

高血压年龄患病专率的数学模型

武汉医学院流行病学教研室 潘小琴

用数学方程式来表达疾病的流行现象，称为流行病学数学模型。它比一般的文字描述方法能更深刻地说明现象，从而使人们对其本质有更深刻的理解。因此，这是流行病学从定性到定量的更进一步的发展。

数学模型在流行病学特别是传染病流行病学中已有颇多应用^[1]，根据Lilienfeld的归纳^[2]，大约有以下四个方面：①揭示尚未被理解的疾病的机转。②阐明影响疾病发生的各种因素之间的关系。③提示应该补充的观察。④帮助说明某种观点的意义。为此，我们试图用建立数学模型的方法来探讨高血压年龄分布的一些规律。

从国内外的资料看来，除极个别局部地区或民族外，高血压患病率普遍随年龄而增长。1979年我国在29个省、市、自治区对4,012,128个城乡人口按照统一的方法和标准进行高血压患病率抽样普查^[3]，其范围之广、人数之多是世界上少有的，我们仔细观察其结果，却发现无论哪一个地区或性别的人群，其15岁至50~60岁的年龄患病专率均呈现基本上类似的趋势。本文试图用指数曲线来模拟这个趋势，并报告其方法和结果。

材料和方法

一、根据1979年湖北省高血压抽样普查资料和“全国1979~1980年高血压抽样普查报告”中一部分有年龄专率的数据，分别绘制成各个人群的点图，并将各点联以直线，以观察其变化趋势。

二、用指数曲线 $\hat{y} = de^{bx}$ 和指数曲线 $\hat{y} = HB^x$ 来配合上述各图中的年龄患病专率曲线。

三、根据配合计算的结果，在画点图的座

标纸上将同一人群的各年龄组患病率的理论值 \hat{y} 也画上去，联成指数曲线。

四、计算理论值 \hat{y} 与观测值 \hat{y} 的相关指数 R^2 ，以检验理论值对实际观测值的拟合度。

现以武汉市15岁以上的男性人群患病专率曲线的配合过程举例如下。

根据观测到的患病率资料，用 $\hat{y} = de^{bx}$ 配合，其计算过程见表1及其说明。从表1得出理论患病率 $\hat{y} = 0.4175e^{0.0678x}$ ，

表1的计算过程说明：

1. 计算回归直线

$$Y' = a + bX$$

$$\text{因 } \Sigma X = 617.5 \quad \Sigma X^2 = 33881.25$$

$$(\Sigma X)^2 / N = (617.5)^2 \div 13 = 29331.25$$

$$l_{xx} = \Sigma X^2 - \frac{1}{N}(\Sigma X)^2 = 33881.25 - 29331.25 = 4550$$

$$\text{因 } \Sigma Y' = 30.5140 \quad \Sigma Y'^2 = 93.9834$$

$$(\Sigma Y')^2 / N = (30.514)^2 \div 13 = 71.6234$$

$$l_{y'y'} = \Sigma Y'^2 - \frac{1}{N}(\Sigma Y')^2 = 93.9834 - 71.6234 = 22.36$$

$$\text{因 } \Sigma XY' = 1758.0411 \quad (\Sigma X)(\Sigma Y') / N = 617.5 \times 30.514 \div 13 = 1449.415$$

$$l_{xy'} = \Sigma XY' - \frac{1}{N}(\Sigma X)(\Sigma Y') =$$

$$1758.0411 - 1449.415 = 308.6261$$

$$\text{回归系数 } b = \frac{l_{xy'}}{l_{xx}} = \frac{308.6261}{4550} = 0.0678$$

$$\text{因 } \bar{X} = 617.5 \div 13 = 47.5 \quad Y' = 30.514 \div 13 = 2.347$$

$$a = \bar{Y}' - \bar{X} \times b = 2.347 - 47.5 \times 0.0678 = -0.8735$$

因此，回归直线为 $Y' = -0.8735 + 0.0678X$

表 1 1979年武汉市15岁以上男性高血压年龄患病率的指数函数曲线计算表 $\hat{y} = de^{bx}$

年龄组	组中值 X	X ²	y%	Y' = lny	Y' ²	XY'	$\hat{y} = 0.4175e^{0.0678x}$
15~	17.5	306.25	0.78	0.2485	0.0617	4.3481	1.37
20~	22.5	506.25	2.19	0.7839	0.6145	17.6358	2.93
25~	27.5	756.25	2.17	0.7747	0.6002	21.3050	2.69
30~	32.5	1056.25	4.89	1.5872	2.5192	51.5837	3.78
35~	37.5	1406.25	5.15	1.6390	2.6863	61.4624	5.25
40~	42.5	1806.25	8.32	2.1187	4.4887	90.0431	7.45
45~	47.5	2256.25	13.38	2.5938	6.7276	123.4630	10.45
50~	52.5	2756.25	18.49	2.9172	8.5102	153.1546	14.67
55~	57.5	3306.25	31.27	3.4427	11.8519	197.9529	20.59
60~	62.5	3906.25	35.05	3.5568	12.6507	222.2985	28.91
65~	67.5	4556.25	41.11	3.7163	13.8105	250.8470	40.57
70~	72.5	5256.25	43.01	3.7614	14.4838	272.7039	56.94
75以上	77.5	6006.25	47.95	3.8708	14.9781	299.9373	79.92
	ΣX	ΣX^2		$\Sigma Y'$	$\Sigma Y'^2$	$\Sigma XY'$	
	617.5	33881.25		30.5140	93.9834	1758.0411	

$$2. \text{求相关系数} = \frac{l_{xy'}}{\sqrt{l_{xx} \cdot l_{y'y'}}}$$

$$= \frac{308.6261}{\sqrt{4550 \times 22.36}} = 0.97$$

3. 变回归直线为指数曲线

因 $a = \ln d$, $\hat{y} = de^{bx}$, 又因 $b = 0.0678$,
 $\ln d = -0.8735$, $d = 0.4175$

故指数曲线为 $\hat{y} = 0.4175e^{0.0678x}$

从表 1 可见, 理论配合的曲线各点与观测值是靠近的, 在 50 岁以前其趋势也是一致的。50 岁以上, 理论曲线似有偏离现象。再计算 15~54 岁的理论估计患病率 \hat{y} 和实际观测患病率 y 的相关指数, $R^2 = 0.90$, 所以, 在 54 岁以前用指数曲线 $\hat{y} = de^{bx}$, 可以很好模拟高血压年龄患病专率的变化趋势。

结 果

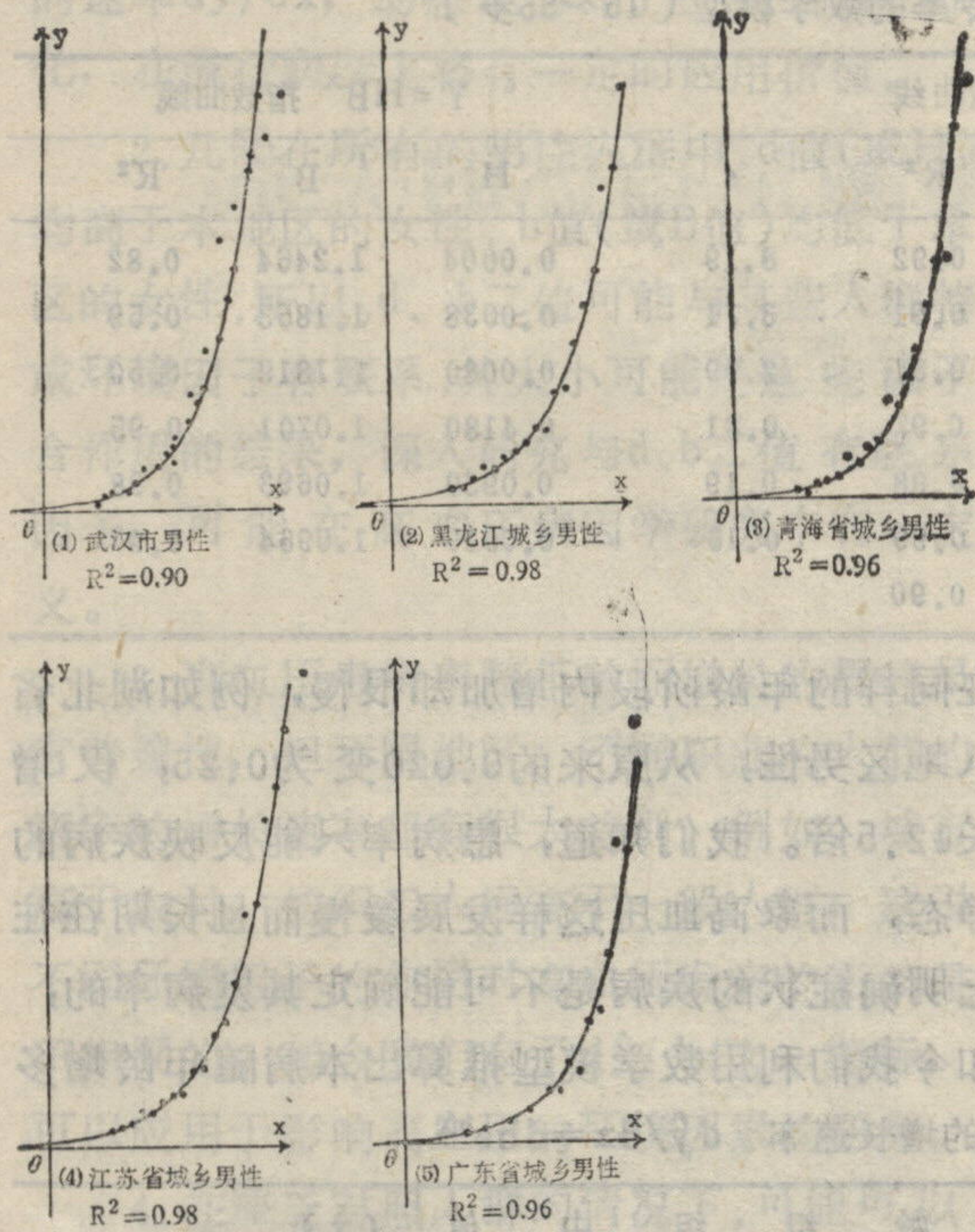
按照以上方法, 我们用两种指数曲线配合不同人群的 54 岁以前的年龄患病专率, 其结果见表 2。因为研究原发性高血压, 应把重点放在青壮年, 以便排除继发因素的影响。所以, 不考虑 55 岁以后的情况, 可能更有利于流行病学病因学研究。从表 2 可见, 27 个人群的 $\hat{y} = de^{bx}$ 指数曲线的 R^2 值均大于 0.9, 说明其拟合度是

良好的, $\hat{y} = HB^x$ 指数曲线的拟合度多数和前一曲线相等, 个别的还要高些, 但有三个群体的患病率曲线拟合不好。见表 2。

从曲线 $\hat{y} = de^{bx}$ 看来, 当 bx 固定时, d 与 y 呈线性关系, d 愈大, y 也愈大; 当 d 和 x 均固定时, b 与 y 呈指数关系, b 值稍大一点, y 值则增大许多; 所以 d 、 b 二个常数可以决定并直接反映一个群体的年龄患病专率的变化趋势, 也就是说, 可以用计算而得的 d 、 b 二值构成各个人群的年龄与高血压患病率关系的数学模型。同样, 如以 $y = HB^x$ 曲线拟合之, 则 H 和 B 二值可以决定并反映人群的年龄患病率的变化趋势, 附图是根据表 2 中 $\hat{y} = HB^x$ 曲线用电子计算机直接绘制的几个例子。

讨 论

一、从以上结果看来, 我国各地区城市或乡村、男性或女性、自然人群或轻、重工业工人, 确诊高血压患病率的年龄分布趋势, 在 50~54 岁以下几乎均符合这些指数曲线的轨迹。所以, 在我国, 高血压患病率随年龄而增加的现象具有规律性, 患病率 (y) 常与年龄 (x) 呈指数关系。由此可以预见, 随着人们平均寿命的延长, 高血压患病率还要大大增高。至于有



附图 男性自然人群年龄患病专率的数学模型 $\hat{y} = HB^x$ 的几个例子 (15~54岁各年龄组)

的人群55岁以上患病率上升速度有减缓倾向，这可能是由于我们的资料属横断面调查结果所致。

二、通过上述数学模型 $\hat{y} = de^{bx}$ 或 $\hat{y} = HB^x$ 我们可以求得微分公式 $d\hat{y}/dx = dbe^{bx}$ 或 $d\hat{y}/dx = HB^x \ln B$ ，它表示年龄(x)每增加1岁时理论患病率 \hat{y} 增加的速率，见表4。 $HB^x \ln B$ 的计算结果与 dbe^{bx} 近似，因限于篇幅省略。

三、过去我们仅用年龄标化率的大小来进行人群间的对比研究，但标化患病率不一定与d、b二值呈平行关系，见表3。而且，标化率不能反映各年龄组的情况，而 $d\hat{y}/dx$ 则可反映患病率随年龄的变化。

从表4还可以看到：

1.湖北省六市女性在17.5岁时， $d\hat{y}/dx$ 仅为0.05，52.5岁时则高达4.7，约增加94倍，天津市纺织系统女性在17.5岁时 $d\hat{y}/dx$ 为0.03，52.5岁时为7.12，约增加237倍，而有些人群

表2 不同人群的高血压年龄患病专率的数学模型 (15~54岁)

人 群	$\hat{y} = de^{bx}$ 指数曲线			$\hat{y} = HB^x$ 指数曲线			
	d	b	R^2	s	H	B	R^2
①湖北省六市城市男性自然人群	0.1313	0.0982	0.98	0.46	0.1312	1.0922	0.98
②" " " " " " " " " " " " 女 " " " " " "	0.0365	0.1277	0.98	1.25	0.0318	1.1431	0.98
③黑龙江省城乡合计男性自然人群	0.2217	0.0733	0.98	0.61	0.2217	1.0760	0.98
④青海省 " " " " " " " " " " " "	0.0654	0.0941	0.98	1.02	0.0654	1.0987	0.96
⑤江苏 " " " " " " " " " " " "	0.0212	0.1175	0.67	2.20	0.1438	1.0813	0.98
⑥广东 " " " " " " " " " " " "	0.0355	0.0881	0.96	0.53	0.0355	1.0921	0.96
⑦北京市市区 男 " " " " " "	0.2796	0.0814	0.96	0.78	0.2796	1.0848	0.97
⑧ " " " " " " " " " " " " 女 " " " " " "	0.0695	0.1123	0.93	6.99	1.1189	0.0695	0.93
⑨ " " " " " " " " " " " " 男 " " " " " "	0.2616	0.0758	0.97	1.39	0.2369	1.0828	0.98
⑩ " " " " " " " " " " " " 女 " " " " " "	0.0950	0.1001	0.95	2.42	0.0943	1.1055	0.95
⑪天津市市区 男 " " " " " "	0.4063	0.0721	0.96	2.01	0.3691	1.0783	0.97
⑫ " " " " " " " " " " " " 女 " " " " " "	0.0363	0.1230	0.94	6.45	0.3629	1.1309	0.94
⑬ " " " " " " " " " " " " 男 " " " " " "	0.0251	0.1015	0.90	0.58	0.0412	1.0968	0.93
⑭ " " " " " " " " " " " " 女 " " " " " "	0.0209	0.1222	0.98	2.36	0.0209	1.1299	0.98
⑮黑龙江省城市 男 " " " " " "	0.1621	0.0825	0.94	0.74	0.1621	1.0860	0.99
⑯ " " " " " " " " " " " " 女 " " " " " "	0.0683	0.1019	0.97	2.30	0.0683	1.1072	0.96
⑰黑龙江省农村 男性自然人群	1.4988	0.1235	0.92	0.77	0.3733	1.0594	0.91
⑱ " " " " " " " " " " " " 女 " " " " " "	1.4341	0.1355	0.92	1.82	0.2982	1.0667	0.95
⑲天津市冶金系统重工业男性职工	0.1667	0.0916	0.97	2.74	0.1667	1.0960	0.97
⑳ " " " " " " " " " " " " 女 " " " " " "	0.0636	0.1081	0.90	5.31	0.0635	1.1142	0.91
㉑ " " " " " " " " " " " " 纺织 " " 轻 " " 男 " " " "	0.2802	0.0836	0.95	2.62	0.2800	1.0872	0.95

续表2 不同人群高血压年龄患病专率的数学模型 (15~55岁)

人 群	$\hat{Y} = de^{bx}$ 指数曲线			$\hat{Y} = HB^x$ 指数曲线			
	d	b	R ²	s	H	B	R ²
②天津市纺织系统轻工业女性职工	0.0075	0.1568	0.92	8.19	0.0004	1.2464	0.82
③湖北省大冶钢厂重“男”工人	0.3521	0.0677	0.91	3.71	0.0038	1.1853	0.59
④武汉市国棉二厂轻“男”职工	1.0865	0.0600	0.81	2.99	0.0080	1.1819	0.50
⑤“城市”自然人群	0.2387	0.0844	0.97	0.81	0.4180	1.0701	0.95
⑥湖北省八地区农村	0.0930	0.0670	0.98	0.19	0.0930	1.0693	0.98
⑦“女”自然人群	0.0510	0.0920	0.99	0.16	0.0510	1.0964	0.99
⑧湖北省武汉市男性	0.4175	0.0678	0.90				

表3 不同地区城乡合计男性人群标化率d、b的比较

	黑龙江	江 苏	青 海	广 东
d 值	0.218	0.142	0.065	0.048
b 值	0.074	0.078	0.094	0.082
年龄标化率%	5.44	4.79	4.07	2.40

在同样的年龄阶段内增加却很慢，例如湖北省八地区男性，从原来的0.020变为0.25，仅增长12.5倍。我们知道，患病率只能反映疾病的静态，而象高血压这样发展缓慢而且长期往往无明确症状的疾病是不可能确定其发病率的，如今我们利用数学模型推算出本病随年龄增多

表4 各个人群确诊高血压患病率的增长速率 $d\hat{y}/dx = dbe^{bx}$

人 群	d	b	年 龄 组 组 中 值 (x)							
			17.5	22.5	27.5	32.5	37.5	42.5	47.5	52.5
湖北省六市男性	0.13	0.09	0.05	0.09	0.14	0.22	0.34	0.53	0.84	1.32
湖北省六市女性	0.04	0.13	0.05	0.10	0.18	0.36	0.68	1.30	2.50	4.79
湖北省八地区男性	0.09	0.07	0.02	0.03	0.04	0.06	0.09	0.12	0.18	0.25
湖北省八地区女性	0.05	0.09	0.02	0.03	0.05	0.08	0.13	0.21	0.32	0.51
黑龙江省城乡合计男性	0.22	0.07	0.05	0.07	0.11	0.15	0.21	0.30	0.43	0.61
青海省城乡合计男性	0.07	0.09	0.03	0.05	0.07	0.12	0.18	0.29	0.45	0.71
江苏省城乡合计男性	0.02	0.12	0.02	0.04	0.07	0.12	0.22	0.39	0.72	1.31
广东省城乡合计男性	0.04	0.09	0.02	0.03	0.04	0.07	0.11	0.16	0.26	0.41
北京市市区男性	0.28	0.08	0.09	0.14	0.20	0.30	0.45	0.67	1.00	1.49
北京市市区女性	0.07	0.11	0.05	0.09	0.16	0.27	0.48	0.83	0.43	2.48
北京市农村男性	0.26	0.08	0.08	0.13	0.19	0.28	0.42	0.62	0.93	1.39
北京市农村女性	0.10	0.10	0.06	0.09	0.16	0.26	0.43	0.70	1.16	1.91
天津市市区男性	0.41	0.07	0.09	0.14	0.20	0.28	0.40	0.56	0.80	1.13
天津市市区女性	0.04	0.12	0.04	0.07	0.13	0.24	0.43	0.79	1.43	2.61
天津市农村男性	0.03	0.10	0.02	0.03	0.05	0.08	0.13	0.21	0.45	0.57
天津市农村女性	0.02	0.12	0.02	0.04	0.07	0.12	0.22	0.39	0.72	1.31
黑龙江省城市男性	0.16	0.08	0.05	0.08	0.12	0.17	0.26	0.38	0.57	0.85
黑龙江省城市女性	0.07	0.10	0.04	0.07	0.11	0.18	0.30	0.49	0.81	1.33
黑龙江省农村男性	0.37	0.06	0.06	0.08	0.10	0.13	0.18	0.24	0.32	0.42
黑龙江省农村女性	0.30	0.07	0.06	0.08	0.12	0.16	0.22	0.31	0.43	0.59
天津市冶金系统男性	0.17	0.09	0.07	0.12	0.18	0.29	0.45	0.70	1.10	1.72
天津市冶金系统女性	0.06	0.11	0.05	0.08	0.14	0.24	0.41	0.71	1.23	2.13
湖北省大冶钢厂男性	0.36	0.07	0.09	0.12	0.17	0.25	0.35	0.49	0.70	0.99
天津市纺织系统男性	0.28	0.08	0.09	0.14	0.20	0.30	0.45	0.67	1.00	1.49
天津市纺织系统女性	0.01	0.16	0.03	0.06	0.13	0.29	0.65	1.44	3.20	7.12
湖北省武汉市国棉二厂男性	1.09	0.06	0.19	0.25	0.34	0.46	0.62	0.84	1.13	1.53
湖北省武汉市城市男性	0.42	0.07	0.10	0.14	0.20	0.29	0.41	0.56	0.82	1.16

的速率 $d\hat{y}/dx$ ，动态地估计了疾病分布的变

化，在流行病学上将有一定的应用价值。

2. 几乎在所有的男性人群中， d 值(或 H 值)

均高于本地区的女性， b 值(或 B 值)均低于本地

区的女性，所以， d 、 b 二值可能与某些人群特征

或环境因子有联系，其大小可能是这些因子综

合作用的结果，深入研究 d 、 b 二值有联系的

因子，可能在高血压病因素研究中有一定意

义。

3. 高血压患病率随年龄而增长的规律虽带

有普遍性，但不同地区、不同职业的人群的患

病率的增长速率却有很大差别。例如：城市远

高于农村，纺织工人远高于一般人群，这说明

不同环境因子的作用对高血压发病的影响是很

不相同的。如今我们有了 $d\hat{y}/dx$ 这一指标，就

可以应用于影响高血压的环境因素的研究。

4. 在缺乏对照人群的情况下，可能可以利

用 $d\hat{y}/dx$ 作为高血压防治效果的评价指标。另

外，我们估计还可能用这个随年龄增长的速率

来预测高血压的发病趋势。

我们对数学模型在心血管病流行病学中的

应用虽还在开始阶段，但是相信它必将日益显

示其实际价值。

摘 要

从1979年我国全国高血压抽样普查结果看来，不

同地区不同性别人群在15~50岁年龄范围内，高血压

患病率普遍呈随年龄增长的类似趋势。本文试用指数

曲线 $\hat{y} = de^{bx}$ 和 $\hat{y} = HBx$ 来模拟各地的这个趋势。所

得的数学模型的拟合度是高的。根据以上模型，可以

出 售 存 刊

本部尚有部分存刊出售，即1984年1~6期。单位购买0.50元一册，开发票；个人购买0.20元一册，为简化手续，不开发票。一律从邮局汇。款汇：北京市昌平区沙河镇(57局)流字五号，中华流行病学杂志编辑部收。汇款后请发一信说明所购期号及册数，收件地址和收件人。

本刊编辑部1985年4月

参 考 文 献

1. 何尚浦：流行病学进展，第256页，第十一章，第一卷，第

1版，人民卫生出版社，北京，1981年

2. Lilienfeld DE: Foundations of Epidemiology, pp

355~363, second Edition by Oxford University

Press, New York, 1980

3. 全国高血压普查协作组：全国1979~1980年高血压抽样普

查报告，内部资料，1982年

(本文承何尚浦副教授、栗毅福老师和黄龙森老师的热

情指导和帮助，特此致谢)

ABSTRACT

According to the results of a nationwide sampling survey in china about the prevalence of hypertension in 1979, we found that within the age range from 15-50, the prevalence of hypertension of various populations, generally, showed a similar tendency of increasing with age. This paper tries to express these changing trends by means of fitting the exponential curve $y = de^{bx}$ or $y = HBx$. The mathematical model we obtained exhibits a high goodness of fit.

Basing on this model, we can calculate $d\hat{y}/dx$ which is the theoretical increasing rate of prevalence of hypertension. Using both $d\hat{y}/dx$ and the characteristic values of d , b or (H, B) in the equation, we suppose that it may be valuable as an index to measure and compare the incidence trends of populations under different conditions. It might help us to find out the relationships between the environmental factors or characteristics of a population and the disease.

的有价值的指标。

推算出高血压患病率随年龄而增的速率 $d\hat{y}/dx$ 。各个

人群的数学模型中特有的 d 、 b 二值(或 H 、 B 二值)及

推导所得的 $d\hat{y}/dx$ ，可能是测量和比较不同条件下的

人群发病趋势和研究环境因子或人群特征与疾病关系