

中国卫生地理信息系统基础数据库的构建

周晓农 胡晓抒 杨国静 孙宁生 汪天平 JB Malone JC McCarroll
林丹丹 洪青标 孙乐平 张治英 徐德忠

【摘要】 目的 构建可用于中国卫生地理信息系统(GIS)应用研究的基础数据库,以合理应用已有的数据资源,并成为一个空间决策系统。方法 采集不同资料来源,经统一标准化后构成基本资料来源。主要的资料来源包括(1)覆盖中国地区的卫星遥感图片库(2)GIS数字化地图库(3)疾病资料库与相关模型库。结果 已获卫星遥感图片库中的植被指数遥感图片、地面温度遥感图片、数字化高程图片、数字化土地利用图片等;GIS数字化地图库中的中国行政区划数字化地图、环境数据矢量地图、人口分布图、气象资料分布图等;疾病资料库与相关模型库中包括了多种疾病的调查报告、疾病防治年报表等资料,以及疾病的传播模型和媒介/中间宿主潜在自然孳生地模型,及恶性肿瘤发病因素相关图。结论 该数据库的构建使不同来源的数据达到统一性,数据库收集的数据量以足能为国内卫生专业人员应用为前提,希望通过更多疾病的应用,使数据库得以完善。

【关键词】 血吸虫病;疟疾;肿瘤;地理信息系统;数据库

Establishment of minimum medical geographic information systems database in China ZHOU Xiao-nong*, HU Xiao-shu, YANG Guo-jing, SUN Ning-sheng, WANG Tian-ping, JB Malone, JC McCarroll, LIN Dan-dan, HONG Qing-biao, SUN Le-ping, ZHANG Zhi-ying, XU De-zhong. *Institute of Parasitic Diseases Control and Prevention, Chinese Center Disease Control and Prevention, Shanghai 200025, China

【Abstract】 Objective To establish a minimum medical geographic information systems(GIS) database as a spatial decision supporting system(SDSS), and to use the database into public health practice in China. **Methods** Spatial data collected from different sources were standardized as decimal degree format, including (1) satellite images covering areas of China; (2) digital maps of China in vector files; (3) diseases database and relevant models. **Results** Necessary satellite images for the database have been collected from NOAA AVHRR, Landsat TM, etc., including the normalized difference vegetation index(NDVI) images from AVHRR, earth surface temperature images from AVHRR, GTOPO30 DEM images from USGS and landuse images from USGS. The digital vector files for GIS analysis were collected including political(county, provinces, country) boundaries file, environmental(drainage, land cover, soil type) vector file, population data and climate data; Data on diseases mainly generated from survey or case reporting. Relevant models on transmission of *Schistosoma japonicum* and *Plasmodium vivax*, and models of *Oncomelania hupensis* and *Anophores sinensis* were developed, and the relevant environmental factors related to incidence of cancers were mapped, to test and verify those database. **Conclusion** The database unified the data from different sources for users. Minimum medical data included in the database could be used in the practice of public health. It is expected that this database be used in a wider range.

【Key words】 Schistosomiasis; Malaria; Carcinoma; Geographic information systems; Database

随着计算机科学及卫星技术的发展和广泛应

基金项目:世界卫生组织/世界银行/联合国开发署/热带病研究与培训特别规划处(970978, 970990)、美国洛克菲勒基金会、国家自然科学基金(30070684)、卫生部疾病控制司的资助及江苏省自然科学基金(BK97210)、江苏省科技厅(BS99026-2)资助项目

作者单位:200025上海,中国疾病预防控制中心寄生虫病预防控制所(周晓农);江苏省疾病预防控制中心(胡晓抒);江苏省寄生虫病防治研究所(杨国静、洪青标、孙乐平);江苏省卫生厅(孙宁生);安徽省寄生虫病防治研究所(汪天平);School of Veterinary Medicine, Louisiana State University, Baton Rouge, LA, USA(JB Malone, JC McCarroll);江西省寄生虫病防治研究所(林丹丹);第四军医大学(张治英、徐德忠)

用,与空间技术有关的地理信息系统(GIS)及遥感技术(RS)已开始应用于疾病预防与控制工作中^[1-3]。而地理信息及遥感数据资料的贫乏已成为大多数用户进一步提高应用水平的瓶颈。为此,“第二次亚太地区应用空间技术于可持续发展的部长级会议”上提出,需重视数据资料的归档、采集、应用及相关鼓励政策的制定,可使更多的用户能容易地、并承受将多种数据应用于发展项目^[4]。为提高我国疾病预防与控制工作进一步应用空间数据资料的能力,我们在血吸虫病GIS数据库建立的基础上^[5,6],

构建了中国卫生 GIS 基础数据库,旨在将与卫生相关的、最基本的 GIS 数据资料收集整理,并构建成为基础数据库,从而更多地为疾病预防与控制服务。

材料与amp;方法

1. GIS 基础数据库构建框架:通过采集不同来源的资料,经统一标准化后构成基本资料来源。主要的资料构成包括三部分:①覆盖中国地区的卫星遥感图片库;②GIS 数字化地图库;③疾病资料库与相关模型。

2. 卫星遥感图片库:该库包括植被指数(NDVI)遥感图片、地面温度遥感图片、数字化高程图片、数字化土地利用图片等。这些数据资料主要来源于现有的国内外资料及从 Internet 网上下载的数据经校正、射影变换后成统一格式。

(1)NDVI 遥感图片:数据源于美国(内政部)地质调查局 EROS 数据中心制作的覆盖全球 AVHRR 遥感图,选择覆盖中国地区的 NDVI 遥感图像,该数据为 10 天 NDVI 复合图。为了便于用户应用,继续对 10 天 NDVI 复合图进行平均值复合,获得每月 NDVI 复合图及全年 NDVI 复合图^[7]。

(2)地面温度遥感图片:数据源于美国地质调查局 EROS 数据中心制作的覆盖全球 AVHRR 遥感图。取覆盖中国地区 AVHRR 遥感图的第 4 频道数据,该数据中的每一象元数值代表地面最高温度值^[7]。

(3)数字化高程图片:数据源于美国地质调查局 EROS 数据中心,该中心的 GTOPO30 高程图(DEM)为二进制 16-bit 的 30-arc-second DEM^[8]。

(4)数字化土地利用图片:数据源于美国地质调查局 EROS 数据中心制作的覆盖全球土地利用遥感图^[9]。

3. GIS 数字化地图:该库主要包括中国行政区划数字化地图、环境数据矢量地图、人口分布图、气象资料分布图等。这些数据资料主要来源于现有的国内外资料及从网上下载的数据经射影变换后成统一格式。

(1)中国行政区划数字化地图:数据源于由美国哥伦比亚大学国际地球信息网(CIESIN)提供的以县为单位中国行政区划数字地图。同时由国家测绘局及美国 ERSI 公司共同开发的 1:1 M 中国数字地图(ArcChina)中,包括了植被分区、公路、铁路等基

本地理信息数据^[10]。

(2)环境数据矢量地图:数据源于 1:1 M 中国数字地图中的不同数据层。这些数据层包括水系图、高程区域图、土地利用图和土壤分布图^[10]。

(3)人口分布图:数据源于 1:1 M 中国数字地图中的人口数据层,该数据层以县为单位构成人口分布图^[10]。

(4)气象资料分布图:数据源于世界粮农组织的 FAOCLIM CD-ROM^[11]。

4. 疾病资料库与相关模型:疾病资料库以血吸虫病、疟疾为例建立,数据资料主要来源于发表于国内的调查报告、疾病防治年报表等资料。相关的模型建立主要为疾病传播模型和媒介/中间宿主潜在自然孳生地模型。同时,恶性肿瘤以江苏省恶性肿瘤调查数据构成,并建立肿瘤发病因素相关图。

结 果

1. GIS 基础数据库构建框架:经统一标准化的基本资料构成了中国卫生 GIS 基础数据库,框架结构见图 1。

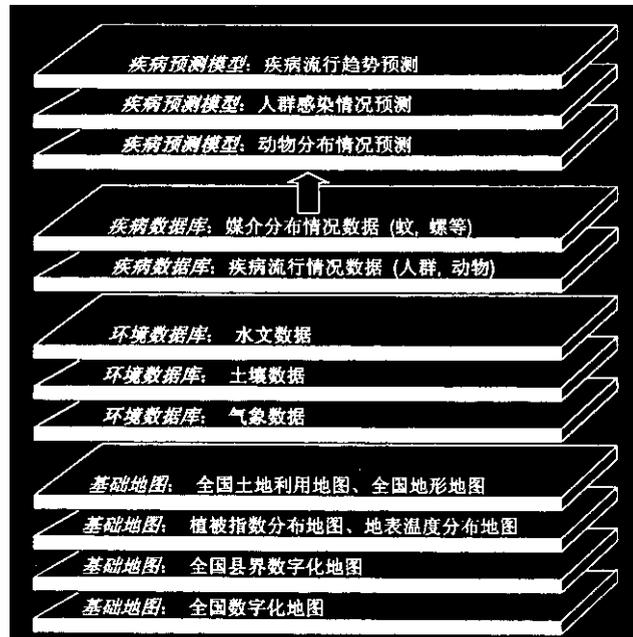


图1 中国卫生 GIS 基础数据库构建框架

2. 卫星遥感图片库:该库包括 NDVI 遥感图片、地面温度遥感图片、数字化高程图片、数字化土地利用图片等。

(1)NDVI 遥感图片:获 1992 年及 1995 年全年 10 天 NDVI 复合图各 36 张,月 NDVI 复合图各 12 张,季 NDVI 复合图各 4 张及全年 NDVI 复合图各

1 张。以上遥感图经偏差校正后,射影变换为 Latitude/Longitude,色彩以 NDVI 值统一设定。遥感图源于 [http://edcdaac.usgs.gov/1 km/1 kmhomepage.html](http://edcdaac.usgs.gov/1km/1kmhomepage.html)

(2)地面温度遥感图片:获 1992 年及 1995 年全年 10 天地面最高温度(T_{max})复合图各 36 张,月 T_{max} 复合图各 12 张,季 T_{max} 复合图各 4 张及全年 T_{max} 复合图各 1 张。射影变换同前,色彩以 T_{max} 值统一设定, T_{max} 数值单位为 $^{\circ}\text{C}$ 。

(3)数字化高程图片:采用全国数字化高程图 1 幅,覆以 NDVI 图后,获立体感较强的彩色高程图。精度为每象元 1 km,射影为 Latitude/Longitude。高程图片源于 <http://support.esri.com/search/KbDocument.asp?dbid=14707aq>

(4)数字化土地利用图片:利用全国数字化土地利用图 1 幅,精度为每象元 1 km,射影为 Latitude/Longitude。图片源于 <http://edcdaac.usgs.gov/glcc/af-int.html>

3. GIS 数字化地图:

(1)中国行政区划数字化地图:利用以县为单位(1990 年版)的中国行政区划数字矢量地图,源于 <http://sedac.ciesin.columbia.edu/china/index.html> 同时获取 1:1 M 中国植被分区数字地图、公路分布数字地图、铁路分布数字地图各 1 幅。

(2)环境数据矢量地图:获 1:1 M 中国水系分布数字地图、高程区域分布数字地图、土地利用数字地图和土壤分布数字地图各 1 幅。

(3)人口分布图:获以县为单位的 1:1M 中国人口分布数字地图。人口数据源于全国第二次抽样调查结果。

(4)气象资料分布图:共获中国 145 个气象站点近 30 年(1940~1970 年)的月平均最高温度、月平均最低温度、月平均温度、月降雨量、月降雨天数、月平均露点、月平均相对湿度、月平均蒸发量、月平均光照、月平均辐射等多项指标数据。

4. 疾病资料库与相关模型:疾病资料库首先针对媒介传播有关的疾病(如血吸虫病、疟疾等)进行数据库的建立。数据源于发表在国内的两次全国血吸虫病流行病学抽样调查报告、流行省血吸虫病防治年报表、流行省疟疾防治年报表等。其中血吸虫病数据库包括病情、螺情、家畜、防治等内容;疟疾数据库包括病情、媒介、防治等内容。尔后,将数据库扩大至肿瘤等慢性疾病的应用研究中。

针对血吸虫病及疟疾流行需经媒介(中间宿主)在自然环境中传播的特点,结合气象等环境资料,以病原体在媒介(中间宿主)中完成发育并繁殖的累积温度为依据,建立疾病传播模型^[12-15]。同时根据媒介(中间宿主)在自然界孳生之特点,应用卫星遥感资料进行空间分析,获媒介(中间宿主)潜在自然孳生地模型。其中血吸虫病传播模型(图 2)、钉螺孳生地模型(图 3)、间日疟传播模型(图 4)及按蚊孳生地模型(图 5)已完成。在慢性病肿瘤方面,已完成江苏省恶性肿瘤的分布图及相关因素的分析(图 6、7)。

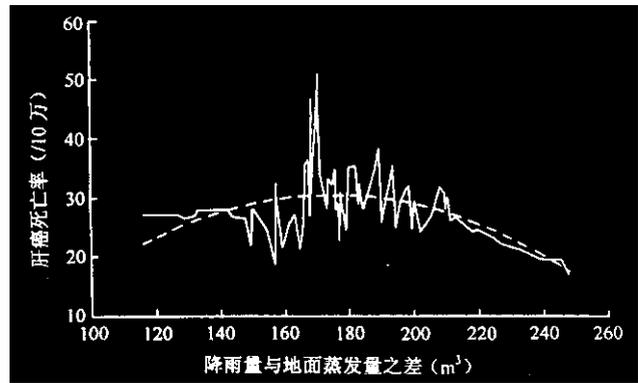


图7 降雨量和地面蒸发量的差值与肝癌死亡率关系的模拟曲线图

讨 论

GIS 及 RS 已用于多种寄生虫病、地方病及部分传染病的应用研究中^[16-18]。随着该项技术的发展,还将用于更多的其他传染病及慢性病中^[19]。但随之出现的共性问题为 GIS 基础数据库的共享应用。目前,这方面的问题已经引起政府部门的重视,但解决的机制与措施尚未成熟。本项目在洛克菲勒基金资助的“全球螺传疾病网络(www.genosisGIS.org)”项目计划中^[20],承担了亚洲区螺传疾病 GIS 基础数据库的一部分进行构建。因此,近年的工作积累,已逐渐构建成为可用于中国卫生 GIS 应用研究的基础数据库。

该数据库的最大特点为数据的统一性。由于不同来源的数据格式不同,对于医学应用者来说使用极不方便,因此,极有必要对不同格式的数据源进行统一。本数据库中的 GIS 数据构建是基于 ArcView 软件平台,同时,本数据库中的遥感资料是基于 ERDAS image 软件平台,并同时可在 ArcView 软件平台上使用。由于 ArcView 软件平

台可在个人计算机上操作,软件简洁,用户界面直观,可视化程度高,支持复杂的空间和属性查询。所以本数据库对目前常用计算机用户来说使用较为方便。

几乎不可能收齐所有的空间属性数据,所以本数据库收集的数据量(包括各种GIS及遥感数据)以满足国内卫生专业人员应用于疾病预防控制需要为限量,在统一设计收集格式的前提下,通过边收集边应用的过程逐步完善。利用这些数据库中的资料,已建立了血吸虫病的传播模型、日本血吸虫中间宿主钉螺孳生地模型、间日疟传播模型、间日疟主要传播媒介中华按蚊孳生地模型,以及恶性肿瘤的分布因素模型等,并对应用前景进行了研究与探索,以逐步应用于疾病防治工作中^[21,22]。本数据库的发展前景主要体现在两个方面,一是扩大卫生领域的应用范围,从媒传(螺传)寄生虫病的应用为基础,逐渐应用于其他媒传传染病,如登革热、肾综合征出血热等,并进而应用于其他急性传染病和慢性非传染病,包括流行性感、肿瘤、AIDS等。另一方面,则通GIS和RS发展,包括应用不同的卫星图像数据、不同的GIS数据源,为不同疾病建立相应的预测模型,在各种技术发展的同时,使本数据库的开发工作日臻完善。

(本文图2~6见插图第4-1页)

参 考 文 献

- 1 Brooker S, Michael E. The potential of geographical information systems and remote sensing in the epidemiology and control of human helminth infections. *Adv Parasitol*, 2000, 47: 245-288.
- 2 Zhou XN, Malone JB, Kristensen TK, et al. Application of geographic information systems and remote sensing to schistosomiasis control in China. *Acta Trop*, 2001, 79:97-106.
- 3 杨国静,周晓农. GIS与RS在寄生虫病防治研究中的应用. *中国寄生虫病防治杂志*, 2001, 14:64-66.
- 4 ESCAP. Report of second ministerial conference on space applications for sustainable development in Asia and the Pacific. In Proceedings of second ministerial conference on space applications for sustainable development in Asia and the Pacific. 15-20 November 1999, New Delhi. United Nations, New York, 2000. 3-4.
- 5 杨国静,周晓农,汪天平,等. 长江下游3省血吸虫病GIS数据库的建立与分析. *中国血吸虫病防治杂志*, 2002, 14:21-24.
- 6 周晓农,孙乐平,姜庆五,等. 全国血吸虫病流行状况的地理信息系统空间分析. *中华流行病学杂志*, 2000, 21:261-263.

- 7 Eidsenshink JC, Faundeen JL. 1-km Avhrr global land dataset: first stages in implementation. *Int J Remote Sensing*, 1992, 10:661-667.
- 8 USGS. Global 30-Arc-second elevation data set; Sioux Falls, South Dakota, U. S. Geological Survey, EROS Data Center, electronic version, 1997, <http://edc.usgs.gov/landdaac/gtopo30>
- 9 USGS. Global land cover characterization; Sioux Falls, South Dakota, U. S. Geological Survey, EROS Data Center, electronic version, 1997, <http://edcdaac.usgs.gov/glcc/glcc.html>
- 10 中华人民共和国国家测绘局. 1:1 000 000 中国数字化地图(DMDC)国际版. 北京:中国地图出版社,1996.129.
- 11 Food Agriculture Orgnization. FAOCLIM 1.2A CD-ROM with worldwide agroclimatic data. User's Manual. Agrometeorology working paper series No. 11. Rome: FAO, 1995.
- 12 周晓农,胡晓抒,孙宁生,等. 地理信息系统应用于血吸虫病的监测 II. 流行程度的预测. *中国血吸虫病防治杂志*, 1999, 11:66-70.
- 13 Zhou XN, Lin DD, Yang HM, et al. Use of landsat TM satellite surveillance data to measure the impact of the 1998 flood on snail intermediate host dispersal in the lower Yangtze River Basin. *Acta Tropica*, 2002, 82:193-198.
- 14 杨国静,周晓农, Malone JB, 等. 江苏省疟疾流行地理信息系统预测模型的研究. *中华预防医学杂志*, 2002, 36:103-105.
- 15 杨国静,周晓农, Malone JB, 等. 多因素空间复合模型预测我国疟疾流行区分布态势. *中国寄生虫学与寄生虫病杂志*, 2002, 20:145-147.
- 16 韩光红,张习坦. 地理信息系统及其在流行病学研究中的应用. *中华流行病学杂志*, 2001, 22:65-67.
- 17 杨长虹,文松. GIS概况及在地方病中的应用简介. *实用寄生虫病杂志*, 1999, 7:33-34.
- 18 陈如桂,杨林生. GIS在卫生与健康领域的应用探索——中国鼠疫医学地理信息系统的设计与建立. *地理科学进展*, 1999, 18:368-372.
- 19 胡晓抒,周晓农,孙宁生,等. 江苏省恶性肿瘤分布态势地理信息系统的空间分析. *中华流行病学杂志*, 2002, 23:73-74.
- 20 Malone JB, Bergquist NR, Huh OK, et al. A global network for the control of snail-borne disease using satellite surveillance and geographic information systems. *Acta Tropica*, 2001, 79:7-12.
- 21 林丹丹,周晓农,刘跃民,等. 鄱阳湖区应用卫星遥感资料预测1998年洪水后钉螺分布状况. *中国血吸虫病防治杂志*, 2002, 14:119-121.
- 22 李文,鲍书明. GIS在肠道线虫感染区域性分布研究中的应用. *中国寄生虫病防治杂志*, 2001, 14:197-199.

(收稿日期:2003-01-03)

(本文编辑:张林东)

