

## • 系列讲座 •

# 临床流行病学

## 第二讲 筛查及其应用（一）

李 辉

**一、筛查的概念：**筛查（screening）是疾病预防的重要手段之一。1951年美国慢性病委员会对筛查提出了如下定义：“通过快速的检验、检查或其它措施，将可能有病但表面上健康的人，同那些可能无病的人区分开来。筛查检验不是诊断检验，仅是一种初步检查，对筛查检验阳性或可疑阳性者，必须进一步进行确诊，以便对确诊病人采取必要的治疗措施。”

筛查最初应用的领域是疾病的二级预防，即早期发现那些处于临床前期或临床初期的病人，以提高治愈率。在这方面应用最广、效果较好的疾病如：儿童青少年期肺结核、成年期高血压、糖尿病及肿瘤，特别是乳腺癌和宫颈癌。

近年来筛查越来越多地应用于疾病的一级预防，即及时发现某些疾病的高危个体，以减缓发病。例如筛查高血压以预防脑卒中；筛查高胆固醇血症以预防冠心病；筛查肥胖儿以预防成人期心血管病等。

近年来筛查还和提高医疗保健服务质量结合起来，以充分利用发展中国家的有限卫生资源，取得最大的疾病控制效果。例如在孕妇中筛查高危产妇，根据其高危评分而将分娩危险性最高的产妇，安排在条件较好的县市级医院分娩，相反分娩危险性低的产妇则留在当地乡卫生院或村卫生室分娩。这种将孕产妇高危者筛查和转诊系统结合起来的管理办法，既合理地分配了卫生资源，又降低产妇死亡率。又如，近年来 WHO 在发展中国家推广“儿童急性呼吸道感染病例管理方案”，也是很好运用了这一原则。该方案提出了仅以儿童呼吸次数为筛查诊断可疑儿童肺炎的标准，而不使用听诊器、胸部 X 线检查和实验室检查的证据。这对细菌性儿童肺炎患者及早得到抗生素治疗，减少死亡有重要意义。这些成功的实例表明筛查方法简单敏感廉价和快速，是

筛查能否成功的关键。

**二、疾病控制中应用筛查的原则和条件：**在疾病控制中，何种疾病应通过筛查手段来控制，需考虑下述一些原则和条件。

（一）疾病应是当地当前最重要的公共卫生问题之一，即劳动力的丧失、资源的消耗、对家庭和社会的影响较大的严重病多发病。因为任何的筛查计划，那怕是最廉价最简单的筛查检测方法，也要耗费大量的卫生资源；在罕见病的筛查中，大量的卫生资源将浪费在无病者或非高危个体身上。

（二）当地应有一定的卫生资源。卫生资源应用到迫切需要解决且能利用有限的卫生资源实现控制的疾病上。例如，先天畸型是某地新生儿一项重要死因，若当地卫生资源极有限，这时不应选择妊娠期 B 超或羊水检查先天畸型儿终止妊娠的筛查手段，作为减少死亡的主要对策，因为当地尚不具备先天畸型儿筛查的条件。

（三）在筛查疾病的患者身上，应有可识别的早期症状、体征及生理、生化、免疫或病理学改变。确定高危个体的高危因素应是疾病的病因因素或因果链中的中间变量，并且应是易于认别的特征。

（四）筛查的疾病应有有效的治疗方法而使筛查阳性者获最好的结局。筛查确诊的病人可被现有医疗手段治愈、防止残废或减缓死亡；高危者停止暴露可避免发病或推迟发病；高危者能得到转诊系统的最佳管理和服务。

（五）筛查检测方法必须简单、敏感度特异度高、无痛、无创伤、群众乐于接受、经济廉价国情所能承担，筛查阳性者应能获很好的医疗保健服务，保证获得进一步确诊、治疗和指导。

（六）筛查所消耗的卫生资源应比因筛查而减少的发病、致残和死亡所节省的资源及避免有害结局而创造的资源的总和低得多。

对那些不具备筛查条件的疾病进行筛查，往往

会造成一系列不良后果。主要表现在下述几方面：

1. 造成巨大的卫生资源的浪费。筛查是在人群中进行的普遍的健康检查和病例搜寻。如果被筛查疾病发生频率和死亡频率都很低时，无论此病有多么严重、过去多么常见，也不应将筛查作为此病的主要控制对策。如某市解放初期的宫颈癌死亡率高达190/10万，此时采取宫颈癌筛查是有其积极意义的。但到90年代，该市宫颈癌死亡率已降至2/10万以下，再坚持使用筛查手段来控制此病，就成为极大的浪费了。因每检查十万名妇女，才可能减少2例宫颈癌死亡，此花费太大即便是发达国家也不愿承受。

2. 对病人或高危人群带来很大的精神压力。对某些医学发展到今天尚无有效治疗方法的疾病进行筛查（如非高危人群筛查艾滋病）时，筛查并未带来真正的好处，反却使病人过早知道自己身患绝症，从而加速病情发展。在无绝对有效治愈手段时，重则可加速死亡；轻的也会给某些心理素质较差的病人，造成很大的精神压力，降低其生活质量甚或轻生。对一些癌症的筛查也可能存在这种弊端。此外，近年来在某些儿童中筛查“高血压”也具有同样的问题。

3. 对社会造成一定压力。在某些职业暴露所致疾病的控制问题中，最明显是矽肺的筛查。在一些地方小企业中，筛查出来的早期矽肺患者成了企业无法负担的包袱。解决矽肺的问题，重在一级预防，而非仅筛查便能解决的。

**三、筛查检测方法的选择及评价：**筛查所采用的检测方法应根据是筛查疾病的可疑者（或临床前期病人），还是筛查疾病的高危个体而有所区别，其检测方法的评价也应有所不同。

(一) 筛查疾病可疑者的检测方法及其评价：筛查疾病可疑者的检测方法应符合快速、简便、经济、安全及真实可靠的标准。从可行性和科学性上来讲，筛查检测方法应更多的考虑可行性，即在经济上群众能负担得起；在安全性上要求检测方法必须是无创伤性、无肉体和精神痛苦，因筛查对象大多数人都是无该病的健康人。这不同于临床诊断所采用的检测方法，后者更着重科学性，即要保证诊断的病人准确无误不应被误诊或漏诊，误诊和漏诊都将使病人丧失治疗的良机。但筛查检测方法应在保证可行性的前提下，尽量提高其科学性。评价其科学性的标准为真实性和可靠性。

#### 1. 真实性 (validity) 的评价：

(1) 评价真实性的指标：真实性又可称为准确性。对一检测方法的真实性的评价使用敏感度、特异

度和正确指数三个指标。

敏感度 (sensitivity) 和假阴性率 (false negative proportion)：敏感度是指总病人数中筛查检测阳性者占的比例，敏感度反映检测方法发现病人的能力；假阴性率是指总病人中筛查阴性者占的比例，假阴性率表示检测方法漏诊病人的情况。敏感度和假阴性率之间的关系可用下式来表达：敏感度=1-假阴性率。

特异度 (specificity) 和假阳性率 (false positive proportion)：特异度是指全部非病人中筛查阴性者占的比例。特异度反映检测方法确定非病人的能力；假阳性率指全部非病人中筛查阳性者所占的比例。特异度=1-假阳性率

正确指数 (youden's index)；正确指数是敏感度与特异度之和减去1。正确指数表示筛查方法发现真正的病人与非病人总的能力。

筛查检测某疾病的病人与非病人的检查结果见表1。

表1 某病病人与非病人筛查结果

筛查结果	病人	非病人	合计
阳性	真阳性(TP)	假阳性(FP)	TP+FP
阴性	假阴性(FN)	真阴性(TN)	FN+TN
合计	TP+FN	FP+TN	TP+FP+FN+TN

注：表中符号代表各种特征的观察人数

根据上述评价筛查检测方法真实性的指标定义，各指标可按下述公式进行计算：

$$\text{敏感度} = \frac{\text{TP}}{\text{TP} + \text{FN}} \times 100\%$$

$$\text{特异度} = \frac{\text{TN}}{\text{TN} + \text{FP}} \times 100\%$$

$$\text{正确指数} = [\frac{\text{TP}}{\text{TP} + \text{FN}} + \frac{\text{TN}}{\text{TN} + \text{FP}}] - 1$$

$$\text{假阳性率} = \frac{\text{FP}}{\text{FP} + \text{TN}} \times 100\%$$

$$\text{假阴性率} = \frac{\text{FN}}{\text{FN} + \text{TP}} \times 100\%$$

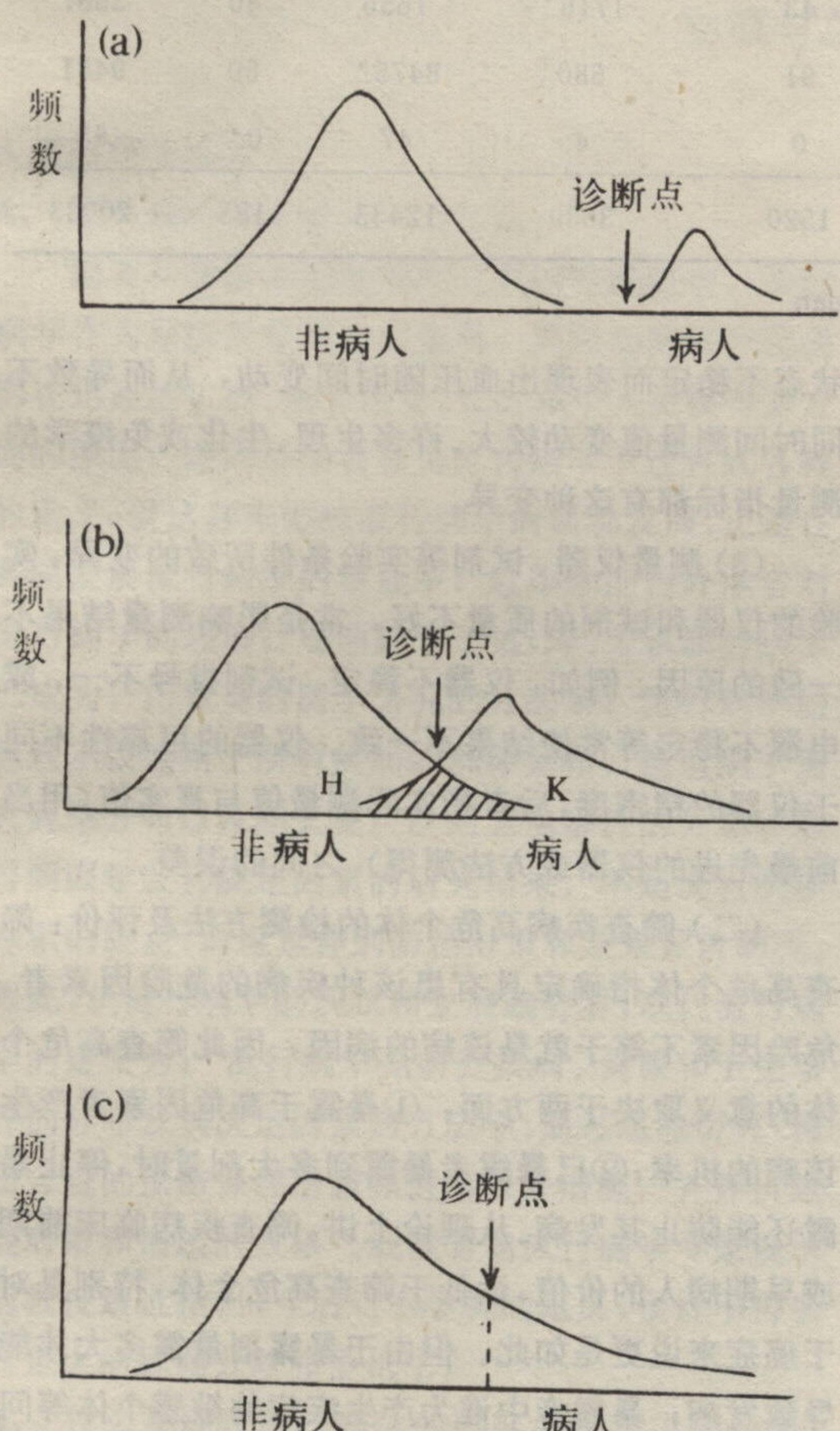
理想化的筛查方法，其敏感度、特异度和正确指数均应接近100%。但实际工作中，因受筛查实施可行性的影响，所选筛查方法的上述三指标很难均达100%，且经常表现为敏感度高者则特异度低，反之亦然的现象。这主要与筛查方法阳性阴性界定值的选择有关。

(2) 评价真实性指标的阳性界定值的选择：在筛查检测方法中，有三种不同的检测结果可用来做诊断：①通过各种化验所获的连续变量测量值；②通过仪器记录的图象；③通过物诊检查记录的结果。

上述三种检测结果中阳性界定值的选择可从统

计学原则和专业要求二方面进行抉择。

①连续性变量测量值的阳性界定原则：筛检指标测量值的分布常见三种类型（附图）。病人与非病人的测量值为两个独立的分布曲线，(a)两条曲线无重叠处；(b)病人与非病人的观测值呈两条相交的分布曲线，两条曲线下有一重叠的部分；(c)病人与非病人的观测值呈一连续性分布曲线。



附图 病人与非病人观测值分布类型

对于附图中(a)的情况，诊断点可选病人中的最小值。此诊断点可使真实性三指标达 100%。

对于附图中(b)的情况，无论诊断点选择在两条分布曲线上的何处，总有假阴性和/或假阳性产生。评价真实性的三指标敏感度、特异度和正确指数均不能同时达到 100%。在图中 HK 两点间，当诊断点向右移时，假阴性率增大，敏感度降低；反之，当诊断点向左移动时，假阳性率增大，特异度减少。当诊断点如移到 K 点时，特异度为 100%，正确指数为敏感

度；反之，当诊断点移到 H 点时，敏感度为 100%，正确指数为特异度。H 和 K 两点间的距离越近，在两点间选择的诊断点所产生的正确的指数就越大。在这种情况下选择诊断点总是顾此失彼，选择何界定值应根据专业要求而定。如筛查的疾病结局严重但早期疗效极好，则诊断点应尽量向 H 点靠近，以便保证病人尽可能的被筛检发现。但假阳性率会大大提高，增加需要进一步确诊的人数，从而使筛查成本提高。因此，向 H 靠近到什么程度，除考虑疾病的严重性外，还应考虑进一步确诊时各种临床检测方法的安全性和花费。若进一步确诊检查会产生别的有害结局且频率较高，显然是不能接受的；当然，进一步做确诊的费用较高时，也是不能接受的。如筛查疾病的结局不很严重，疗效也不太好，且进一步确诊检测方法的安全性不高、花费较大时，诊断点应向 K 点靠近，以尽量避免非病人被诊断为假阳性。但此时假阴性率会增大，敏感度降低，降低了筛查的意义。因此，向 K 点靠近到什么程度，应视卫生资源状况和进一步确诊检测方法的安全性而定。

②图型资料阳性界定的原则：图型资料主要指筛检方法采用 X 线、红外线扫描，以及超声图像等检查所获结果，其阳性的诊断也有类似的上述问题。这类检查使用了阳性、可疑阳性及阴性作为测量的结局，因图型的判别尚缺乏转换为定量指标的方法，目前都是以临床经验为依据，因此阳性与可疑阳性之间或可疑阳性与阴性之间的界定，也受部分人为因素的影响，特别是可疑阳性诊断标准的掌握，也应根据疾病的严重程度、治疗的费用和进一步检查方法的安全性来考虑阳性界定值。

2. 可靠性 (reliability) 的评价：可靠性又称可重复性。可靠性指某一筛检方法重复测量同一受试者时，所获结果的一致性。影响检测结果不一致性的原因是多方面的，包括：

(1) 测量员的测量变异：测量变异可来自测量员间，也可来自测量员本身。

①不同测量员间的测量变异：不同的测量员同时测量一批检测标本时，常因测量员间检测技术或能力的不一致而导致测量结果的不一致。例如人群高血压调查时，参加调查的血压测量员间听第四音的能力不一样；或捕捉收缩压和舒张压读数的能力不一样；或捆扎血压计袖带的松紧掌握不一样等类似的原因，造成彼此间测量同一受检者的结果不一致。现以两名放射科医生对 20033 张 X 线胸片的诊断结果的比较，计算两者诊断的一致性（表 2）。

表 2 两名放射科医生诊断 X 线胸片结果的一致性比较

放射科医生 A 诊断结果	放射科医师 B 诊断结果						
	可疑肿物	其它明显肺部异常	心血管异常	无明显异常	正常	不详	合计
可疑肿物	61*	16	1	9	8	0	95
其它明显肺部异常	70	1320*	63	861	367	33	2714
心血管异常	19	151	1322*	369	1880	62	3803
无明显异常	25	407	43	1716*	1656	40	3887
正常	28	157	91	680	8475*	50	9481
不详	0	2	0	4	47	0*	53
合计	203	2053	1520	3639	12433	185	20033

\* 为两医师诊断结果一致者；资料来源：Lilienfeld and Kordan

两医师诊断一致性百分比计算如下：

$$\text{一致性百分比} (\%) = \frac{61+1320+1322+1716+8475}{20033-(185+53)} \times 100\% = 65.1\%$$

此结果表明，两医师诊断的 20033 张 X 线片的结果，仅有 65.1% 是一致的，而 34.9% 均做了各自不相同的诊断。

② 同一名测量员不同测量次数间的测量变异：同一测量员在不同的测量时间测量同一批标本时，因测量员的自身稳定性不好，同样会造成测量结果的不一致，现仍以 Lilienfeld 和 Kordan 的资料为例，说明这种变异对结果的影响（表 3）。

表 3 同一名测量员二次 X 线胸片诊断结果

自身比较	一致性百分比 (%)	
	分五类诊断	分二类诊断
医师 A	55.1	91.3
医师 B	46.2	80.8

表 3 中按分五类诊断和二类诊断读片，医师 A 两次不同时间诊断结果的一致性分别为 55.1% 和 91.3%；医师 B 则为 46.2% 和 80.8%。

③ 受试对象的变异：受试者变异指同一指标在同一受试者身上重复测量时，测量结果表现为不一致。例如，血压测量时常遇到受试者生理状态、精神

状态不稳定而表现出血压随时间变动，从而导致不同时间测量值变动较大。许多生理、生化或免疫学的测量指标都有这种变异。

④ 测量仪器、试剂等实验条件所致的变异：实验的仪器和试剂的质量不好，常是影响测量结果不一致的原因。例如，仪器不稳定、试剂批号不一，或电源不稳定等常使结果不一致。仪器的可靠性不同于仪器的精密度，后者指实际测量值与真实值（用当前最先进的仪器或方法测得）之间的误差。

(二) 筛查疾病高危个体的检测方法及评价：筛查高危个体指确定具有患该种疾病的危险因素者。危险因素不等于就是该病的病因，因此筛查高危个体的意义取决于两方面：①暴露于高危因素者产生该病的机率；②已暴露者暴露到多大剂量时，停止暴露还能防止其发病。从理论上讲，筛查疾病临床前期或早期病人的价值，远低于筛查高危个体，特别是对于癌症来说更是如此。但由于暴露剂量需多大才能导致发病；暴露者中谁为产生疾病的敏感个体等问题，在多数疾病中尚无定论，因而很难评价高危个体筛查的效果。目前通过测量血压以确定脑卒中高危个体；通过测量血胆固醇以确定冠心病高危个体是较为成功的，在预防有害结局的发生上取得了一定的效果。

有关高危个体筛查中，筛检方法的评价与前述筛检疾病可疑者的评价相同。