*Am. J. Trop. Med. Hyg.*, 61(3), 1999, pp. 356–360


# APPLICATION OF RISK ANALYSIS AND GEOGRAPHIC INFORMATION SYSTEM TECHNOLOGIES TO THE PREVENTION OF DIARRHEAL DISEASES IN NIGERIA

PHILIP C. NJEMANZE, JOSEPHINE ANOZIE, JACINTHA O. IHENACHO, MARCIA J. RUSSELL, AND AMARACHUKWU B. UWAEZIOZI

*Institute of Space Medicine and Terrestrial Science, and Institute of Non-invasive Imaging for Parasitology, International Institutes of Advanced Research and Training, Chidicon Medical Center Owerri, Imo State, Nigeria; Imo State Water Corporation, Imo State Government, Owerri, Imo State, Nigeria*

*Abstract.* Among the poor in developing countries, up to 20% of an infant's life experience may include diarrhea. This problem is spatially related to the lack of potable water at different sites. This project used risk analysis (RA) methods and geographic information system (GIS) technologies to evaluate the health impact of water source. Maps of Imo State, Nigeria were converted into digital form using ARC/INFO GIS software, and the resulting coverages included geology, hydrology, towns, and villages. A total of 11,537 diarrheal cases were reported. Thirty-nine water sources were evaluated. A computer modeling approach called probabilistic layer analysis (PLA) spatially displayed the water source at layers of geology, hydrology, population, environmental pollution, and electricity according to a color-coded five-point ranking. The water sources were categorized into A, B, and C based on the cumulative scores $< 10$ for A, 10–19 for B, and $> 19$ for C. *T*-test showed revealed significant differences in diarrheal disease incidence between categories A, B, and C with mean $\pm$ SEM values of $1.612 \pm 0.325$, $6.257 \pm 0.408$, and $15.608 \pm 2.151$, respectively. The differences were significant between categories A and B ($P = 0.0000022$), A and C ($P = 0.0000188$), and B and C ($P = 0.0011348$). The PLA enabled estimation of the probability of the risk of diarrheal diseases occurring at each layer and solutions to eliminate these risks.

Diarrhea is an increase in daily stool weight above 200 g per day and an increase in stool liquidity and frequency. Watery diarrhea is caused by several microorganisms; the most prevalent causative agents include *Vibrio cholerae, Escherichia coli, Giardia lamblia,* and *Salmonella typhi.* In Africa, Asia, and South America, acute diarrheal illnesses are not only a leading cause of morbidity in children, producing an estimated 1 billion cases per year, but also are responsible for 4–6 million deaths per year, or 12,600 deaths per day. In some areas in sub-Saharan Africa, more than 50% of the deaths of children are directly attributable to acute diarrheal illnesses.[1,2] This problem is most evident in Nigeria, Africa's most populous country.[3]

The objective of the present project was to devise a model that will eventually lead to reduction in the incidence of diarrheal diseases related to water-borne infections. This project used a new approach of risk analysis and geographic information system (RA/GIS) technologies to perform health and environmental impact assessment of water source (treated surface water and tube wells or bore holes).

The GIS provides computerized capture, storage, management, analysis, retrieval, and display of spatial and descriptive data that are geographically referenced to a common coordinate system. It is composed of a database (spatial and attribute), data input (digitized scan), cartographic display system, database management, and geographic analysis systems (e.g., overlay process and buffer zone creation). The spatial database contains information held in the form of digital coordinates, which describe spatial features.[4]

Risk analysis includes risk assessment, risk management, and risk communication. Risk assessment is the process of identifying a hazard and evaluating the risk of a specific hazard, either in absolute or relative terms. Risk assessment provides estimates of the probability of occurrence of unwanted future events. This information is used by the risk management teams to choose the best course of action. This approach has been previously applied in environmental assessment.[5]

## MATERIALS AND METHODS

The population of the area studied was 2,311,675 persons 0–85 years of age in 18 local government areas of Imo State, Nigeria. The populations in the various areas did not differ significantly with regard to socioeconomic status, educational level and access to health care based on a self-reported questionnaire. The reported diarrheal diseases were obtained from the annual disease surveillance report from the various health centers in the local government areas studied. Digital coverages were created using ARC/INFO GIS software (Environmental Systems Research Institute, Inc., Redlands, CA). These coverages were derived from existing 1:350,000 scale maps of Imo State. The resulting GIS layers included geology, hydrologic features, towns, and villages. All coverages were converted from state plane coordinates to the universal transverse mercator coordinate system and projection. The digitized geology was color-coded according to soil types. The true spatial relationship of the geology, hydrology, towns, and villages was provided using an overlay of the layers in the GIS database. The geology coverage was used to determine suitability of the various areas for placement of bore holes. The hydrologic features were used initially to determine the proximity of rivers and streams to towns and villages studied. The use of the water source by inhabitants of the towns and villages was latter confirmed by field data. The means of distribution of water was considered reticulated when water distribution was by an extensive pipe-borne water network or satellite when water was collected for use from the site of a bore hole. The latter entailed walking short distances to the site of the bore hole by users.

To examine the contribution of the various layers of the database to the overall probability (P) of a hazard occurring, each data layer was sequentially considered by what we here refer to as probabilistic layer analysis (PLA). The PLA is a method of risk assessment that begins with identification of a hazard endpoint and sequentially weighing the factors contributing to this endpoint or its probability at each layer of

Table 1
Diarrheal disease incidence rate (DDIR) in the communities served by categories A, B, and C water sources in Imo State, Nigeria

| Category A | DDIR | Category B | DDIR | Category C | DDIR |
|---|---|---|---|---|---|
| Aguneze | 1.504 | Awarra, Oguta | 7.669 | Agbala | 35.294 |
| Amala | 0.166 | Eziudo | 6.836 | Anara | 14.692 |
| Amorie | 2.648 | Mbutu | 4.439 | Ejemekwuru | 24.065 |
| Awarra, Ohaji-Egbema | 0.981 | Nwaoriubi | 6.851 | Eke Nguru | 8.446 |
| Dikenafi | 0.382 | Obodoukwu | 3.078 | Eziama-Ntu | 12.615 |
| Egbema | 0.093 | Obiangwu | 6.032 | Ezi-Orsu | 10.357 |
| Enyiogugu | 1.884 | Ogbaku | 7 | Eziagbogu | 11.528 |
| Lorji | 2.649 | Oguta, Egbuoma | 5.734 | Eziama-Obaire | 9.781 |
| Nnarambia | 0.972 | Oke Uvuru | 8.731 | Ife | 9.805 |
| Ohaji Egbema | 1.851 | Okpala | 5.74 | Izombe | 11.861 |
| Owerri | 1.236 | Ulakwo | 6.364 | Onicha | 16.731 |
| Umuohi | 4.206 | Umuhu | 5.298 | Orsu-Obodo | 24.039 |
| Umuohiagu | 2.395 | Umuneke | 7.576 | Osina | 13.697 |

a spatially referenced database. This in principle relates to probabilistic scenario analysis.[5–9] Probabilistic modeling has been applied in other fields.[6–12] However, spatial features are integrated in the PLA.

A computer model comprising a two-dimensional representation of each layer (geology, hydrology, population, environmental, and electricity) was drawn and stacked respectively in sequence. The data from the GIS coverages, field collected data, and population were color-coded in a five-point scale (from 1 to 5) of magnitude for each layer, respectively. For example, on the geologic layer, the most favorable was considered water available at a good depth of subterranean aquifers that diluted to a safe concentration any fecal contamination, excluded other known sources of toxicity, and at the same time was cost-effective. Each water source was assigned an applicable value at each layer. A line was projected through the various layers at sites corresponding to the features of a particular water source, starting with an initiating event (IE) and terminating with an end point (EP) considered as the incidence rate of diarrheal disease.

The relative risk (RR) and attributable risk (AR) due to factors operational at each layer were computed as follows: RR = diarrheal rate among those exposed to source/diarrheal rate among those unexposed to source = $P(D/S)/P(D/\bar{S})$; AR is defined as the proportion of diarrheal disease cases attributable to the source or as the proportional decrease in incidence of diarrheal disease if the entire population was not affected by the source. $AR = P(D/S) - P(D/\bar{S})/P(D/S) = RR - 1/RR$ where P(D) = the prevalence of diarrheal disease, P(S) = proportion of people exposed to the source, P(S/D) = proportion of patients with diarrhea exposed to the source, $(S/\bar{D})$ = proportion of people free of diarrhea exposed to the source, P(D/S) = proportion of patients exposed to the source who have diarrhea; $P(D/\bar{S})$ = proportion of patients not exposed to the source but who have diarrhea, $P(\bar{D}/\bar{S})$ = proportion of people not exposed to source and are free of diarrhea.

The population figures were obtained from the National Population Commission based on head count of 1991 and the projections for 1996 for each community studied.[3] Discriminant function analysis using the software package STATISTICA for Macintosh™ (Statsoft, Tulsa, OK) was applied to determine which variables discriminate between the preselected groups. Data were expressed as the mean ± SEM, and significance was accepted at the 5% level. Other statistical analysis of the epidemiologic data of diarrheal disease prevalence was based on the probability concept as given above.[13]

## RESULTS

A total of 11,537 diarrheal cases were reported in the period January to December 1996. There was an annual incidence rate of 5 per 1,000. The male to female ratio was 1:1. Children less than five years of age comprised 46.6% of the patients. The study included 39 community water sources. Of these, one was a treated surface water source. The others were bore holes that were either hand-operated tube wells or pump-operated bore holes. The diarrheal disease incidence in the communities served by the water sources are summarized in Table 1 and categorized into A, B, and C.

Discriminant function analysis revealed significant differences in diarrheal disease incidence between categories A, B, and C, with means ± SEM values of 1.612 ± 0.325, 6.257 ± 0.408, and 15.608 ± 2.151 respectively. The differences were significant between categories A and B ($P$ = 0.0000022), A and C ($P$ = 0.0000118), and B and C ($P$ = 0.0011348). Thus, category A is associated with low risk for diarrheal disease, category B with moderate risk, and category C with high risk. There was no difference related to the means of distribution of the water (satellite versus reticulated; $P$ = 0.61).

Figure 1 shows the models derived from the PLA approach using three layers. A category A water source is situated in a geologically most favorable area with a very low probability (p1) of contributing to the overall diarrheal disease hazard. It serves an area that is very densely populated with people living a safe distance from the water source, reducing the risk of human waste contamination, thus a low probability (p2). The water source is functional most of the time, thereby reducing the need for the usually polluted alternative water sources from rivers and streams; thus, a low probability (p3). The overall probability of the hazard is a low PA = (p1)(p2) (p3).

Category B and C water sources were analyzed in a similar fashion. For people using a category B water source, alternative sources of water (rivers, streams, and lakes) are easily accessible, especially in periods of non-function of the

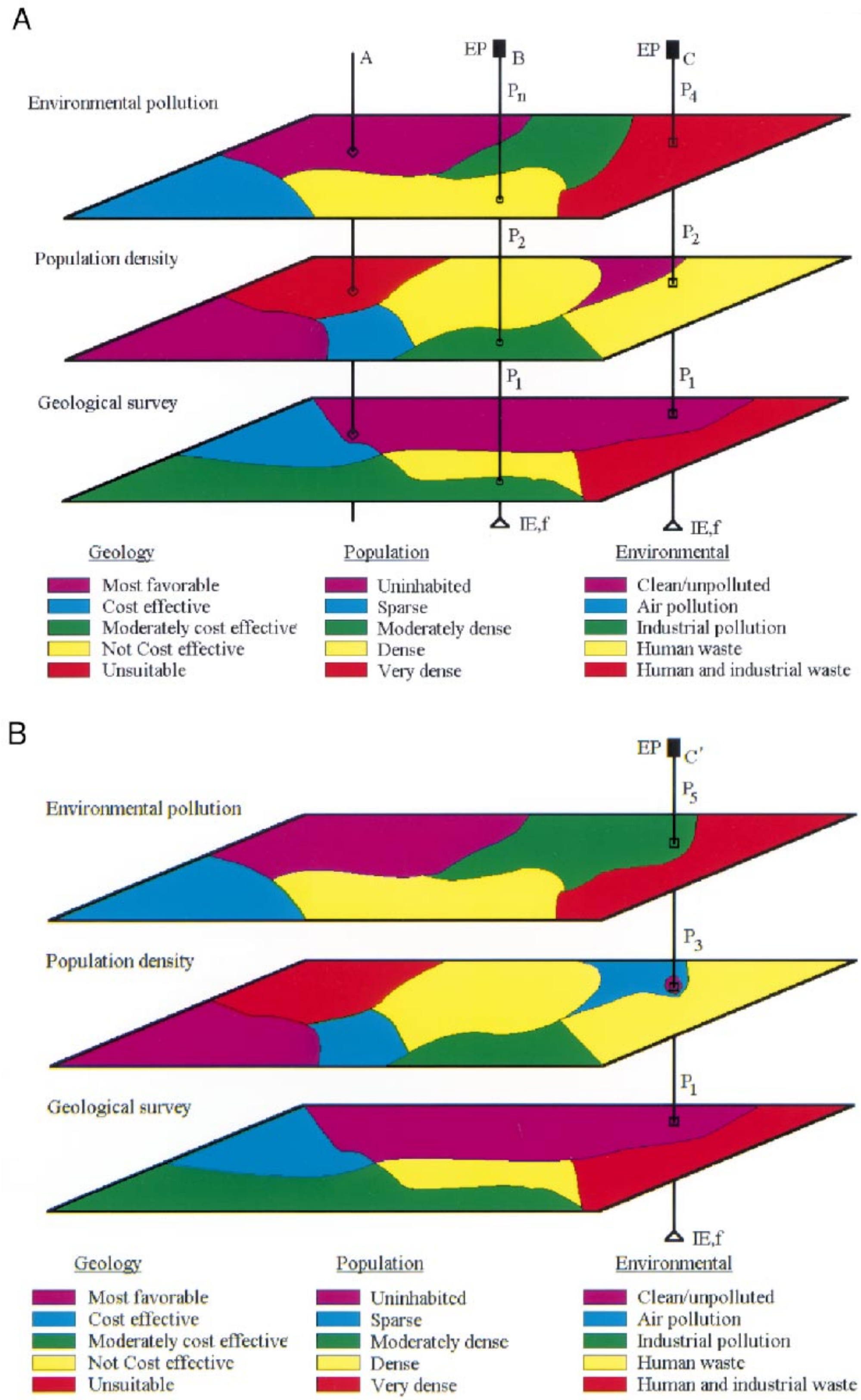


FIGURE 1. **A,** probabilistic layer analysis (PLA) application to the three categories (A, B, and C) of water sources (see Results for an explanation). **B,** PLA approach to modify factors acting on category C to reduce diarrheal diseases incidence in C′. EP = endpoint; IE = initiating event.

water source, leading to multiple contributing factors (pn) such that pn > p2 > p1; PB = (p1)(p2)(pn). For a category C water source, easy access to untreated surface water alternatives and other environmental hazards due to human and industrial activities increased the probability (p4) of the hazard. The overall probability of the hazard is high; PC = (p1)(p2)(p4).

Recommendations were made to local authorities to reduce overall risk by measures aimed at modifying conditions at one or more layers (Figure 1B). For example, category C water source requires creating a buffer zone around the water source. This implies measures as simple as raising the platform of the hand-operated tube well so that people cannot clean their laundry around the well, to building restrictions that require an offset of several meters from the water source. These measures limiting unwanted human activities around the site will modify the impact of the population layer, and the environmental layer will be free of human waste, such that p3 < p2 and p5 < p4. This reduces the overall probability PC′ = (p1)(p3)(p5). Other recommendations included perimeter fencing of the contaminated alternative water source to reduce access, relocation of animal slaughter houses to areas remote from the source to prevent contamination by animal waste, periodic flushing of pipes or replacement of old pipes for reticulated networks, and regular water treatment. The implementation of these recommendations and monitoring of effects have yet to be evaluated in most communities.

## DISCUSSION

The results show that application of spatially referenced RA methods helped to convert the statistical and tabular data into meaningful easy-to-understand information for decision-making. These applications in the planning of water resources development policy are a novel idea, applied in this manner for the first time in literature. The integration of the PLA in the GIS database is a further development in risk assessment of environmental and health impacts with potentially wide applications.

Water is a scarce resource in most African countries. This is further exacerbated by much wasting of water through poor water-use efficiency. Poor planning of water resource facilities has led to a situation in which there is no restriction to ground water mining (drilling wells); some areas have too much water and others have none. In this regard, the PLA is most appropriate for planning agricultural, health, environmental, mining, and other development programs in Africa. Examining the intersectoral relationships is much easier from a regional perspective when RA/GIS is applied. For example, the relationship between water resources, agriculture, and health becomes evident when information at the corresponding layers of the GIS database is examined.

In Nigeria, as in most developing countries, the problem of data collection is compounded by rapid rates of change in administrative boundaries and city size, and bureaucratic and political factors. The existing maps are at least a decade old. There are numerous agencies involved in water resource management, yet there is no comprehensive data pool. However, while it may seem attractive to create a comprehensive centralized national GIS database for planners to use for PLA applications, such a system will be plagued by inefficiency that is typical of such government-backed schemes. Moreover, the affordable technology may not support such a massive scale program. It is more practical and easy to manage a GIS database for water resource management at state levels. This condition will ensure that indigenous expertise can have effective control of this useful technology. However, a centralized GIS database for PLA applications can be adopted for specific larger water resources development programs. To overcome some of these limitations, use of recently acquired high-resolution remote sensing images may help provide detailed up-to-date maps. In addition, normalized vegetation index satellite images from very high resolution radiometers may be useful in estimating overall water needs for plants. This may help quantify, in addition to rainfall data, the water sources required in a particular region.

There is need for further studies using multiple endpoints including biologic and chemical factors affecting water purity. The recent tragedy in Bangladesh related to arsenic poisoning from an underground water source[14] further emphasize the need for PLA with multiple endpoints including toxicology to be used in future designs of public water projects. The ability to apply PLA to computer models permits diagnosis, monitoring, and project evaluation in the pre-implementation stage of water projects.

Other areas of application of PLA include environmental impact analysis of toxic wastes for health-related endpoints, e.g., oil spills in southeastern Nigeria or in the case of the Exxon Valdez accident in Alaska. Forecasts of health impact can be made for new facilities and for evaluation of existing facilities that work with nuclear materials. Planned air strikes, as in the Kosovo conflict, may acquire a new dimension if the choice of targets was analyzed using PLA to evaluate the effects the target facilities will have on health-related endpoints. The construction of roads and bridges can be better assessed on how they will impact on the transportation layer in a PLA model. Market analysis of goods and services in a geographic area can be better appreciated with this model. The PLA can be applied in several areas not directly related to the present work; these include environmental impact assessments, and construction, industrial, commercial, and military issues. Further studies are required in these areas to demonstrate that GIS with PLA model applications will have potential benefits.

In summary, PLA used in conjunction with GIS technologies have much potential application in developmental programs in African countries, particularly for risk assessment in the health sector. Its application to development of water resources and evaluation of health impact on diarrheal diseases remains to be fully demonstrated. This demonstration must come from scientists and policy makers of developing countries themselves if it is to be meaningful to the development process.[15]

Acknowledgments: Philip C. Njemanze was a Third World Foundation Award Scholar at the Center for Health Applications of Aerospace Related Technologies (CHAART) at the National Aeronautic and Space Administration (NASA) Ames Research Center.

Financial support: Funding for this work was provided by the Chidicon Medical Research Foundation.

Authors' addresses: Philip C. Njemanze and Marcia J. Russell, Institute of Space Medicine and Terrestrial Science, International Institutes of Advanced Research and Training, Chidicon Medical Center, No. 1 Uratta Road, Owerri, PO Box 302, Imo State, Nigeria. Josephine Anozie and Jacintha O. Ihenacho, Institute of Non-invasive Imaging for Parasitology, International Institutes of Advanced Research and Training, Chidicon Medical Center Owerri, Imo State, Nigeria. Amarachukwu B. Uwaeziozi, Water Corporation, Imo State Government, Owerri, Imo State, Nigeria.

## REFERENCES

1. World Health Organization, 1993. *Tropical Disease Research.* Geneva: World Health Organization.
2. Field M, ed, 1991. *Diarrheal Diseases.* New York: Elsevier.
3. National Population Commission, 1996. *Final Results of 1991 Population Census of Nigeria.*
4. Loslier L, 1995. Geographic information systems (GIS) from a health perspective. Wijeyaratne P, ed. *GIS for Health and the Environment.* Ottawa, Ontario, Canada: International Development Research Center, 13–20.
5. National Research Council, 1983. *Risk Assessment in the Federal Government: Managing the Process (The Red Book).* Washington, DC: National Academy Press.
6. Hammersley JM, Handscomb DC, 1964. *Monto Carlo Methods.* New York: John Wiley & Sons.
7. Kaplan S, Garrick BJ, 1981. On the quantitative definition of risk. *Risk Anal 1:* 11–27.
8. Miller L, McElavaine MD, McDowell RM, Ahl AS, 1993. Developing a quantitative risk assessment process. *Rev Sci Technol Office Int Epizootics 2:* 1153–1163.
9. McElavaine RM, McDowell, Fite RW, Miller L, 1993. An assessment of risk of foreign animal disease introduction into the United States of America through garbage from Alaska cruise ships. *Rev Sci Technol Office Int Epizootics 2:* 1165–1173.
10. Ahl AS, 1994. From farm to table; applications of risk analysis to the safety of foods of animal origin: technology. *J Franklin Inst 331A:* 77–82.
11. Hertz DB, 1964. Risk analysis in capital investment. *Harv Bus Rev 42:* 95–108.
12. Ahl AS, 1996. The application of probabilistic scenario analysis for risk assessment of animal health in international trade. Camus E, House JA, Uilenberg G, eds. *Vector-Borne Pathogens: International Trade and Tropical Animal Diseases. Ann NY Acad Sci 791:* 255–268.
13. World Health Organization, 1992. *Health Research Methodology.* Manila: World Health Organization.
14. Chatterje P, 1999. Water of life or death. *Development and Cooperation* (*March/April*): 28.
15. Njemanze PC, 1996. Satellite technology and aerospace related warning systems (STARWARS) for disease control: a strategy for disease prevention in developing countries of Africa, South America and Asia-Pacific region. *Jpn J Aerosp Environ Med 33:* 117–130.