

# 多环芳烃污染的健康效应及其社会代价

## ——基于 EEDH 模型的研究方案设计

燕书欣<sup>1</sup> 游威<sup>12</sup> 房晨<sup>1</sup> 马啸天<sup>1</sup>

(<sup>1</sup> 北京大学环境科学与工程学院, 北京 100871; <sup>2</sup> 北京大学光华管理学院, 北京 100871)

**摘要** 随着社会的发展, 日常生活和工业生产上各种燃料的使用不断增加, 大量多环芳烃(简称 PAHs) 被排放而进入环境中。由于 PAHs 中多种物质的强烈致癌性和诱变性, 再加上排放量的大幅度增加, PAHs 污染已经成为人类的一个重要健康威胁。由此本文希望对于 PAHs 污染的健康效应及其社会代价进行一定的评估。我们从污染现状出发, 基于对环境暴露学, 环境毒理学和环境流行病学领域的研究以及自身设计的针对长期低剂量暴露的队列研究实验方案, 建立了 EEDH (排放-暴露-剂量-健康) 模型, 目标利用此模型设计了对于个体, 项目, 宏观三个层面的“PAHs 污染的健康效应及社会代价”进行评估的研究方案, 并在项目层面举出实例展现模型的应用。

**关键词** 多环芳烃 EEDH 模型 健康效应 社会代价

# The health effects and social costs of polycyclic aromatic hydrocarbons pollution

## ——Research plan design based on EEDH model

YAN Shuxin<sup>1</sup> YOU Wei<sup>12</sup> FANG Chen<sup>1</sup> MA Xiaotian<sup>1</sup>

(<sup>1</sup> College of Environmental Sciences and Engineering, Peking University, Beijing 100871; <sup>2</sup> Guanghua School of Management, Peking University, Beijing 100871)

**Abstract:** With the development of society, the use of various fuels in daily life and industrial production continues to increase, and a large number of polycyclic aromatic hydrocarbons (referred to as PAHs) are discharged into the environment. Due to the strong carcinogenicity and mutagenicity of many substances in PAHs, coupled with the substantial increase in emissions, PAHs pollution has become an important health threat to humans. Therefore, this article hopes to make a certain assessment of the health effects of PAHs pollution and its social costs. Starting from the pollution status, we established EEDH (emission-exposure-dose-health) based on research in the fields of environmental exposure, environmental toxicology, and environmental epidemiology, as well as our own design of cohort study protocol for long-term low-dose exposure. Our goal is using this model to design a research plan for evaluating the "health effects and social costs of PAHs pollution" at the three levels of individuals, projects, and macros, and present examples at the project level to demonstrate the application of the model.

**Key words:** polycyclic aromatic hydrocarbons (PAHs), EEDH model, health effects, social costs

## 目录

1	问题概述.....	3
2	PAHs 污染现状 .....	4
2.1	三大主要排放源污染情况.....	4
2.1.1	全国 PAHs 年排放量的估算.....	4
2.1.2	一级逸度模型对 PAHs 不同介质内浓度的估算.....	6
2.1.3	更为复杂的情况——多介质-归趋模拟.....	7
2.2	室内排放源污染情况.....	8
2.2.1	生活燃料燃烧: .....	8
2.2.2	吸烟及二手烟暴露: .....	9
3	环境暴露学部分: 外暴露水平与内剂量标志物.....	9
3.1	外暴露水平的评价——BGI 个体采样法.....	9
3.2	内剂量的评价——生物标志物法.....	10
3.3	暴露水平-剂量浓度的定量关系.....	11
4	毒理学.....	12
4.1	PAHs 致癌作用的生物化学与分子毒理学机理 .....	12
4.2	<b>B[a]P</b> 神经发育毒性 AOP.....	13
5	环境流行病学研究.....	13
5.1	横截面研究.....	14
5.1.1	数据描述 .....	14
5.1.2	PAHs 浓度的影响因素 .....	14
5.1.3	PAHs 浓度与疾病的关联.....	16
5.2	队列研究.....	18
5.2.1	研究方案设计 .....	18
5.2.2	因果关系分析 .....	19
5.2.3	寿命损失核算 .....	19
5.2.4	PAHs 浓度回落的恢复效应.....	20
5.2.5	暴露组学研究的探索 .....	20
6	案例分析部分.....	20

6.1	污染情况计算.....	20
6.2	暴露模型.....	21
6.3	健康效应模型.....	22
7	结论.....	22
8	分工情况.....	22

## 1 问题概述

多环芳烃（Polycyclic Aromatic Hydrocarbons，简称 PAHs）是两个或两个以上苯环或环戊二烯稠合而成的化合物，目前已知的 PAHs 及其衍生物有几百种。随着现代工业的兴起，交通工具、燃煤炼焦、垃圾焚烧等将大量 PAHs 排放至大气中，据统计，2004 年 66%以上的 PAHs 源于生物质的燃烧<sup>[1]</sup>（图 1）。另外人们日常生活中吸烟、烹饪、煤炉取暖等行为也是 PAHs 的重要排放源。

PAHs 可以通过各种途径（如皮肤吸收、经肺吸入等）进入人体内，对人类具有致癌、致畸、致突变的“三致”效应，在遗传、生殖方面也有一定的毒性，并且可以损害中枢神经、干扰内分泌系统，对人类的生存和繁衍构成威胁。

由此可知，PAHs 排放量巨大且与人类健康关系密切，可能会对人体健康效应造成巨大威胁，值得我们对其排放、扩散、暴露以及健康效应加以研究。

本文第一部分首先从全国尺度分析了 PAHs 的排放量与其在不同环境介质当中的扩散情况，第二部分则研究了外暴露与内暴露水平的评价方式以及暴露水平与剂量之间的关系，然后在第三部分我们从毒理学的角度介绍了 PAHs 对于致癌以及神经发育的影响，在第四部分我们从环境流行病学研究的角度探讨了如何从横断面研究与队列研究两条路径进行实验设计，并给出实验方案。最后，我们以北京大学校园燃煤供暖为研究案例，在此情境下对上述的四个部分的研究进行了细化的描述与展示。我们的研究方案框架如图 2 所示。

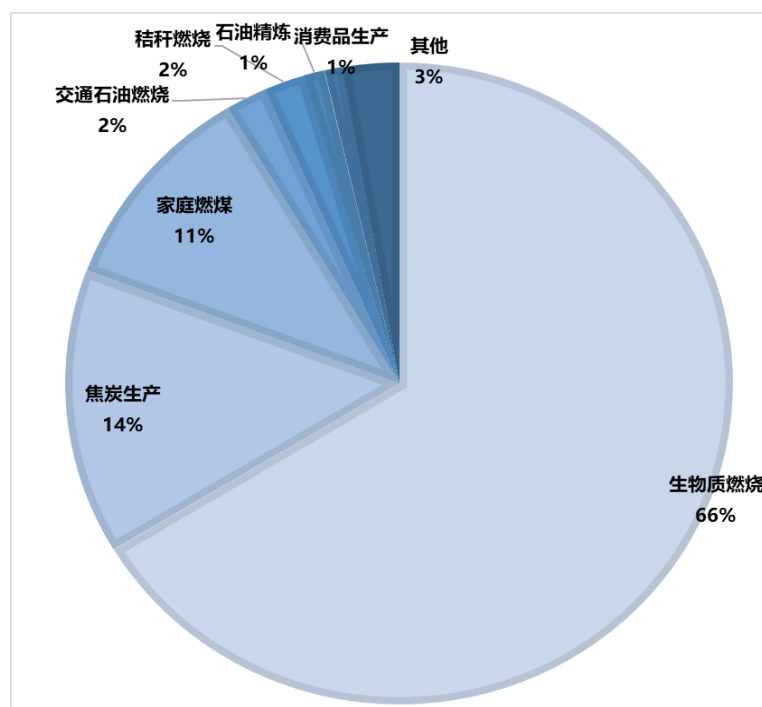


图 1 2004 年中国大气多环芳烃排放清单

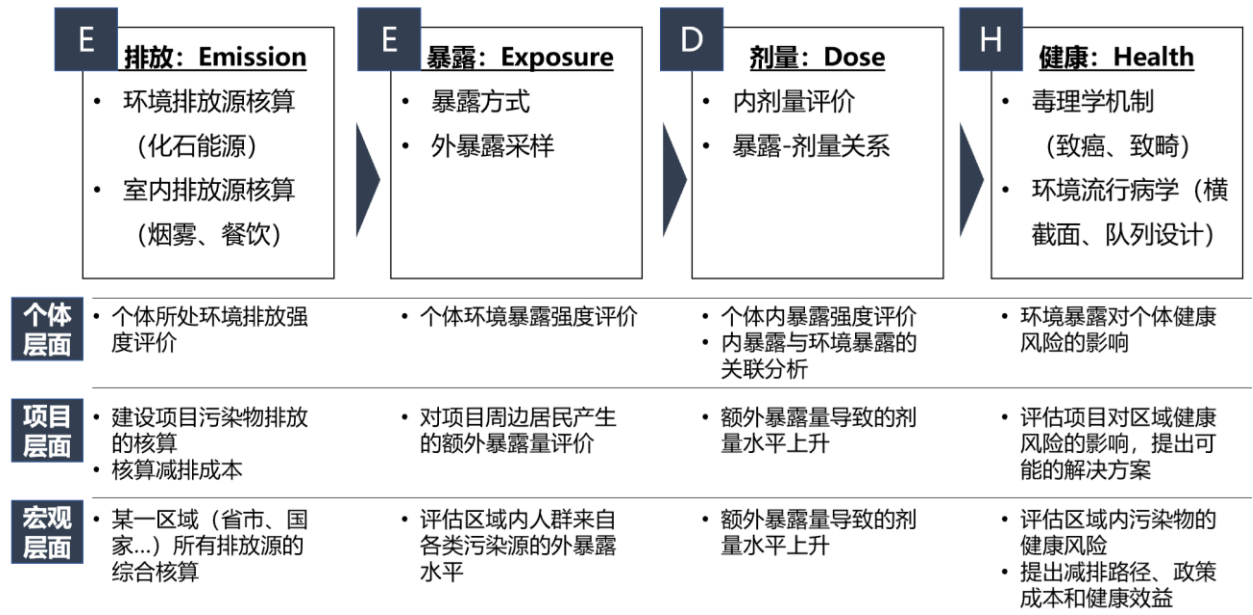


图 2 研究多环芳烃污染问题的 EEDH 模型结构框架

## 2 PAHs 污染现状

### 2.1 三大主要排放源污染情况

#### 2.1.1 全国 PAHs 年排放量的估算

估算 PAHs 的年排放量对于在时空尺度上建立排放清单, 深入了解污染的源头有着重要意义<sup>[2]</sup>。由于 PAHs 主要以人为源的排放为主, 如化石燃料的不完全燃烧, 机动车排放等, 因此这里采用排放因子法的模型分别计算不同人为源排放对 PAHs 污染量的贡献。

研究单考虑了国内 PAHs 的三大主要排放源(工业燃煤, 炼焦用煤, 火力发电用煤)与相应的排放因子, 综合分析计算了各省 16 种亲体 PAHs(来自美国国家环保局 USEPA 推荐的优先污染物清单)各自的排放量, 以全国 30 个省级行政单位为研究的尺度(除去西藏与港澳台), 选取各省在 2017 年的燃煤数据, 对其排放的总 PAHs 量进行估算。

PAHs 排放量的计算公式具体如下<sup>[2]</sup>:

$$Q_T(t) = \sum_i Q_i(t) \quad (1)$$

$$Q_i(t) = \sum_j Q_{i,j}(t) \quad (2)$$

$$Q_{i,j}(t) = \sum_k Q_{i,j,k}(t) \quad (3)$$

$$Q_{i,j,k}(t) = F_{i,j,k}(t)EF_{i,j,k}(t) \times 10^7 \quad (4)$$

为保证估算结果的准确性, 这里选择的不同省份的燃料消耗量来自于国家统计局 2017 年的数据, 排放因子  $EF$  的值来自于许珊珊等于 2005 年综合以往文献总结得到的经自然对数变换和逆变换后计算得到的排放因子  $EF$  中位数 (M) 与标准偏差 (S.E.) 表<sup>[4]</sup>。

式中,  $Q$  为 PAHs 的排放量, t/a;  $F$  为不同情境燃料消耗量,  $10^3$ t/a;  $EF$  (Emission Factor)

为排放因子， $10^{-4}\text{g/t}$ ； $t$ 为年份（2017）； $i$ 为地区（省份）， $i = 1, 2, 3 \dots 30$ ； $j$ 为能源消耗类型（分别为工业、炼焦、火力发电三个场景）， $j = 1, 2, 3$ ； $k$ 为 PAHs 的种类， $k = 1, 2 \dots 16$ 。

估算结果显示：2017 年由于工业、炼焦与火力发电而导致煤炭燃烧排放的 16 种亲体 PAHs 的全国总量约为 1.58 万吨，其中 7 种致癌性 PAHs 的排放总量约为 5251 吨，占总量的 33.2% 左右。

另外，从 PAHs 的种类来看，萘、菲、芘、苯并[a]芘等在 PAHs 污染物中占据了主要的地位（图 3）；从地区角度来看，山西、山东、河北、陕西等作为重工业较为发达，规模较为庞大的省份在全国范围内 PAHs 的排放中贡献比例巨大（图 4）。

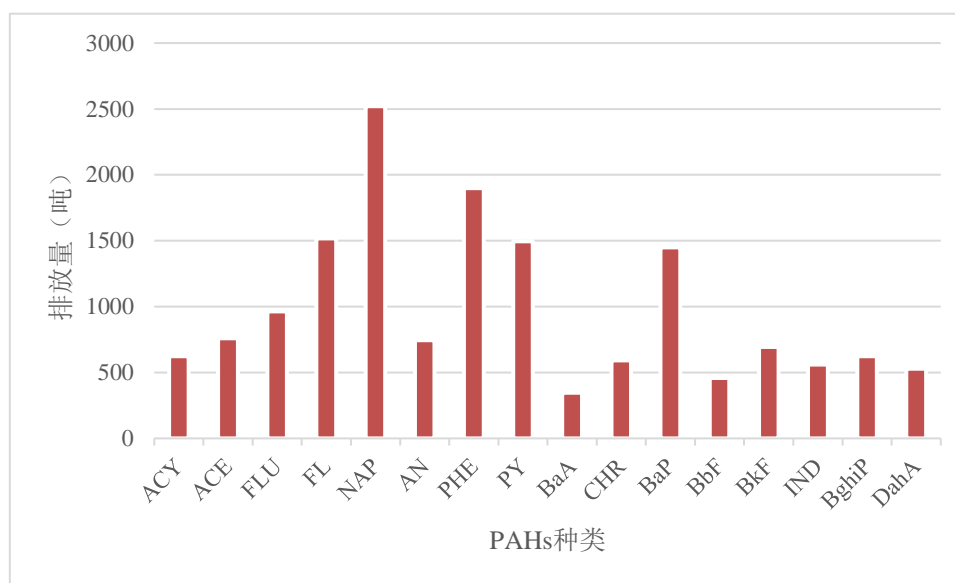


图 3 全国范围的 16 种 PAHs 的排放量分配<sup>1</sup>

[1] 注：ACY 为苊烯；ACE 为苊；FLU 为苈；FL 为苊烯；NAP 为萘；AN 为蒽；PHE 为菲；PY 为芘；BaA 为苯并（a）蒽\*；CHR 为屈\*；BaP 为苯并（a）芘；BbF 为苯并（b）苊烯\*；BkF 为苯并（k）苊烯\*；IND 为茚并[123-c, d]苊\*；BghiP 为苯并[g, h, i]苊；DahA 为二苯并[a, h]苊\*（其中 7 种标有\*的 PAHs 表示该种多环芳烃有较强致癌性）

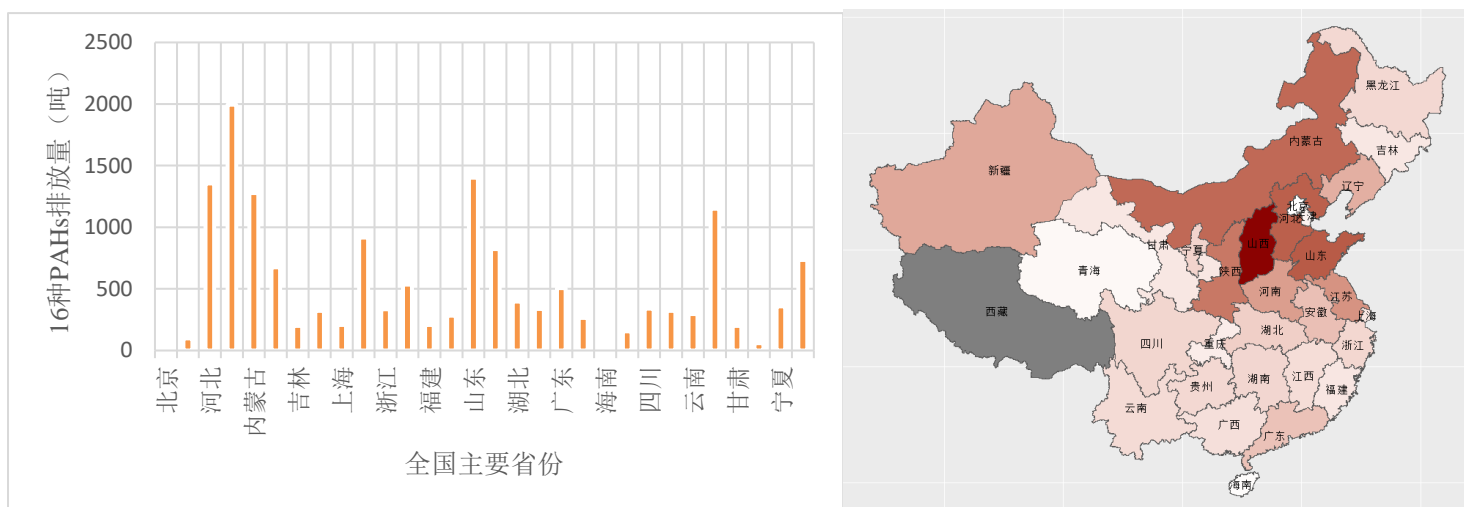


图 4 不同省份 2017 年 PAHs 的排放量统计（左侧为柱状表示右侧为地图可视化）

### 2.1.2 一级逸度模型对 PAHs 不同介质内浓度的估算

加拿大学者 Mackay 等人建立的逸度模型以其结构简单、整体性好等特点成为了模拟有机污染物环境行为的优秀模型之一，由于持久性有机污染物（POPs）具有很强的迁移性，因此该模型在研究 PAHs 的多介质污染时也被科学家们广泛应用。

逸度（fugacity）是一种衡量物质“逃离”介质的趋势的一种物理量，它可作为判断不同相之间平衡的准则，如对于理想气体而言，逸度就等于分压。而对于某种特定化学物质在某种环境介质内的浓度  $c$ ， $c = Zf$ （ $f$  为逸度，单位：Pa， $Z$  为逸度容量，单位： $\text{mol}/(\text{m}^3 \cdot \text{Pa})$ ），其中逸度容量  $Z$  主要取决于化学品、介质的特性以及温度。因此我们可以根据 PAHs 的理化性质对  $Z$  值进行估算，从而计算评估环境中该化学品在各个相中的浓度、总量与百分比，更好地了解扩散与传输的机理。

考虑本报告在实际操作过程中的可行性，这里采用经简化的一级逸度模型，主要通过分析不同介质之间的相平衡进行多介质分配的计算。

在环境体系中分析污染物在四种介质之间的相对传输：空气（A）、水（W）、底泥（S）、鱼（F），对应的体积与 PAHs 的浓度而分别为  $V_1$ 、 $V_2$ 、 $V_3$ 、 $V_4$  以及  $c_1$ 、 $c_2$ 、 $c_3$ 、 $c_4$ ，则可以根据质量守恒关系列出如下方程：

$$M = c_1V_1 + c_2V_2 + c_3V_3 + c_4V_4$$

（ $M$ ：环境中 PAHs 的总质量，可用该区域 PAHs 的总排放量来近似估计）

将  $c = Zf$  代入上述方程即可得到下面的等式：

$$M = Z_A f_1 V_1 + Z_W f_2 V_2 + Z_S f_3 V_3 + Z_F f_4 V_4$$

则当四种介质之间的传输与迁移达到平衡时，PAHs 在每种介质内的逸度均相等，满足  $f_1 = f_2 = f_3 = f_4 = f$ ，则此时可以化简得出  $f$  的表达式（即逸度方程）：

$$f = M / (Z_A V_1 + Z_W V_2 + Z_S V_3 + Z_F V_4)$$

接下来给出不同介质中，逸度容量  $Z$  值的计算公式：

$$\text{空气: } Z_A = \frac{1}{RT}$$

(R 为理想气体状态方程的气体常数, 数值为  $8.314(\text{Pa} \cdot \text{m}^3)/(\text{mol} \cdot \text{K})$ , T 为温度(K))

$$\text{水: } Z_W = \frac{1}{H} = \frac{C^S}{P^S}$$

(H 为亨利常数( $\text{Pa} \cdot \text{m}^3/\text{mol}$ ),  $C^S$  是水中的溶解度( $\text{mol}/\text{m}^3$ ),  $P^S$  是蒸气压(Pa))

$$\text{土壤: } Z_S = y_{OC} K_{OC} Z_W$$

( $y_{OC}$  为有机碳的质量分数(g/g),  $K_{OC}$  为有机碳-水分配系数( $\text{m}^3/\text{mol}$ ))

$$\text{鱼: } Z_F = L K_{OW} Z_W$$

(L 为类脂含量(g/g),  $K_{OW}$  为辛醇-水分配系数)

之后即可分别算出 PAHs 在各个介质内的浓度:

$$c_1 = Z_A f$$

$$c_2 = Z_W f$$

$$c_3 = Z_S f$$

$$c_4 = Z_F f$$

最后还可根据  $M_i = c_i V_i$  得出每种介质内含有的 PAHs 的总量, 从而可以进一步分析污染物在环境内的迁移模式与分布情况。[5]

### 2.1.3 更为复杂的情况——多介质-归趋模拟

在实际科研的应用中, 一级逸度模型由于存在一定的缺陷, 使用得并不多。在其基础上发展的其他等级的逸度模型则更接近于实际的污染物扩散传输情形。如: 二级逸度模型考虑了物质本身在环境体系下的降解与消耗, 三级逸度模型将环境若处于非平衡下发生迁移的情况纳入分析, 四级逸度模型则拓展到了非稳态情形。

其中, 多介质-归趋模拟可在三级逸度模型的基础上对污染物在不同介质之间的迁移扩散的规律与最终平衡的浓度值进行模拟, 能从更宏观的角度把握污染物的环境行为, 从而更有针对性地进行治理与约束。

以一个城市为研究尺度, 模型将城市内的环境分为若干个主相——大气、水体、土壤、沉积物、悬浮颗粒物、植物、鱼体等等, 并将道路与建筑物整体抽象为一片大面积的不透水层, 则可以根据稳态假设和质量的平衡关系建立不同主相之间的迁移量的平衡表达式。具体表达式如下:

$$\text{大气相: } E_1 + f_2 D_{21} + f_3 D_{31} + f_7 D_{71} + f_8 D_{81} = f_1 (D_{12} + D_{13} + D_{17} + D_{18} + D_{ST} + D_{R1} + D_{A1}) \quad (1)$$

$$\begin{aligned} \text{水体相: } E_2 + f_1 D_{12} + f_3 D_{32} + f_4 D_{42} + f_5 D_{52} + f_6 D_{62} + f_7 D_{72} \\ = f_2 (D_{21} + D_{24} + D_{25} + D_{26} + D_{R2} + D_{A2}) \end{aligned} \quad (2)$$

$$\begin{aligned} \text{水体相: } E_2 + f_1 D_{12} + f_3 D_{32} + f_4 D_{42} + f_5 D_{52} + f_6 D_{62} + f_7 D_{72} \\ = f_2 (D_{21} + D_{24} + D_{25} + D_{26} + D_{R2} + D_{A2}) \end{aligned} \quad (2)$$

$$\text{土壤相: } E_3 + f_1 D_{13} + f_8 D_{83} = f_3 (D_{31} + D_{32} + D_{38} + D_{LS} + D_{R3}) \quad (3)$$

$$\text{沉积物相: } E_4 + f_2 D_{24} + f_5 D_{54} = f_4 (D_{42} + D_{45} + D_{BX} + D_{R4} + D_{A4}) \quad (4)$$

$$\text{悬浮颗粒物相: } E_5 + f_2 D_{25} + f_4 D_{45} = f_5 (D_{52} + D_{54} + D_{R5} + D_{A5}) \quad (5)$$

$$\text{鱼相: } E_6 + f_2 D_{26} = f_6 (D_{62} + D_{R6} + D_{A6}) \quad (6)$$

$$\text{有机膜相: } E_7 + f_1 D_{17} = f_7 (D_{71} + D_{72} + D_{R7} + D_{A7}) \quad (7)$$

$$\text{植物相: } E_8 + f_1 D_{18} + f_3 D_{38} = f_8 (D_{81} + D_{83} + D_{R8} + D_{A8}) \quad (8)$$

式中:  $E_i$  为污染物向环境相 i 的排放速率 ( $\text{mol} \cdot \text{h}^{-1}$ );

$f_i$  表示污染物在环境相 i 中的逸度 (Pa);

$D_{ij}$  表示污染物从环境相 i 到环境相 j 的相间迁移速率 ( $\text{mol} \cdot \text{h}^{-1} \cdot \text{Pa}^{-1}$ )



$D_{ST}$  表示空气扩散至平流层的迁移速率 ( $\text{mol} \cdot \text{h}^{-1} \cdot \text{Pa}^{-1}$ )

$D_{LS}$  表示土壤过滤至地下水的迁移速率 ( $\text{mol} \cdot \text{h}^{-1} \cdot \text{Pa}^{-1}$ )

$D_{BX}$  表示底泥掩埋迁移速率 ( $\text{mol} \cdot \text{h}^{-1} \cdot \text{Pa}^{-1}$ )

$D_{Ri}$  表示污染物在环境相  $i$  中的降解速率 ( $\text{mol} \cdot \text{h}^{-1} \cdot \text{Pa}^{-1}$ )

$D_{Ai}$  表示污染物从环境相  $i$  中平流输出的速率 ( $\text{mol} \cdot \text{h}^{-1} \cdot \text{Pa}^{-1}$ )

之后在城市模型中引入污染,此时需要考虑 PAHs 进入城市的具体输入方式(废水、废气的点源面源排放),以及其在环境之间的交换过程(包括气-土、气-水界面的交换,地表径流侵蚀,生物富集,水中颗粒物沉降/再悬浮等等),并考虑最后的输出过程(大气或水的平流输出,环境主相中的降解等)。将各个过程中涉及到的具体理化参数输入模型,设定系统的稳态条件,即可算出各环境主相中 PAHs 估算的平衡浓度值。最后进行参数敏感性的分析与验证,进一步提升模型的准确度<sup>[6]</sup>。

经过上述三级逸度模型运算最终可以得出 PAHs 在兰州城区不同介质内的浓度,以大气为例,浓度主要位于  $10^{-13} \sim 10^{-10} \text{mol} \cdot \text{m}^{-3}$  的范围内,与实际测量的结果差异不大。

逸度模型的真实性与准确度需要大量的环境参数作为支撑,因此目前主要运用在 POPs 在城市、污灌区、湖泊、河流等尺度较小的范围内扩散与迁移的模拟研究<sup>[7]</sup>,难以在全国规模内得出 PAHs 的背景值。另外该模型仍处在不断完善的阶段,应用范围还不够广泛,准确度还不够高。我们在之后的暴露分析中仍以 PAHs 的实际测量值为研究对象。

## 2.2 室内排放源污染情况

### 2.2.1 生活燃料燃烧:

在我国,由于供暖,餐饮等需求,在家庭中使用生活燃料的情况十分普遍,2017 年我国居民生活能源消费量为 57459 万吨标准煤,人均能源消费量为 414 千克标准煤<sup>[8]</sup>。很多研究证实了生活燃料燃烧会导致 PAHs 的大量排放,且由于这种排放方式会使得人们暴露多环芳烃更加直接,所以生活燃料燃烧已经成为重要的 PAHs 排放源,也有很多相关的研究。

现有的流行病学研究证实了室内燃料燃烧与 PAHs 排放以及肺癌等疾病的相关性。在局地有大量使用烟煤的习惯,则有可能导致肺癌发病率较高<sup>[9]</sup>。在家庭中使用煤炉也会导致呼吸系统疾病的相对危险度上升。这些疾病风险与家庭中燃料燃烧时排放的 PAHs 有关,中国预防医学科学院曾进行室内  $B[a]P$  污染水平与人群肺癌死亡率之间关系的定量研究,结果反映二者之间存在明显的暴露-反应关系<sup>[10]</sup>。

在室内进行 PAHs 排放相关的研究,需要计算室内燃料的排放因子,有研究给出了对于烟煤,无烟煤,秸秆,木柴等数十种家庭燃料 PAHs 排放因子的估计<sup>[11]</sup>,各种研究得到的家庭燃煤排放因子有一定差异,与煤的品质,燃烧方式都有一定相关性<sup>[12]</sup>,但大都为室外的 2-3 倍<sup>[13]</sup>。具体计算方法大致如下:

采用滤膜法进行采集,根据滤膜样品中的 PAHs 含量,结合烟气采样分流比和燃煤质量计算出每个煤样的 PAHs 排放因子<sup>[12]</sup>:

$$EF_{PAHs}(\text{mg} \cdot \text{kg}^{-1}) = M_{PAHs} / (R \cdot W)$$

(其中  $M_{PAHs}$  ( $\mu\text{g}$ ) 是滤膜样品中 PAHs 含量,  $R$  和  $W$  ( $\text{g}$ ) 分别是烟气采样分流比和燃煤质量)

通过实验计算得到生活燃料的 PAHs 排放因子后,就能较好地估算各种生活燃料的 PAHs 排放量,进而较好地估计暴露风险。

基于这些,也有人对于供暖等相关问题以及某些人群进行针对性的流行病学研究,进一步证实了室内生活燃料燃烧带来的 PAHs 排放,以及进一步的致癌风险。<sup>[14]</sup>

### 2.2.2 吸烟及二手烟暴露：

在我国，烟民众多，且由于公共场所禁烟效果不是特别理想，很多公共场所存在吸烟的问题。烟草的燃烧同样是一个重要的 PAHs 排放源，吸烟者和二手烟暴露者都有暴露风险。

我国的吸烟率居高不下，同时，许多场所的二手烟暴露也十分严重，2015 年中国成人烟草调查报告显示，57.1% 的调查对象在家中看到有人吸烟，室内工作场所中，这一数据也高达 54.3%。<sup>[15]</sup>虽然这一数据相比于 2010 年有一定下降，但室内的吸烟以及与此相关的二手烟暴露依然十分严重。

吸烟会带来 PAHs 排放以及进一步的疾病风险，有人对于卷烟烟气中 PAHs 含量进行测定，发现卷烟烟气中多种 PAHs 中的物质浓度较高。<sup>[16]</sup>当有吸烟者时，室内 PAHs 浓度会上升，这可以通过颗粒物上的 PAHs 浓度进行定量的分析<sup>[17]</sup>，此时在室内的吸烟者或二手烟暴露者就会面临一定浓度的 PAHs 暴露，这可以通过其尿液中的 PAHs 代谢物浓度进行定量分析，研究显示吸烟者的尿液中 PAHs 代谢物浓度会显著高于不吸烟者。<sup>[18]</sup>

基于这些，某些地区的相关流行病学研究也有所开展，针对中东青少年群体的一项研究就表明了烟草暴露与呼吸系统疾病的相关性。<sup>[19]</sup>

综合来看，室内排放的 PAHs 虽然在总量上相对于其他排放源并不多，但由于与人的接触更为直接，所以会产生很多的疾病风险，针对这些风险还需要进一步的研究。

## 3 环境暴露学部分：外暴露水平与内剂量标志物

### 3.1 外暴露水平的评价——BGI 个体采样法<sup>[20]</sup>

1992 年，美国国家环保局（USEPA）发布了《暴露评价技术导则》，其中将“暴露”定义为人体可见边界（如皮肤、口和鼻腔）接触化学物质的过程。进行暴露评价的目的在于对一定范围人群接触某种污染物的程度或可能性进行测算，以进一步建立与后续参数之间的联系，形成如下图所示的暴露-剂量体系。

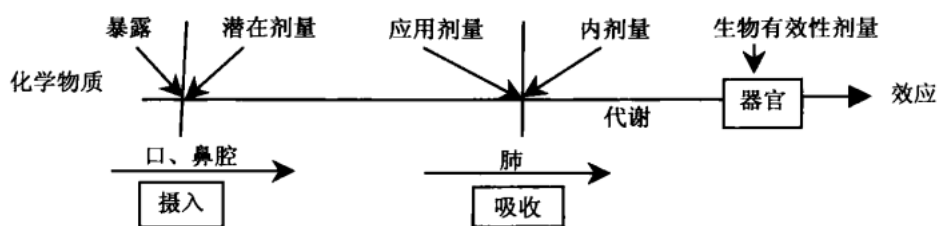
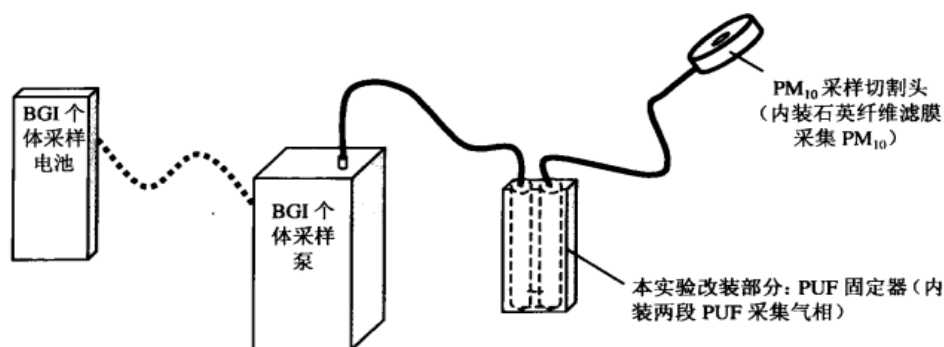


图 5 暴露和剂量的定义示意图<sup>[20]</sup>

对典型人群进行环境暴露学监测（直接测量）是评价外暴露水平的重要方法。考虑到人在空间之间的切换和移动，个体对污染物的暴露水平有很大差异，目前主要用于环境空气中 PAHs 污染物监测的固定空气采样泵，在环境暴露采样中呈现出很大的局限性。为准确定量评价人体环境暴露污染物的浓度和剂量，需用个体采样器 24 小时跟踪采集每名受试者的呼吸样品，例如下图所示的 BGI 个体采样器，由采样泵、电池、采样管路和采样头四部分组成，总重约 2 千克。采样器的创新之处在于在采样头后加两段 PUF 采集气相样品防止漏气，以避免长期以来忽略气相中 PAHs 分布所形成的系统误差。

图 6 BGI 个体空气采样装置示意图<sup>[20]</sup>

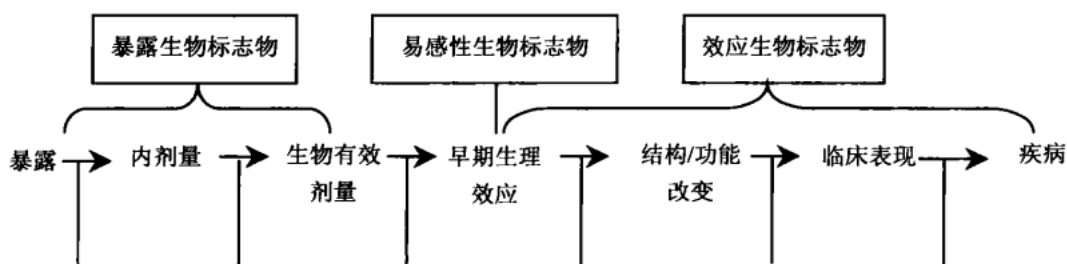
在参试的 24 小时内,每位受试者都要随身佩戴个体采样器,日常行动时采样头挂在胸前,睡觉时放在枕边,甚至洗澡时也要在防潮的前提下放在离口鼻最近的地方。采样流速约 4L/min,采样开始时应进行流量校准,采样结束时进行流量测定,以保证样品体积的准确性。采样后 PUF 和石英滤膜应避光放在零下四度的冰箱中,之后尽快提取和分析。注意整个操作过程都要在避光条件下进行,防止发生光解。

从已有实验数据来看,14 种 PAHs 的暴露浓度数据相关性很高,虽然人体通过呼吸往往同时暴露于多种 PAHs 的混合物,但 PAHs 化合物之间具有一定的关系,在呼吸暴露中,用单个化合物(例如 B[a]P 或 Py)作为人体呼吸暴露 14 种 PAHs 的代表是可行且合理的。

### 3.2 内剂量的评价——生物标志物法<sup>[20]</sup>

根据美国国家环保局的定义,“剂量”指当物质穿过有机体的外边界后与代谢过程或者生物受体发生相互作用的量。换句话说,“剂量”是“暴露”的进一步延伸,是直接影响生物体生命活动的表征量。

在从暴露到剂量再到疾病的历程中,生物标志物扮演了重要的过渡和指示作用,作为定量要素沟通了模型的整个量化过程,本文主要测量和应用暴露生物标志物。

图 7 暴露与疾病关系中各个阶段的生物标志物<sup>[20]</sup>

暴露(内剂量)生物标志物是指通过监测生物体体液或组织中的某种化合物或其代谢物来指示生物对某种化合物的暴露水平的标志物。1-羟基芘(1-OH-Py)是被广泛用于评价 PAHs 暴露的代谢物,因为芘在 PAHs 混合物中含量较高,并且 1-OH-Py 是芘在哺乳动物体内的主要代谢产物,也就具有较高的可靠性,被主流研究公认为评价个体 PAHs 暴露的最相关参数。

由于尿样的易获得性和易保存性,人体尿液中的 1-羟基芘浓度成为表征内暴露水平的主要指标。对尿中 1-羟基芘的分析方法已经比较成熟,以 HPLC/FLD 方法应用最广,该方法准确灵敏、检出限低、可操作性强。

在开展内剂量生物标志物法研究过程中,还有一些问题需要注意。首先,1-羟基芘是芘的

代谢产物，并非直接表征，所以在代谢过程中的个体差异可能会在一定程度上影响其代表性，需要大样本稀释个体差异；其次，研究表明尿液中 1-羟基芘的半衰期为 4-48 小时不等，环境暴露与尿样采集之间的时间差是浓度的重要影响因子，这是实验过程中需要严格控制的变量，其中 24 小时尿样能比较好地反映 PAHs 暴露与 1-羟基芘浓度的关系；第三，从代谢动力学的角度看，环境暴露的影响具有持久性，尿中 1-羟基芘的浓度在最大暴露后很快达到峰值，但在两三天内不会下降到暴露前的状态。

### 3.3 暴露水平-剂量浓度的定量关系

为获得暴露水平-剂量水平的定量关系，使得在后续案例研究中仅由暴露水平的计算值即可代入得出剂量值，而不用对大样本进行复杂的实际监测，本文援引段小丽等研究者对东北某工业城市进行的先导性研究的 1-羟基芘和 Py 实测数据和二者的线性拟合关系。<sup>[21]</sup>

以 Py 日暴露量的对数值为横坐标，以晨尿 1-羟基芘肌酐校正浓度的对数值为纵坐标，则晨尿 1-羟基芘肌酐校正浓度随 Py 日暴露量变化趋势的散点图如下所示，图中的 100 个点表示 100 名受试者，同时标出了线性回归直线和 95%置信区间。

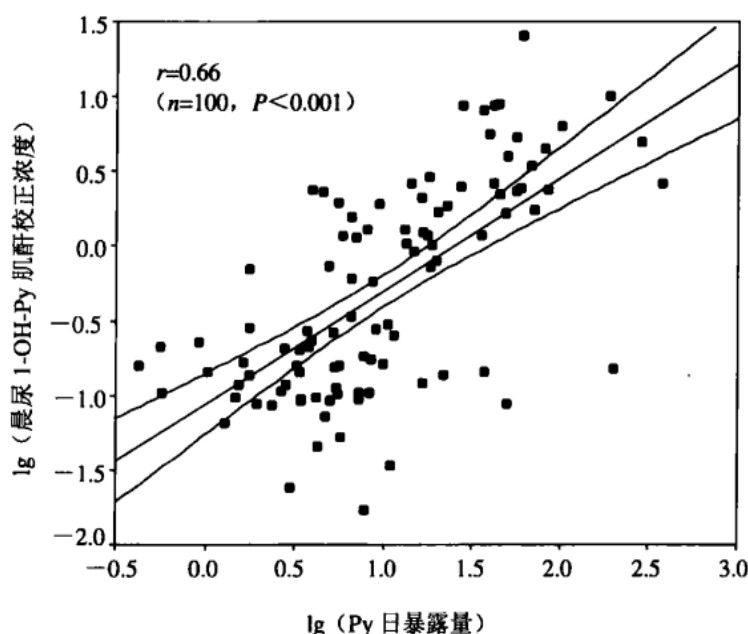


图 8 晨尿 1-OH-Py 与 Py 日暴露量散点图<sup>[21]</sup>

对实验结果进行显著性检验，年龄、体质指数变量的显著性 P 分别为 0.99 和 0.79，相关性可以排除；而 Py 日暴露量的对数值的显著性 P 小于 0.001，是高度相关变量。根据筛选所得变量，建立晨尿 1-羟基芘肌酐校正浓度与 Py 日暴露量的定量关系，定量关系为下式：

$$\lg(C) = 0.76 \times \lg(D_{Py}) + 1.06$$

其中，C 为晨尿 1-羟基芘的肌酐校正浓度，单位： $\mu\text{mol/mol}$ ； $D_{Py}$  为 Py 日暴露量，单位： $\mu\text{g/d}$ 。公式拟合度  $R^2 = 0.43$ ，达到了极显著性水平。由此说明，尿中 1-羟基芘的代谢浓度与人体 Py 的暴露量具有显著的定量关系。

虽然前文提到了用某一种 PAHs 代表 14 种总体暴露水平的合理性，事实上 1-羟基芘对于暴露水平的显著线性关系并不限于 Py，下表展示了晨尿 1-羟基芘与 14 种 PAHs 日暴露量的线性回归方程参数，全面提供了由暴露-剂量模型的换算可行性。

表 1 晨尿 1-OH-Py 与 PAHs 日暴露量的线性回归方程参数<sup>2</sup> [20]

	PAHs 化合物	日暴露量范围/ ( $\mu\text{g/d}$ )	相关系数 $r$	式 (5-4) 参数		拟合度 $R^2$
				$k$	$b$	
1	萘	4.5~673	0.66	0.48	1.77	0.44
2	二氢苊	0.2~306	0.80	0.76	1.09	0.64
3	芴	1.3~57	0.82	0.47	1.12	0.67
4	菲	0.1~144	0.57	0.68	1.37	0.32
5	蒽	0.08~34	0.66	0.41	0.69	0.44
6	荧蒽	0.6~459	0.35	0.25	1.40	0.12
7	芘	0.4~377	0.06	0.57	1.17	0.43
8	苯并(a)蒽	0.05~183	0.68	0.74	0.73	0.46
9	䓛	0.3~235	0.69	0.66	0.83	0.48
10	苯并(b)荧蒽	0.3~107	0.74	0.67	0.54	0.55
11	苯并(k)荧蒽	0.01~66	0.75	0.75	0.27	0.56
12	苯并(a)芘	0.2~156	0.75	0.73	0.63	0.56
13	苯并(ghi)芘	0.2~112	0.68	0.59	0.53	0.46
14	$\Sigma\text{BaP}_{\text{eqi}}$	0.4~202	0.75	0.69	0.80	0.56

## 4 毒理学

多环芳烃中的许多化学物具有强致癌性和诱变性，关于其致癌作用的机理，最常见的就是在生物化学与分子毒理学角度进行的研究。另外，多环芳烃还具有致畸性和神经发育毒性等。文章将会首先介绍多环芳烃致癌作用的生物化学与分子毒理学机理，再通过  $B[a]P$  的神经发育毒性，建立多种组织水平上的致毒路径，介绍一个新兴的污染物危害评估研究方法。

### 4.1 PAHs 致癌作用的生物化学与分子毒理学机理

多环芳烃致癌作用的生物化学与分子毒理学机理的核心思想是：PAHs 经过体内代谢活化成为亲电子终致癌物以后，可与细胞 DNA 分子共价结合，使 DNA 烷基化，形成 DNA 化合物，导致遗传物质发生改变。其主要的过程如下：

**PAHs 的体内转化：**PAHs 在体内首先经过微粒体混合功能氧化酶的催化氧化，生成具有强致癌性的，亲电子的 PAHs 环氧化物。PAHs 经过环氧化物水化酶催化，转化为 PAHs 二氢二醇衍生物，其中有一些失去致癌活性，而也有一些具有了很强的致癌活性，比如 7,8-二醇- $B[a]P$  环氧化物，会继续氧化生成更强的致癌物 4,5-二醇-9,10-环氧化物<sup>[22]</sup>。

比如一种重要的多环芳烃  $B[a]P$  经过上述过程生成  $B[a]P$ -7,8-二醇-9,10-环氧化物，是  $B[a]P$  的主要终致癌物，可以与生物大分子，如 DNA 的亲核基团发生共价结合，形成 DNA 加合物，引起基因突变，进而在促癌因素的协同作用下引起细胞癌变。

[2] 注：表格中

(1)  $\Sigma\text{BaP}_{\text{eqi}}$  表示 13 种 PAHs 的 BaP 当量暴露量之和；

(2) 表中所有  $r$ 、 $k$ 、 $b$ 、 $R^2$  均达到了显著性水平 ( $n = 100$ ,  $P < 0.001$ )。

## 4.2 B[a]P神经发育毒性 AOP

在对于一个污染物进行环境危害评估的时候，传统的毒理学方法需要使用大量实验动物，费时费力，成本较高。计算毒理学作为新型的策略，通过毒理/生理学信号通路的激活发现化学品毒性和效应机制，进而明确一种分子起始事件和结局路径(AOP)。

对于多环芳烃的毒理学研究，同样可以通过这种手段进行模拟，利用模型加以测试，从而预测健康损害效应。我们以B[a]P的神经发育毒性为例，来构建这样的AOP。

关于B[a]P的神经发育毒性机制，目前主要有以下几个方面：

- 有行为动力学数据结果显示，B[a]P代谢产物能够使得纹状体，海马超氧化物歧化酶和谷胱甘肽过氧化酶的活性降低 15%-70%，而过氧化物酶的活性增高至 68%，说明B[a]P可能抑制大脑的抗氧化系统，引起氧化应激，进而引发急性神经毒性。

- B[a]P能够激活芳香烃受体，进而诱导 CYP1A1 的基因表达，并使该蛋白含量上升，CYP1A1mRNA 和蛋白与神经细胞凋亡高度相关。

- B[a]P的代谢产物能够降低 NMDA 受体亚单位 mRNA 和蛋白的表达，妊娠期B[a]P的暴露可以使得个体发育早期中 NMDA 受体亚单位调节降低，NMDA 受体的表达降低会诱发神经细胞凋亡，进而使得海马体或大脑皮质功能受损，可能使得个体终身行为异常。

- 妊娠期暴露B[a]P可能使得个体发育早期脑内 Sp-1 结合物在发育活跃的脑组织区域的表达增高，进而使得突触，髓鞘损伤，引发学习记忆等高级认知功能损害。

由此，我们构建出B[a]P神经发育毒性的 AOP，基于此，可以利用高通量测试数据建立一个较为完整的毒性测试模型。

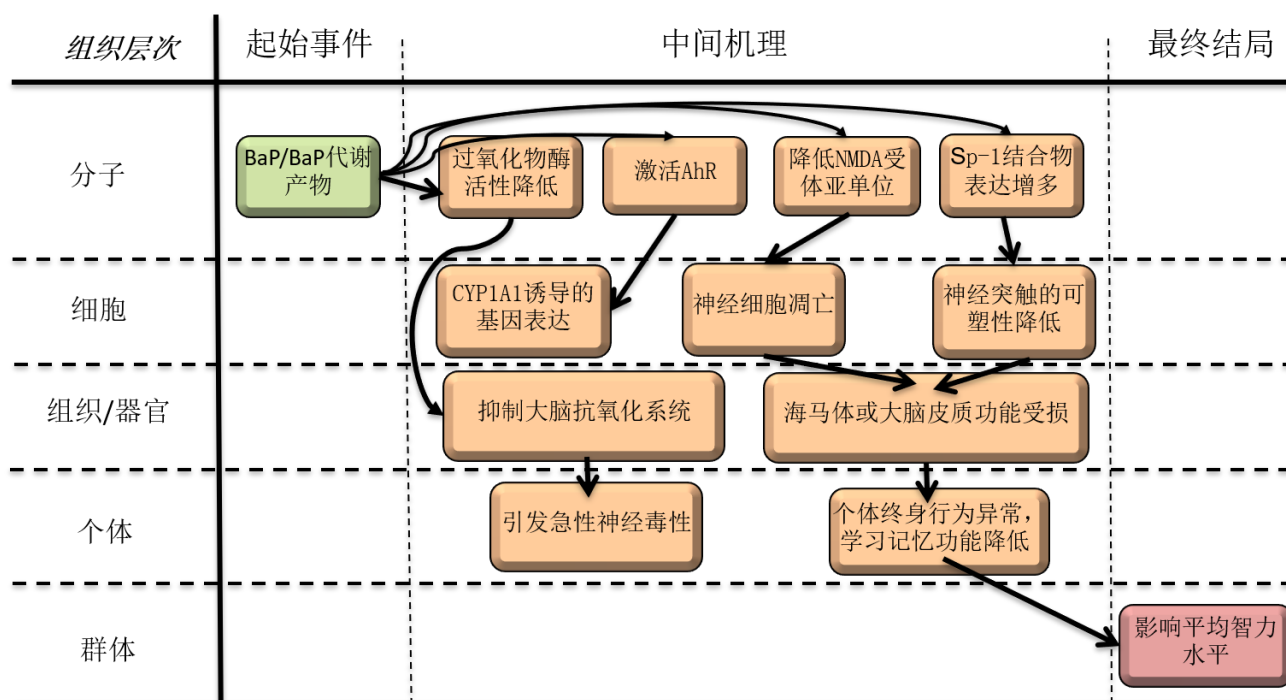


图9 B[a]P神经发育毒性有害结局路径 (AOP)

## 5 环境流行病学研究

## 5.1 横截面研究

### 5.1.1 数据描述

为了探究人体 PAHs 暴露水平与各项生理指征的相关性，我们选用美国疾病预防控制中心（CDC）国家健康数据中心（NCHS）开展的 2013-2014 年全国健康与营养测试研究（NHANES）数据库<sup>[23]</sup>。该数据库调研了 10175 个个体样本，包括其人口学数据、营养学数据、外科体检数据、血尿常规数据、用药数据和个人问卷。在该调查中，随机选取 2637 个样本，进行了尿样多环芳烃含量的检查。使用的数据变量说明详见附表。

在该研究中调查了尿样中的 7 类多环芳烃生物指示物，具体如表 2 所示。在后续的分析当中，PAHs 浓度均根据尿肌酐浓度进行标准化，并取标准化浓度的对数作为因变量。

表 2 NHANES 调查的尿样多环芳烃项目概况

变量名	数据类型	单位	数据概述
URXP01	连续	ng/L	1-萘酚，尿液
URXP02	连续	ng/L	2-萘酚，尿液
URXP03	连续	ng/L	3-茚酚，尿液
URXP04	连续	ng/L	2-茚酚，尿液
URXP06	连续	ng/L	1-菲酚，尿液
URXP10	连续	ng/L	1-芘酚，尿液
URXP25	连续	ng/L	2-/3-菲酚，尿液

如图 10 所示，所有类型的多环芳烃含量之间均存在一定正相关关系，而除 1-萘酚外，其他类型多环芳烃的浓度具有显著的相关性，这也表明这些多环芳烃很有可能具有比较一致的来源。

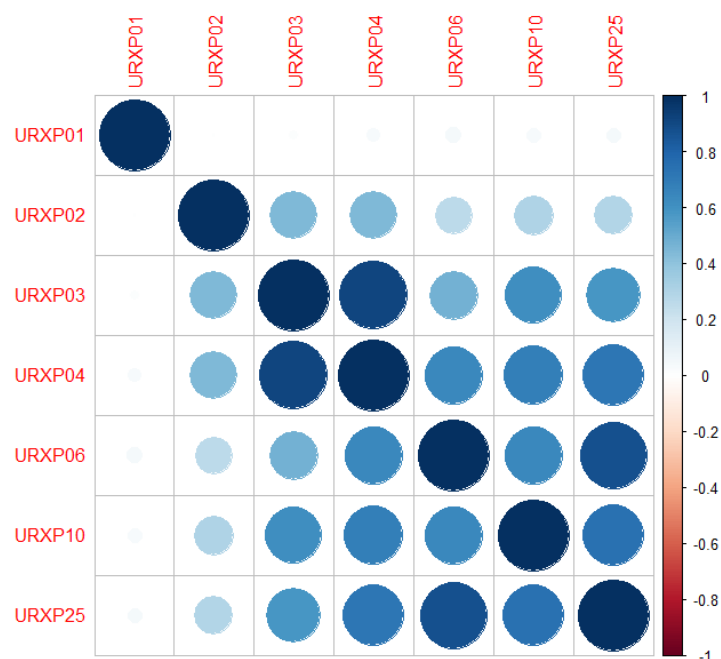


图 10 各类多环芳烃指示物之间的相关性（蓝色指示正相关）

### 5.1.2 PAHs 浓度的影响因素

以 1-萘酚和 1-菲酚作为代表性的生物指示物，取浓度的对数之后与受试者的吸烟频率、每日吸烟量和二手烟暴露频率进行对照绘制图 11，可见吸烟频率和二手烟暴露较高的个体，体



内多环芳烃指示物的含量也相对较高。

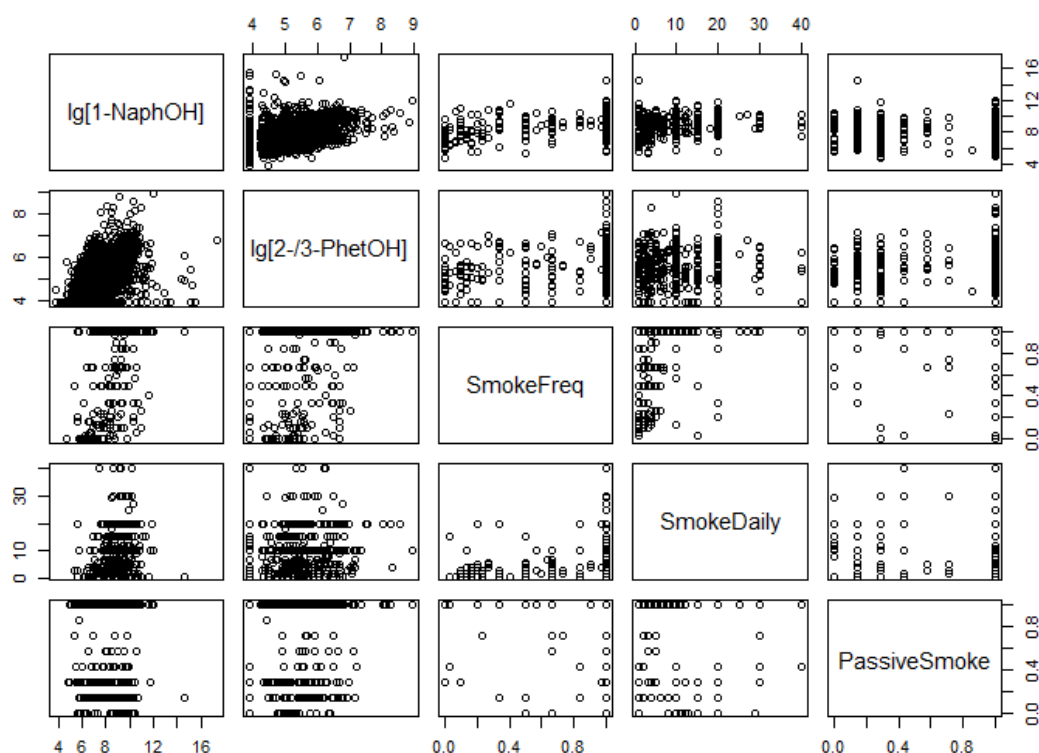


图 11 多环芳烃指示物与吸烟行为的相关性散点图

我们以 1-萘酚和 1-芘酚分别为因变量，对吸烟量、吸烟频率、单日吸烟数量、被动吸烟（二手烟）情况分别进行回归，得到回归结果如表 3 所示。可见，两类多环芳烃指示物与吸烟量、吸烟频次、单日吸烟数目、二手烟均呈现显著的正相关关系，但其中解释力最好的是过去 30 天内的吸烟频次这一项，表明多环芳烃在体内的代谢可能具有一定积累效应，因此更经常吸烟的人体内多环芳烃指示物的浓度更显著。与此同时，即使排除了自身吸烟的影响，非吸烟人群中二手烟暴露也与体内多环芳烃指示物含量呈现显著的正相关关系，表明暴露于二手烟可能也会提升对多环芳烃的暴露。

表 3 多环芳烃指示物与吸烟行为的多元回归分析

	<i>Dependent variable:</i>			
	1-NaphOH (1)	1-PyreOH (2)	1-NaphOH (3)	1-PyreOH (4)
Smoke Volume	0.121*** (0.005)	0.044*** (0.003)		
Smoke Frequency			1.914*** (0.119)	0.496*** (0.076)
Smoke Daily			0.019** (0.008)	0.018*** (0.005)
Passive Smoke	0.482*** (0.087)	0.316*** (0.053)	0.281*** (0.083)	0.264*** (0.053)



Constant	2.591*** (0.024)	0.331*** (0.015)	2.518*** (0.024)	0.311*** (0.015)
Observations	2,637	2,637	2,637	2,637
R <sup>2</sup>	0.221	0.108	0.295	0.123
Adjusted R <sup>2</sup>	0.221	0.107	0.294	0.122
Residual Std. Error	1.169 (df = 2634)	0.719 (df = 2634)	1.113 (df = 2633)	0.713 (df = 2633)
F Statistic	374.481*** (df = 2; 2634)	159.059*** (df = 2; 2634)	366.562*** (df = 3; 2633)	123.422*** (df = 3; 2633)

Note:

\*p<0.1; \*\*p<0.05; \*\*\*p<0.01

### 5.1.3 PAHs 浓度与疾病的关联

接下来，我们考察了各种疾病与多环芳烃指示物浓度之间的关联关系。

下图（图 12）显示了一组 1-萘酚的分组概率密度分布。从年龄分布上看，年龄越大，体内 1-萘酚的含量分布峰值越高。从性别分布上来看，男性体内 1-萘酚水平略低于女性，但差异不大。从疾病分布上看，癌症/高血压患者体内 1-萘酚水平普遍高于非患者。

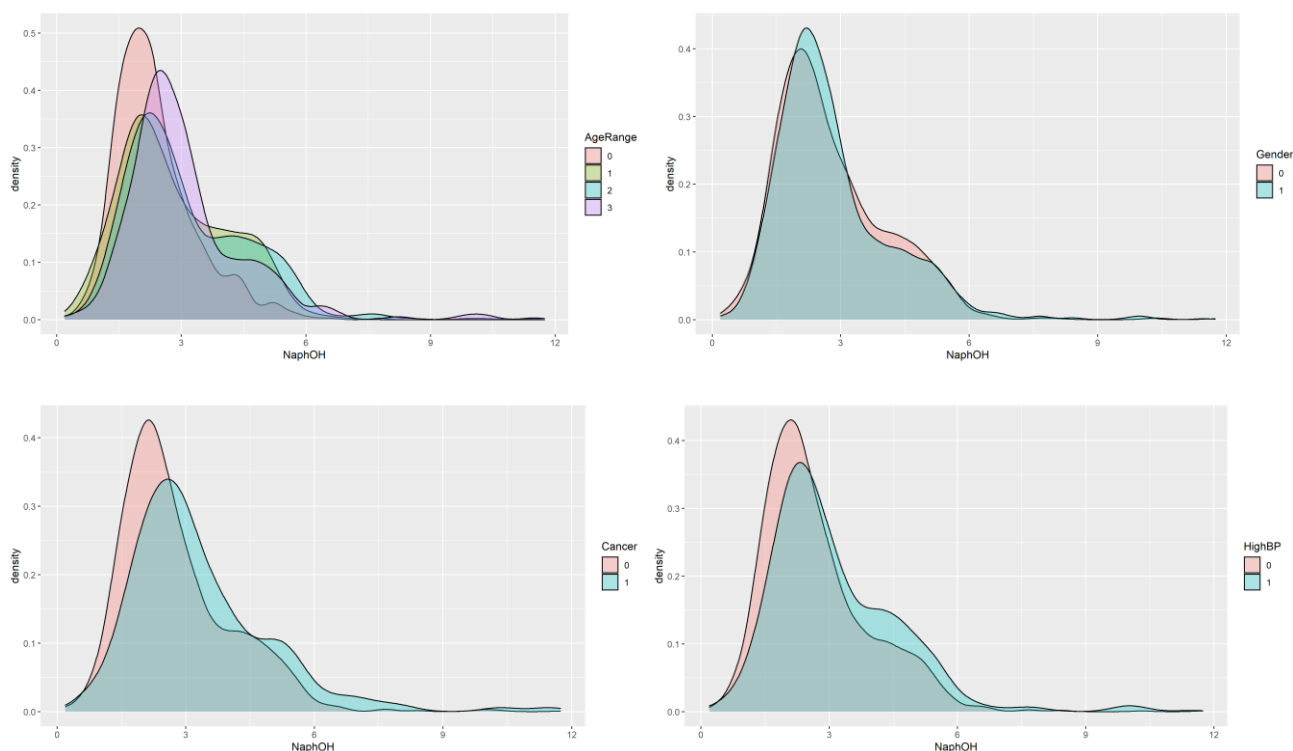


图 12：多环芳烃指示物（1-萘酚）对数浓度与年龄（左上，0-3 分别代表 0-20、20-40、40-60、60-80 岁），性别（左下，0/1 分别代表男、女）和典型疾病（右上：癌症；右下：高血压；0/1 代表非患病和患病）的概率密度分布图

我们以患病与否（1/0）作为因变量，以 PAHs 指示物为自变量作逻辑回归，结果发现 PAHs 指示物中的 1-萘酚与各类疾病的报告率有较显著的正相关关系。如表 4 所示，1-萘酚对癌症的相关性最强，其次依次为胸部不适、高血压、哮喘、认知功能和糖尿病。其他变量不变下，肌酐调整尿样 1-萘酚浓度每升高至原来的  $e$ （约 2.71）倍，患癌风险上升 25.3%。

表 4 多环芳烃指示物与典型疾病关联的多元逻辑回归分析

	<i>Dependent variable:</i>						
	Recognition	Chest Discomfort	Asthma	High Blood Pressure	High Cholesterol	Cancer	Diabetes
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
1-NaphOH	0.194*** (0.043)	0.250*** (0.043)	0.223*** (0.040)	0.235*** (0.033)	0.161*** (0.033)	0.253*** (0.050)	0.096** (0.045)
Constant	-2.777*** (0.150)	-2.969*** (0.152)	-2.658*** (0.139)	-1.783*** (0.108)	-1.599*** (0.107)	-3.463*** (0.182)	-2.407*** (0.146)
Observations	2,636	2,636	2,636	2,636	2,636	2,636	2,636
Log Likelihood	-844.113	-828.761	-950.135	-1,455.573	-1,450.412	-608.309	-890.171
Akaike Inf. Crit.	1,692.226	1,661.523	1,904.271	2,915.146	2,904.824	1,220.619	1,784.343

Note:

\*p&lt;0.1; \*\*p&lt;0.05; \*\*\*p&lt;0.01

为了排除其他类型污染物的影响，我们控制受试者年龄，并将血样和尿样中的微量元素、重金属、免疫水平等变量加入到控制变量中分别回归，并进行变量筛选。经过筛选后，PAHs 变量仍然充分显著，表明其独立致病证据比较充分。

我们注意到部分 PAHs 衍生物的系数为负，这可能意味着 PAHs 浓度之间存在未被描述的微观结构，产生了伪回归现象。但所有 PAHs 系数之和始终为正，这表明排除微观结构影响下，PAHs 作为一个整体对疾病仍然具有正的贡献。

表 5 多环芳烃指示物与典型疾病关联的多元逻辑回归分析（含控制变量）

	<i>Dependent variable:</i>					
	Cancer		High Blood Pressure		Asthma	
	(1)	(2)	(3)	(4)	(5)	(6)
1-NaphOH	0.257*** (0.069)	0.217*** (0.061)	0.236*** (0.056)	0.239*** (0.055)	0.138** (0.061)	0.135** (0.059)
2-NaphOH	-0.060 (0.121)		0.036 (0.077)		-0.004 (0.094)	
3-FluoOH	0.275 (0.262)		-0.306* (0.165)	-0.317* (0.164)	-0.281 (0.202)	-0.323 (0.197)
2-FluoOH	-0.384 (0.331)		0.429** (0.208)	0.462** (0.202)	0.482* (0.259)	0.549** (0.246)
1-PhetOH	0.652*** (0.239)	0.421** (0.173)	0.365** (0.153)	0.359** (0.151)	0.165 (0.182)	

1-PyreOH	-0.405** (0.193)	-0.533*** (0.167)	-0.241** (0.119)	-0.238** (0.118)	-0.099 (0.146)	
2-/3-PhetOH	-0.310 (0.289)		-0.316* (0.189)	-0.341* (0.187)	-0.256 (0.221)	-0.214 (0.149)
Constant	-9.252 (6.349)	-7.135 (5.928)	-1.130 (4.121)	0.399 (3.924)	-5.925 (5.012)	-3.688 (4.540)
Controls	Y	Y	Y	Y	Y	Y
Step Regression	N	Y	N	Y	N	Y
Observations	2,051	2,051	2,051	2,051	2,051	2,051
Log Likelihood	-464.963	-469.810	-987.501	-989.513	-706.253	-710.157
Akaike Inf. Crit.	1,001.925	975.619	2,047.002	2,031.026	1,484.507	1,462.314
Note: *p<0.1; **p<0.05; ***p<0.01						

## 5.2 队列研究

基于现有的横截面数据，不足以回答我们对 PAHs 健康效应的全部疑问，而目前我国也缺少一个以 PAHs 暴露为主要研究对象的队列研究。为此，我们进一步设计了如下的队列研究方案。这里我们将介绍研究方案的设计思路和概况，详细的研究方案请参见附录说明。

### 5.2.1 研究方案设计

#### (1) 地点选取

我们选取大庆、沈阳、北京、天津、晋城、银川、乌鲁木齐、西安、南京、杭州、苏州、广州、香港、南宁、武汉、合肥 16 个城市队列作为研究现场。这 16 个城市地理上涵盖了东北、京津冀、西北、长三角、珠三角、华中等重点区域，对应不同的区域发展水平和污染控制情况，在产业布局、能源结构、环保政策上各有特色，无论是跨区域比较还是区域内比较都有较好的代表性和典型性。

#### (2) 抽样方法

采用随机整群抽样方法，随机选取各城市中 3 个一级监测站点直线距离 1 公里内的社区进行编号，形成抽样框架；以社区为单位随机抽样，每个站点周围抽取 5 个社区，根据各社区户籍资料将所选社区符合条件的居民纳入调查范围，并开展 2 年为周期的随访调查，建立前瞻性队列。在每个城市监测站点覆盖范围以外的城区和乡村非工业区分别随机设置 1 个临时监测站点，用相同方法抽取周边社区，建立安慰剂序列和背景序列。安慰剂序列的建立主要是为了验证监测站点周边的大气污染情况与非监测站点周边不存在系统性差异，而背景序列是为了排除城市和工业污染源，得到多环芳烃的区域背景值。

在保证样本多样性方面，重点考虑性别分组（男、女）、年龄分组（5-20 岁、21-65 岁、66 岁及以上）、教育程度分组（文盲、小学学历、中学学历、本专科及以上）、职业暴露分组（是否重工业一线工人），使队列结构丰富合理，为后续数据分析打好基础。

#### (3) 数据收集

在每 2 年的随访调查当中，我们采用调查问卷、定期体检、环境监测相结合的方法，采集多角度、多层次的数据。

问卷数据：家庭及成员基本信息、医疗覆盖情况、能源消耗情况、生活习惯、饮食习惯、通勤方式、工作环境、家族病史、过往患病情况、药物使用等，2 年追踪调查一次。

体检数据：每 2 年安排一次补助体检，除基本体检项目、血尿常规外，补充测量血重金属、尿 PAHs 衍生物、癌指示物数据。

环境数据：社区周边大气监测情况，水源样本和土壤样本的污染情况，补充监测环境 PAHs 的浓度。

污染设施：社区周边 10km 范围内主要污染设施的信息，如化工园区、垃圾填埋场和焚烧厂等。

### 5.2.2 因果关系分析

基于这一大规模的队列研究设计，我们首先希望能够建立起更加详细的 PAHs 暴露情况与癌症等各类疾病发病的因果关系，尤其是长期低浓度暴露与疾病的因果关系。为此，我们需要排除一些常见的混杂因素，如其他类型的污染物暴露、工作性质等。

基于此，我们考虑采用基于多元逻辑回归的倾向得分匹配(PSM, Propensity Score Matching)方法<sup>[24]</sup>。首先基于现有的分组(PAHs 暴露程度高低的 0/1 因变量，记为 $PAHs_i$ ，其中 $i$ 为个体编号)，用各种混杂因素(记为 $Confound_i$ ，为高维变量)首先对因变量进行逻辑回归：

$$PAHs_i = Probit(\alpha + \beta \cdot Confound_i) + \varepsilon_i$$

其中Probit为联系方程，形式为 $Prob(x) = (e^{-x} - 1)^{-1}$ ；也可以使用其他类型的联系方程，如logit方程。得到各个混杂因素对现有因变量的贡献度 $\beta$ 和截距项 $\alpha$ 后，可以计算得到各个个体的倾向得分，记为 $PS_i$ 。

$$PS_i = Probit(\alpha + \beta \cdot Confound_i)$$

则 $PS_i$ 反映了第 $i$ 个个体基于其混杂因素的患癌倾向。我们可以认为，具有相近 $PS_i$ 值的个体，其基于混杂因素的患癌倾向基本相近，从而可以将 $PS_i$ 值相近而患癌情况不同的个体两两配对，纳入新的组别，纳入低暴露组和高暴露组的个体分别称为 $K_l$ 和 $K_h$ 。配对后，进行平衡性评价，确定匹配结果满足平衡性要求后，再研究自变量多环芳烃暴露高低对实际患癌情况(记为 $Cancer_i$ )的影响。引入平均实验效果(ATE, Average treatment effect)的概念：

$$ATE = E(Cancer_i | i \in K_h) - E(Cancer_i | i \in K_l)$$

则ATE就是高暴露组相对于低暴露组增加的患癌风险。若ATE显著为正，则 PAHs 暴露影响患癌情况的假设成立。

若将 $PAHs_i = c_i$ 用连续变量代替，则可以得到每一步多环芳烃暴露上升对于患癌风险的提升概率，可以估计如下的概率密度分布：

$$ATE(c) = \lim_{\Delta c \rightarrow 0} [E(Cancer_i | i \in K_{c+\Delta c}) - E(Cancer_i | i \in K_c)]$$

则 $ATE(x)$ 具有概率密度分布的特点，能够更精确地衡量每一单位的暴露上升对患癌风险的提升，相较于现在常用的线性模型有较大的提升。

### 5.2.3 寿命损失核算

在环境背景值的基础上，我们考虑各类生活习惯对个人 PAHs 暴露的影响，包括吸烟、二手烟、家庭燃煤、通勤等，从而使得我们的结论可以推广到更大的范围，根据生活区域背景值和个人或家庭的生活方式，估测不同区域、不同人群的 PAHs 暴露水平。

在 PAHs 暴露水平的基准上，我们引入两种不同的致癌效应评估和寿命损失核算方法：

#### (1) 线性模型：USEPA 模型

美国环保署使用的这一模型简单地假设 PAHs 暴露与致癌效应之间为线性关系，现有绝大多数文献以此方法进行核算。首先，根据 PAHs 暴露水平算出致癌危险度：

$$R = q \times ADD$$

其中 $R$ 为致癌的终身超额危险度， $q$ 为致癌因素的致癌强度系数（例如 $q_{BaP} = 7.3(\text{kg} \cdot \text{d}) \cdot \text{mg}^{-1}$ ）， $ADD$ 为致癌因素的额外日均暴露水平。

之后将致癌危险度转化为预期寿命损失和货币化健康成本。根据测算<sup>[27]</sup>，癌症超额发病率 $10^{-5}$ 所对应的成年人的预期寿命损失为62.16 min，则个体的平均预期寿命损失为：

$$\text{loss} = 62.16 \text{ min} \times R \div 10^{-5}$$

## （2）基于患癌风险对 PAHs 暴露概率分布方程的核算

USEPA 虽然简便易算，但简单的线性假设忽略了 PAHs 暴露浓度与患癌风险之间实际上可能存在非线性联系。为此，我们引入上一节中算出的平均效果函数 $ATE(c)$ ，并将其从 0 开始积分，得到效果函数：

$$Effect(c_0) = \int_{c=0}^{c_0} ATE(c)dc$$

对于特定人群来说，设人群中平均浓度暴露水平为 $c$ 时，人群的暴露水平分布为 $f(c)$ ，这一分布通常可以假设为对数正态分布。则人群的累计致癌风险水平为：

$$R = \int_{R^+} Effect(c) \cdot f(c)dc$$

可以用与 USEPA 模型相同的方法算出预期的寿命损失值。

### 5.2.4 PAHs 浓度回落的恢复效应

除此之外，我们也同样关注在一些由于经济结构转型而 PAHs 暴露水平有所回落的社区，其患癌风险是否能够有所改善。我们将着重关注一些由于外生事件导致 PAHs 背景浓度回落的社区和城市，如焦化厂技改、搬迁等事件；运用准实验研究，考虑 PAHs 背景浓度降低对癌症发病率变化的恢复效应。

在数据处理方面，我们将主要采用双重差分（DID，Difference in difference）方法，将外生事件冲击的影响从患癌风险的变化趋势中分离出来，准确地衡量这一恢复效应的出现时间、恢复速度和平衡点。

### 5.2.5 暴露组学研究的探索

同传统的暴露学实验不同，暴露组学数据通过收集人体在一段相对长期限内所有的暴露数据和疾病数据，衡量全部暴露与疾病风险之间的关联性。从数据特征上看，暴露组学数据与基因组学数据相似，都具有维度高、样本量少（ $n/p = O(1)$ ）、跨期相关等特点，传统的回归分析和时间序列分析方法均很难适用。在此情况下，引入适用于高维序列的协方差分析方法十分必要<sup>[28]</sup>。带宽自适应方法作为高维数据分析中协方差矩阵估计的一种手段，能够解决传统时间序列分析方法过参数化的问题，提升估计的稳健性和对小样本数据的适应性。

## 6 案例分析部分

我们以北京大学的供暖季为案例分析若使用燃煤发电供暖释放的 PAHs 会对校园内的师生造成多大的健康损失，选取的物质是 PAHs 中影响程度较为显著的标志性物质苯并[a]芘（BaP）。

### 6.1 污染情况计算

首先，北大供暖季的燃煤量大概为27000t，这里利用排放因子法可以计算出燃烧后一共会产生的苯并[a]芘的质量：

$$Q_T(t) = EF(t) \times F(t) \times 10^7 \approx 0.009t$$

之后我们搜索获得了有关北大燕园校区的基本数据<sup>[29]</sup>:

$$V_A \approx 4.67 \times 10^8 \text{m}^3$$

$$V_W \approx 86520.6 \text{m}^3$$

$$V_S \approx 4.5 \times 10^3 \text{m}^3$$

$$V_F \approx 5 \text{m}^3$$

接下来利用 B[a]P 的理化性质计算其在校区内扩散后达到平衡时的浓度。

根据文献中的数据, 我们可以查出 B[a]P 的物理与化学性质参数<sup>[30]</sup>: 摩尔质量  $M = 252.3 \text{g/mol}$ 、水中溶解度  $C^S = 0.0038 \text{g/m}^3$ 、饱和蒸气压  $P^S = 7 \times 10^{-7} \text{Pa}$ 、辛醇-水分配系数  $K_{OW} = 10^{6.04}$ ; 另外根据文献数据可知其他参数的一般取值<sup>[31]</sup>: 鱼的类脂含量  $L = 0.048 \text{g/g}$ 、温度  $T = 25^\circ\text{C} = 298 \text{K}$ 、土壤的有机碳质量分数  $y_{OC} = 0.02 \text{g/g}$ 、有机碳-水分配系数  $K_{OC} = 41000 \text{m}^3/\text{mol}$ 。之后代入一级逸度模型即可进行求解:

逸度容量  $Z$  的计算结果为:

$$Z_A \approx 4.04 \times 10^{-4} \text{mol}/(\text{m}^3 \cdot \text{Pa})$$

$$Z_W \approx 21.52 \text{mol}/(\text{m}^3 \cdot \text{Pa})$$

$$Z_S \approx 1.76 \times 10^4 \text{mol}/(\text{m}^3 \cdot \text{Pa})$$

$$Z_F \approx 1.13 \times 10^6 \text{mol}/(\text{m}^3 \cdot \text{Pa})$$

苯并[a]芘在不同介质内的浓度值为:

$$c_1 \approx 1.66 \times 10^{-10} \text{mol/m}^3$$

$$c_2 \approx 8.8 \times 10^{-6} \text{mol/m}^3$$

$$c_3 \approx 7.2 \times 10^{-3} \text{mol/m}^3$$

$$c_4 \approx 0.46 \text{mol/m}^3$$

求解结果如图 13 所示:

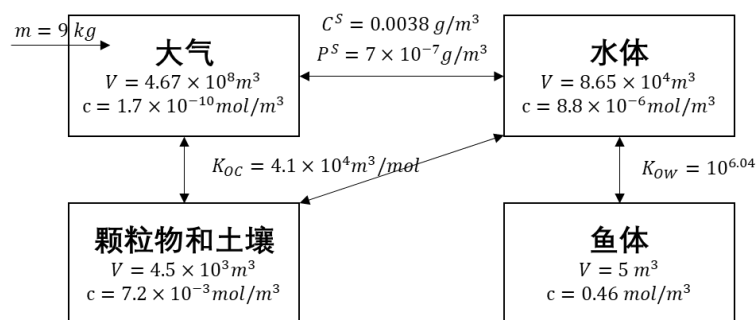


图 13 逸度模型对 BaP 传输扩散模式的示意图

## 6.2 暴露模型

考虑到新建煤炉为燕园校区大气、水体、土壤和鱼均带来苯并[a]芘浓度增量, 我们考虑一个典型学生的暴露场景。暴露于大气中的苯并[a]芘。按照每人每日呼吸  $20 \text{m}^3$  空气计算, 每人每日增加累计苯并[a]芘暴露量  $8.3 \times 10^{-2} \mu\text{g}$ 。暴露于颗粒物携带苯并[a]芘。按照北京冬季平均  $PM_{2.5}$  浓度为  $60 \mu\text{g}/\text{m}^3$ , 每人每日增加 PAHs 暴露量  $1.1 \times 10^{-3} \mu\text{g}$ 。平均增量苯并[a]芘污染为  $8.4 \times 10^{-2} \mu\text{g}/\text{d}$ , 大约相当于 4 支卷烟的苯并[a]芘含量。

此处不考虑鱼体富集的苯并[a]芘对人的影响, 但注意到由于鱼体内脂质富集, 1 克纯鱼油带来的苯并芘暴露量就会达到 107 微克。

### 6.3 健康效应模型

将额外暴露水平代入 USEPA 的致癌效应评估方程：

$$R = q \times ADD$$

其中 $R$ 为致癌的终身超额危险度， $q$ 为致癌强度系数（ $q_{BaP} = 7.3 (kg \cdot d)/mg$ ）， $ADD$ 为额外日均暴露。若根据癌症超额发病率 $10^{-5}$ 所对应的成年人的预期寿命损失为 62.16 min，则这一额外的苯并[a]芘暴露，每供暖季平均对每位同学造成 2 分钟的寿命损失。如果考虑货币化，则该煤炉每年仅因苯并[a]芘额外暴露造成的健康损失约等于 9.7 万元，若花 50 万元建设一尾气完全燃烧处理器，则大约 5 年可以收回健康成本，为同学们带来长久的福祉。

## 7 结论

从案例中我们可以看出，经过我们的模型对于这样一个项目进行的计算，多环芳烃污染确实可以产生较大的健康危害和社会代价。至此，我们构建了一个较为完整的关于“多环芳烃污染的健康效应和社会代价评估”的研究体系，在此进行一个简单的总结，并提出对于未来研究方向的展望。

本报告中主要完成了以下工作：

- 宏观层面，分析了当前我国多环芳烃排放量跨省、分物种的分布情况。
- 中观和个体层面，采用逸度模型等方法核算排放量与暴露水平之间的联系，并总结个体暴露计量方法。
- 总结多环芳烃的暴露-计量关系和毒理学机制。
- 通过现有横截面数据进行数据探索，分析多环芳烃的来源和健康影响。
- 设计队列研究方案及相应的数据处理方法。

未来研究方向展望：

- 在省市层面推广基于污染物源解析的多环芳烃排放和积累核算工作，效仿“棕地”机制建立多环芳烃污染分布的数据系统。
- 探索在敏感区域将多环芳烃纳入常态化监测体系。
- 建立有关多环芳烃的队列研究设计，深入研究多环芳烃长期低剂量暴露的健康效应及恢复效应。
- 应用暴露组学方法，提升多环芳烃健康效应计量的准确性和针对性。

## 8 分工情况

小组成员	分工情况	贡献占比
燕书欣	3 环境暴露学部分：外暴露水平与内剂量标志物；5.2.1 研究方案设计（部分）	25%
房晨	1、问题概述；2.1 三大主要排放源污染情况；6.1 污染情况计算	25%
马啸天	摘要与关键词；2.2 室内排放源污染情况；4 毒理学；5.2.1 研究方案设计（部分）；7 结论；	25%
游威	5.1 横截面研究；5.2.2 因果关系分析；5.2.3 寿命损失核算；5.2.4 PAHs 浓度回落的恢复效应；5.2.5 暴露组学研究的探索；6.2 暴露模型；6.3 健康效应模型	25%

## 参 考 文 献

- [1] Zhang Y , Tao S . Global atmospheric emission inventory of polycyclic aromatic hydrocarbons (PAHs) for 2004 - ScienceDirect[J]. Atmospheric Environment, 2009, 43( 4):812-819.
- [2] 刘伟亚,刘敏,杨毅,陆敏,侯立军,于英鹏,汪青.上海市多环芳烃排放清单构建及排放趋势预测[J].长江流域资源与环境,2015,24(06):1003-1011.
- [3] 蒋秋静,李跃宇,胡新新,卢彬,陶澍,王戎.太原市多环芳烃(PAHs)排放清单与分布特征分析[J].中国环境科学,2013,33(01):14-20.
- [4] 许姗姗,刘文新,陶澍.全国多环芳烃年排放量估算[J].农业环境科学学报,2005(03):476-479.
- [5] MACKAY, DONALD. Multimedia Environmental Models: The Fugacity Approach[M]. 2001.
- [6] 董继元,王金玉,张本忠,王式功,尚可政,王胜利.兰州城区多环芳烃的多介质归趋模拟研究[J].环境科学学报,2013,33(02):570-578.
- [7] 陈春丽,杨洋,戴星照,李志龙,何宗健.鄱阳湖区 PAHs 的多介质迁移和归趋模拟[J].环境科学研究,2016,29(02):218-226.
- [8] 2017 年中国统计年鉴: <http://www.stats.gov.cn/tjsj/ndsj/2017/indexch.htm>
- [9] 何兴舟, 蓝青, 杨儒道,等. 宣威肺癌危险因素研究概述(1979~1993)[J]. 卫生研究, 1995(4).
- [10] 何兴舟, 周晓铁. 室内空气污染的健康效应[J]. 环境与健康杂志, 1991, 008(001):17-20.
- [11] 沈国锋. 室内固体燃料燃烧产生的碳颗粒物和多环芳烃的排放因子及影响因素[D].北京大学,2012.
- [12] 海婷婷, 陈颖军, 王艳,等. 民用燃煤源中多环芳烃排放因子实测及其影响因素研究[J]. 环境科学, 2013, 34(007):2533-2538.
- [13] 段小丽. 多环芳烃污染的人体暴露和健康风险评价方法[M]. 中国环境科学出版社, 2011.
- [14] Liu Y , Qin N , Liang W , et al. Polycycl. Aromatic Hydrocarbon Exposure of Children in Typical Household Coal Combustion Environments: Seasonal Variations, Sources, and Carcinogenic Risks[J]. International Journal of Environmental Research and Public Health, 2020, 17(18).
- [15] 中国疾病预防控制中心发布 2015 中国成人烟草调查报告 (2015-12-28):  
[http://www.chinacdc.cn/zxdt/201512/t20151228\\_123960.htm](http://www.chinacdc.cn/zxdt/201512/t20151228_123960.htm)
- [16] Ding Y S , Ashley D L , Watson C H . Determination of 10 carcinogenic polycyclic aromatic hydrocarbons in mainstream cigarette smoke.[J]. J Agric Food Chem, 2007, 55(15):5966-5973.
- [17] Slezakova K , Castro D , Pereira M C , et al. Influence of tobacco smoke on carcinogenic PAH composition in indoor PM10 and PM2.5[J]. Atmospheric Environment, 2009, 43(40):6376-6382.
- [18] Suwan-Ampai P , Navas-Acien A , Strickland P T , et al. Involuntary tobacco smoke exposure and urinary levels of polycyclic aromatic hydrocarbons in the United States, 1999 to 2002.[J]. Cancer Epidemiol Biomarkers Prev, 2009, 18(3):884-893.
- [19] Boalayan H , Abdualrasool M , Al-Shanfari S , et al. Exposure to environmental tobacco smoke and prevalence of asthma among adolescents in a middle eastern country[J]. BMC Public Health, 2020, 20(1).
- [20] 段小丽. 多环芳烃污染的人体暴露和健康风险评价方法[M]. 中国环境科学出版社, 2011.
- [21] 段小丽, 魏复盛, 张军峰, 等. 人尿中 1-羟基芘与环境暴露多环芳烃的定量关系. 环境化学, 2005,24 (1): 89-91.
- [22] Shiizaki K, Kawanishi M, Yagi T. Modulation of benzo[a]pyrene - DNA adduct formation by CYP1 inducer and inhibitor[J]. Genes and Environment, 2017.
- [23] Centers for Disease Control and Prevention. National Health and Nutrition Examination Survey - NHANES datasets from 2013-2014[DB/OL]. <https://www.kaggle.com/cdc/national-health-and-nutrition-examination-survey> (2017-01-27)[2020-12-01].



- [24] Rosenbaum P R, Rubin D B. The central role of the propensity score in observational studies for causal effects[J]. *Biometrika*, 1983.
- [25] 金银龙,李永红,常君瑞,等.我国五城市大气多环芳烃污染水平及健康风险评价[J].*环境与健康杂志*,2011,28(9):758-761.
- [26] 王进军,武晓燕,古晓娜,等.我国某地区居住环境空气 B(a)P 致肺癌的健康风险评价[J].*环境与职业医学*,2014,31(10):803-805.
- [27] Yang Y, Hu J, Tao S. Loss of life expectancy analysis for cancer risk in Tianjin area[J]. *环境科学*. 2005;26:168-172.
- [28] Chen S X, Qin Y. A two-sample test for high-dimensional data with applications to gene-set testing[J]. *The Annals of statistics*. 2010;38:808-835.
- [29] 陶怡乐,张晨枫,徐琳,温东辉.北京大学未名湖水质与底泥菌群结构的时空变化[J].*北京大学学报(自然科学版)*,2017,53(06):1150-1160.
- [30] 李静,吕永龙,王铁宇,焦文涛,史雅娟,罗维,汪光,王斌.苯并(a)芘的环境多介质迁移和归宿模拟[J].*环境工程学报*,2008(02):279-284.
- [31] 唐纳德·麦凯. 环境多介质模型·逸度方法 第二版[M]// 环境多介质模型·逸度方法, 第二版. 化学工业出版社, 2007.

## ● 附表：模型变量列表

测试	变量名	数据类型	单位	数据概述
编号	SEQN	整数	-	受试者编号
人 口 学 数 据	RIAGENDR	分类	-	性别
	RIDAGEYR	整数	a	年龄
	DMDMARTL	分类	-	婚姻状况
	RIDRETH1	分类	-	人种
	RIDEXPRG	分类	-	是否怀孕
	INDFMPIR	连续	-	家庭总收入与当地贫困线之比
体 检 数据	BPXCHR	连续	min <sup>-1</sup>	心率
	BPXDI1	连续	mmHg	舒张压
	BPXSY1	连续	mmHg	收缩压
	BPXPULS	逻辑	-	是否有心律不齐
	BMXHT	连续	cm	身高
	BMXWT	连续	kg	体重
	BMXBMI	连续	kg/m <sup>2</sup>	BMI 指数
	BMXWAIST	连续	cm	腰围
血 尿 常 规 数据	LBXAPB	连续	mg/dL	载脂蛋白
	LBXHGB	连续	g/dL	血红蛋白
	LBDBANO	连续	1000 cells/ μ L	嗜碱性粒细胞数量
	LBDEONO	连续	1000 cells/ μ L	嗜酸性粒细胞数量
	LBDLYMNO	连续	1000 cells/ μ L	淋巴细胞数量
	LBDMONO	连续	1000 cells/ μ L	单核白血球数量
	LBDNENO	连续	1000 cells/ μ L	嗜中性粒细胞数量
	LBXPLTSI	连续	1000 cells/ μ L	血小板计数
	LBDHDD	连续	mg/dL	直接高密度胆固醇
	LBXTC	连续	mg/dL	总胆固醇
	LBXGH	连续	%	糖化血红蛋白
	LBDSALSI	连续	g/L	白蛋白
	LBDSGBSI	连续	g/L	球蛋白
	LBDSBUSI	连续	mmol/L	血液尿素氮
	LBDSCASI	连续	mmol/L	总钙
	LBDSCHSI	连续	mmol/L	胆固醇
	LBDSCRSI	连续	mmol/L	肌酐

	LBDSGLSI	连续	mmol/L	血清葡萄糖
	LBDSIRSI	连续	$\mu$ mol/L	铁
	LBDSPHSI	连续	mmol/L	磷
	LBDSTBDI	连续	$\mu$ mol/L	总胆红素
	LBDSTPSI	连续	g/L	总蛋白质
	LBDSTRSI	连续	mmol/L	甘油三酯
	LBDSUASI	连续	$\mu$ mol/L	尿酸
	LBXSC3SI	连续	mmol/L	碳酸氢盐
	LBXSCK	连续	IU/L	肌酸磷酸激酶 (CPK)
	LBXSCLSI	连续	mmol/L	氯
	LBXSKSI	连续	mmol/L	钾
	LBXSLDSI	连续	U/L	乳酸脱氢酶
	LBXSNASI	连续	mmol/L	钠
	LBXSOSI	连续	mmol/kg	渗透度
	LBDSCUSI	连续	$\mu$ mol/L	血清铜
	LBDSESI	连续	$\mu$ mol/L	血清硒
	LBDSZNSI	连续	$\mu$ mol/L	血清锌
	URXUCR	连续	mg/dL	尿肌酐
	LBDINSI	连续	pmol/L	胰岛素
多 环 芳 烃 含 量	URXP01	连续	ng/L	1-萘酚
	URXP02	连续	ng/L	2-萘酚
	URXP03	连续	ng/L	3-萘酚
	URXP04	连续	ng/L	2-萘酚
	URXP06	连续	ng/L	1-菲酚
	URXP10	连续	ng/L	1-苊酚
	URXP25	连续	ng/L	2-和 3-菲酚
问 卷 调 查	DLQ040	逻辑	-	是否有注意力不集中、记忆障碍
	CDQ001	逻辑	-	是否有胸部不适
	CDQ010	逻辑	-	爬楼梯时, 是否感到明显气喘
	BPQ020	逻辑	-	是否有高血压
	BPQ080	逻辑	-	是否有高胆固醇
	MCQ220	逻辑	-	是否曾确诊癌症或恶性肿瘤
	MCQ230a	分类	-	曾确诊的癌症类型
	MCQ230b	分类	-	曾确诊的癌症类型
	MCQ230c	分类	-	曾确诊的癌症类型
	MCQ230d	分类	-	曾确诊的癌症类型
	DIQ010	逻辑	-	是否确诊糖尿病
	SMD480	整数	-	7 日内家中有人吸烟的日数
	SMD641	整数	-	本人过去 30 天内吸烟天数
	SMD650	整数	-	本人每日吸烟数量

	SMQ858	逻辑	-	室内工作场合是否有二手烟暴露
--	--------	----	---	----------------