

卷烟烟气危害性指数研究

谢剑平¹, 刘惠民¹, 朱茂祥², 钟科军³, 戴亚⁴, 杜文⁵, 谢复伟¹, 缪明明⁶, 邓家云⁷,
聂聪¹, 邵学广⁸, 胡清源¹, 姚小军⁹, 李忠¹⁰, 张晓兵¹

1. 中国烟草总公司郑州烟草研究院, 郑州高新技术产业开发区枫杨街2号 450001

2. 军事医学科学院放射与辐射医学研究所, 北京市太平路27号 100850

3. 常德卷烟厂, 湖南省常德市洞庭大道西段199号 415000

4. 重庆烟草工业有限责任公司, 重庆市南岸区南城大道197号 400060

5. 长沙卷烟厂, 长沙市劳动中路426号 410014

6. 红塔烟草(集团)有限公司, 云南省玉溪市凤凰路112号 653100

7. 湖北中烟工业公司, 武汉市汉阳区十升路22号 430051

8. 南开大学, 天津市南开区卫津路94号 300071

9. 兰州大学, 兰州市天水南路222号 730000

10. 云南烟草研究院, 昆明市高新技术产业开发区科医路41号 650106

关键词: 卷烟; 主流烟气; 有害成分; 毒理学分析; 相关性分析; 危害性指数

摘要: 为评价卷烟产品的危害性, 研究建立了一种新的卷烟烟气危害性评价方法。用于分析评价的163个卷烟样品购自于中国市场。在分析卷烟主流烟气中29种有害成分(包括4种TSNAs、3种PAHs、8种羰基化合物、7种酚类物质、HCN、NQ、NO_x、NH₃、CQ烟碱和焦油等)以及4种毒理学指标(小鼠吸入急毒试验、细胞毒性试验、Ames试验和细胞微核试验)的基础上, 建立了烟气有害成分与毒理学指标的函数关系。通过采用无信息变量删除法和遗传算法, 筛选出了最具代表性的7种卷烟烟气有害成分, 即CQ、HCN、NNK、NH₃、B[a]P、苯酚和巴豆醛。多元线性回归分析表明, 7种有害成分对毒理学指标影响显著, 相关模型留一交叉验证的相对误差小于40%, 对小鼠吸入急毒试验、细胞毒性试验、Ames试验和细胞微核试验的R²值分别为0.524、0.595、0.504和0.571。

基于卷烟主流烟气中7种有害成分的释放量, 建立了一种新的卷烟烟气危害性指数:

$$H = \frac{Y_{CO}}{C_1} + \frac{Y_{HCN}}{C_2} + \frac{Y_{NNK}}{C_3} + \frac{Y_{NH_3}}{C_4} + \frac{Y_{B[a]P}}{C_5} + \frac{Y_{PHE}}{C_6} + \frac{Y_{CHO}}{C_7}$$

式中: Y——卷烟主流烟气有害成分释放量; C₁~C₇——参考值。

应用研究建立的危害性指数对上述163个卷烟危害性进行了评价, 结果表明, 危害性指数排序不同于卷烟焦油量排序。

中图分类号: TS472 文献标识码: A 文章编号: 1002-0861(2009)02-0005-11

Development of a novel hazard index of mainstream cigarette smoke and its application on risk evaluation of cigarette products

XIE JIAN-PING(1), LIU HUI-MIN(1), ZHU MAO-XIANG(2), ZHONG KE-JUN(3), DAI YA(4), DU WEN(5), XIE FU-WEI(1), MIAO MING-MING(6), DENG JIA-YUN(7), NIE CONG(1), SHAO XUE-

基金项目: 科技部资助项目“卷烟安全性综合评价指标体系研究”(2005DIB1J158), 国家烟草专卖局资助项目“卷烟危害性指标体系研究”(110200601041)。

作者简介: 谢剑平(1959—), 硕士, 研究员, 博导, 主要从事烟草化学研究与科研管理。E-mail: xiejp@ztj.com.cn

收稿日期: 2009-02-01

责任编辑: 刘立全 E-mail: llq@tobaccoinfo.com.cn 电话: 0371-67672637

1. Zhengzhou Tobacco Research Institute of CNTC Zhengzhou 450001, Henan, China
2. Institute of Radiation Medicine of the Academy of Military Medical Sciences, Beijing 100850
3. Changde Cigarette Factory, Changde 415000, Hunan, China
4. Chongqing Tobacco Industrial Co., Ltd, Chongqing 400060, China
5. Changsha Cigarette Factory, Changsha 410014, China
6. Hongta Tobacco Group Co., Ltd, Yuxi 653100, Yunnan, China
7. China Tobacco Hubei Industrial Company, Wuhan 430051, Hubei, China
8. Nankai University, Tianjin 300071, China
9. Lanzhou University, Lanzhou 730000, China
10. Yunnan Research Institute of Tobacco Science, Kunming 650106, Yunnan, China

Keywords: Cigarettes; Mainstream smoke; Toxic chemicals; Toxicology; Correlation; Risk evaluation; Hazard index

Abstract: A novel hazard index of mainstream cigarette smoke for risk evaluation of cigarette products was developed. 163 commercial cigarettes, including international and Chinese brands, American blend and Virginia blend cigarettes, purchased in China were investigated. The yields of 29 toxic chemicals (4 TSNA's, 3 PAH's, 8 carbonyls, 7 phenolic compounds, HCN, NO, NO₂, NH₃, CO, Nic and Tar) in mainstream cigarette smoke were determined, and the toxicity of cigarette smoke was measured employing a bacterial mutagenicity assay (Ames *Salmonella* mutagenicity assay) with cigarette smoke condensate (CSC), the MTT cytotoxicity assay with CSC, the micronucleus assay in vivo with CSC, and a mouse inhalation study.

In order to establish the functional relationships between smoke yields and smoke toxicity, the correlation between the yields of 29 toxic chemicals and 4 toxicity assay results were investigated. The key toxic chemicals were screened with uninformative variable elimination (UVE) and genetic algorithm. The results indicate that the contributions of 7 toxic chemicals (CO, HCN, NNK, NH₃, B[a]P, Phenol and crotonaldehyde) to the toxicity of mainstream cigarette smoke were significant. The correlation models between 7 key chemicals and 4 toxicity assay response were developed by multiple linear regression analysis. The R square for Ames assay, MTT assay, in vivo micronucleus assay, and mouse inhalation study were 0.504, 0.595, 0.571, and 0.524, respectively. The relative errors of leave-one-out cross-validation for all models were less than 40%.

Based on the yields of 7 toxic chemicals in mainstream cigarette smoke, a novel hazard index was developed to evaluate the risk of mainstream cigarette smoke; the calculation is as follows:

$$H = \frac{Y_{CO}}{C_1} + \frac{Y_{HCN}}{C_2} + \frac{Y_{NNK}}{C_3} + \frac{Y_{NH_3}}{C_4} + \frac{Y_{B[a]P}}{C_5} + \frac{Y_{PHE}}{C_6} + \frac{Y_{CRO}}{C_7}$$

Where: Y—Yields of toxic chemicals in mainstream smoke; C₁ ~ C₇—Constant

Applying this index, the hazard of mainstream cigarette smoke of main cigarette products from the China market was evaluated. The results indicate that the product ranking according to the hazard index is different from the ranking according to tar deliveries, so the hazard index may be a more reasonable way to evaluate the risk of cigarette products.

1 引言

卷烟烟气是一种极其复杂的混合物,它是在卷烟抽吸过程中由烟草燃烧、裂解和蒸馏而产生的^[1]。早期的烟草烟气分析主要集中在烟草烟气中的化学成分与卷烟产品质量关系方面的研究上,近年来,随着社会各界

对吸烟与健康关系的日益重视,烟草和烟气危害性的研究逐渐成为烟草化学和毒理学研究的主要方向之一。

1.1 烟草和烟气中的主要有害成分

在烟草和烟气中鉴定出的 5000 多种化合物中,有害和致癌成分仅 100 多种。1954 年,Cooper 等人从卷烟烟气中分离鉴定出苯并[a]芘^[2],这是烟草烟气中

鉴定出的第一种致癌成分。从那时起烟草烟气中有害成分的分析鉴定研究得到了广泛的开展,不断有新的有害成分被鉴定出来。第一份卷烟烟气有害成分名单收录在1964年美国公共健康服务部咨询委员会有关吸烟与健康的报告^[3]中。自第一份名单出现后,几乎每年都有新的主流烟气有害物质名单发表,或对以前发表过名单的修订再公布。最著名的是Hoffmann和Hecht1990年发表的“43种成分”名单^[4]。随着研究的发展,最近几年Baker和Proctor^[5]、Hoffmann^[6-7]等、Smith^[8-10]等也分别发表了补充的和修订的烟气有害物质名单。2002年,Rodgman和Green^[11]对烟气中已报道的有害成分进行了总结,认为在卷烟烟气中共存在149种有害成分。

1998年,加拿大政府通过立法,要求卷烟生产商定期检测卷烟主流烟气中46种有害成分的释放量^[12-13],并将结果向社会公布。这一名单实际是一个修正的Hoffmann名单,在世界范围内造成了很大的影响,名单中的有害成分得到了医学界和烟草行业的普遍认可。

1.2 卷烟烟气毒性评价研究概况

1.2.1 卷烟烟气有害成分(纯品)致癌性评价结果

国际癌症研究机构(IARC)依据纯化学品的试验结果对加拿大检测名单中有害成分的致癌性评价结果^[14],认为在46种有害成分中,11种成分为人体致癌物质(第1类),5种成分为可疑的人体致癌成分(2B类),9种成分为致癌性不明确的化合物(第3类)。因此,对于纯化学品来说,以上的25种化学成分应当为卷烟烟气中最重要的有害成分。

烟草烟气是极其复杂的混合物,这些有害成分之间相互作用,会对以上化学成分的实际毒性产生影响^[15-17],由单一物质所做的毒理学检测结果几乎是不可能准确外推的。因此,以单纯使用纯化学物质得到的毒性评价结果来评价卷烟烟气这种复杂体系的危害性并不科学。另外,这些有害成分在烟气中的释放量差异极大,如果不考虑某种有害成分在烟气中的实际释放量,而只是定性地对其进行危害性评估,没有实际意义。

1.2.2 卷烟主流烟气危害性毒理学评价方法

1999年美国食品药品监督管理局(FDA)要求美国科学院医学研究所(IOM)制定一套科学的方法,用于评估长期使用药物产品或者烟草替代产品(用于降低而不是消除烟草危害)的安全性和效果。研究结果认为,毒理学评价数据可以在一定程度上支持“低危害产品”^[19]的评价。

CORESTA(国际烟草科学研究合作中心)烟草烟气

体外毒性测试工作组^[18],根据国际认可的体外毒性测试准则并加以修改,以适应烟草烟气的性质和独特特性,提出了3种毒理学测试方法,即:细菌诱变分析——艾姆斯门氏菌诱变分析(Ames Test, Ames试验)、细胞遗传学突变的哺乳动物细胞分析——微核分析(Micronucleus Test)、染色体畸变分析或I5178Y小鼠淋巴瘤分析、哺乳动物细胞系细胞毒性分析——中性红细胞毒性分析(Neutral Red Test, 中性红试验)^[20]。目前国际上已经普遍接受这三种毒理学测试方法。

1.3 卷烟烟气危害性评价存在的问题

人们已经普遍接受了这样一种观点,即通过降低卷烟烟气中有害成分的释放量可以有效地降低卷烟对人体的危害,于是各种各样的“低危害”卷烟产品不断出现。但是,如何评价“低危害”卷烟产品,或者到底什么卷烟是“低危害”卷烟,国内外尚没有形成统一的标准。目前,国内外普遍以焦油、烟碱、一氧化碳等主流烟气常规化学指标作为评判标准。但卷烟焦油中只有极少量的物质是有害成分,因此单纯用这些指标评价就不够科学、全面、客观。近年来,对卷烟产品的危害性评价一般均进行有害成分分析和毒理学测试,但有害成分分析往往只偏重于少数几项有害成分,体外毒理学指标测定以及动物体致癌、致突变试验的选择较为随意。出现了检测指标较单一,评价方法较混乱的情况,缺乏对卷烟危害性的综合评价,尤其缺乏卷烟危害性的整体评价方法,因此评价结果缺乏较强的说服力,不能客观真实地反映卷烟产品的危害性。同时,也导致减害技术的研究和评价只偏重于某一种或几种有害成分的减少,忽略了卷烟烟气危害性的整体降低。

要解决上述问题,就需要建立新的、比较全面客观的卷烟危害性综合评价指标,以期为社会公众提供准确翔实的卷烟烟气危害性数据,引导烟草行业开发低危害卷烟,整体降低卷烟烟气危害性,减少吸烟对消费者健康造成的危害。

2 实验

2.1 研究目的和技术路线

通过对卷烟主流烟气中主要的29种有害成分和主流烟气体外毒理学指标的研究,结合卷烟烟气的急性致死性试验,进行相关性分析,确定卷烟主流烟气中最主要的有害成分,建立一种新的卷烟主流烟气危害性评价指标,从整体上对卷烟烟气危害性进行评价。

2.2 分析指标的确定

2.2.1 烟气化学成分指标的确定

在加拿大检测名单的46种有害成分中,重金属主

要是烟草生长过程中由环境吸收而来,与其它植物相比没有特殊性,并且从烟草到烟气的传输率也相对较低,因此各国对烟草中重金属的重视程度相对较低。芳香胺类化合物在烟气中的释放量非常低,目前的测试技术难于准确测定;挥发性有机化合物尽管在烟气中的释放量比较高,但由于它们的挥发性太强,很难选择适宜的烟气捕集方法,分析结果的变异性非常大。其它的 29 种有害成分(表 1)目前具有比较成熟的测定方法,可确保得到准确的测定数据,并且毒性较强、在主流烟气中的释放量较大,适于本研究。

2.2.2 毒理学指标的确定

对于毒理学分析,在 CORESTA 推荐的三项毒理学指标(Ames 试验、微核分析和细胞毒性试验)之外,增加了卷烟烟气动动物急性毒性实验,以增强卷烟烟气毒理学评价的全面性和合理性。

2.3 分析方法

在所测定的 29 种有害成分和 4 种毒理学指标中,焦油、烟碱和 CO 采用国际标准方法,其余方法采用了郑州烟草研究院研究建立的分析方法(表 2 和表 3)。

焦油、烟碱和 CO ISO4387《卷烟 用常规分析用吸烟机测定总粒相物和焦油》、ISO10315《卷烟 总粒相物中烟碱的测定方法 气相色谱法》、ISO10362-1《卷烟 总粒相物中水分的测定方法 气相色谱法》、ISO8454《卷烟烟气质相中一氧化碳的测定 非散射红外法》。

2.4 卷烟样品

共收集卷烟样品 163 种,其中焦油量为 1~5 mg 样品 5 个,占 3.1%; 6~10 mg 样品 24 个,占 14.7%; 11~13 mg 样品 31 个,占 19.0%; 14~15 mg 样品 103 个,占 63.2%。

表 1 有害成分分析列表^①

常规成分	焦油、烟碱、CO
无机成分	HCN、NH ₃ 、NQ、NO _x
多环芳烃	苯并[a]芘、苯并[a]蒽、Chrysene
烟草特有亚硝胺	NNN、NAT、NAB、NNK
挥发性羰基物	甲醛、乙醛、丙酮、丙烯醛、丙醛、巴豆醛、2-丁酮、丁醛
挥发酚	对苯二酚、间苯二酚、邻苯二酚、苯酚、对甲酚、间-甲酚、邻甲酚

注:①在实际分析测试中,对甲酚和间-甲酚很难分离,在计算时将两者合并,因此实际测试的是 28 种成分。

3 数据分析与讨论

3.1 主流烟气有害成分释放量及毒理学指标分布概况

图 1 为 163 种卷烟主流烟气部分有害成分(焦油、CO、HCN、NNK、NH₃、B[a]P、苯酚和巴豆醛)释放量的分布图。大部分有害成分的分布比较均匀,各样品有害成分释放量主要集中在总平均值的附近,在低端和高端分布的样品数量基本一致。例如焦油主要分布在 12~16 mg/支之间,CO 主要分布在 12~17 mg/支之间,HCN 主要分布在 87~156 μg/支之间,NH₃ 主要分布在 5.5~9.7 μg/支之间,B[a]P 主要分布在 5.6~9.9 ng/支之间,苯酚主要分布在 9~27 μg/支之间,巴豆醛主要分布在 42~67 μg/支之间。而 NNK 的分布与其它有害成分有较大的差异,分布偏于低端,绝大部分卷烟样品烟气中 NNK 释放量低于 10 ng/支,少量样

表 2 卷烟烟气有害成分分析方法

分析物	捕集方法	检测方法
3 种多环芳烃	剑桥滤片	GC-MS
4 种 TSNA ^s	剑桥滤片	GC-TEA
7 种酚类成分	剑桥滤片	HPLC-FLD
8 种羰基化合物	DNPH 处理的剑桥滤片	HPLC-DAD
氨	剑桥滤片和气体吸收瓶(0.005 mol/L HCl)	IC
氢氰酸	剑桥滤片和气体吸收瓶(0.1 mol/L 的 NaOH)	连续流动分析仪
氮氧化物	逐口气相物分析	化学发光法

表 3 卷烟烟气毒理学检测方法

毒理学指标	染毒方式	检测方法	结果形式
细胞毒性	CSC ^①	MTT 法	细胞增殖死亡的平均剂量(D ₃₇)
细菌致突变性	CSC*	微量波动 AMES 法	50% 细菌回复突变的剂量(M ₃₀)
诱发小鼠微核形成率	小鼠体内 CSC* 染毒	抗体荧光标记法	每 1000 个成熟红细胞中的微核细胞数(‰)
动物吸入急性毒性	全烟气	—	动物半数死亡时间(T _{1/2})

注:① CSC 为全烟气的乙酸乙酯和细胞培养液的吸收溶液。

表 4 卷烟样品盒标焦油量统计

卷烟类型	盒标焦油量 (mg)			
	1~5	6~10	11~13	14~15
国内烤烟型	0	4	17	95
国内混合型	2	7	5	6
国外烤烟型	1	6	6	1
国外混合型	2	7	3	1
总计	5	24	31	103

品的 NNK 释放量相当高。这是由于烤烟型卷烟和混合型卷烟在 TSNA 释放量上明显的差异造成的。

图 2 为 163 种卷烟烟气 4 种毒理学指标分布图, 动物急毒和微核试验结果的分布较为均匀, 而细胞毒性试验和 Ames 试验结果的分布比较偏向低端。

3.2 影响烟气毒理学指标最主要有害成分的确定

以实验测定数据 (即 163 个卷烟样品主流烟气中的 28 种主要有害成分, 4 项毒理学指标) 为基础, 应用各种数理统计方法, 考察烟气中 28 项主要有害成分与 4 项毒理学指标之间的关系, 并对 28 种有害成分进行筛选, 找到影响卷烟主流烟气危害性最主要的化学成分指标。化学成分的筛选分别采用了无信息变量删除 (UVE-uninformative variable elimination) 方法^[21-22]、遗传算法^[23]和 LG-GA 算法 (改良遗传算法) 通过考察变量 (即化学成分) 对定量模型稳定性或预测能力的影响, 分别对 28 种化学成分进行评价和筛选。定量模型采用多元线性回归模型。

为了保证模型的可靠性, 在研究过程中对所建立的模型均进行了统计检验, 并采用留一法交叉验证 (LOOCV-leave one out cross validation) 对模型的预测能力进行了考察。

最后, 综合考察 3 种统计方法得到的变量筛选结果, 并结合毒理学指标与化学成分之间以及 28 种化学成分之间的相关关系, 得到对烟气毒理学指标影响最大的有害化学成分。

3.2.1 测试数据的相关性分析

3.2.1.1 毒理学指标与化学成分之间的线性相关性

图 3~图 6 分别为 4 项毒理学指标与 28 种有害成分之间的相关性。可以看出: 4 项毒理学指标与化学成分之间均具有一定的相关性, 其中, 微核试验与 28 种化学成分的相关性最强。

3.2.1.2 化学成分之间的相关分析

初步分析表明, 化学成分之间具有较强的相关性, 并且可以明显地按照化合物的种类分为几组, 即: ①焦

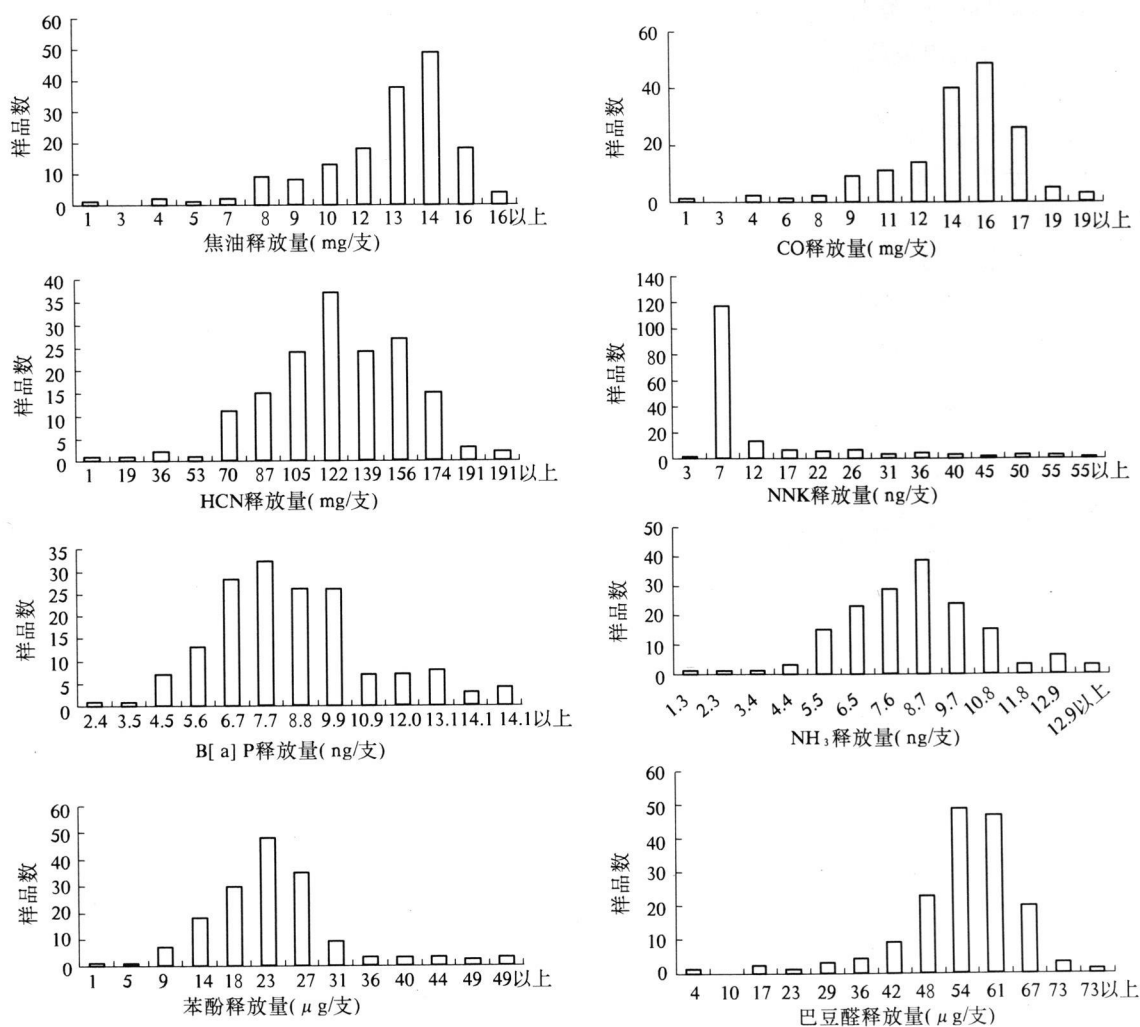


图 1 163 种卷烟样品烟气部分有害成分释放量分布图

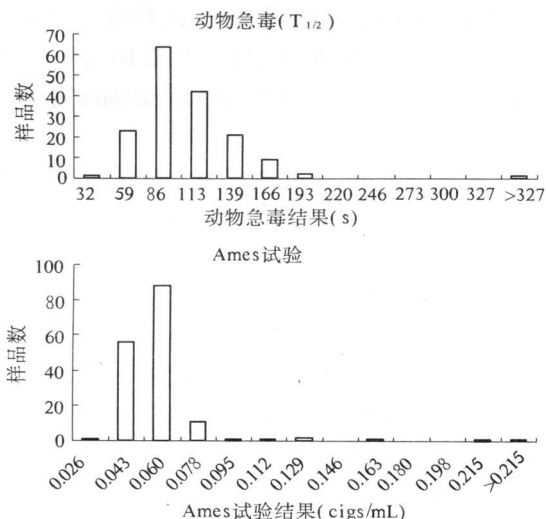
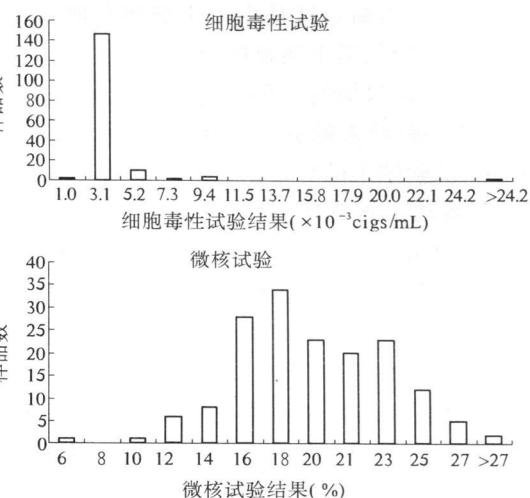


图 2 163种卷烟烟气毒理学指标分布图

油、烟碱、CO(常规成分)；②HCN；③NNN NAT NAB NNK(亚硝胺类)；④NH₃；⑤B₁ q P B₁ q A Chrysene(多环芳烃类)；⑥对苯二酚、间苯二酚、邻苯二酚、苯酚、对-间甲酚、邻甲酚(酚类)；⑦NQ NO_x(氮氧化物)；⑧甲醛、乙醛、丙酮、丙烯醛、丙醛、巴豆醛、2-丁酮、丁醛(醛、酮类)。

因此，分别对第①、③、⑤、⑥和⑧类化学成分之间的相关性进行了分析。图 7~图 11 分别为焦油、Chrysene NAT邻苯二酚、乙醛与其它化学成分之间的相关性。可以看出，焦油与 CO及酚类物质之间存在着高度相关；多环芳烃之间存在着高度相关；亚硝胺类化合物和含氮类化合物之间存在高度相关；酚类之间存



在着高度相关；醛类和酮类之间存在着高度相关。

3.2.2 无信息变量删除(UVE)方法建模结果

3.2.2.1 UVE方法的原理

UVE方法的基本原理是根据每个变量的稳定性对

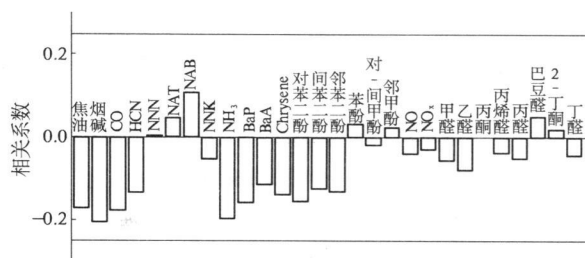


图 5 Ames 试验与 28 种化学成分之间的相关性

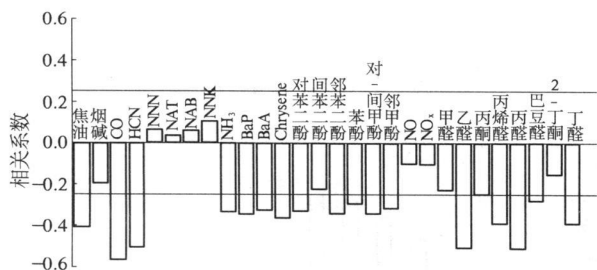


图 3 动物急毒与 28 种化学成分之间的相关性

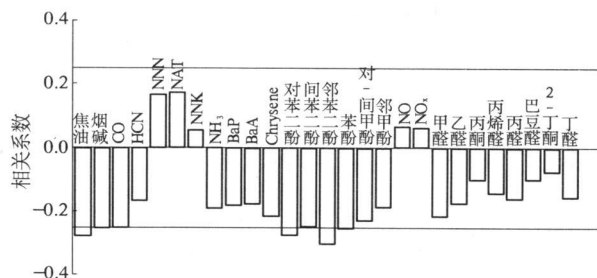


图 4 细胞毒性试验与 28 种化学成分之间的相关性

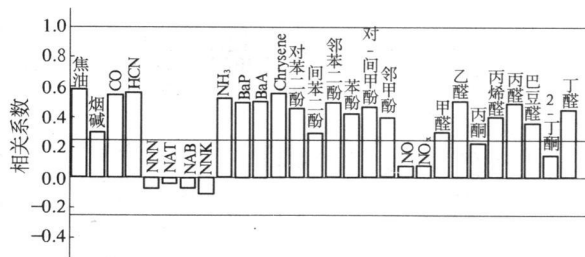


图 6 微核试验与 28 种化学成分之间的相关性

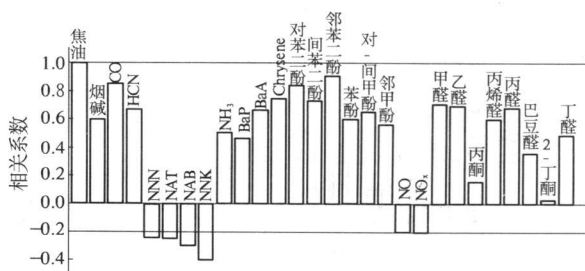


图 7 焦油与其它组分之间的相关性

变量进行评价,并根据稳定性大小进行变量的筛选。
变量的稳定性由下式计算:

$$s_j = \frac{\text{mean}(\beta_j)}{\text{std}(\beta_j)}, \quad j=1, \dots, p$$

式中: p ——变量个数; β ——多元线性回归模型的系数矩阵,由留一法交叉验证得到,每一列对应一个变量,每一行表示一个模型。 $\text{mean}(\beta_j)$ 和 $\text{std}(\beta_j)$ ——变量对应的 n (模型个数)个模型的回归系数 β_j 的平均值和标准偏差。

由稳定性的计算公式可以看出,稳定性表征了每个

变量在交叉验证过程中系数值的变化。显然,平均值大且标准偏差小的系数所对应的变量较稳定,而平均值小且标准偏差大的系数所对应的变量则不稳定。因此,稳定性高的变量(即化学成分)回归系数较大且在交叉验证过程中变化较小,而稳定性低的变量则回归系数较小且在交叉验证过程中变化较大。所以,稳定性较高的变量对模型具有关键作用,而不稳定的变量对模型的贡献具有随机性。

在本研究中,通过 UVE方法对 28种化学成分的稳定性的评价,然后根据稳定性的大小筛选重要的化学成分。

3.2.2.2 数据预处理

为了得到可靠的计算结果,对原始测量数据进行预处理。首先,在 28种化学成分中,由于各成分的量纲及变化幅度不同,其绝对值大小相差很大。为了消除量纲和变化幅度不同带来的影响,对测量数据进行了调整,即各列化学成分的释放量分别除以其对应的最大值。另外,采用了一种基于交叉验证的统计方法进行了奇异样本的识别,结果表明:烤烟型卷烟样本中的 K050和 K116为奇异样本,混合型卷烟样本中的 H004、H011、H021、J016和 J019为奇异样本。

3.2.2.3 UVE方法的计算过程

分别基于扣除奇异样本后的 114个烤烟样本和 156个全部卷烟样本,采用 3-fold交叉验证方法计算每个变量(即化学成分)的稳定性。即对每个毒理学指标计算 3次,其中 2/3的样本建模,剩余的 1/3样本用于预测。4个毒理学指标共计算 12次,每次计算筛选出稳定性排在前 10位的化学成分,然后按照化学成分在 12次计算中出现的频数进行统计。为了更加强微核试验和 AMES试验两个毒理学指标的结果,在形成结论时重点考虑了此两个指标的计算结果。图 12为烤烟型样品处理结果示意图。

为了进一步简化所保留的化学成分,对选中的同类化学成分进行了相关性分析,对于相关性较强(相关系数大于 0.8)的化学成分进一步删除。

结合上述分析,烤烟型卷烟模型的最终变量筛选结果为:CO、HCN、NNK、BaP、Chrysene、邻苯二酚、对-间甲酚、邻甲酚、甲醛、乙醛、丙醛、丙酮、巴豆醛、2-丁酮、丁醛。

全部样本的最终变量筛选结果为:CO、HCN、NNK、NH₃、Chrysene/BaP、邻苯二酚、对-间甲酚、丙酮、丙醛、巴豆醛。

3.2.3 遗传算法 (Genetic Algorithm)结果

3.2.3.1 遗传算法的原理

遗传算法是 20世纪 70年代初兴起的一种新型数

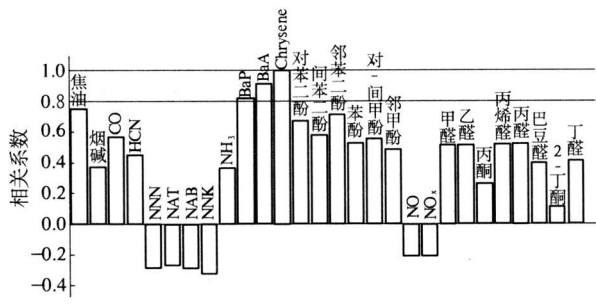


图 8 Chrysene 与其它组分之间的相关性

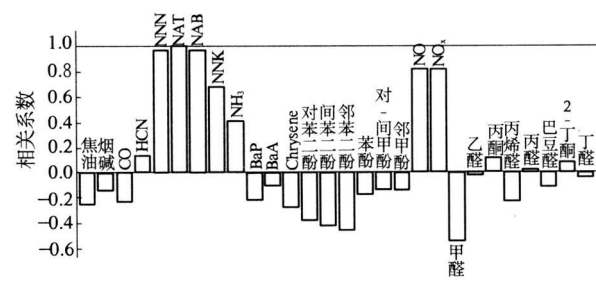


图 9 NAT 与其它组分之间的相关性

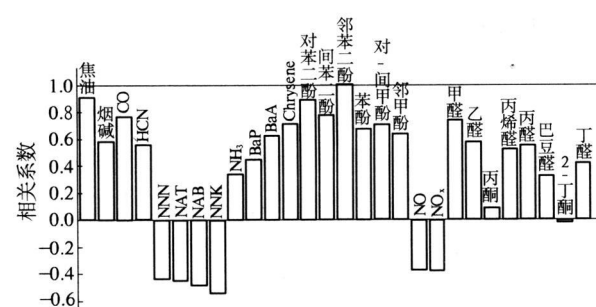


图 10 邻苯二酚与其它有害组分之间的相关性

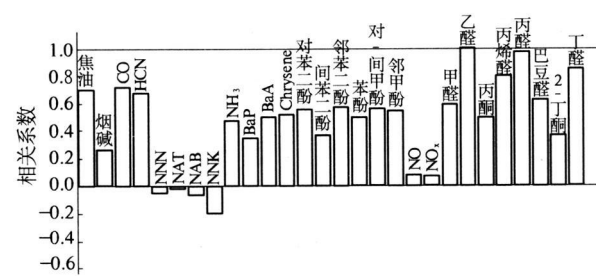


图 11 乙醛与其它有害组分之间的相关性

理统计方法,它基于对生物界自然选择和自然遗传机制的模拟来解决问题,是一种具有高度的并行、随机和自适应性的搜索方法。它主要用于处理最优化问题和机器学习,特别适合于处理高度复杂的高维搜索问题。遗传算法在各领域中的应用研究始于 20 世纪 80 年代,目前遗传算法的应用已非常广泛,如图象处理、参数计算、模式识别、网络设计、气象分析、组合优化等等。在化学领域中,遗传算法已成功地应用于参数优化、光谱模拟和光谱解析、多变量分析、分子构象分析等。

在本研究中,优化的目标为选择对模型具有重要性的化学成分,可能的答案为各种化学成分的一种组合。因此,采用一个 m 位 (m 为变量数) 的 0、1 编码即可表示一种组合,用 0 表示该变量未被选中,而 1 表示该变量被选中。评价的标准为采用所选择的变量建立的多元回归模型的预测能力,即交叉验证结果与测量值的相关系数或称为确定系数 (R^2)。

3.2.3.2 数据预处理

(1) 奇异样本的剔除。利用遗传算法建立简单预测模型,将预测值偏差比较大的样本作为奇异值剔除。最终确定的样本为 K028、K050、K063、K116、H004、J016、J023、J026。剩余样本数 155 个,其中包含 112 个烤烟样本,43 种混合型卷烟。

(2) 交叉项的问题。数据处理分析过程中,考虑引入交叉项后进行变量选择,但计算结果表明交叉项对于毒理学指标影响不大。因此,本项研究只考虑了二

次项和三次项,用于非线性因素的校正,即:在变量筛选过程中,不是对 28 种化学成分进行选择,而是对 28 种化学成分以及它们的平方、立方 (共 84 项) 进行选择。

3.2.3.3 变量筛选结果

根据 4 种毒理学指标建模结果,对 28 项有害成分指标进行筛选,以所建立模型 R^2 最高的模型为主要依据,烤烟型样品所筛选出的化学成分为:焦油^{*}, CO, HCN, NNN, NAT, NAB^{*}, NNK^{*}, NH₃, Chrysene, B[a]a^{*}、对-间甲酚、苯酚、邻甲酚、对苯二酚^{*}、邻苯二酚^{*}、NQ, NO₂, 2-丁酮、丙醛、巴豆醛,其中注有星号的化学成分为只出现在动物急毒和细胞毒性试验两个毒理学指标模型中的化学成分。为了重点考虑 Ames 试验和微核试验的筛选结果,标有星号的化学成分在简化模型时将优先被删除。

为了进一步简化所保留的化学成分,对选中的同类化学成分进行了相关性分析,对于相关性较强 (相关系数大于 0.80) 的化学成分可以进一步删除。

烤烟模型的最终变量筛选结果为: CO, HCN, NAT, NH₃, Chrysene, 苯酚、丙醛、巴豆醛。

同样可以得到全部卷烟模型的最终变量筛选结果为: CO, HCN, NNK, NH₃, B[a]a^{*} 苯酚、对苯二酚、巴豆醛。

3.2.4 改进遗传算法 (IG-GA)

3.2.4.1 遗传回归建模方法 (IG-GA) 原理

LongGelabHG 软件是上海烟草 (集团) 公司技术中心张龙根先生根据自创的筛选优化计算方法自编的利用 Microsoft Excel 运行的遗传回归计算软件,可用于较复杂系统的建模和预测。根据统计分析数据对象性质,将数据进行数据变换,使复杂的多元非线性关系数据转化为可以线性分析的数据;然后进行以最小二乘法为基础的多元线性回归分析;判断回归结果的误差和可靠性,并以遗传算法的方法对回归方案进行优化计算,选择优化方案,最后尽可能选取可能是最佳的方案进行建模和预测。再进行回归假设前提相符性验证和内部交叉验证、外部验证,以观察模型的正确性和预测误差的大小。采用建立的数据模型可以进行参数估计。软件可以选择多种不同方案进行自动计算,建立预测能力较强的分析模型。进而利用模型估计,可分析出一些明确的参数关系。

本项研究根据 163 个各种品牌规格卷烟样品的检测数据,就烟气的焦油、TSNA、PAH 和酚等 28 个烟气化学组分指标对微核试验、细菌半数致突变剂量、细胞毒性试验和动物急毒 4 个生物试验指标的关系用

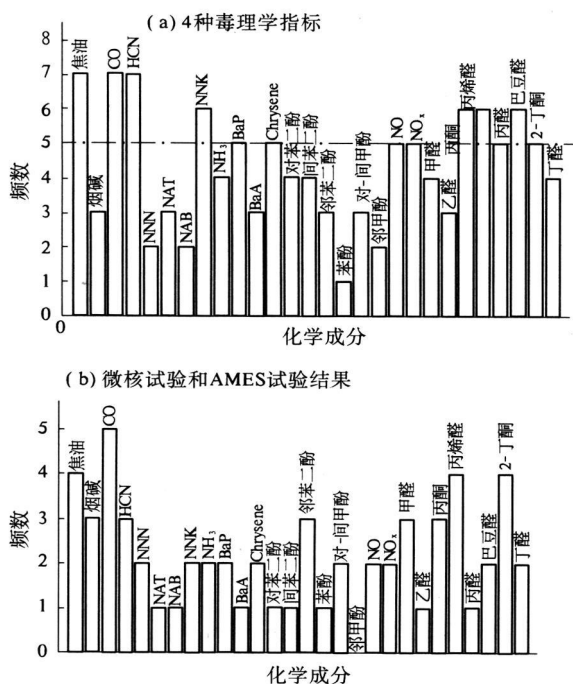


图 12 烤烟样本模型中不同变量的出现频次

LongGElabHC软件分别进行遗传回归分析,建立了预测模型,取得了较好结果,可供进一步研究分析参考和实际应用。为突出重要指标并避免叙述烦琐,以下主要对“微核试验”和“Ames试验”进行分析。

3.2.4.2 变量的筛选

首先根据 28 项指标所建立模型所得到的各自变量与因变量的对应关系计算各自变量线性拟合标准化斜率,并按斜率绝对值排序记分,计算各自变量对应“细菌半数致突变剂量”和“微核试验”的排序得分之和,最后按得分之和排序,取前 7 项对应的自变量。结果为:对 一间甲酚、巴豆醛、苯酚、2-丁酮、邻甲酚、NQ 丙醛。该变量组的模型标为“wh156 彡 k”。如果分别取“细菌半数致突变剂量”和“微核试验”排序分(排序分从大到小)前 7 项,并计算各自变量得分之和,再按得分之和排序,取前 7 项,得到结果为:对 一间甲酚、NO_x 巴豆醛、NQ Chrysene HCN 苯酚。该变量组的模型标为“wh156 彡 l₂”。

因此,对于全部样本模型,综合“wh156 彡 k”和“wh156 彡 l₂”的结果,筛选结果为:HCN Chrysene NQ NO_x 苯酚、邻甲酚、对 一间甲酚、丙醛、巴豆醛、2-丁酮。

3.2.5 对 4 种毒理学指标贡献最大的有害成分指标

表 5 列出了 3 种统计方法方法对 28 种化学成分筛选结果。可以看出三种方法所得到的结果基本一致。

图 13 为三种方法筛选结果出现频次图,从中可以看到,出现频次最高的化学成分为 CO HCN NNK NH₃、B₁ q P 苯酚、巴豆醛、Chrysene 和丙醛,结合相关性分析的结果,并重点考虑全部样本的筛选结果,对筛选的化学成分进一步简化,删除 Chrysene 和丙醛,可形成如下结论:对烟气毒理学指标影响最为重要的有害化学成分是:CO HCN NNK NH₃、B₁ q P 苯酚、巴豆醛。

3.2.6 数据统计结果的验证

为了进一步验证所得结论的可靠性,采用遗传算法对所选择的 7 种化学成分以及它们的倒数项、平方项、立方项、交叉项进行了模型验证。验证结果表明,所有的预测结果误差都小于 40%。对于 4 种毒理学指标预测模型的 R² 分别为:0.524、0.595、0.504、0.571。

吸入急毒: $Y = 5.38 \times 10^1 C_{H \cdot q \cdot P} - 1.93 \times 10^{-2} C_{HCN} C_{NNK} - 1.75 \times 10^{-1} C_{NNK} C_{H \cdot q \cdot P} + 1.42 \times 10^{-1} C_{NNK} C_{\text{苯酚}} + 5.12 \times 10^{-2} C_{\text{苯酚}} C_{\text{巴豆醛}} - 6.46 C_{H \cdot q \cdot P}^2 - 4.26 \times 10^{-2} C_{\text{苯酚}}^2 - 6.47 \times 10^{-3} C_{CO} + 7.23 \times 10^{-4} C_{NNK}^3 + 2.38 \times 10^{-1} C_{H \cdot q \cdot P}^3 + 3.20 \times 10^2 C_{\text{巴豆醛}}^{-1} - 4.53$
 $R^2 = 0.524 \quad R_{CV}^2 = 0.660 \quad RMSE_{CV} = 11.98$

细胞毒性: $Y = 7.31 \times 10^{-5} C_{\text{巴豆醛}} - 2.28 \times 10^{-4} C_{NH_3} + 5.05 \times 10^{-7} C_{HCN} C_{\text{苯酚}} - 6.15 \times 10^{-7} C_{HCN} C_{\text{巴豆醛}} - 7.74 \times 10^{-6} C_{NNK} C_{H \cdot q \cdot P} + 8.73 \times 10^{-6} C_{NNK} C_{NH_3} + 2.15 \times 10^{-5} C_{H \cdot q \cdot P} C_{NH_3} + 1.63 \times 10^{-6} C_{NNK}^2 - 4.41 \times 10^{-8} C_{NNK}^3 - 2.68 \times 10^{-8} C_{\text{苯酚}}^3 + 7.46 \times 10^{-3} C_{H \cdot q \cdot P}^{-1} + 1.12 \times 10^{-2} C_{\text{苯酚}}^{-1} - 2.84 \times 10^{-4}$

表 5 三种统计方法筛选结果总结

方法	样本	筛选结果
UVE	烤烟	CO HCN NNK B ₁ q P NQ 丙酮、丙醛、巴豆醛、邻苯二酚 *
	全部	CO HCN NH ₃ 、NNK Chrysene/B ₁ q P*, 邻苯二酚、对 一间甲酚、丙酮、丙醛、巴豆醛
GA	烤烟	CO HCN NAT NH ₃ 、Chrysene 苯酚、丙醛、巴豆醛
	全部	CO HCN NNK NH ₃ 、B ₁ q P 苯酚、对苯二酚、巴豆醛
LG—GA	全部	HCN Chrysene NQ NO _x 苯酚、邻甲酚、对 一间甲酚、丙醛、巴豆醛、2-丁酮

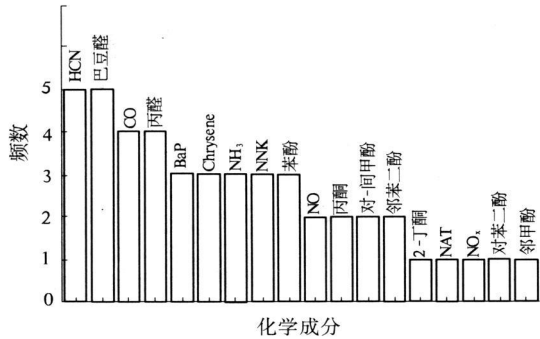


图 13 三种方法筛选结果总结

$R^2 = 0.595 \quad R_{CV}^2 = 0.702 \quad RMSE_{CV} = 3.25 \times 10^{-4}$
AMES: $Y = -3.33 \times 10^{-3} C_{NNK} + 4.60 \times 10^{-3} C_{B_1 \cdot q \cdot P} - 6.73 \times 10^{-6} C_{HCN} C_{NNK} - 1.20 \times 10^{-5} C_{HCN} C_{\text{巴豆醛}} - 8.60 \times 10^{-5} C_{NNK} C_{H \cdot q \cdot P} + 1.91 \times 10^{-4} C_{NNK} C_{\text{苯酚}} - 1.64 \times 10^{-4} C_{H \cdot q \cdot P} C_{\text{苯酚}} + 7.82 \times 10^{-5} C_{\text{苯酚}} C_{\text{巴豆醛}} + 3.43 \times 10^{-5} C_{NNK}^2 - 1.78 C_{HCN}^{-1} + 6.49 \times 10^{-1} C_{\text{苯}}^{-1} + 2.16 \times 10^{-2}$
 $R^2 = 0.504 \quad R_{CV}^2 = 0.582 \quad RMSE_{CV} = 6.72 \times 10^{-3}$
微核: $Y = -3.05 \times 10 C_{H \cdot q \cdot P} - 3.72 \times 10^{-3} C_{NNK} C_{\text{苯酚}} + 3.33 \times 10^{-3} C_{\text{苯酚}} C_{\text{巴豆醛}} + 1.64 \times 10^{-2} C_{CO} + 2.70 C_{H \cdot q \cdot P} + 3.74 \times 10^{-2} C_{NH_3} - 8.18 \times 10^{-2} C_{H \cdot q \cdot P}^3 - 5.37 C_{NNK}^{-1} - 2.35 \times 10^2 C_{H \cdot q \cdot P}^{-1} + 1.55 \times 10^2$
 $R^2 = 0.571 \quad R_{CV}^2 = 0.711 \quad RMSE_{CV} = 1.58$

4 卷烟危害性定量评价方案

4.1 卷烟主流烟气危害性定量评价指数的提出

根据以上的数据分析,确定了影响卷烟危害性最重要的7项有害成分指标,即CO、HCN、NNK、NH₃、B[a]P、苯酚、巴豆醛。

为综合评价卷烟的危害性,有必要研究确定一种评价指数来表征卷烟烟气对人体的危害性大小。从不同角度,可以提出各种各样的评价指数建立方法。任何一种评价方法能够为各方所普遍接受而没有任何异议是一件非常困难的事情。

由于4项毒理学指标分别代表了卷烟烟气生物学效应不同的评价方面,因此评价方案必须考虑4项毒理学指标。同时,7项有害成分指标对4项毒理学指标的影响各不相同,各成分之间很难设定权重;7项成分纯品的毒理学数据并不完整,但由于含量差异大,即使完整也难于设定权重。

为了综合考察7种有害成分释放量排序和4种毒理学指标排序的相关性,我们将163种卷烟在每一种有害成分释放量上进行排序,某样品7种成分的得分相加即为该样品的排序和。4种毒理学指标也按同样方法计算排序和。有害成分排序和与毒理学指标排序和的相关关系如图14所示。

图14的结果表明,7种有害成分释放量排序之和与4种毒理学指标排序之和存在显著相关关系,说明7种指标等权相加的结果可以反映4种毒理学指标的综合结果。基于以上结果,提出了如下式所示的卷烟危害性评价指数:

$$H=\frac{Y_{CO}}{C_1}+\frac{Y_{HCN}}{C_2}+\frac{Y_{NNK}}{C_3}+\frac{Y_{NH_3}}{C_4}+\frac{Y_{B[a]P}}{C_5}+\frac{Y_{PHE}}{C_6}+\frac{Y_{CHO}}{C_7}$$

式中:Y——卷烟主流烟气有害成分释放量;C₁~C₇——参考值。

卷烟危害性评价指数由CO、HCN、NNK、NH₃、B[a]P、苯酚、巴豆醛7个分项组成,每个分项值为某一卷烟的实测值与参考值之比,7个分项值之和即为该卷烟的危害性评价指数值。值越大,则危害性越大,值越小,则危害性越小。每项有害成分指标的参考值即本研究测定的163个样品的平均值。

4.2 163个卷烟危害性评价结果

163个样品的评价指数值分布如图15所示。图15的结果表明,研究测试的163个卷烟的主流烟气危害性指数分布在2.18~17.72之间,大部分卷烟危害性指数在6~10之间,全部卷烟危害性指数平均值为7.75。

4.3 卷烟危害性指数排序与焦油排序比较

卷烟焦油量是比较常用的卷烟烟气危害性评价指标,但焦油是一种复杂的混合物,其中绝大部分成分不具有生物活性,因此采用焦油作为危害性评价指标有比较大的缺陷。图16为危害性指数和主流烟气焦油的相关性图,图17为163种卷烟焦油排序和危害性指数排序结果的比较。

图16的结果表明卷烟主流烟气的危害性指数和焦油基本没有线性相关性,从163种卷烟的排序结果来看,同一种卷烟的焦油排序和危害性指数排序结果也有较大的不同。

5 结论

(1)对163个卷烟产品的主流烟气29种有害成分释放量和4项毒理学指标进行了分析研究,系统掌握了目前卷烟产品有害成分释放量和毒理学指标的分布状况。

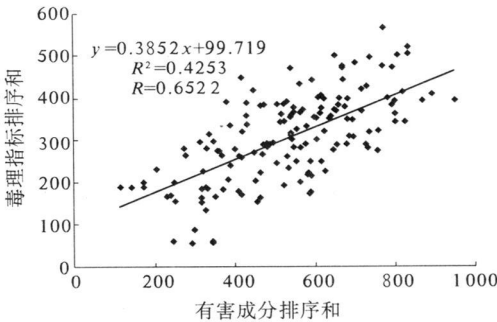


图14 4项毒理学指标排序和与28项有害成分排序和的关系

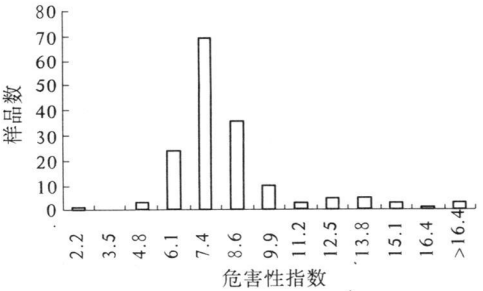


图15 163个样品的评价指数分布图

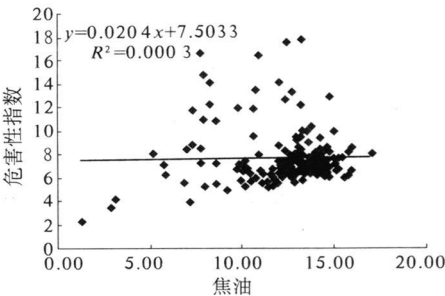


图16 危害性指数和焦油相关性分析

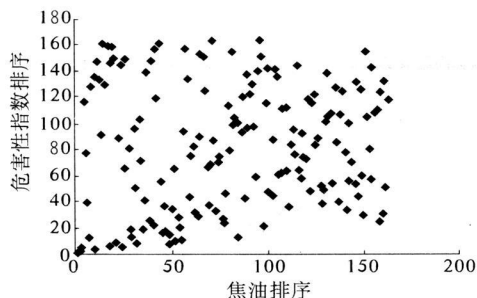


图 17 163种卷烟焦油排序和危害性指数排序结果的比较

(2) 采用多种数理统计方法研究了卷烟主流烟气中主要的 29 种有害成分释放量与 4 项毒理学指标之间的相关性, 采用遗传算法建立有害成分与毒理学指标之间量化数学模型, 首次确定了对卷烟主流烟气危害性影响最大的 7 项有害成分指标: CO, HCN, NNK, NH₃, B[a]P, 苯酚、巴豆醛。上述 7 项有害成分指标与 4 项毒理学指标数学模型的 R^2 值分别为: 0.524, 0.595, 0.504, 0.571。

(3) 以卷烟主流烟气中上述 7 种有害成分指标为基础, 综合考虑 7 种有害成分对卷烟烟气危害性的贡献, 确定每种有害成分释放量参考值, 建立了一种新的卷烟主流烟气危害性评价指数。

致谢: 项目研究实施过程中得到国家烟草专卖局科技司的大力支持, 特此致谢。

参考文献

- [1] Baker R R, Pereira J R da Silva, Smith G. The effect of tobacco ingredients on smoke chemistry. Part I: Flavourings and additives [J]. Food and Chemical Toxicology 2004, 42: 3-37.
- [2] Cooper R L, Lindsey A J, Waller R E. The presence of 3,4-benzopyrene in cigarette smoke [J]. Chem Ind 1954, 46: 1418.
- [3] United States Public Health Service. Smoking and Health. Report of the Advisory Committee to the Surgeon General of the Public Health Service [R]. Publ No. (PHS) 1103, 1964.
- [4] Hoffmann D, Hecht S S. Advances in tobacco carcinogenesis [M] // Chemical carcinogenesis and mutagenesis. I Cooper C S, Grover P L, Springer Verlag London, UK, 1990, 63-102.
- [5] Baker R R, Proctor C J. 2001 - A smoke odyssey [C]. CORESTA Smoke Techno Meeting, Xian, China, 2001.
- [6] Hoffmann D, Hoffmann J, El-Bayoumy K. The less harmful cigarette: A controversial issue. A Tribute to Ernst L. Wynder [J]. Chem Res Toxicol 2001, 14: 767-790.
- [7] Hoffmann D, Hoffmann J. The changing cigarette. Chemical studies and bioassays [M] // Risks associated with smoking cigarettes with low machine-measured yields of tar and nicotine.

- the NCI Smoking and tobacco control Monograph 13. Burns DM, Benowitz N L, Bethesda MD, 2001, 159-191.
- [8] Smith C J, Livingston S D, Doolittle D J. An international survey of "IARC group 1 carcinogens" reported in mainstream cigarette smoke [J]. Food Chem Toxicol 1997, 35: 1107-1130.
- [9] Smith C J, Perfetti T A, Rundle M A, et al. "IARC group 2A carcinogens" reported in cigarette mainstream smoke [J]. Food Chem Toxicol 2000, 39: 371-383.
- [10] Smith C J, Perfetti T A, Rundle M A, et al. "IARC group 2B carcinogens" reported in cigarette mainstream smoke [J]. Food Chem Toxicol 2001, 39: 183-205.
- [11] Rodgman A, Green C R. Toxic chemicals in cigarette mainstream smoke - Hazard and health [J]. Beitr Tabakf. Int 2003, 20(8): 481-545.
- [12] <http://www.hc-sc.gc.ca/hl-vs/tobacco/legislation/reg/industry/index.html>
- [13] Chen P X, Moldoveanu S C. Mainstream smoke chemical analyses for 2R4F Kentucky Reference Cigarette [J]. Beitr Tabakf. Int 2003, 20(7): 448-458.
- [14] <http://www.iarc.fr/ENG/Databases/index.php>
- [15] Wattenberg L W, Coccia J B. Inhibition of 4-methylnitrosamino-1-(3-pyridyl)-1-butanone carcinogenesis in mice by D-limonene and citrus fruit oil [J]. Carcinogenesis 1991, 12: 115-117.
- [16] Lee C K, Fulp C W, Bombick D W, et al. Inhibition of mutagenicity of N-nitrosamines by tobacco smoke and its constituents [J]. Mutat Res 1996, 367: 83-92.
- [17] Brown B, Avary J T, Lee C K, et al. The effect of tobacco smoke nicotine and cotinine on the mutagenicity of 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanol (NNAL) [J]. 55th Tobacco Science Research Conference, Program Booklet and Abstracts 2001, 55(10): 26.
- [18] http://www.coresta.org/Reports/MT_TF_Report_Particular_Matter_ToX.Pdf
- [19] Stratton K, Shetty P, Wallace R, et al. Clearing the Smoke: Assessing the Science Base for Tobacco Harm Reduction [M]. Washington D C: National Academy Press, 2000.
- [20] Doolittle D J, Massey E D. CORESTA Task Force On In Vitro Toxicity Testing of Tobacco Smoke [C]. 烟草制品及卷烟添加剂安全评价与低潜在性危害卷烟产品开发国际研讨会, 2004.
- [21] Put R, Daszykowski M, Baczek T, et al. Quantitative analysis of mitochondrial protein expression in methylnitrosamides by two-dimensional difference gel electrophoresis [J]. J Proteome Research 2006, 5: 1618-1625.
- [22] Cennier V, Massart D L, deNoord O E, et al. Nation of uninformative variables for multivariate calibration [J]. Anal Chem 1996, 68: 3851-3858.
- [23] Rogers D, Hopfinger A J. Application of genetic function approximation to quantitative structure-activity relationships and quantitative structure-property relationships [J]. J Chem Inf Comp Sci 1994, 34: 854-866.