

· 管理纵横 ·

国家自然科学基金项目的函评意见质量现况与分析

——以2019年医学科学部两学科的面上、青年科学基金项目为例

刘 昀^{1,2} 梁 雪^{1,3} 廖丹颖^{1,4} 江虎军^{1*}

1. 国家自然科学基金委员会 医学科学部, 北京 100085
2. 西安交通大学 第二附属医院, 西安 710004
3. 广州医科大学 附属第五医院, 广州 510700
4. 华中科技大学 同济医学院附属协和医院, 武汉 430022

[摘要] 以2019年医学科学部呼吸、血液两个学科受理的面上项目和青年科学基金项目的函评意见为研究对象,就函评意见的专业性和从函评意见中的收获对项目申请人进行了问卷调查,同时学科管理人员也对函评意见专业性进行了评价。其结果在一定意义上反映了国家自然科学基金项目函评意见的基本情况,表明保持和进一步提高同行评议质量仍然是国家自然科学基金项目同行评议工作中的重要任务。

[关键词] 科学基金项目;函评意见;质量评价

国家自然科学基金项目(以下简称“基金项目”)的同行评议制度已实施了33年,针对函评意见质量的评价分析却少有报道^[1],其原因可能在于缺少合适的评价指标和评议人。虽然基金项目评审的公正性已经得到申请者的广泛认可,但近十年来项目受理数量出现了快速增长,科学基金管理人员的数量并没有相应增加,这导致其工作强度远远超过国际同行,也给项目评审质量带来风险^[2]。为了客观认识基金项目函评意见的质量现状,以提高今后函评意见质量,对目前函评意见质量进行分析和评价十分有必要。

我们以2019年医学科学部呼吸、血液这两个学科受理的面上项目和青年科学基金项目(以下简称“青年项目”)两类项目的函评意见为研究对象,全文反馈每一项目的3份函评意见给申请者,并对申请者进行问卷调查。同时学科管理人员也从管理者的角度对项目的函评意见进行了评价分析,力求从不同角度对函评意见的质量进行评价。

1 项目受理情况

呼吸学科2019年接收面上、青年项目申请数分



江虎军 研究员,博士,现任国家自然科学基金委员会医学科学部科学一处处长。



刘 昀 西安交通大学第二附属医院呼吸与危重症医学科副主任,主任医师,博士生导师,主要研究方向:支气管哮喘及变态反应性疾病发病机制的研究,2018年至2020年在国家自然科学基金委员会医学科学部任呼吸血液流动项目主任。

别为683项、770项,初筛分别为9项、17项;血液学科接收面上、青年项目申请数分别为730项、897项,初筛分别为4项、1项(表1)。所有受理的项目均指派给三位同行专家进行函评,函评意见回收率为100%,共收回函评意见9147份。

2 申请人对函评意见调查结果分析

基金项目评审结束后,项目的函评意见全文反馈

表1 2019年呼吸、血液学科面上、青年项目函评项目数

项目类别	申请数 (项)	初筛数 (项)	受理数 (项)	发函数 (项)	回函数 (项)
呼吸	面上	683	9	674	2 022
	青年	770	17	753	2 259
血液	面上	730	4	726	2 178
	青年	897	1	896	2 688
合计	3 080	31	3 049	9 147	9 147

给每一位申请人,同时向申请人发出了函评意见调查问卷。问卷要求项目申请人从如下两个方面评价所反馈的函评意见:一是函评意见是否从专业角度进行评价;二是申请人从反馈的函评意见中是否有所收获,即设置评价指标“函评意见的专业性”与“从函评意见中的收获”。

为方便申请人进行评价,以提高问卷调查回收率,上述两个指标均设计成下拉菜单式选项,申请人不需要提供任何文字评价,但要求申请人客观、公正地对3位专家的函评意见进行评价,同时要求在一周之内返回调查问卷。回收调查问卷后,对相关数据进行统计分析。

2.1 调查问卷发放回收情况

呼吸、血液两学科2019年受理的面上、青年项目的函评意见全部被反馈给了申请人。但呼吸学科面上项目问卷回收率为16.47%,青年项目问卷回收率仅为6.51%;血液学科面上项目问卷回收率为20.52%,青年项目问卷回收率仅为7.48%(表2)。两个学科的问卷回收率之间无统计学差异($P>0.05$),但面上项目与青年基金的问卷回收率差异有统计学意义($P<0.05$)。

尽管设计了方便的调查问卷,但问卷回收率依然不高,面上项目的问卷回收率稍好于青年项目。可以看出,申请人对函评意见的反评估积极性不高。进一步分析发现,已获资助申请人问卷回收率为34.93%,未获资助的为8.54%,两组之间的回收率差异有统计学意义($P<0.05$),这种回收率差异或许与项目申请人是否获资助有关。

2.2 申请人对“函评意见的专业性”的评价

调查问卷中,评价指标“函评意见的专业性”分为三档,分别是A:“专业性较好或以上”;B:“专业性一般”;C:“专业性不足或较差”。呼吸学科申请人对该指标评价为A、B、C的百分率依次是78.33%、14.58%、7.08%,血液学科申请人对该指标的评价

表2 呼吸、血液学科调查问卷发放回收情况统计

项目类别	受理数 (份)	发放数 (份)	回收数 (份)	回收率 (%)
面上	674	674	111	16.47
呼吸	753	753	49	6.51
小计	1 427	1 427	160	11.21
面上	726	726	149	20.52
血液	896	896	67	7.48
小计	1 622	1 622	216	13.32
合计	3 049	3 049	376	12.33

为A、B、C的百分率依次是77.01%、13.43%、9.57%。两学科之间申请人对函评意见的专业性评价无统计学差异($P>0.05$)。总体上申请人认为约80%的函评专家意见的专业性较好或以上,约20%的函评专家意见的专业性一般或较差(表3),虽然绝大多数的申请人对函评意见的专业性给予了好的评价,但也要看到依然有约20%的申请人对函评意见不满。

若按项目类别进行统计,两学科面上项目申请人对该指标评价为A、B、C的百分率依次是74.10%、16.15%、9.74%,青年项目申请人对该指标评价为A、B、C的百分率依次是85.34%、8.91%、5.74%;青年项目申请人认为函评意见专业性更好(申请人对函评意见专业性评价在两类项目之间有统计学差异, $P<0.05$)(表3)。

若按申请者是否获得资助进行统计,两学科已获资助项目申请人对该指标评价为A、B、C的百分率依次是95.21%、4.14%、0.65%,未获资助项目申请人对该指标评价为A、B、C的百分率依次是65.47%、20.62%、13.90%。已获资助项目申请人认为函评意见专业性明显比未获资助申请人好(两组之间有统计学差异, $P<0.05$)(表3)。显然,未获资助申请人对函评意见不满的百分率较高。

2.3 申请人对“从函评意见中的收获”的评价

调查问卷中,评价指标“从函评意见中的收获”也分为三档,分别是A:“收获较大”;B:“收获一般”;C:“几乎没有收获”。呼吸学科申请人对该指标评价为A、B、C的百分率依次是72.33%、18.66%、9.01%,血液学科申请人对该指标评价为A、B、C的百分率依次是70.05%、18.43%、11.52%。总体上,

表3 申请人对指标“函评意见的专业性”各档次评价所占百分率(%)

函评意见的专业性	按学科		按项目类别		按项目是否获资助	
	呼吸	血液	面上	青年	获资助	未获资助
专业性较好或以上(A)	78.33	77.01	74.10	85.34	95.21	65.47
专业性一般(B)	14.58	13.43	16.15	8.91	4.14	20.62
专业性不足或较差(C)	7.08	9.57	9.74	5.74	0.65	13.90

申请人认为从约70%的函评意见中的收获较大,约30%的函评意见中的收获一般或较差(两学科之间无统计学差异, $P > 0.05$) (表4),这说明在申请人从函评意见中的收获方面仍有较大的提升空间。

若按项目类别进行统计,两学科面上项目申请人对该指标评价为A、B、C的百分率依次是65.90%、21.92%、12.18%,青年项目申请人对该指标评价为A、B、C的百分率依次是82.47%、10.92%、6.61%。青年项目申请人认为“从函评意见中的收获”明显优于面上项目申请人对该指标的评价(两组之间有统计学差异, $P < 0.05$) (表4),这与上一个评价指标的结论是一致的。

若按申请者是否获得资助进行统计,两学科已获资助项目申请人对该指标评价为A、B、C的百分率依次是89.54%、9.15%、1.31%,未获资助项目申请人对该指标评价为A、B、C的百分率依次是58.30%、24.96%、16.74%。已获资助项目申请人认为从约90%函评意见中的收获较大,约10%的函评意见中的收获为一般或没有;但在未获资助项目申请人中,不到60%的申请人认为函评意见中收获较大,超过40%的申请人认为函评意见中的收获一般或没有(两组之间有统计学差异 $P < 0.05$) (表4)。在未获资助申请人中对函评意见不满意的百分率较高,这不排除申请人因未获资助原因所带来的不满,但作为项目管理者,还必须进一步提高函评质量,特别是让未获资助的申请人能从函评意见中有所收获,提高下次申请书的质量。

3 项目管理人员对函评意见专业性的评价

学科项目管理人员同时对呼吸、血液两学科面上、青年项目所有回收的函评专家意见逐条进行分析和评价。为了便于学科工作人员尽可能准确、客观地进行判断,只评价函评意见是从专业角度进行的评议还是套话式的评议,即函评意见的专业性是好还是不足。

学科项目管理人员对呼吸、血液两学科回收的总共9147份函评意见进行了分析评价,函评意见专业性好的占88.98%,专业性不足的占11.02%。其中,血液学科专业性好的占92.09%,专业性不足的占7.91%;呼吸学科专业性好的占85.45%,专业性不足的占14.55%(两学科间有统计学差异, $P < 0.05$,表5),血液学科函评意见专业性要高于呼吸学科,这种差异可能与学科特点有关。

按项目类别统计,面上项目函评意见专业性好的占88.48%,专业性不足的占11.52%;青年项目函评意见专业性好的占89.41%,专业性不足的占10.59%(两组之间无统计学差异, $P > 0.05$,表5)。

按项目是否获得资助统计,获资助项目的函评意见专业性好的占86.82%,专业性不足的占13.18%;未获资助项目函评意见专业性好的占89.34%,专业性不足的占10.66%。未获资助函评意见专业性评价略微优于获资助项目函评意见(两组之间无统计学差异, $P < 0.05$)。这或许是函评专

表4 申请人对指标“从函评意见中的收获”各档次评价所占百分率(%)

从函评意见中的收获	按学科		按项目类别		按项目是否获资助	
	呼吸	血液	面上	青年	获资助	未获资助
收获较大(A)	72.33	70.05	65.90	82.47	89.54	58.30
收获一般(B)	18.66	18.43	21.92	10.92	9.15	24.96
几乎没有收获(C)	9.01	11.52	12.18	6.61	1.31	16.74

表 5 学科对函评意见专业性评价“好或不足”所占百分率(%)

函评意见专业性	按学科		按项目类别		按项目是否获资助	
	呼吸	血液	面上	青年	获资助	未获资助
专业性较好或以上(A)	85.45	92.09	88.48	89.41	86.82	89.34
专业性不足或较差(B)	14.55	7.91	11.52	10.59	13.18	10.66

家在给出“不予资助”意见时,会从专业的角度给出尽可能充分的理由。而未获资助申请人对函评意见的专业性评价时,申请人认为只有 65.47% 的函评意见的专业性较好或以上,与学科评价函评意见有统计学差异($P < 0.05$),其原因或许与评价尺度的把握有关,申请人或许会从更苛刻的专业角度进行评价,而学科项目管理人员则只是判断函评意见是否从专业角度进行评价,当然也可能与未获资助申请者的心理因素有关。

按函评评议表中的“资助意见”分类统计(“函评资助意见”分“优先资助”、“可资助”与“不予资助”)。资助意见为“优先资助”且函评意见专业性好的占 83.98%,专业性不足的占 16.03%;资助意见为“可资助”且函评意见专业性好的占 83.83%,专业性不足的占 16.17%;资助意见为“不予资助”且函评意见专业性好的占 92.77%,专业性不足的占 7.24%。资助意见为“优先资助”与“可资助”之间没有统计学差异($P < 0.05$),二者与“不予资助”之间有统计学差异($P < 0.05$)。很明显,资助意见为“不予资助”的函评意见专业性评价较好的占比较高,达到约 93%,这也说明基金项目的同行评议专家给出“不予资助”意见时更为慎重,力求从专业的角度给出不予资助的理由(表 6)。

4 结论与讨论

对函评意见质量进行评估是一件有难度的事^[3],如何选择评议人和如何设置评价指标事关整个评价的科学性和公正性。所以在让申请人对反馈的函评意见进行评价的同时,学科管理人员也对所有

函评意见进行评价,希望能从不同角度分析并评价函评意见的质量,以了解函评意见质量的整体情况。在设置评价指标中力求指标易于判断、易于把握,而且方便评议人进行评价。

这次对医学科学部呼吸、血液两学科受理的面上、青年项目的函评意见质量的评价与整个国家自然科学基金委员会所受理的项目相比,只能反映一个很小范围内的大致情况,尚不具有普遍性。而且不同的学科领域的情况会有各自的特点,比如有的学科主要以基础研究为主,有的学科则主要以应用基础为主,函评专家的组成必然有差异,在评审项目时评审专家的侧重点会有所不同,因此申请人或学科项目管理人员评价函评意见时也会存在差异。

申请人与学科项目管理人员对函评意见的专业性均进行了评价,申请人的评价在学科间和项目类别间不存在统计学差异,但如果按项目是否获得资助统计看,获资助项目与未获资助项目之间在专业性评价方面存在明显的差异,未获资助项目申请人认为函评专家评审意见专业性明显劣于获资助申请人的评价,说明是否获得资助所产生的心理因素可能影响到了申请人的评价。而学科项目管理人员在对函评意见专业性进行评价时,在项目类别间和项目是否获资助之间,均没有统计差异。但在学科间存在差异,正如前面所述,可能与学科的特点有关,也可能与学科管理人员在评价时所把握的尺度差异有关。可以肯定的是,学科项目管理人员的评价不存在心理因素的干扰,所以比较客观些,但从专业的角度看,存在不及申请人判断专业程度高的可能性。

总体上看,申请人对函评意见的专业性认为“较好或以上”的所占比例接近 80%,认为“一般或不足”的超过 20%;对是否从函评意见中的收获看,“收获较大以上”的所占比例约 70%，“收获一般或没有收获”的约 30%。这些数据与 2011 年的《国家自然科学基金资助与管理绩效国际评估报告》中提供的数据比较接近,但满意度相对偏低^[2]。一方面说明基金项目的同行评议质量绝大多数是不错的,但是函评质量还有较大的提升空间;另一方面也说明,虽然已过去 8 年,现在的基金项目的同行评议质

表 6 按“函评资助意见”统计函评意见专业性“好或不足”所占百分率(%)

函评意见专业性	函评资助意见		
	“优先资助”	“可资助”	“不予资助”
专业性较好或以上(A)	83.98	83.83	92.77
专业性不足或较差(B)	16.03	16.17	7.24

量并没有得到进一步的提高,甚至有所下降。

我们曾针对国家自然科学基金委员会2014—2019年间同行评议专家对所评议项目的熟悉程度做过一次调研,选择“熟悉”的专家的百分率基本维持在63%~65%,属于所评议项目的小同行专家;约35%~37%的通讯评议专家对所评议项目的熟悉程度为“部分熟悉”和“不熟悉”,应不属于所评议项目的小同行专家。是否是所评议项目的小同行专家对能否从专业上给出科学、客观的评价十分重要,特别是申请人能否从函评意见中有所收获,这些数据也初步印证了申请人对函评意见的专业性并从中获益的评价的基本情况。因此,进一步提高函评专家的小同行百分率是提高同行评议质量的一项重要内容。当然,可能还有其他一些因素同样会影响函评意见的质量^[4],比如评审专家的学术水平、评审负荷以及与被评议项目申请人之间的利益冲突或相关等。

通过上述对呼吸、血液两个学科受理的项目的函评意见分析,说明保持和进一步提高同行评议质量仍然是基金项目同行评议中的一项很重要的任务,特别是在项目受理量不断增加而管理人员编制不能相应增加的情况下,更显得紧迫。

参 考 文 献

- [1] 朱志文, 于晟. 对同行评议质量与公正性的探讨. 地球科学进展, 1998, 13(1): 81—84.
- [2] 国家自然科学基金资助与管理绩效国际评估报告(2011). [2019-11-20]. http://www.nsf.gov.cn/nsfc/cen/pgbg/fj/fj20111102_02.pdf
- [3] Blockeel C, Drakopoulos P, Polyzos NP, et al. Review the 'peer review'. Reproductive Biomedicine Online, 2017, 35(6): 747—749.
- [4] 江虎军, 徐岩英, 朱蔚彤, 等. 同行评议制度的公正性与局限性. 中国科学基金, 2019, 33(4): 403—406.

The Studies on the Quality of Reviewers' Comments for Proposals Received by NSFC —by Analyzing Reviewers' Comments for Proposals Received in 2019 from Two Disciplines of Department of Health Sciences

Liu Yun^{1,2} Liang Xue^{1,3} Liao Danying^{1,4} Jiang Hujun^{1*}

1. Department of Health Sciences, National Natural Science Foundation of China, Beijing 100085

2. The Second Hospital Affiliated to Xi'an Jiaotong University, Xi'an 710004

3. The Fifth Hospital Affiliated to Guangzhou Medical University, Guangzhou 510700

4. Union Hospital Affiliated to Tongji Medical College, Huazhong Science and Technology University, Wuhan 430022

Abstract With the rapid growth of the number of proposals received by National Natural Science Foundation of China (NSFC), the quality of peer review show a trend to decline. So it is very significant for not only applicants but grant managers to improve the quality of peer review. We ask applicants to evaluate the value of reviewers' comments for their proposals from academic angle, while grant managers also evaluate all of reviewers' comments in a similar way, in order to outline the current general situation of peer review in NSFC. Our results indicate that the quality of peer review is far from perfect although generally good, and to improve the quality of peer review is still a major task for NSFC.

Keywords National Natural Science Foundation; reviewers' comments; quality evaluation

(责任编辑 齐昆鹏)

* Corresponding Author, Email: jianghj@nsfc.gov.cn