

文章编号: 2095-1663(2011)05-0054-06

建立医学领域研究型博士学位评价标准的探讨

——应用德尔菲法的尝试

陈建俞 刘少雪

(上海交通大学高等教育研究院, 上海 200240)

摘要:医学领域研究型博士学位评价是对博士生经过3~5年学习之后形成的最终学习成果进行价值判断。为了探究医学领域研究型博士学位评价标准,我们采用德尔菲(Delphi)法,对22名专家进行两轮咨询。研究结果初步确立了医学领域研究型博士学位评价体系,包含4大要素、25项评价指标。本研究不仅为社会评价医学领域研究型博士学位提供统一的参考标准和导向,还为高校授予医学领域研究型博士学位提供一套实用的工具。

关键词:研究型博士学位;医学领域;评价标准;德尔菲法

中图分类号: G643.7

文献标识码: A

一、建立博士学位评价标准的必要性和可能性

《中华人民共和国学位条例》对博士学位的授予标准作了一般性的规定,但一直缺少一个可操作性的具体标准。在具体评价实践活动中,博士学位评价标准大多通过导师以及行内专家把握,即主观性标准强,客观性标准弱。随着我国博士学位授予量的不断增加、具有博士学位授予权的学校逐渐增多,建立相对客观并易于把握的博士学位评价标准,已成为一种社会需求。

随着经济全球化,高等教育呈国际化的发展趋势,2005年欧洲各国教育部长在挪威卑尔根(Bergen)召开会议,确定了欧洲高等教育区的学位标准和质量保证准则,其目的是为了指导欧洲各国规范学位教育,提高学位教育水平,同时促进各国之间学位互认,便于高级人才在欧洲各国之间,乃至欧

洲和世界各地之间的流动。由此给我们一个启示,建立一个具有国际可比性的博士学位评价标准体系,保证我国博士学位在世界范围内具有相当的可接受性,是当前我国博士生教育改革和完善我国博士学位评价体系的重要动力之一。

目前社会各行各业设置的标准有很多,设置的方法渠道也有多种。博士学位评价标准所针对的是主观性特别强、差异性特别大的人和教育,其评价标准的建立必然面临其内在本质上的挑战。其挑战主要来自于:①博士学位质量是否可以评价?②如果可以评价的话,那应该由谁来评价?③如何保证和监督评价?学位是标志被授予者的受教育程度和学术水平达到规定标准的学术称号,博士是学位的最高一级,因此博士学位本身就是一种标准。对于博士生来说,博士学位标准就是博士生为了获得学位必须达到的学习成果水平;对于博士学位授予机构来说,博士学位标准就是判断博士学位候选人是否

收稿日期:2011-07-18

作者简介:陈建俞(1970—),女,江苏张家港人,上海交通大学高等教育研究院博士研究生,上海交通大学医学院科技发展有限公司助理研究员。

刘少雪(1967—),女,山东莱阳人,上海交通大学高等教育研究院教授,博士。

达到学位授予的最低要求;对于用人单位来说,博士学位标准就是判断博士学位拥有者是否达到了博士学位名称相应学术水准的依据;对于政府主管部门来说,博士学位标准就是确保不同博士学位授予机构所授予的博士学位达到统一的基本要求。因此,虽然不同评价主体所持的质量观不同,但评价内容是相同的,即博士学位获得者应达到某种程度的知识能力程度和特征。

鉴于不同博士学位类型不同学科领域的区别,本研究试图以医学学科为例,对医学领域研究型博士学位获得者在知识能力方面应具有的特征提出一个较为全面的可采纳的统一标准。

二、基本假设

1. 世界各国和地区的医学学位体系虽然不尽相同,但对于医学研究型博士学位(PhD)这一类型这一层次都有一个共同基准,即都以博士生获得一定的专业知识和研究技能,能够独立从事实验室研究为培养目标。因此,评价医学研究型博士学位的核心要素应该是博士生所拥有的医学专业知识以及从事研究的实验技术能力。

2. 博士学位是一种学术等级,代表着在一个特定学术领域内的最高程度的学术水平,因此对博士学位的评价标准应该掌握在相关学科领域学术同行专家手里。也只有学术同行专家才能认定博士生的知识程度和能力水平。

德尔菲(Delphi)法是一种综合专家经验统计

判断的方法。它是通过广泛征求专家意见,经反复多次的信息交流和反馈修正,使专家的意见逐步趋向一致,最后根据专家的综合意见,对评价对象做出评价的一种定量与定性相结合的方法。20世纪60年代美国兰德公司首次将德尔菲法用于技术预测。由于德尔菲法具有匿名性、反馈性和趋同性三大特点,其应用范围后来迅速扩大,被广泛地应用到各个领域的综合评价实践中。本研究将应用此方法来确定医学领域研究型博士学位评价标准的各级评价指标及其权重。

三、研究步骤

1. 初步拟定评价指标体系

本研究通过对2008年英国质量保障机构(the Quality Assurance Agency, QAA)公布的高等教育资格框架(FHEQ),国际生物化学与分子生物学联合会(IUBMB)教育委员会建议的分子生物科学博士学位标准,国际免疫学联合会(IUIS)教育委员会建议的免疫学博士学位标准和《中华人民共和国学位条例》等国内外文献资料进行分析研究,在结合部分专家意见的基础上,提出了初始医学领域研究型博士学位评价指标体系(见表1)。拟定的评价指标体系分为两级,一级指标分别为知识、研究能力、态度和职业技能4个,二级指标共10个,分别为本学科专业知识、本学科相关的知识、创造解释新知识、判断能力、技术能力、交际能力、发表能力、科学道德、学习技能和职业发展。

表1 两轮拟定的评价指标体系

第一轮		第二轮	
一级指标	二级指标	一级指标	二级指标
知识 K	本学科专业知识 K1	知识 K	医学专业知识 K1
	本学科相关的知识 K2		医学相关的知识 K2
	创造解释新知识 K3		创造解释新知识 K3
研究能力 R	判断能力 R1	研究能力 R	判断能力 R1
	技术能力 R2		技术能力 R2
	交际能力 R3		交流能力 R3
	发表能力 R4	态度 A	科学道德 A1
态度 A	科学道德 A1		科学热情 A2
职业技能 P	学习技能 P1	职业技能 P	学习技能 P1
	职业发展 P2		职业发展 P2

2. 选择咨询专家

由于函询单位的科教部门往往对本单位专家的情况了如指掌,所以本研究在与函询单位的科教部

门充分沟通之后,委托他们来确定专家名单,发放并回收专家咨询表,以保证德尔菲法咨询的应答率,保证反馈的及时性。依据代表性与权威性的原则、学

术专家与管理专家相结合的原则,本研究共选择了 22 名专家,分别来自于上海交通大学医学院、复旦大学、第二军医大学、华东师范大学、瑞金医院、新华医院、第九人民医院、第一人民医院、华山医院等 4 所大学 6 个医院,其中医学专业领域的专家占 54.5%,研究生教育管理专业领域的专家占 13.7%,担任一定领导职务并具有医学研究生教育管理经验的专家占 31.8%。

3. 实施与统计

本研究于 2010 年 3 月~2010 年 5 月共进行了两轮专家咨询。首先,根据初步确立的评价指标体系,拟订并发送第一轮专家咨询函。其中第一轮专家咨询函中介绍了本次研究的目的、研究背景以及国内外相关的学位标准,供专家参考。接着,剔除第一轮回收的不合格咨询表,根据第一轮咨询表结果和专家提出的意见,修改并完善第二轮医学领域研究型博士学位评价指标体系(见表 1)。最后,发送第二轮专家咨询函,请各位专家对指标重新进行评分,回收第二轮咨询表。对于两轮专家反馈的结果,我们运用 SPSS19.0 统计软件,就专家的积极性和权威性,专家意见的集中程度和协调程度进行相关统计学分析。

四、结果分析

1. 专家的积极性

第一轮发出专家咨询函 22 份,回收有效咨询表 20 份,支持系数 $W_1=90.9\%$,其中对本轮咨询表提出修改意见的专家有 15 人,占本轮专家总人数的 75%;第二轮发出专家咨询函 21 份,回收有效咨询表 20 份,支持系数 $W_2=95.2\%$,其中对本轮咨询表提出修改意见的专家有 8 人,占本轮专家总人数的 40%。说明专家对本研究具有很高的积极性。

2. 专家的权威性

专家的权威性一般由专家自身知识水平和专家对评价指标体系熟悉度两方面因素决定。结合本研究咨询专家的情况分析,工作年限 20 年以上的专家 16 人(80%),正高级职称专家 17 人(85%)。从专家对博士学位评价指标体系的熟悉程度的自评情况来看,10 人(50%)很熟悉,9 人(45%)比较熟悉,熟悉程度一般的仅为 1 人(5%)。因此本研究专家的权威程度和可靠程度较高。

3. 专家意见的集中程度

本研究采用专家直接评分法来计算各指标权重,权重的总和为 100%。权重系数反映了指标的重要性大小,系数越大表示指标越重要。平均数(Mean)和中位数(Median)是统计学中最常用的统计量,用来描述集中趋势的一组指标。

从附表 1 可以看出,专家对一级指标知识(K)和研究技能(R)在博士学位评价标准中的重要性意见较为集中。相对于第一轮权重咨询结果,第二轮专家们对一级指标态度(A)的重视程度要高于职业技能(P)。同时,对两轮权重咨询结果进行方差分析,结果发现一级指标和二级指标的 P 值分别为 1.000 和 0.932,均大于 0.05,因而认为两轮专家意见无统计学差异,专家对指标权重的咨询意见趋于一致。

4. 专家意见的协调程度

变异系数 CV 是代表专家对指标评价相对波动大小的重要指标,变异系数越小,说明专家意见的协调程度越高。变异系数的计算公式为: $C \cdot V = \frac{S}{\bar{x}} \times 100\%$, $\bar{x}$:专家对各指标赋予的权重值的平均值; S :专家对各指标赋予的权重值的标准差。根据国内几项大型德尔菲法在卫生系统的应用研究,经过 2~3 轮函询协调后,协调系数一般在 0.5 范围内波动。

从两轮专家对指标评价的量化结果来看,专家意见基本协调,指标权重得到收敛(见附表 1~3)。在第一轮专家咨询结果中,一级指标的变异系数为 15.89%~34.49%,二级指标的变异系数为 28.53%~39.04%。通过修改调整后,在第二轮专家咨询结果中,一级指标的变异系数为 13.66%~25.99%,二级指标的变异系数为 23.50%~31.11%,三级指标的变异系数为 24.06%~45.37%。经卡方检验 $\chi^2=65.154, P < 0.000$,本研究两轮协调系数有统计学意义,评价结果可取。

5. 评价指标体系

经过两轮专家咨询,最终形成了医学领域研究型博士学位评价的 4 个一级指标、10 个二级指标、25 个三级指标及其权重。从表 2 可以看出,根据权重大小,研究型博士学位评价指标可分为核心指标、重要指标和一般指标三大类。

在一级评价指标体系中,“知识 K”和“研究能力 R”为核心指标,在 4 个一级指标体系中所占的权重最大,分别为 0.3075 和 0.3690,这与研究型博士学位的本质内涵相一致。研究型博士学位本身就是一

种因具备开展学术研究、产生新知识能力而被授予的学位。另外一级指标“态度 A”为重要指标,“职业技能 P”为一般指标。因为主动、求实、严谨、刻苦钻研的科学态度是决定博士生能否在就读期间取得科研成果的重要内因。

在二级评价指标体系中,权重大于 0.13 的“医学专业知识 K1”和“技术能力 R2”为核心指标,权重大于 0.08 的“医学相关的知识 K2”、“创造解释新知识 K3”、“判断能力 R1”、“交流能力 R3”“科学道德 A1”和“学习技能 P1”为重要指标,而权重小于 0.08 的“科学热情 A2”和“职业发展 P2”为一般指标。其中二级指标“技术能力”的权重明显高于其他能力指标,这充分体现了医学学科领域的特殊性,医学领域研究型博士应具有设计实验并进行有效的独

立研究的能力。在当今科学界屡次出现做假欺骗行为的情况下,二级指标“科学道德 A1”更有重要的现实意义和时代价值。

在三级评价指标体系中,核心指标为权重大于 0.05 的指标,如“医学专业基础理论、原理与概念 K11”、“新角度、新假设、新方法、新发现或新成果 K3”、“甄别值得研究的问题的能力 R12”和“独立设计实验方案的能力 R22”。重要指标为权重介于 0.03~0.05 之间的指标。一般指标为权重小于 0.03 的指标。三级评价指标体系中权重最高的指标是“新角度、新假设、新方法、新发现或新成果 K3”。诚然,能开展独立研究并对知识做出原创性的贡献,是授予研究型博士学位的最终根本,也是评价研究型博士学位的核心要素。

表 2 我国医学领域研究型博士学位评价指标体系

一级指标及权重系数	二级指标及权重系数	三级指标	权重系数
知识 K (0.3075)	医学专业知识 K1 (0.1374)	医学专业基础理论、原理与概念 K11	0.0534
		医学实验基础与实验方法 K12	0.0384
		医学前沿知识与技术 K13	0.0455
	医学相关的知识 K2 (0.0853)	心理学、法律学、伦理学等人文方面的知识 K21	0.0409
		与研究课题相关的仪器设备与材料的知识 K22	0.0443
	创造解释新知识 K3 (0.0849)	新角度、新假设、新方法、新发现或新成果 K3	0.0849
研究能力 R (0.3690)	判断能力 R1 (0.1259)	批判性阅读文献的能力 R11	0.0376
		甄别值得研究的问题的能力 R12	0.0507
		评价个人及其他研究工作的能力 R13	0.0376
	技术能力 R2 (0.1319)	使用信息技术的能力 R21	0.0374
		独立设计实验方案的能力 R22	0.0551
		与研究课题相关的实验技术操作能力 R23	0.0394
	交流能力 R3 (0.1113)	口头与书面的表达能力 R31	0.0392
		团队协作能力 R32	0.0346
		在国际一流刊物上发表论文的能力 R33	0.0374
态度 A (0.1635)	科学道德 A1 (0.0858)	遵守学术道德规范 A11	0.0441
		坚守学术诚信 A12	0.0416
	科学热情 A2 (0.0778)	探索求知的好奇心 A21	0.0396
		求真务实的进取心 A22	0.0381
职业技能 P (0.1600)	学习技能 P1 (0.0835)	外语运用能力 P11	0.0278
		获取新知识的能力 P12	0.0269
		信息采集、归纳、分析及应用的能力 P13	0.0278
	职业发展 P2 (0.0765)	处理复杂的伦理和职业问题的能力 P21	0.0284
		敢于承担责任的能力 P22	0.0255
		心理承受与调节能力 P23	0.0228

五、讨论与说明

1. 是否需要列入“职业技能”一级指标？

在第一轮专家咨询结果中发现,有些专家认为研究型博士主要靠“研究能力”来获得职业发展机会,没有单列一级指标“职业技能”的必要性。然而随着社会经济的发展以及博士数量的激增,目前医

学领域研究型博士就业呈现新的趋向。高等院校和科研院所这些传统意义上的就业单位已不再是医学研究型博士的唯一选择,进入政府、医疗卫生机构、企业等非学术性机构就职的比例逐年上升。实际上,20世纪90年代起,英美等国相关机构就开始注意到博士学位内涵在发生改变,博士生教育的核心是通过原创性研究促进知识发展,与此同时,博士生培养必须符合劳动力市场需要,而不仅仅是满足学术需要。1995年美国自然科学、工程和公共政策委员会(the Committee on Science, Engineering and Public Policy, COSEPUP)在其报告《重塑培养科学家和工程师的研究生教育》中提出,研究生培养计划应提供多样化的培训。1993年英国科学技术办公室(Office of Science and Technology, OST)也表达了其对传统博士学位狭窄就业面的关注,“政府欢迎研究生课程的变化。但值得关注的是研究型博士不能很好地满足学术研究之外的职业要求或者不能满足工业研究实验室的要求”。1997年英国迪尔英报告(Dearing Report)认为,传统上博士研究只关注学位论文的产出,这已经导致不能识别获得广泛技能的需求。鉴于上述理由,本研究在发放第二轮专家咨询表中特附上相关资料说明,结果反馈显示第

二轮专家对设列“职业技能”一级指标的意见趋于一致,大家都认为我国医学院校确实应采取各种积极措施,增加实际技能培训,以满足医学研究型博士从事多方面工作的需求,以便更好地应对社会的挑战。

2. 增列“科学热情”二级指标的意图是什么?

科学研究是一个积极探索生产新知识的过程,也是一个漫长、曲折、复杂而且充满酸甜苦辣的过程。因此,作为研究者的博士生,如果没有一定的研究兴趣、热情和动力,是很难坚持下来的。在第一轮专家咨询中,有些专家提出激发博士对科学问题的好奇心和探究热情,激发博士对科学研究的创新活力,是提高博士生学术水平和研究能力的重要保证。因此在第二轮专家咨询表中我们新增列了一个二级指标“科学热情”。

3. 需进一步说明的问题

本文论述了医学领域研究型博士学位的一般评价标准。虽然本文已论及许多相关因素,但此评价标准还有待进一步完善,特别是各指标间的相关性研究目前尚未进行。尽管如此,目前已有的这些标准至少可以作为探索此类问题的一个出发点,而不应该看作是一种定论,对指导具体评价实践活动具有一定的借鉴意义。

附表1 一级指标咨询结果

指标 编号	第一轮				第二轮			
	均数 Mean	中位数 Median	标准差 Std Dev	变异系数 CV(%)	均数 Mean	中位数 Median	标准差 Std Dev	变异系数 CV(%)
K	30.75	30.00	6.13	19.93	30.75	30.00	4.38	14.24
R	37.50	40.00	5.96	15.89	36.90	40.00	5.04	13.66
A	14.50	15.00	4.56	31.45	16.35	18.50	4.25	25.99
P	17.25	17.50	5.95	34.49	16.00	15.00	3.84	24.00

附表2 二级指标咨询结果

指标 编号	第一轮				第二轮			
	均数 Mean	中位数 Median	标准差 Std Dev	变异系数 CV(%)	均数 Mean	中位数 Median	标准差 Std Dev	变异系数 CV(%)
K1	11.90	10.00	4.33	36.39	13.74	13.88	3.26	23.73
K2	8.38	9.50	2.59	30.91	8.53	9.00	2.61	30.60
K3	10.48	10.00	2.99	28.53	8.49	8.00	2.18	25.68
R1	11.04	10.00	3.20	28.99	12.59	12.00	3.61	28.67
R2	11.59	10.00	3.66	31.58	13.19	12.63	3.10	23.50
R3	6.89	6.38	2.69	39.04	11.13	10.00	2.86	25.70
R4	8.41	8.00	2.74	32.58	8.58	9.00	2.14	24.94
A1	13.70	11.00	4.92	35.91	7.78	8.00	2.42	31.11
P1	9.25	10.00	3.61	39.03	8.35	8.50	2.05	24.55
P2	7.50	7.00	2.71	36.13	7.65	7.75	2.23	29.15

附表 3 三级指标咨询结果

指标编号	均数 Mean	中位数 Median	标准差 Std Dev	变异系数 CV (%)
K11	5.3400	5.0000	2.0793	38.94
K12	3.8438	3.8750	0.9611	25.00
K13	4.5538	4.4625	1.5111	33.18
K21	4.0950	4.2500	1.2280	29.99
K22	4.4300	4.7500	1.8874	42.60
K3	8.4875	8.0000	2.1832	25.72
R11	3.7644	3.5000	1.5313	40.68
R12	5.0669	5.0000	1.5004	29.61
R13	3.7563	3.0000	1.7042	45.37
R21	3.7350	3.7125	1.3758	36.84
R22	5.5119	5.0000	1.7869	32.42
R23	3.9406	3.8750	1.0767	27.32
R31	3.9188	4.0000	1.3043	33.28
R32	3.4638	3.0000	1.3521	39.04
R33	3.7425	3.0750	1.6184	43.24
A11	4.4125	4.5000	1.0615	24.06
A12	4.1625	4.0000	1.1536	27.71
A21	3.9625	4.0000	1.4830	37.43
A22	3.8125	4.0000	1.4162	37.15
P11	2.7825	2.7500	0.9295	33.41
P12	2.6875	3.0000	0.7691	28.62
P13	2.7800	3.0000	0.8377	30.13
P21	2.8450	3.0000	1.1482	40.36
P22	2.5500	2.3250	0.8474	33.23
P23	2.2800	2.0000	0.8848	38.81

参考文献:

- [1] 曾光,李辉.现代流行病学方法与利用[M].北京医科大学中国协和医科大学联合出版社,1996:250-270.
- [2] 吴欣娟,谢瑶洁,马丽莉,等.医院护士长量化考核指标体系的初步研究[J].中华护理杂志,2008,43(10):872-875.
- [3] 朱小凤,曾智,饶莉,等.应用 Delphi 法建立四川省农村卫生适宜技术筛评指标体系[J].现代预防医学,2009,36(2):201-202
- [4] 袁青,黄淇敏.应用德尔菲法筛选医院中层管理人员评价指标的研究[J].中国医院管理,2009,29(7):9-12.
- [5] 1999 年国际生物化学与分子生物学联合会(IUBMB)教育委员会建议,分子生物科学博士学位标准.清华大学生物科学与技术系和研究生院译.

Discussion of Evaluation Criteria for a Research Doctoral Degree in Medicine — An Attempt at Utilizing the Delphi Technique

CHEN Jian-yu, LIU Shao-xue

(Institute of Higher Education, Shanghai Jiaotong University, Shanghai 200240)

Abstract: The evaluation for a research doctoral degree in medicine is a judgment of the results of 3-5 years of study. We investigated the evaluation criteria with two rounds of consultation with 22 experts by means of the Delphi technique. It was found that the criteria consist of four main elements with 25 indicators of evaluation. It is hoped that our findings will provide useful guidance as well as a practical tool for medical degree evaluations.

Keywords: research doctoral degree; medical field; evaluation criteria; Delphi technique