构成比

$$= \frac{\text{某一组成部分的观察单位数}}{\text{事物各组成部分的观察单位总数}} \times 100\%$$

构成比的特点：各组成部分的构成比之和为 100%；某一部分所占比重增大，其他部分也会相应地减少。

相对比：又称比，是两个有联系的指标之比，常以倍数或百分数（%）表示。如果 A 指标大于 B 指标，结果用倍数表示，反之用百分数表示。计算公式：

$$\text{相对比} = A/B (\times 100\%)$$

相比较的两个指标可以性质相同，也可以不同；计算相对比时，A、B 两指标既可以是绝对数，也可以是相对数。

率：又称为频数指标，是指某种现象在一定条件下，实际发生的观察单位数与可能发生该现象的总观察单位数之比，用来说明某种现象发生的频率大小或强度。常用百分率（%）、千分率（‰）、万分率（1/万）、十万分率（1/10万）等表示。计算公式为：

$$\text{率} = \frac{\text{发生某现象的观察单位数}}{\text{可能发生该现象的观察单位数}} \times 100\%$$

率与构成比不同之处在于：率的指标的大小仅决定于某种现象的发生数和可能发生该现象的总数，不受其他指标的影响，并且各率之和一般不为 1。

(2) 应用相对数的注意事项

构成比和率是两种性质不同的指标，构成比只能说明事物内部各组成部分的比重，不能说明某现象发生的频率或强度。

在计算相对数时，分母不宜过小，分母过小会使相对数波动较大。只有观察单位足够时，计算的相对数才比较稳定，才能正确反映实际情况。

对观察单位数不等的几个率，不能直接相加求其平均率，而应用合计的数据来计

算。

资料应具有可比性。决定率或构成比等相对数大小的因素是多方面的。所以进行两个或多个率的比较时，除了所研究的因素以外，还有其他重要的影响因素应相同或相近，要在相同的条件下对比。

由于样本率和构成比都存在着抽样误差，所以比较时必须进行差别的假设检验。

9. 分类指标的假设检验

(1) 率的抽样误差与标准误

由于抽样的原因造成的样本率与总体率之间的差别叫做率的抽样误差。率的抽样误差的大小可以用率的标准误来衡量。率的标准误的计算公式为：

$$\sigma_p = \sqrt{\pi(1-\pi)/n}$$

如果不知道总体率，可用下式求样本率的标准误：

$$S_p = \sqrt{P(1-P)/n} = \sqrt{P \cdot q/n}$$

式中 S_p 为样本率的标准误， P 为样本的阳性率， q 为样本的阴性率。 $q = 1 - P$ 。 n 为样本量。

(2) 总体率可信区间的估计

公式：

$$\text{通式： } P \pm U_{\alpha} S_p$$

总体率的 95% 可信区间是： $P \pm 1.96 S_p$

总体率的 99% 可信区间是： $P \pm 2.58 S_p$

适用条件：要求样本量 n 不能太小，样本率 P 不能太接近 0% 或 100% 即： $nP > 5$ 且 $n(1-P) > 5$ 。

总体率的 95% 可信区间是指有 95% 的把握说总体率在估计的区间范围内，区间估计错误的可能性是 5%。

(3) χ^2 检验

χ^2 检验可用于两个及两个以上率或构成比的比较；两分类变量相关关系分析。

① χ^2 检验的基本公式和基本原理。

基本公式：

$$\chi^2 = \sum \frac{(A - T)^2}{T}$$

自由度：