A Geographic Information System for Characterizing Exposure to Agent Orange and Other Herbicides in Vietnam

Jeanne Mager Stellman,[1] Steven D. Stellman,[2,3] Tracy Weber,[1] Carrie Tomasallo,[1] Andrew B. Stellman,[4] and Richard Christian, Jr.[5]

[1]Department of Health Policy and Management and [2]Department of Epidemiology, Mailman School of Public Health, Columbia University, New York, New York, USA; [3]Institute for Cancer Prevention, Valhalla, New York, USA; [4]Foundation for Worker, Veteran, and Environmental Health, Inc., Brooklyn, NY, USA; [5]Lt. Col. U.S. Army (retired) and former Director U.S. Army and Joint Services Environmental Support Group, Washington DC, USA

Between 1961 and 1971, U.S. military forces dispersed more than 19 million gallons of phenoxy and other herbicidal agents in the Republic of Vietnam, including more than 12 million gallons of dioxin-contaminated Agent Orange, yet only comparatively limited epidemiologic and environmental research has been carried out on the distribution and health effects of this contamination. As part of a response to a National Academy of Sciences' request for development of exposure methodologies for carrying out epidemiologic research, a conceptual framework for estimating exposure opportunity to herbicides and a geographic information system (GIS) have been developed. The GIS is based on a relational database system that integrates extensive data resources on dispersal of herbicides (e.g., HERBS records of Ranch Hand aircraft flight paths, gallonage, and chemical agent), locations of military units and bases, dynamic movement of combat troops in Vietnam, and locations of civilian population centers. The GIS can provide a variety of proximity counts for exposure to 9,141 herbicide application missions. In addition, the GIS can be used to generate a quantitative exposure opportunity index that accounts for quantity of herbicide sprayed, distance, and environmental decay of a toxic factor such as dioxin, and is flexible enough to permit substitution of other mathematical exposure models by the user. The GIS thus provides a basis for estimation of herbicide exposure for use in large-scale epidemiologic studies. To facilitate widespread use of the GIS, a user-friendly software package was developed to permit researchers to assign exposure opportunity indexes to troops, locations, or individuals. *Key words:* Agent Orange, cacodylic acid, defoliants, 2,4-dichlorophenoxy acetic acid, exposure opportunity, geographic information system, GIS, herbicides, military, picloram, 2,4,5-trichlorophenoxy acetic acid, Vietnam. *Environ Health Perspect* 111:321–328 (2003). doi:10.1289/ehp.5755 available via *http://dx.doi.org/* [Online 1 November 2002]

Between 1961 and 1971, the United States Armed Forces in South Vietnam used nearly 19.5 million gallons of chemical herbicides for tactical defoliation and crop destruction. Several types and combinations of chemicals were used. The mixtures were nicknamed by the color of the identification stripe that appeared on their chemical storage drums (National Research Council 1974). The three most common mixtures were Agent Orange [esters of 2,4-dichlorophenoxyacetic acid (2,4-D) and 2,4,5-trichlorophenoxyacetic acid (2,4,5-T)], Agent White (triisopropanolamine salts of 2,4-D and picloram), and an arsenical called Agent Blue (cacodylic acid). It has been estimated that Agent Orange was contaminated with < 0.05–100 ppm of dioxin (Darrow 1971). The U.S. Air Force in its Operation Ranch Hand dispersed between 97% and 98% of all the herbicides used. Figure 1 shows a typical Ranch Hand mission composed of three C-123 aircraft (each known as a "sortie") flying in typical tight formation and usually spraying at an altitude of 150 feet. Spraying was only carried out in clear weather, with wind speeds less than 8–10 knots and no temperature inversion present.

An estimated 3.2 million American men and women served in the Armed Forces in Vietnam, many of whom were assigned to areas defoliated by herbicides such as Agent Orange (Stellman et al. 1988a). Neither the extent of exposure nor long-range health effects are fully known after 30 years. The Agent Orange Act of 1991 (P.L. 102-4) directed the Secretary of Veterans Affairs to request that the National Academy of Sciences (NAS) conduct a comprehensive review and evaluation of available scientific and medical information about the health effects of Agent Orange and other herbicides. The Academy convened a committee of experts at the Institute of Medicine (IOM), which since 1994 has issued biennial reports that have associated a number of veteran illnesses with herbicide exposure, based on data drawn primarily from nonveteran observations (IOM 1994, 1996, 1998, 2001). The IOM noted that studies of Vietnam veterans were urgently needed but first required development of exposure reconstruction models that could become the basis of the new epidemiologic studies (IOM 1994). In 1996, a request for proposals for such exposure modeling was issued by NAS in a contract with the Department of Veteran Affairs. The system described herein has been developed by the project awarded the NAS subcontract.

Conceptual Framework for Use of a Geographic Information System in Vietnam Studies

We describe here a geographic information system (GIS) that we developed for this study to address requirements of both epidemiologic and environmental studies for a methodology whereby chronologic listings of the dates, amounts, and chemical details of herbicide spray missions can be transformed into vectors consisting of various measures of exposure opportunity. A GIS is a powerful tool that can be used to facilitate exposure assessment by combining and integrating a variety of data resources, such as those encountered in environmental and epidemiologic studies. Studies of veterans in particular will require that data on military assignments and duties be extracted, analyzed, and transformed into a format compatible with the assessment of herbicide exposure opportunity, and the conceptual framework must further be able to take into account potential confounders and covariates of exposure, such as combat stress and occupational hazards, which have been shown to be associated with physical and mental health outcomes (Stellman et al. 1988b).

Our GIS approach to assessment of herbicide exposure in Vietnam uses a framework of records-based exposure reconstruction methodology that is becoming increasingly common in environmental and epidemiologic

studies. Recent approaches have included such diverse types of studies as estimation of herbicide application in small area tracts in the U.S. Midwest (Ward et al. 2000), residential exposure to agricultural and commercial chemicals on Cape Cod, Massachusetts (Brody et al. 2002), and "pesticide use density" in studies of childhood cancer in California (Gunier et al. 2001). Additional applications are constantly appearing, many of which may be found by linking through a federal website (GIS and Public Health 2003).

The specific function of our Vietnam GIS is to represent spatial and temporal relationships between instances of herbicide

Figure 1. Photo of three C-123 aircraft sorties dispersing herbicides during a Ranch Hand mission over South Vietnam.

application and the locations of exposed individuals, military units, or other defined populations. The Vietnam GIS incorporates several extensive data resources, which include *a*) distribution of herbicide spray based on a cleaned and updated spray record (HERBS file); *b*) Air Force Ranch Hand project and target maps; *c*) databases of locations of military troops; *d*) databases of Vietnamese population centers such as villages and hamlets; and *e*) data on Vietnam land characteristics such as soil typology, waterways, and ecologic parameters. The GIS incorporates spatial and temporal proximity algorithms that operationalize mathematical models to permit assignment of quantitative exposure opportunity indexes (EOIs) to military units or individuals. This functionality makes it a useful tool for epidemiologic investigations of health outcomes.

Description of the GIS

The GIS is a relational database whose component tables contain exposure data, population or troop occupancy data, geographic data, and other information relevant to exposure estimation (Stellman et al. 2002). The GIS is built around two interrelated concepts: *a*) geographic partitioning of Vietnam into $0.01° \times 0.01°$ "square" grids, and *b*) association of each grid with counts of "hits" from the individual herbicide missions that occurred inside the grid or near it, as well as with estimates of exposure opportunity that are based on more elaborate exposure models. The grid system covers all of South Vietnam (176,060 grids), the island of Phu Quoc, and sprayed areas of Laos and Cambodia.

Using the grid point as the linkage key, the GIS incorporates geocodable external reference data resources (i.e., those that can be expressed in longitude/latitude coordinates). Examples are listed in Table 1. They include locations of inhabited places in South Vietnam during the

Vietnam War that have been cataloged in a gazetteer, and on which we have carried out extensive quality control. The coordinates of many constructed locations, such as roadways and utilities, are contained in commercial data sources, such as ADC WorldMap (American Digital Cartography, Inc., Appleton, WI, USA). This resource also provides coordinates of a variety of land features, including elevations and land contours, rivers and streams, mountains and highlands, coastal areas and mangrove forests, and bays and estuaries. By linking to their coordinates (expressed as grid points), exposure estimates can be restricted to or stratified by topographic features. This may be useful if different exposure models are posited for different land types or soil typology, which affects the ground retention time of herbicides (Stellman et al. 2001).

We have also developed a number of special-purpose tables of locations of military bases, base camps, landing zones, and airfields. A unique table, described below, contains the perimeter coordinates of more than 400 targets that were used for carrying out Operation Ranch Hand.

Herbicide Spray Database

The GIS makes use of a comprehensive file of Ranch Hand herbicide spray applications that was originally created by the Department of Defense and is known as the HERBS file (Data Management Agency 1970). A supplementary database called the Services HERBS file, which includes data on nonaircraft herbicide applications as well as previously unrecorded Ranch Hand missions, was released by the U.S. Army and Joint Services Environmental Support Group [now known as the U.S. Armed Services Center for Research of Unit Records (CRUR)] (U.S. Army & Joint Services Environmental Support Group 1985). The combined HERBS database describes nearly 19.5 million gallons of

Table 1. Examples of types of location data available in the GIS to be linked with EOIs.

Type of activity or data	Original sources	GIS Adaptation
Civilian habitations: towns, villages, hamlets, plantations	Gazetteer of Vietnam, Hamlet Evaluation System[a]	Cleaned and incorporated as GIS tables
Vietnam land and water features	FAO map (paper, 40 × 27 in) of Vietnam soil typologies[b]	Scanned (National Reprographics, New York, NY); vectorized (Digital GeoTechnologies, Washington, CT); then processed via point-in-polygon application to create soil region–grid linkages
	ADC WorldMap[c], other commercial sources: land topography, rivers, streams as of about 1970	Unchanged
Civil structures	ADC WorldMap: roadways, utility lines, rail lines, canals, airfields	Unchanged
Military structures: base camps, landing zones, air fields	NARA documentation (paper)	Data entered, proofed, and consistency checked, incorporated into GIS tables
Troop locations	Troop strength lists, Army Post Office lists (~100,000 paper records), assorted military records	Data entered, proofed, and consistency checked, incorporated into GIS tables
Operation Ranch Hand targets	NARA paper documents and folders	Digitized[d] or manually data entered and incorporated into GIS tables
Herbicide storage, transport, and unplanned dispersal	Services HERBS file	Corrected[e] and integrated into HERBS file and GIS

Abbreviations: FAO, U.N. Food and Agricultural Organization; NARA, U.S. National Archives and Records Administration.
[a]Initial data were uncleaned, in an obsolete computer format and on obsolete computer media. [b]Moormann (1961). [c]American Digital Cartography, Inc., Appleton, WI. [d]R2V Software, Able Software, Inc., Belmont, MA. [e]In collaboration with CRUR, Department of Defense.

herbicides, as broken down in Table 2. We have carried out extensive quality control on all spray data, which we validated by meticulous scrutiny of primary sources available from CRUR, the U.S. National Archives, and other sources.

The HERBS file and all military herbicide operations are organized by mission. The Stellman-NAS version of the HERBS file (S-NAS-HERBS) contains data for 9,141 spray missions for all modes of delivery, carried out by both the Air Force (Ranch Hand) and the Army. A Ranch Hand mission consisted of one or more fixed-wing or helicopter sorties, dispersing a specified amount and type of herbicide along a specified route on a specified day. The Services HERBS file also represents as missions the ground applications carried out by truck or backpack spraying or other non-aerial means.

Most mission records in the S-NAS-HERBS file contain data on the chemical agent used, the military purpose of the mission, its date, the number of aircraft flown, the area sprayed, and the coordinates of the path taken by the aircraft or other applicator. Some missions may have null data for fields pertaining to aircraft parameters. Helicopter, truck, or backpack missions differed from aerial spraying in that much smaller amounts of herbicides were used in different patterns and with smaller potential dispersion than from the C-123 fixed-wing aircraft.

The HERBS file provides the actual flight paths taken by Ranch Hand aircraft as they carried out their spray missions. Alphanumeric indicators show the locations at which the aircraft switched directions or turned off and on their spray nozzles. The data are thus structured in a way that emphasizes the continuity of flight of the fixed-wing aircraft that carried out most of the Operation Ranch Hand missions. (Connectivity may not apply to ground perimeter spraying, which generally went from guard post to guard post around the base camp.) To convey this information, each mission within the HERBS file is organized as a sequence of "vertices" that were the starting, turning, and stopping points of spray aircraft as they carried out a mission. Figure 2 illustrates defoliation along a roadway. The flight path consists of four connected segments in which the plane begins spraying at point 1A and continues spraying until it reaches point 1E. Points 1B, 1C, and 1D are intermediate "turning" points at which the plane changes direction while continuing to spray. The dashed line represents a 1-km envelope about the spray path. Planes frequently flew multiple paths in a single mission; additional vertices would be designated 2A, 2B, 3A, 3B, and so on. Spray nozzles would be turned off between ending one path and beginning another.

The spray path legs of C-123 missions averaged 8.1 km [95% confidence interval (CI), 8.0–8.2], and the total spray route was on average 16.8 (95% CI, 16.6–17.0) km in length, dispersing 900–1,000 gallons per sortie. Complete spray flight paths exist for the great majority of Ranch Hand missions. Some early 1965 HERBS fixed-wing missions contain records with a single coordinate, generally representing the starting point or center-of-mass of a mission. We have developed schemata to impute the likely flight paths for many of the fixed-wing missions based on flights flown over the same target. Imputed coordinates are always identified as such in the database so that they may be included or not, at the user's option.

By contrast, perimeter spraying may indeed have consisted of a sequence of unconnected coordinates; these "point sources" reflect the manner in which perimeter spraying was carried out. They represent approximately 534,000 gallons of herbicide, or 2.8% of the 18.6 million gallons dispersed by fixed-wing aircraft for which we have records.

Ranch Hand targets. One important source of military data is the targets that comprised Ranch Hand herbicide projects. As described elsewhere (Stellman et al. In press), we have identified 428 known targets which we digitized from maps or transcribed from coordinates found in the Air Force Operation Ranch Hand project folders that we retrieved at the U.S. National Archives. Each target is represented in the GIS database as a uniquely identified polygon. A point-in-polygon application was developed that identifies every Vietnam grid point that fell within each target. A target "containment" table keys a total of 25,296 grids to the targets that contain them.

The targets represent an independent and previously unrecognized data resource that provides a framework for understanding the spray program. We believe this database contains about 60% of all approved targets, although some targets were approved but never sprayed. All Ranch Hand missions were flown within previously approved targets, and spray paths—represented in the HERBS file as straight lines—generally fall within the target polygons, although spray drift outside the target areas was always possible (missions were generally flown in favorable weather with winds below 8–10 knots). The two databases thus provide mutual corroboration of the geographic coordinates of aerial spray missions.

Herbicide storage, transport, and unplanned dispersal. The Services HERBS file contains a field that describes so-called "incidents," such as an aircraft dumping its load or crashing. A GIS query will yield information as to whether a grid contains or is proximate to such events or to known leaks or other "incidents." Our database also contains information on the locations (airfields or two storage depots) where herbicides were found in bulk.

Soil typology. One specialized feature of the GIS is a soil typology database. The retention times of phenoxy herbicides and dioxin vary with the type of soil, and they disappear over time from different soils at different rates (National Research Council 1974). The GIS herbicide exposure models make use of this fact through their ability to assign different time constants to a first-order exponential decay term in the time factor, depending on the soil type for a particular location. We incorporated a grid-keyed soil table based on our digitization of a soil typology map that was prepared from a 1960 soil survey carried out for the Vietnam Agricultural Ministry by the U.N. Food and Agricultural Organization (Moormann 1961). Researchers wishing to use other than exponential decay will be able to develop their own models within the GIS.

U.S. Troop Location Data

The GIS was developed to facilitate estimation of exposure of U.S. military personnel to herbicides. With the assistance of CRUR, we developed a database of location histories of practically all combat arms support and combat support units assigned to Vietnam.

Table 2. Quantities of documented herbicides used in Vietnam.

Agent	Gallons
Orange	12,066,840
White	5,430,462
Blue	1,252,541
Pink	13,291
Purple	500,018
"Unspecified"[a]	227,538
Total	19,490,690

[a]About 8% (776) of all recorded missions did not specify a specific herbicide agent. Of these, 474 were perimeter sprays that involved comparatively small volumes and were most likely Agent Orange. The remainder of the missions with "unknown" herbicide also had a high probability of being Agent Orange. These data do not include the small quantities of dinoxol and trinoxol tested in 1961 in Vietnam (Stellman et al. In press).
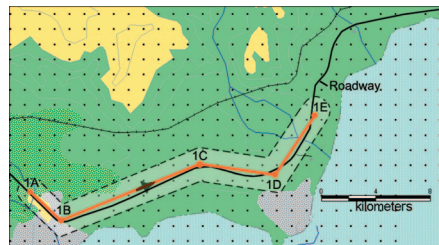

Figure 2. Typical Ranch Hand flight path (orange line) as represented in the HERBS file. Aircraft sprayed Agent Orange defoliant continuously along roadway from 1A to 1B to 1C to 1D to 1E. The dashed line is a 1-km envelope around the spray path. Dots are grid points spaced at 0.01° in longitude and latitude. Each dot represents the origin (southwest corner) of a grid. Hits are referred to the center point of the grid.

The majority of the military personnel served in units that moved infrequently or not at all and whose locations are well documented. Daily location data are available for a substantial portion of the remaining combat units, particularly for combat battalions assigned to a heavily sprayed area often referred to as the Iron Triangle in Military Region III, as designated by the U.S. Military Assistance Command–Vietnam. These are units that the CRUR had tracked for the period 1966–1969. We have consolidated these data, cleaned them, and entered them into grid-keyed tables. This Military Unit Database continues to be refined and expanded.

Vietnam Civilian Population Data

A large number of civilian locations, some with specific monthly population figures, are available from various electronic and paper files held in the U.S. National Archives (Carter and Ellis 1976). We have retrieved these data and have carried out extensive quality control. A gazetteer of known populated areas and their specific longitude and latitude is one source. Another is the Hamlet Evaluation System, which is a compendium of data used by program directors in Washington, DC, and Saigon for evaluating the degree of "pacification" of the population in the hamlets and villages of South Vietnam. Monthly surveys were taken to assign a score for perceived level of sympathy with the government of South Vietnam at one end of the scale to complete sympathy with the Vietcong insurgents at the other. The files provide extensive population estimates at the hamlet level of detail.

A similar sequence of data linkage steps can be followed to estimate herbicide EOIs for civilian population centers or particular areas in Vietnam for specific time periods using locations in the gazetteer and Hamlet Evaluation System tables that are contained in the GIS.

Exposure Opportunity Index Table

One function of the GIS is to enable the user to calculate a herbicide exposure opportunity index (EOI) for any entity (e.g., individual, military unit, village) whose location and residence dates are known based on historical reconstruction of spraying activities. An EOI model is a quantitative description of an individual's proximity in time and space to the herbicide application. It is not a toxicologic measure, but may play the part of an external or "presented" dosage within a more elaborate biologic model that might include route of entry, target organs, and metabolic activity. The GIS is model independent but does

include an EOI table based on a model we developed (Stellman and Stellman 1986). A relevant property of the EOI we use is that it is quantitative on a ratio scale: An EOI of 1,000 represents twice as much exposure opportunity as an EOI of 500. It is also additive over time: An individual who has an EOI of 1,000 from one source and 500 from another source has a total EOI of 1,500. These properties greatly facilitate database manipulation. A variety of EOI models can be constructed, and sophisticated GIS users could substitute or add different modeling algorithms to this GIS, which is flexible and can accommodate models of arbitrary complexity.

The GIS produces as its output an "exposure opportunity vector" in an easy-to-read format. This is simply an array that contains the calculated EOI as well as the number of "hits" that occurred within 0.5, 1, 2, and 5 km of the center of the GIS grid. EOI data are stored in the "exposure opportunity" table within the GIS. The table contains three essential elements: *a*) the grid-point key, consisting of an integer that uniquely defines the grid; *b*) a mission identifier, consisting of an integer that uniquely defines the mission; and *c*) an exposure vector composed of the EOI and the hit counts, which are described more fully below.

The mission number links to the HERBS file that permits exposure estimates to be confined to a particular herbicide, a range of application dates, and type of application (e.g., fixed-wing aircraft). It thus permits aggregation of EOIs by different types of herbicidal agent, and over different time periods.

The size of the database obviously depends on how many missions contribute to the EOI of each grid. As a practical matter, the database is restricted to exposure arising from missions for which the grid fell within 5 km of any herbicide application; there are approximately 1.45 million such records, with 196,735 at 1 km and 92,635 at 0.5 km.

Both the "hits" and EOIs are always referred to the center point of the 0.01° × 0.01° grid, whether for individuals, military units, hamlets, villages, or any other entity. Grid areas range from 1.18 to 1.22 km^2, depending on latitude, so the center point can be at most about 800 m from the corner of a grid. Where relevant, areas are adjusted to take into account latitude-related variation, and the adjustment factor is a keyed table in the GIS.

Hits Score

A "hit" is defined as an instance of a herbicide spray application falling within a prescribed distance of a receptor point (RP), the location for which exposure opportunity is to be estimated. By definition, a hit has an associated radius, so we speak of hits within

0.5 km, 1 km, 2 km, and so forth, of an RP, which for the GIS is always the center point of a grid. The 0.5-km hits have been expanded to include 3,005 grids for which the point of closest approach of a spray mission fell within the periphery of the grid beyond a 0.5-km radius of the grid's center. These peripheral grids, together with grids whose centers were within 0.5 km of the mission, are counted as "direct hits" to the given grids.

The direct hits model takes into account the continuous flight path of each aerial spray mission. For the flight path shown in Figure 2, for example, any RP falling within the dotted contour (including the semicircles at the two ends) is credited with one hit within 1 km. Contours such as those in Figure 2 are constructed about each flight path for distances of 0.5, 1, 2, and 5 km. This approach is nearly identical to that used by NAS in its first appraisal of the ecologic effects of herbicides used in Vietnam (National Research Council 1974). These four proximity count variables are contained in the keyed exposure opportunity GIS table.

The 5-km limit is an arbitrary figure that merely ensures that computation time is not wasted on distant, irrelevant sprays. The hit scores describe exposure in a fundamental way that is based on trigonometry and not dependent on any model.

The HERBS file includes a small number of "incidents" such as leaks, crashes, and dumps. In computing exposure opportunity, these missions are treated identically to normal missions. The user can cross-reference the GIS "incidents" table and decide how best to interpret scores arising from incidents.

Continuous Exposure Models: Direct and Indirect Exposure

In addition to models that count only discrete hits when they actually occur, we have developed a continuous EOI model that takes both current and past spraying into account and incorporates data about herbicide quantity, distance, and time (Stellman and Stellman 1986; Stellman et al. 1988a). This method was used in our previous studies of health of American Legionnaires (Stellman et al. 1988b) and in a study of dioxin levels in adipose tissue of Vietnamese (Verger et al. 1994). Both discrete and continuous EOI models use all available information on the flight-path structure of the HERBS file.

Time is an essential characteristic of the continuous EOI model. We distinguish between "direct" and "indirect" exposure as illustrated in Figure 3. Any person or entity that is present on the day of spray would be considered to have "direct exposure." Those entering a sprayed location after that time, or those remaining in the location after

having been directly exposed, would also be considered to receive "indirect exposure," that is, exposure to any residual herbicide or dioxin that is present.

Exposure opportunity model E4. For every grid point that received at least one direct hit within 5 km, we calculate a function, E4, of both direct and indirect exposure. The EOIs in the GIS use the E4 model. E4 is a refinement of our previously published model, E3 (Stellman and Stellman 1986). The E4 model is the product of three factors: quantity of herbicide sprayed, an individual's distance from spray paths, and residence time at an exposed location:

$$E4 = \text{quantity factor} \times \text{distance factor} \\ \times \text{time factor}$$

The quantity factor is the number of gallons of herbicide sprayed, which is known for the great majority of spray missions. Each "leg" of a spray run (Figure 2 shows a mission with four legs) is treated as an independent source of herbicide exposure and is assigned a gallonage in proportion to its length. E4 for a given mission is the sum of the component E4 values for all its legs. For direct hits, a term is added to reflect the greater likelihood of toxic absorption through all routes of entry and is proportional to the actual amount of toxic substance delivered in a given mission. The distance factor assumes that exposure is inversely proportional to the distance from the spray. It thus gives higher weight to closer sprays. The time factor uses a first-order exponential decay function to take into account the fact that the herbicide (and any toxic constituents) decays continuously from the time of application.

Cohort residence time and herbicide environmental half-life. "Residence time" is the period during which an individual, military unit, and so forth, occupied a given location in Vietnam during the war period.
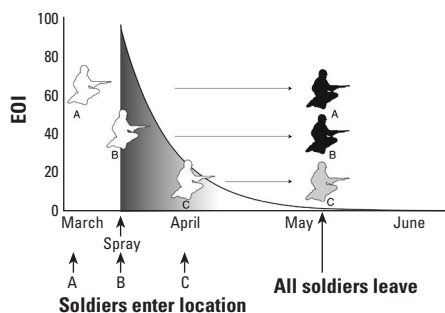

Figure 3. Direct and indirect modes of exposure to herbicide spray. The area under the curve represents the value of the EOI at a given location. Soldiers A and B are present during a direct hit. Soldier C enters the location at a later time and experiences only an indirect exposure opportunity to residual herbicide.

Any query for which exposure is sought requires an "in-date" and an "out-date" that define the residence time. For direct hit models, this pair of dates is used simply to restrict the missions to those sprays that occurred during the interval of residence at that location. The EOI that is stored in the GIS exposure table for every grid point is computed from the date of spray through 31 December 1971. Because the time factor is independent of quantity and distance, the EOI can be reweighted to reflect any specific pair of in/out dates for which an EOI is to be estimated. A different half-life also can be introduced in the adjustment. Thus, a researcher could continue the decay extrapolation if it is desirable to incorporate dates beyond the end of the herbicide program. Ecologic studies might leave the in/out dates intact but vary the half-lives to reflect different ecologic conditions.

Coordinate Precision and Robustness of GIS

It is important that exposure rankings of individual grids be insensitive to shifts in the location within the grid at which exposure is calculated, because the relative exposure rankings of the grids will ultimately become the basis for assigning exposures to military units, troops, or populations in epidemiologic analyses. Grid rankings depend in turn on the fact that exposure opportunity at any point in Vietnam is computed at the center of the grid that contains it. Every location within the grid is thereby assigned the same exposure score, including its corners, which may be about 800 m from the center.

To test the robustness of the GIS, EOIs were computed at the centers of every grid that fell within 1 km of a spray mission for June 1969, one of the most heavily sprayed months, to provide a wide range of exposures. EOIs were then computed at the same points shifted 0.005° east (half a grid or about 500 m). The data from these two runs were matched on grid point and mission. There were 158,565 grid–mission combinations. The rank-order correlation between the unshifted and shifted exposure scores, computed as Spearman's rho, was 0.972 ($p < 0.001$), indicating that very little change in ranking of exposures is introduced by shifting the locus of exposure calculation from the middle to the edge of a grid.

As an additional test of robustness, the original scores and the scores after shifting were computed for June 1969, for every grid in Vietnam, and summed over missions to yield a total score for each grid. Both sets of scores were sorted into 10 ordered groups, and their category ranks were compared to determine the numbers of grid points whose ranks changed. Of 176,060 grid points spaced at

0.01°, which cover the entire land area of South Vietnam, there were 18,308 grid points with nonzero scores, of which only 2,842 (1.6% of all grid points) experienced a change in exposure rank. In other words, 98.4% of the land area retained its original exposure ranking. Furthermore, the great majority of those 2,842 grids that changed (2,574, or 90.6%) shifted by only a single rank, with 1,424 (50%) shifting to a higher exposure and 1,418 (50%) to a lower exposure. That is, the few category shifts were not to systematically higher or lower levels. Were exposure ranks to be collapsed into fewer categories, most of these grid shifts would result in no change in the assigned category.

Illustrations of the GIS in Action

Hits close to spray paths. To illustrate how the GIS can be used, consider two Ranch Hand missions that were carried out along the coastal waterways south of Saigon on 1 January 1966. A total of 345 of South Vietnam's 176,060 grid points fell within 5 km of these runs, 106 of which were within 2 km, 49 within 1 km, and 23 within 0.5 km. Figure 4 shows the flight paths for these two missions. A simple query in the GIS will determine which grid points were within 0.5, 1, 2, or 5 km of the missions. Conversely, given a grid point, one can identify which mission flight paths were within 0.5, 1, 2, or 5 km. In the illustration, the 5-km "hits" are represented by the small black dots, and the large red, green, and blue dots represent grids within 0.5, 1, and 2 km respectively of the flight path.

Characterization of exposure to Vietnam. The grid structure allows us to characterize exposure variation of broad land areas. Figure 5 shows a choropleth plot of the entire country during June 1969, which we have chosen as our reference month because of its intense herbicide spray history. The map was constructed by computing the E4 score at every grid point in the country. Each E4 value so computed represents the exposure that would have been experienced by an individual who resided continuously at that point from 1 June through 30 June or, equivalently, the June exposure for any portion of a village or hamlet with those coordinates. The E4 score takes into account both direct and indirect exposures. Color is used to code relative levels of herbicide exposure and is keyed to 10 ranks increasing in intensity from dark blue (lowest) to red (highest); the full color scale is shown in Figure 5. The E4 score uses information from all previous sprayings that occurred at each location, based on a first-order decay curve with a half-life of 30 days. Thus, the most intense red "hot spots" generally represent sprays that occurred during the month, and the dark blue regions of low

intensity generally reflect exposures to the residue of sprays that occurred in previous months. It is important to note that the relative EOIs rise dramatically with their rank, so that areas shown as red on the map were much greater sources of exposure opportunity during June 1969 than were blue areas.

Calculation of EOIs for noncombat troops. Using the GIS, we estimated the number of "hits" and E4 scores for army units whose missions did not require frequent movement from one location to another. We call these units "stable" and have constructed a database of known locations for these noncombat units over the course of the conflict. We linked the data on the location of the stable units in June 1969 to the spray data for that month and to a separate data source, the troop strength table, which provides the number of troops present for the great majority of military units in all services present. We have used the troop strength entries that span the months May through September 1969 (no data are available for June alone, but the troop strengths remain fairly stable).

We identified 1,957 different "stable" Army units present in Vietnam during June 1969 and 2,095 different grid points that were occupied by these support units (i.e., 138 units shifted from one stable location to another during the month). Of the 2,095 "occupied" points, 937 were not sprayed that month and also had zero residual spray from prior months. We were able to match 815 of the 1,045 units that occupied locations with nonzero scores and could thus estimate the number of people assigned to these units exposed, either directly or indirectly, during June 1969. Figure 6 shows the distribution of the logarithm of E4 in for the 142,583 troops

with nonzero E4 scores who were stationed at 815 Army posts during that month. The distribution is approximately log-normal and spans several orders of magnitude, two properties that support the utility of this measure in epidemiologic studies of Army veterans who belonged to stable units.

Automation of the GIS. The GIS used here was developed by us and uses Microsoft Access to store data tables and to implement queries. It contains a master table of hits and E4 scores that have been precalculated for each grid–mission combination. These are the EOIs used by the GIS. Stored E4 scores are based on residence at each location throughout the entire era of spraying (1 January 1961 through 31 December 1971) and use a 30-day half-life. Many queries that would typically be required for an epidemiologic study involve a complex succession of steps that include data subsetting and mathematical transformations of date parameters to different residence intervals and/or half-lives.

To facilitate use of the GIS by researchers, we have designed and are in an advanced development stage of a unique user-friendly software system that implements the GIS and that might serve as a useful model for other epidemiologic software for GIS-based analyses. A function of this system is to permit transformation of date and location inputs into EOI outputs with a minimum of user intervention, and with the facility to impose a variety of selection options such as restriction to specific types of herbicides, ranges of dates, or types of missions. The user inputs dates and locations of residence and receives the EOIs as an output table. The default half-life of 30 days can be changed to any desired value. For spot checking, the GIS can also

display the input locations on a thumbnail map of Vietnam, but users will usually wish to import the locations and exposure scores into their own mapping software.

In an epidemiologic study of military troops, the exposure assessment process typically includes the following steps: *a*) input individual veterans' military unit histories as a series of records, each containing the unit identification code, location, and residence time interval for one unit; *b*) specify a set of selection criteria (e.g., specific herbicide types, mission types); *c*) obtain EOIs from the EOI table for each location, using selection criteria as query restrictions; and *d*) adjust EOIs at each location for residence dates and aggregate for each military unit.

The software system is designed to receive military unit histories as its input and carry out the foregoing sequence of operations with a minimum of involvement by the user, producing a set of EOIs and hits at 0.5, 1, 2, and 5 km that can then be used to calculate risks of health outcomes in epidemiologic studies. The system enables estimation of herbicide exposure for both military troops and Vietnamese civilian populations. The software system is represented in Figure 7.

Discussion

We have described a GIS that is suitable for estimation of exposure opportunity of military troops and civilian populations and has been developed in a user-friendly format for use in large-scale epidemiologic studies. It makes use of extensive data on herbicide application, with a level of detail and precision not often available for historical exposure reconstruction.
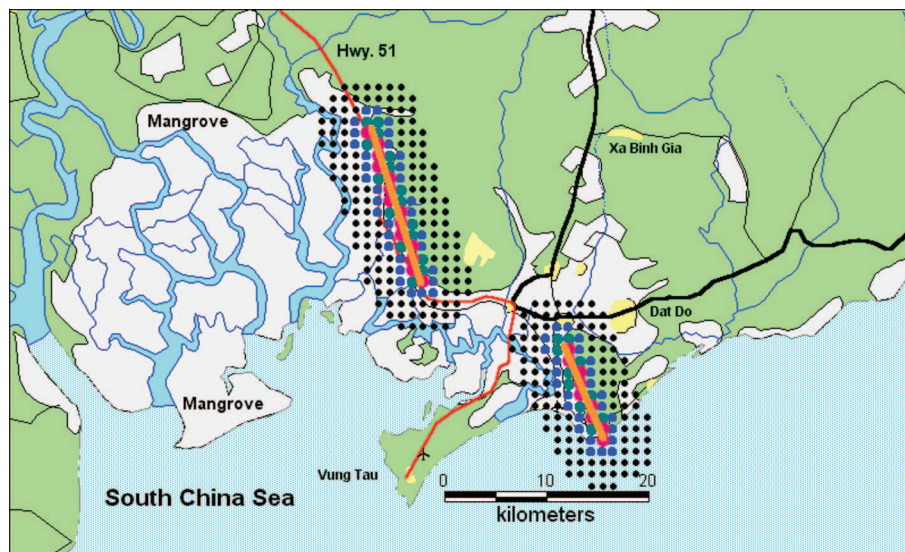

Figure 4. Two Agent Orange herbicide missions (shown as orange lines) flown 1 January 1966 in the Rung Sat Special Zone. The figure shows all grid points whose centers fell within 5 km (small black dots) of the flight paths. Blue, green, and red dots are centers of grids that fell within 2, 1, and 0.5 km, respectively.
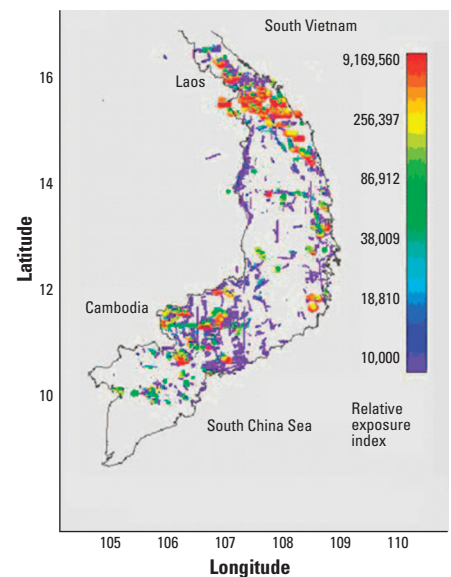

Figure 5. Choropleth plot of Vietnam for June 1969 showing intensity of exposure using the E4 model. A similar map showing the location of "hits" can readily be created.

The utility of the GIS is not limited to epidemiologic studies, however. Researchers who wish to study past or current ecologic conditions in Vietnam (ascertained, say, via satellite imaging) or to compare levels of contamination from biologic or soil sampling could use the GIS to estimate the original exposure levels at the time of spraying.

The utility of the GIS as an exposure reconstruction instrument depends, of course, on how accurately the E4 and hit scores can act as proxies for true toxicologic exposure. Proximity to an environmental insult is a widely used concept and various schema such as job-exposure matrices rely on this conceptual premise (Blair et al. 2001). In the case of exposure to herbicides in Vietnam, which began 40 years ago and ended 30 years ago, no other reliable measure is available for large-scale epidemiologic studies.

It is clear from our analyses of HERBS data and Ranch Hand Project maps that Vietnam was not uniformly sprayed but rather shows substantial geographic and temporal variation in herbicide application. When juxtaposed with locations of residential areas and military troops, the spray patterns are sufficient to justify epidemiologic studies on military and civilian populations as well studies of environmental and ecologic damage.

Both types of applications are facilitated by the keyed grid-point structure of the GIS. Each grid can be linked to the spray missions that overflew it, the type and quantity of herbicide deposited, the distance from the spray run, containment within one or more Ranch Hand targets, and land features such as topography or soil type. Grid-keyed tables of population centers (hamlets and villages), troop locations, and military installations are also included.

One advantage to a grid-based approach that models exposure separately from troop location is that it separates the errors, because each factor is modeled independently. This approach is reasonable because the factors that affect the accuracy and precision of spray mission data may not be the same as the factors that affect the accuracy and precision of troop location data. For example, environmental decay parameters, spray dispersion patterns, and inaccurate gallonage data affect the quality of the spray data but are not relevant to troop location data.

Another advantage is that more elaborate and potentially more accurate models can be built by using empirical knowledge of the variation of EOI over a localized area. For example, if we consider the spray history of all fire support bases within the command of a particular Army division and examine the spray patterns for those units in which record keeping was known to be done well, we can then model the grid locations of other fire support bases using the data set from the "good" record set.

A third advantage is a pragmatic one in that the effect of various military unit movement models can be very readily tested, because the EOI calculations are reduced to straightforward products and sums. To obtain the exposure for a military unit, we place it each day in the grid it occupied and then aggregate the exposure over time. Accurate locations and residence dates, however, may not be available for many combat troops. Rather, locations of headquarters companies or other battalion-based data may be all that are available from unit histories or data files. The GIS is being adapted to accommodate "fuzzy" location and date information. Suppose the location is not precisely known for one day, but it is likely that the unit was not more than one grid away. Then we might assign the average E4 score for all surrounding grids. More generally, if there are different probabilities of

being in different elements, we can assign the expected value over those elements. On a given day, adjacent elements are likely to have very similar exposure levels, so many different probability distributions will yield similar values.

The E4 score, which assumes a reciprocal distance and a first-order decay, is included in the GIS as a convenience to investigators. However, the GIS itself is not dependent on any particular model. Many further EOI refinements are being considered, such as introduction of second- or higher-order decay, use of multiple time constants for areas that were sprayed on more than one occasion, and considerations of weather as reported in U.S. Air Force Daily Air Activity Reports, which are available for about 60% of Ranch Hand missions or from historic weather databases. One refinement uses data in the HERBS file on the number of fixed-wing aircraft that flew a given mission. It was routine Ranch Hand procedure for multiple aircraft to disperse their spray over laterally contiguous swaths (a swath is a finite width of spray dispersion), as shown in Figure 1. Our model assumes that the flanking aircraft took flight paths that were parallel to the central aircraft and flew at the same altitude with a horizontal separation of 80 m. Other wind drift models may be available.

The GIS brings together a variety of diverse databases on herbicide spraying, geography, and population and troop location that can enable pursuit of epidemiologic and environmental studies. It is built on the power of a relational database system and provides the ability to do otherwise complex exposure model calculations with rapid, straightforward arithmetic procedures. Lack of coherent data and lack of an exposure reconstruction model no longer need be the major impediments they have been in
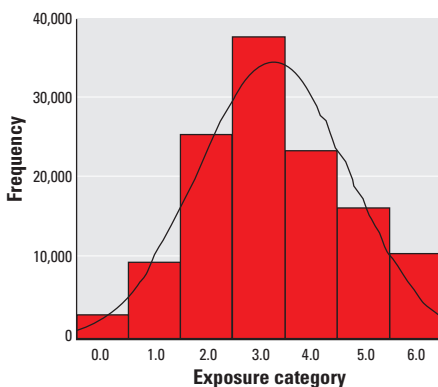

Figure 6. Frequency distribution of the logarithm of the E4 model score for 142,583 troops in 815 units designated as "stable" or "stable with mobile elements," June 1969.
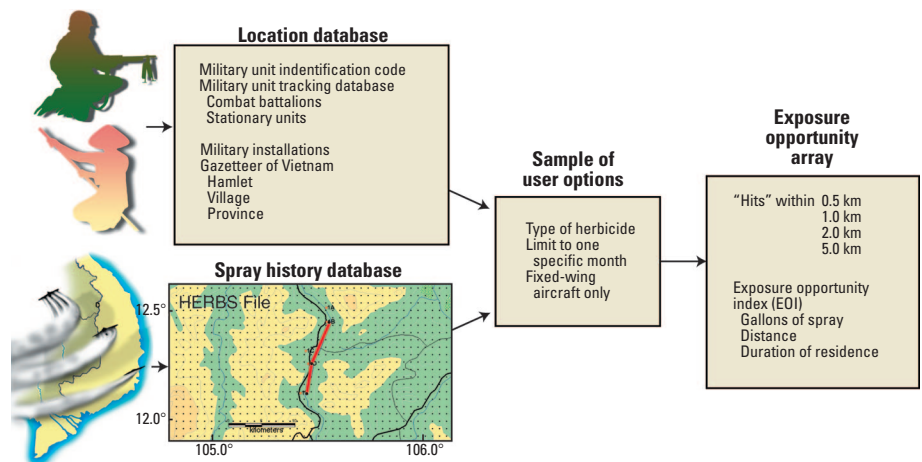

Figure 7. Schematic illustration of the GIS. Databases are linked via the "grid point," which indexes a grid covering the land area of South Vietnam at 0.01° intervals. Individual queries may be restricted to specific types of herbicides, dates, or other variables. Output consists of an array of proximity counts plus the EOIs.

the past to research on health of Vietnam veterans and the residents of Vietnam itself.

Considerable attention has been given to the design of the GIS input and output formats that permit a user to input a table of dates and locations and receive as output a table of exposure variables. With little or no additional processing, outputs that are created in this format can serve as input to most popular statistical software and cartographic packages. With the exception of the E4 model score, data in the supplied tables are not model dependent, making the system highly transferable to other applications. Even broader applications are accessible to researchers who can use SQL or a similar query language to manipulate the geographic exposure information, and the methodology itself may be adapted to other situations in which it is desirable to combine records-based exposure opportunity data with locations of potentially exposed individuals or groups.

REFERENCES

Agent Orange Act of 1991. 1991. Public Law 102-4.
Blair A, Zheng T, Linos A, Stewart PA, Zhang YW, Cantor KP. 2001. Occupation and leukemia: a population-based case-control study in Iowa and Minnesota. Am J Ind Med 40:3–14.
Brody JG, Vorhees DJ, Melly SJ, Swedis SR, Drivas PJ, Rudel RA. 2002. Using GIS and historical records to reconstruct residential exposure to large-scale pesticide application. J Expos Anal Environ Epidemiol 12:1–17.
Carter GA, Ellis JW Jr. 1976. User's Guide to Southeast Asia Combat Data. R-1815-ARPA. Santa Monica, CA:RAND Corporation.
Darrow RA (ed). September 1971. Historical, Logistical, Political and Technical Aspects of the Herbicide/Defoliant Program 1967–1971. A Resume of the Activities of the Subcommittee on Defoliation/Anticrop Systems for the Joint Technical Coordinating Group/Chemical-Biological. Ft. Detrick, MD:Vegetation Control Committee.
Data Management Agency, USMACV. 1970. Herbicide Report System (HERBS). Document No. DARU07. San Francisco:Headquarters, U.S. Military Assistance Command, Vietnam.
GIS and Public Health. Hyattsville, MD:National Center for Health Statistics. Available: http://www.cdc.gov/nchs/gis.htm [accessed 20 January 2003].
Gunier RB, Harnley ME, Reynolds P, Hertz A, Von Behren J. 2001. Agricultural pesticide use in California: pesticide prioritization, use densities, and population distributions for a childhood cancer study. Environ Health Perspect 109:1071–1078.
Institute of Medicine Committee to Review the Health Effects in Vietnam Veterans of Exposure to Herbicides. 1994. Veterans and Agent Orange: Health Effects of Herbicides Used in Vietnam. Washington, DC:National Academy of Sciences Press.
———. 1996. Veterans and Agent Orange: Update. Washington, DC:National Academy of Sciences Press.
———. 1998. Veterans and Agent Orange: Update. Washington, DC:National Academy of Sciences Press.
———. 2001. Veterans and Agent Orange: Update 2000. Washington, DC:National Academy Press.
Moormann FR. 1961. Republic of Viet Nam, General Soil Map. National Geographic Service of Viet Nam Map. Saigon:Ministry of Agriculture.
National Research Council Committee on the Effects of Herbicides in South Vietnam. 1974. The Effects of Herbicides in South Vietnam: Part A. Washington, DC:National Academy of Sciences.
Stellman JM, Stellman SD, Christian R, Weber T, Tomasallo C. In Press. Extent and patterns of use of agent orange and other herbicides in Vietnam. Nature.
Stellman SD, Stellman JM. 1986. Estimation of exposure to Agent Orange and other defoliants among American troops in Vietnam. A methodological approach. Am J Ind Med 9:305–321.
Stellman SD, Stellman JM, Christian R, Weber T. 2001. Influence of soil-specific dioxin decay rates on estimates of exposure to residual phenoxy herbicides in Vietnam. Presented at International Society of Exposure Analysis, 5 November 2001, Charleston, SC.
Stellman SD, Stellman JM, Sommer JF Jr. 1988a. Combat and herbicide exposures in Vietnam among a sample of American Legionnaires. Environ Res 47:112–128.
———. 1988b. Health and reproductive effects of combat and herbicide exposure in Vietnam among American Legionnaires. Environ Res 47:150–174.
Stellman AB, Stellman JM, Stellman SD. 2002. Herbicide Exposure Assessment—Vietnam. User's Manual. Brooklyn, NY:Foundation for Worker, Veteran, and Environmental Health, Inc.
U.S. Army & Joint Services Environmental Support Group. 1985. Services Herbs Tape—A Record of Helicopter and Ground Spraying Missions, Aborts, Leaks, and Incidents. Washington, DC:Department of the Army Adjutant General—Environmental Support Group.
Verger P, Cordier S, Thuy LTB, Bard D, Dai LC, Phiet PH, et al. 1994. Correlation between dioxin levels in adipose tissue and estimated exposure to Agent Orange in South Vietnamese Residents. Environ Res 65:226–242.
Ward MH, Nuckols JR, Weigel SJ, Maxwell SK, Cantor KP, Miller RS. 2000. Identifying populations potentially exposed to agricultural pesticides using remote sensing and a geographic information system. Environ Health Perspect 108:5–12.