PREDICTION OF HEALTH LEVELS BY REMOTE SENSING

E-4

By Marjorie Rush, Ph.D., University of Texas School of Public Health, and
Sally Vernon, M.A., University of Texas School of Public Health

ABSTRACT

Measures of the environment derived from remote sensing were compared to census population/housing measures in their ability to discriminate among health status areas in two urban communities. Three hypotheses were developed to explore the relationships between environmental and health data. Univariate and multiple step-wise linear regression analyses were performed on data from two sample areas in Houston and Galveston, Texas. Environmental data gathered by remote sensing were found to equal or surpass census data in predicting rates of health outcomes. Remote sensing offers the advantages of data collection for any chosen area or time interval, flexibilities not allowed by the decennial census.

INTRODUCTION

Population health status is a response to multi-factoral circumstances, one of which is the urban environmental set. Several authors believe that states of health and disease are linked to environmental factors and express attempts to adapt to changes in the environment (e.g., DuBos, 1965 and Stallones, 1972). Certain of these factors are revealed by remote sensing in the form of aerial photography and are readily discernible by an image analyst. Poverty and density are two environmental conditions, which can be delineated on aerial photographs (Mumbower and Donoghue, 1967; Bowden, 1968; Wellar, 1968; Manji, 1968; Mullens, 1969; and Horton and Marble, 1970). These conditions have been found to be related to health by previous investigators (e.g., Loring, 1964; Lieberman and Duhl, 1964; Schmitt, 1966; Faris and Dunham, 1967; Galle, et al, 1972; and National Health Survey, 1972). The purpose of this study was to determine the applicability of a new health data source--remote sensing--in assessing health levels in two communities. Such data may expedite placement of health intervention, monitoring, or treatment units. This remote sensing capability was demonstrated in studies of both Houston and Galveston, Texas (Rush and Vernon, 1973 and Rush, Goldstein, and Hsi, 1974).

HYPOTHESES

Based on a review of environmental health literature, an empirical generalization was formulated: land use and residential quality of an area are associated with the health status of the population residing in that area. Three hypotheses were developed:

1. Environmental variables revealed by remote sensing can predict mortality and morbidity rates.
2. Population and housing variables reported by the Census can predict mortality and morbidity rates.
3. There is no difference in the health predictive strengths between environmental variables and census variables.

METHODOLOGY

Subareas within Houston and Galveston were described by three sets of data. The independent variables consisted of two sets of ecological data--environmental data interpreted from photography and population/housing data from the 1970 Census. Dependent variables were selected measures of population health responses. Health data were compared with the spatial distributions of visual, physical, environmental characteristics identified from photography. The association of census population/housing data with the health variables was also measured to serve as a standard for comparison of the predictive strength of remotely perceived environmental data.

Data for both cities were collected by census block groups. Thirty-seven block groups in Houston were chosen for a representational cross-section of socio-economic residential levels and industry. In Galveston, a 50% areal sample of 70 block groups was chosen representing discrete urban subareas with homogeneous environmental characteristics.

Environmental data were determined from low level color photography. For Houston the scales were 1:6,000 and 1:12,000 while for Galveston it was 1:24,000. The viewing scale for photo interpretation in both studies was 1:6,000. Regular color photographs were taken of Houston while color infra-red film was used in Galveston. Interpretation of land uses and quality was performed by lot, then recorded by block on a map overlay by the photo interpreter. Quantification by square feet of each land use was carried out by a point counting technique for Houston and by a scaled grid for Galveston. Although the measurement techniques differed, testing indicated no significant differences in results between techniques.

In the Houston study, eleven land use categories were identified and residential land use was further broken down into density and quality classes making a total of 22 categories (Table I). Three residential density classes were used: high (under 60' frontage), medium (60' - 90') and low (+90'). Apartments formed a separate category. Four quality designations were made by the interpreter based on subjective evaluation of house and lot size, curbs and gutters, sidewalks, garages, streets, foliage and mixed land uses.

In the Galveston study, nine land use categories were identified. Two categories of multi-family housing replaced the apartment category in Houston. The quality variables used for a subjective evaluation in both studies were also assessed separately by percentage of block coverage for Galveston.

Health outcomes common to both studies were mortality, tuberculosis, shigella and salmonella, infectious hepatitis and meningitis (Table II). For Houston, first offense juvenile delinquency referrals and mental health referrals were also used. The Galveston study utilized venereal disease (syphilis and gonorrhea), hypertension, and cardiac arrest/myocardial infarction as additional health outcomes. Data were gathered for a time period of from two to eight years except for mortality in both studies, juvenile delinquency in Houston, and venereal disease (VD) and tuberculosis (TB) in Galveston which were collected for one year. These behavioral and morbidity data are subject to errors of underreporting, unsystematic reporting, and problems of definition generally recognized.

Morbidity and behavioral data were converted to crude rates per 1,000 population. Houston mortality data were standardized by sex and three age groups through the indirect method, and the difference between this rate and the crude rate for each block group was used for analysis. For Galveston three standardized age-adjusted rates were utilized. Only three age groups were available on the block group level, and this constituted a limitation in the analysis. Incidence rates were calculated for all morbidity data except TB for which a prevalence rate was used for Galveston.

Univariate and multiple step-wise linear regression analyses were performed on the data. The percentages of explained variation (R^2) were compared after ten steps in the regression equations.

RESULTS AND ANALYSIS

HYPOTHESIS 1: Environmental variables revealed by remote sensing can predict mortality and morbidity rates.

This hypothesis was supported by data from both cities. All 16 multiple correlation coefficients between the environmental variables and health outcomes reached statistical significance of at least .05 (Table III). Fifty percent or more of the variation in the health outcomes was explained in seven cases after ten steps of the regression equations.

In the Houston study, the environmental variables explained 50% or more of the variation for shigella/salmonella (68%), infectious hepatitis/meningitis (50%), juvenile delinquency (57%), and mental health referrals (66%). In Galveston the health outcomes for which the environmental data set explained the most variation were TB (82%), meningitis (73%), and VD (74%).

The data collected by remote sensing were subdivided into density and quality variables for descriptive purposes. These correspond to the environmental conditions of density and poverty, previously demonstrated by others to be health related.

In both studies, census measures of density were internal measures of overcrowding while the physical environmental variables measured external density. In the Houston study, the environmental variables representing density ("schools," "residential medium density," and "apartments") explained the most variability for mental health referrals (35%) and juvenile delinquency (21%) while for mortality and TB they explained somewhat less (10% and 12% respectively). The environmental density variables in the Galveston study were not strong predictors for any of the health outcomes.

The environmental variables representing quality seemed to be stronger predictors of the infectious disease categories in Houston. For infectious hepatitis/meningitis, nonresidential land uses designating quality were the most important. "Unimproved land" entered first and accounted for 16% of the variation. This land use, particularly in low income neighborhoods, is frequently associated with the presence of litter and other unsanitary environmental conditions. "Streets" was the next variable to enter the equation explaining an additional 17% of the variability. Streets, like unimproved land may be a collection point for litter and rubbish which foster the spread of these communicable diseases. In lower income neighborhoods where front and back yards are practically nonexistent, streets and unimproved land may serve as play areas for children. Residential quality was the most important explanatory variable for shigella/salmonella. Both "fair" and"excellent" quality were inversely related to this disease rate explaining 35% of the variability. The three highest quality categories acted similarly in their relationship to all health outcomes. As expected, only poor quality housing showed a positive relationship to health.

The conclusions reached in measuring quality with environmental variables in the Galveston project were generally similar to those reached in the Houston project and were consistent with the literature, i.e., that poverty neighborhoods measured by poor housing and poor environmental conditions seem to be the setting for higher disease rates than neighborhoods of middle and upper income levels. For TB, the quality variable "parking lots" accounted for 57% of the variance while the "number of square feet per dwelling unit," which could be considered either a density or quality variable, accounted for an additional 13% of the variance. "Litter," "industry," and "multi-family residence" explained 63% of the variations in the VD data.

In addition to the infectious disease categories, mortality also showed a relationship to environmental quality variables in the regression models for Galveston. The first three variables accounting for 36% of the variation in mortality under age 18 were "industry," "parking lots" and "narrow lot frontages." For mortality, ages 18-61, the variables accounting for 35% of the variation were "house size," "industry" and "vacant land."

In Houston, remote sensing environmental variables representing both quality and density were important while in Galveston those representing quality showed stronger predictive ability. The pattern of association between the quality variables and the health outcomes in both Houston and Galveston was generally similar, i.e., quality variables were good predictors of infectious disease rates, with the addition of mortality in Galveston.

HYPOTHESIS 2: Population and housing variables reported by the Census can predict mortality and morbidity rates.

This hypothesis was also given support with twelve of the sixteen multiple correlation coefficients reaching statistical significance of .05. In seven cases 50% or more of the variation in health outcomes was explained.

The health variables which were explained by the census variables differed somewhat between the two studies. In Houston after ten steps 50% or more of the variations in the TB (65%), shigella/salmonella (72%), mortality (50%), and juvenile delinquency (55%), data were acounted for by census variables while in Galveston census data predicted at least 50% of the TB (54%), cardiac arrest/myocardial infarction (50%), and VD (71%) variations.

Census variables also consisted of both environmental quality and density measures.

Density as measured by census variables was not an important explaining variable for any of the health outcomes in the Houston study. Census derived quality measures appeared to be slightly stronger predictors of health levels than density measures for both cities. Ten of the sixteen first place variables entering each equation could be considered quality indicators. However, of the first two variables entering the equations half were labeled quality and half density. The four strongest census predictors were "% Black," "% 1 person households," "average value of owner occupied dwellings" and "% rental units." These indicators were the strongest predictors of mortality, TB, and mental health referrals for Houston while they were good predictors of VD, TB, hypertension and meningitis in Galveston.

The mix in the equations of quality and density variables from both remotely perceived and census data reflects the strong relationship between environmental quality and density. They are interacting variables which support one another in their relationship to health outcomes.

HYPOTHESIS 3: There is no difference in the health predictive strengths between environmental variables and census variables.

In eleven of the sixteen cases variation explained by environmental data was greater than that by census data. In only one of the five cases where census variables explained more variation was the difference in predictive strength greater than 10%.

It can be seen from Table II that the environmental variables in the Houston study had a higher level of association with three of the dependent health variables. In two cases--infectious hepatitis/meningitis and mental health referrals--the differences were substantially higher, 26% and 22% respectively. In only one of three cases where the census variables had greater predictive strength was the difference great--20% in the case of TB. For the other two--mortality and shigella/salmonella--the differences were slight, 8% and 4% respectively. Overall, however, in the Houston study the results showed about equal predictive strength for both environmental and census data sets.

In the Galveston study the environmental variables outperformed the census variables (Table III). The results showed that for eight of the ten dependent health variables, environmental variables accounted for a greater level of association and predictive strength than census variables. This suggests that environmental variables may be considered surrogates for the usual health correlates.

In only two cases, hypertension and cardiac arrest/myocardial infarction, did the predictive strength of the census model exceed that of the environmental model and in both of these cases the differences were slight, amounting to 4% and 5% respectively. This result is significant in that both health variables are chronic diseases and had not been expected to show strong association with the physical environment. It appears that even for chronic diseases the physical setting may act as a surrogate for social characteristics when examining the ecological distribution of these diseases.

The predictive strength of environmental variables in explaining mortality variation for all age groups in Galveston reached 42%, 49% and 26% respectively. These mortality models all reached significance of 5%, while only one age category (18-61) reached 5% significance in the census model. Since the mortality data is the most valid data set in the group of dependent variables, these data support the utilization of physical environmental variables in ecological studies of mortality, a frequently used measure of health status.

CONCLUSION

Environmental data derived from aerial photographs were found to equal or surpass census data in predicting health levels of urban subareas in the two communities studied. However, the environmental surrogates of health outcomes need to be validated in other communities. Remote sensing may offer an alternative data source to health planners. It may be both a faster and more current source than present ground survey methods in delimiting public health problem areas.

ACKNOWLEDGEMENTS

This investigation was partially funded by Contract NAS 9-12823, National Aero-.:utics and Space Administration, Johnson Space Center, Houston, Texas. Dr. Bartholomew P Hsi, Associate Professor of Biometry, University of Texas School of Public Health served as Statistical Consultant. Ms. Sally Vernon served as Project Director for the Houston study and Ms. Janice Goldstein served as Project Director for the Galveston study. Mr. Calvin B. Olsen performed the image analysis on both cities.

REFERENCES

Bowden, L. W.
1968 Multi-sensor Signatures of Urban Morphology, Function and Evolution. Technical Report No. 2, U.S.D.A. Contract No. 14-08-0001-10674. Departments of Geography, University of California at Los Angeles and Riverside.

DuBos, R.
1965 Man Adapting. New Haven: Yale U. Press.

Faris, R. E. L. and H. W. Dunham
1967 Mental Disorders in Urban Areas. Chicago: University of Chicago Press.

Galle, O. R., W. R. Gove and J. M. McPherson
1972 "Population Density and Pathology: What are the Relations for Man?" Science 176, April 7: 23-30.

Horton, F. E. and D. F. Marble
1970 "Housing Quality in Urban Areas: Data Acquisition and Classification through the Analysis of Remote Sensor Imagery," in Second Annual Earth Resources Aircraft Program Status Review Vol. I, Part 15, Houston, Texas Manned Spacecraft Center: 1-13.

Lieberman, E. J. and L. J. Duhl
1964 "Physical and Mental Health in the City," The Annals of the American Academy 352: 13-24.

Loring, W. C.
1964 "Residential Environment: Nexus of Personal Interactions and Healthful Development," Journal of Health and Human Behavior 5, Winter: 166-169.

Manji, A. S.
1968 The Uses of Conventional Aerial Photography in Urban Areas: A Review. Technical letter NASA-131, Washington, D.C.: U.S. Geological Survey.

Mullens, R. H., Jr.
1969 "Analysis of Urban Residential Environments Using Color Infrared Aerial Photography: An Examination of Socio-economic Variables and Physical Characteristics of Selected Areas in the L.A. Basin," Interagency Report NASA-153.

Mumbower, L. E. and J. Donoghue
1967 "Urban Poverty Study," Photogrammetric Engineering 33: 610-618.

National Health Survey
1972 "Health Characteristics of Low-Income Persons." National Health Survey, Series 10, No. 74, Mary Lou Bauer, (ed.) U.S. Dept. of H.E.W., Health Services and Mental Health Administration, Rockville, Md., (July).

Rush, M. N. and S. Vernon
1973 Urban Environmental Health Applications of Remote Sensing, Contract NAS 9-12823, The University of Texas Health Science Center at Houston School of Public Health, December, 1973, (Unpublished).

Rush, M. N., J. Goldstein and B. P. Hsi
1974 Urban Environmental Health and Remote Sensing, Galveston, Texas, Contract NAS 9-12823, The University of Texas Health Science Center at Houston School of Public Health, November, 1974, (Unpublished).

Schmitt, R. C.
1966 "Density, Health, and Social Disorganization," Journal of the American Institute of Planners 32 (Jan.): 38-40.

Stallones, R. A.
1972 "Community Health," Science 175, No. 4024 (25 February).

Wellar, B. S.
1968 Generation of Housing Quality Data from Multi-band Aerial Photographs. Interagency Report NASA-119, Washington, D.C.: U.S. Geological Survey.

TABLE I. - LAND USE CATEGORIES HOUSTON AND GALVESTON

HOUSTON		GALVESTON	
1-12	Residential-Single Family Quality: Excellent, Good, Fair, Poor Density: High, Medium, Low	1 2 3	Residential-Single Family Multi-family 1-3 story Multi-family +3 stories
13	Apartments		
14	Commercial	4	Commercial
15	Industrial	5	Industrial
16	Unimproved	6	Vacant & Unimproved
17	Parks & Recreation Areas	7	Open Space
18	Cemetaries		
19	Churches	8	Community Facilities
20	Schools & Educational facilities		
21	Hospitals & Health-related facilities		
22	Streets	9	Parking Lots

TABLE II. - HEALTH OUTCOMES HOUSTON AND GALVESTON

HOUSTON	GALVESTON
Mortality-sex and age group adjusted (3)	Mortality - ages 17 and under Mortality - ages 18-61 Mortality - ages 62 and over
Tuberculosis (incidence)	Tuberculosis (prevalence)
Shigella & Salmonella	Shigella & Salmonella
Infectious hepatitis & Meningitis	Infectious hepatitis Meningitis
Juvenile delinquency Mental health referrals	
	Venereal disease Hypertension Cardiac arrest/myocardial infarction

TABLE III. - EXPLAINED VARIATION (CUMULATIVE MULTIPLE CORRELATION COEFFICIENT SQUARED) OF FIRST 10 VARIABLES IN STEP-WISE REGRESSION HOUSTON AND GALVESTON

Health Outcome	Houston		Galveston	
	Environmental Variables	Census Variables	Environmental Variables	Census Variables
	R^2	R^2	R^2	R^2
Tuberculosis	.45	.65	.82	.54
Shigella/Salmonella	.68	.72	.42	.25
Hepatitis			.28	.16*
Meningitis			.73	.29
Inf. Hep/Men	.50	.24*		
Mortality	.42	.50		
Under Age 18			.42	.20*
Ages 18-61			.49	.35
Over 61			.26	.12
Cardiac Arrest/MI			.45	.50
Venereal Disease			.74	.71
Hypertension			.45	.49
Juvenile Delinquency	.57	.55		
Mental Health Referrals	.66	.44		

*All variables were significant at the .05 level except those marked with this symbol