

# 住院医师临床能力计算机辅助考试系统

广州军区广州总医院 黄家驹 吴伟斌 龚志梅 由广堂 何炳武

**摘要:** 本文介绍了住院医师临床能力计算机辅助考试系统研究的缘由, 系统结构、原理和特点。

笔者试图探索应用计算机辅助考试的方法, 来测量应试者的临床能力, 使毕业后教育中临床能力的测量逐步走向标准化、制度化。

## 一、系统结构

该系统采用模块化结构编程, 主要分为五个子模块:

1. 设置运行环境参数, 改变参数的值, 使系统适应不同的运行环境。

2. 建立和维护试题库, 包括对本文格式自动检测、纠错, 辅助文件的建立、修改、目录的管理以及试题的打印输出等。

3. 考试管理。主要包括考卷的生成、考生数据文件存放路径的建立、考前练习、考试过程管理、评分、辅助教学、答卷输出、考生成绩单显示和打印等。

4. 考绩管理及分析。主要包括多类考生数据的连接、考生成绩排序、百分位表及频数分布图的显示等。

5. 试题综合统计分析。主要包括数据统计、考生答题情况分析、试题答案分析、试题质量综合分析等。

## 二、系统的运行环境

主机: IBM 兼容的 T, AT(286、386、486) 微机;

显示图像的配置: TVGA(RAM>512KB) 显示适配器和 VGA 彩色或单色显示器;

不显示图像的配置: EGA、CEGA、HERCLUS 显示适配器和与之匹配的显示器

## 三、工作原理

系统工作原理(见图 1), 是由医学专家把临床的实际病例编写成具有固定格式的程序化试题, 将经过试测的试题建立成题库。试题的文字部分以文本格式存放, 答案及其它参数以随机文件格式存放。这样既达到了试

题答案的保密要求, 同时文字编辑处理又简单方便, 而且几乎没有冗余空间, 占用外存相对较小。生成考卷采用“自选法”和“随机法”二种方式, “自选法”以点菜的方式选择试题, “随机法”由计算机依照给予的条件(专业+试题难度+题目数量), 在大量同等条件的试题中随机挑选出所需的试题。计算机自动按先易后难的原则排列试卷, 文字和图像资料均显示在显示器上, 应试者直接以人一机交互的方式进行考试, 只需在计算机键盘上选择按动备选答案相应的数字和个别特定代号, 随意修改, 操作简便。考试过程自动计时、评分, 考试结束即输出考试分数, 包括对各种能力测试情况。考试结束后, 可以作考绩分析和试题分析, 将信息反馈给有关人员。

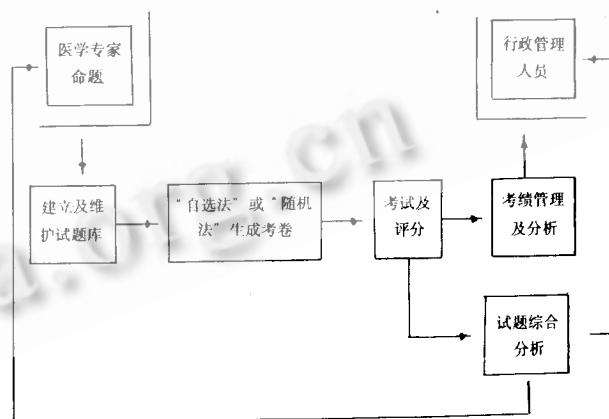


图 1 系统工作流程

## 四、系统特点

1. 提出了适应计算机辅助考试的临床能力分类方法。临床能力是由多方面所组成的, 哪些能力的测量适用计算机辅助考试, 至今尚未见有文献报告。著名心理学家 Bloom 对教育目标的分类提出了认知、精神运动和情感三个领域。根据这一原则, 临床能力的评价则属于

非认知领域行为的评价,世界各国在医学考试中有不同的评价范围和分类方法,其中以 WHO 专家委员会及美国医学考试委员会制定的方案比较具有代表性。但是,在进行临床能力测量时,仍需对以上分类方案作更加明确的限定。我们从评价分类与测量技术相统一的角度出发,将临床能力的评价分成两大类。第一类动作技能,其包括四个方面:①病史采集能力;②全面体检能力;③正确执行医疗操作的技能;④继续医疗和护理能力。第二类是决策能力,其包括三个方面:①运用诊断性辅助检查的能力;②诊断和鉴别诊断的能力;③处置(急诊、日常、随访)能力。本系统主要是解决临床决策能力的测量问题。

2. 实现了建立题库—生成试卷—考试—评分—试题分析—数据管理全过程自动化。本系统可实现试题录入、编辑、维护;提供“自选法”和“随机法”二种编排考卷方式;采取了人一机交互的方式进行考试,建立了一种临床能力测量的新途径,丰富了现行临床能力的测量方法,考试过程中计算机自动评分,分数可直接被用于各项分析,使考试全过程避免了应试者之外人员的干与,既大大提高了客观性,又提高了效率。

3. 解决了多种医学影像图片同步实时显示。医学影像图片在临床上应用非常重要,如 CT 片、ECT 片,核磁共振、X 光片、超声图像和心电图等,它们都需要有高清晰度多灰度级的显示效果,才能达到考试要求。本系统可以显示两种模式的高清晰度影像,一是 16 级灰度图像,如用于心电图等灰度要求低的图片,它的特点是占用外存较小(一般在 10KB 以下)。二是 256 级灰度图像,显示效果与原片无异,它占用外存稍大(一般在 50KB 左右)。实时显示影像图片,为考生创造了一个模拟实际的考试环境。

4. 利用“考绩报告”作信息反馈的窗口。本系统考生答题结束,及时评分,便可报告每个考生的能力达标值,根据达标值可以对某个或某一批考生的各种预测能力进行分析,从中对考生的能力缺陷和差异作出诊断,使系统起到模拟“医师诊病”的作用,这一点使用传统的考试方法是难以实现的。

5. 采用了以病例为基础,程序化的多解选择题。为测验鉴别、解决和处理病人问题的临床能力,1961 年美

国全国医学考试委员会首次设计和应用这种试题。这种试题的结构,是由一个病历摘要和若干组检查提示、提问、备选答案所组成。其突出的特点之一,是围绕着某个病例而逐步引申出与该病例有关的一系列临床问题。另一特点是备选答案有正确、错误和无效三种答案,选了正确答案得分,选了错误答案扣分,无效答案选与不选均不得也不扣分。

## 五、试用结果

系统经过广泛的试用,体现了以下优点:

1. 对住院医师临床能力具有较高的评价作用。我们分别对三所医院 32 名住院医师的平时临床工作能力(综合考核)与考试成绩进行相关性(等级相关)研究,其结果为:  $r_R = 0.97$ ,  $t = 13.83$ ,  $(t_{(12)}0.01 = 3.055)$ ;  $r_R = 0.82$ ,  $t = 4.31(t_{(9)}0.01 = 3.250)$ ;  $r_R = 0.81$ ,  $t = 5.325(t_{(5)}0.01 = 4.032)$ ,  $P$  均  $< 0.01$ 。从总体上看,应试者的考试成绩与平时临床工作能力之间存在着较高的正相关。可见,该考试方法有着较高的达标关联效度。说明。考试能反应应试者平日的临床能力。

2. 具有较高的信度。对一组 36 名普通外科考生的试测样本进行了试题质量分析。考生采用相同的试题,共考了五个大题,其中有 36 道小题,试题分析的结果是:试题难度  $P = 0.3 \sim 0.7$  占 83%, 0.8 以上占 17%, 无 0.2 以下的;试题区分度  $r = 0.3$  以上占 89%, 0.2 以下占 11%;标准度  $\Delta = 11.8$ ,  $\alpha$  系数 = 0.88 ~ 0.95 占四大题,  $\alpha$  系数 = 0.778 的只占一大题。从对试题的难度、区分度和内部一致性信度的分析,反映了考试的结果是可靠的。

3. 系统提供了程序化的文字资料,并实时显示医学影像图片,使考试方法较好地模拟了临床实际,而笔测则难以实现。

4. 试题难度相同,评分标准一致,克服了病例难易因素和人为因素的差异,大大提高了考试的信度,这明显优于床边考试。

5. 考试过程采用人一机对话的方式,考生独立完成,系统自动记时(超时仍未完成则自动终止考试),从根本上杜绝了考生作弊的现象。