

**Item ID Number** 05654

☐

**Not Scanned**

**Author**

**Corporate Author**

**Report/Article Title** Correspondence and reports relating to the CDC Agent  
Orange Study, 1985

**Journal/Book Title**

**Year** 0000

**Month/Day**

**Color**

☐

**Number of Images** 0

**Description Notes**



DEPARTMENT OF THE ARMY  
US ARMY & JOINT SERVICES ENVIRONMENTAL SUPPORT GROUP  
1730 K STREET N.W. ROOM 210  
WASHINGTON, DC 20006-3868

REPLY TO  
ATTENTION OF

December 4, 1985

DAAG-ESG

Dr. Vernon Houk  
Director, Center for Environmental Health  
Centers for Disease Control  
1600 Clifton Road, NE  
Atlanta, Georgia 30333

Dear Dr. Houk:

After analyzing and evaluating the AOP interim report to OTA by our epidemiologist, a scientist and military analyst we found the report did not contain clear guidance in the determination of possible exposure opportunity. The material and data contained in this report, especially the tables, are set up to demonstrate a single point of view only. It would have been helpful to have the reader understand all the aspects of research, especially the use of data from combat records. The comments that follow are submitted for your determination in having the best available analysis of all sides of this complex issue. Further, where errors are noted, it is requested that the report be amended/corrected. The epidemiologist that reviewed the report requests that we see the raw data in order to evaluate the tables. Nowhere in the report was it mentioned that the methods for capturing and recording of all the data were dictated by AOP and we fully complied. The AOP liaison officer spent two years with the data abstractors in the completion of 18 battalions.

The following are specific comments relative to Interim Report No. 2:

a. AOP Statement (Section II - Exposure Opportunity and Combat Selection, paragraph 1, page 2). " We have a list of 65 combat battalions that served for at least 18 months in III Corps during 1967 and 1968. This list, called the AOP master list of battalions,...."

ESG Comment - ESG at the request of AOP was directly responsible for identification of the 65 units that were se-

lected for the study. On 16 November 1983 we provided the master list of 123 units which operated in the III Corps tactical zone in 1967 & 1968.

b. AOP Statement (Section II -- Exposure Opportunity and Cohort Selection, paragraph 4, subparagraph 2, page 2). "The present selection criteria or a slight extension of these will allow us to identify 17,000 U.S. Army veterans...."

ESG Comment - We sent AOP 7,500 qualified personnel data abstraction forms (under AOP's new criteria) in which only 1,500 study subjects finally qualified for the study. The view here is that the criteria should change only when it becomes apparent AOP will not receive enough study subjects. This changing of criteria is not considered a slight change when 70% of the already qualified study subjects become disqualified.

c. AOP Statement (B. Selection of men from combat line companies, paragraph 1 & 2, page 11). "Additionally, our projections show that in order to obtain the required number of men (17,000 qualified veterans), it would be necessary to select men from almost all of the units on the list and the selection criteria may have to be changed."

ESG Comment: ESG has presently qualified 12,500 study subjects for the study. The new criteria would eliminate approximately 8,500 of these study subjects. AOP has added significant criteria changes to the study which eliminate large numbers of veterans. Before we eliminate study subjects we request Science Panel review, because of all the work that has gone into this phase of this program.

d. AOP Change: "A veteran must only serve his tour in 1967 & 1968."

ESG Comment: Every veteran we have qualified that had served one or more days in 1966 or 1969 is disqualified, even if the veteran experienced several herbicide exposures while in a tracked unit during 1967 or 1968. We would like AOP to explain to the Science Panel, why study subjects who may register herbicide exposure hits in 1967 or 68 should be disqualified. This eliminates significant numbers of study subjects. A total of 828 study subjects were disqualified from the first subgroup at AOP after qualification and abstraction at ESG. This change of criteria would require ESG to review an additional 60,000 records for the study in

order to qualify enough study subjects. This change will disqualify the majority of the many study subjects that were previously qualified by ESG and sent to AOP.

e. AOP Change: " Entire tour spent in units for which location data is being collected."

ESG Comment: ESG is currently qualifying veterans who served in a tracked unit but also transferred to units that are not being tracked. Even if a veteran experienced several herbicide exposures while assigned to a tracked unit, he will be disqualified by AOP if he transferred to a non-tracked unit. The reasoning AOP uses to disqualify these veterans who were exposed has not been made clear. We would like AOP to explain to the Science Panel, why study subjects who may register multiple herbicide exposure hits while assigned to tracked units and later assigned to non-tracked units should be disqualified. This again eliminates high numbers of study subjects. A total of 1,871 study subjects were disqualified at AOP after qualification at ESG from the first subgroup. The disqualification of these veterans would require ESG to review additional records each month.

The disqualifying of such large numbers of veterans could result in tracking of additional battalions in order for AOP to qualify 17,000 veterans for the study. ESG is currently reviewing over 6,000 records a month in order to qualify 2,500 study subjects for use by AOP in which 70% of the records are later disqualified at AOP. This is a major increase in the quota every month. We completed the abstraction for the Vietnam Experience Study in which ESG qualified 1,433 study subjects per month for AOP without difficulty. This is a significant increase to ESG when compared to the Vietnam Experience study. Every change impacts heavily on extraction operations and cannot help but affect time scheduling from ESG's standpoint.

f. AOP Statement (Section III - Quality and Completeness of Location Information, paragraph 3, page 14). "Limited reproducibility conducted by AOP lead us to conclude that the data are incomplete and inaccurate."

ESG Comment: The word "limited" might be a description of the quality control functions AOP has performed to date. The initial AOP quality control report dated 21 March 1985 states (see TAB A), "Assuming standard procedures for abstraction,

these differences are due to either error in ESG's abstraction, error in AOP's or both. I scrutinized the documents, the same for ESG and AOP, and found the latter to be the case. Since the purpose of this report is to review ESG's abstraction of troop locations from Army documents, the discussion that follows is limited to ESG's errors only. Please note that "error" subsumes a variety of possible mistakes and oversights and together they refer only to the standard of abstraction outlined by AOP."

This report verifies that AOP made errors in their QC reabstraction process but still only stressed ESG's abstraction process. AOP used the term "possible errors" because they cannot be completely sure they are errors.

The Epidemiologist, Scientist and Analyst have not had a chance to analyze these points AOP extracted from the records. It should be noted that AOP reabstracted or edited their own abstractions countless times finding numerous discrepancies from their original abstraction. There is no yardstick. The rules changed and there is no way to determine who is correct.

g. AOP Statement (A. Completeness and accuracy of available location data, page 14). "The ability to reproduce data is essential to establish the integrity and credibility of a scientific study."

ESG Comment: It should be clear that there is nothing comparable to laboratory data findings in the grid locations and the recording of military data in journals and other like records. They were never intended to be used for epidemiological studies. The definition of a grid coordinate means a reference point on the ground. One has to ask the question who has more expertise and credibility concerning U.S. Army combat troop data, the U.S. Army or AOP?

h. AOP Statement (A. Completeness and accuracy of location data, page 14)".... members of the AOP staff began reabstracting data previously abstracted by ESG on selected units. These AOP staff members were trained by the ESG abstraction supervisor prior to beginning the process."

ESG Comment: The facts are that one member of AOP Staff received training from the ESG supervisor for three days. It would be desirable if we could make an expert on the battalion tracking process in three days but this is just not possible. It is apparent the AOP abstractors encountered numerous problems and difficulties when they performed their

quality control mission. AOP had to go over their own abstraction countless times. AOP's latest QC report in July contained errors.

i. AOP Statement (A. Completeness and accuracy of available location data, paragraph 1, page 14). " We now have completed data for four months on three different battalions and present the results in table 7."

ESG Comment: AOP has attempted to provide a rather one sided view of the data. AOP has provided charts which show AOP's initial reabstraction compared to ESG's data. AOP does not state in the report that they reviewed their abstraction and found numerous errors in their original abstraction. AOP did not provide their edited or reabstracted tables to the Science Panel. These edited and reabstracted tables were originally reported in AOP's first Quality Control Report dated 21 March 1985.

j. AOP Statement (A. Completeness and accuracy of available location data, paragraph 1, page 14). " We compare the difference between the average of the locations found in the two independent abstractions of the data on the battalions, the abstraction originally supplied by ESG and the re-abstraction completed by AOP. If the number of points representing unit locations and the general area in which they occurred were similar for the two independent abstractions, we would expect this distance between the centroids to be small. This, however, was not the case for many of the days within the period studied."

ESG Comment: First of all, AOP has not provided the correct data from their own Quality Control Report (Reference statement above). There are many reasons why these two abstractions disagree: (1) ESG or AOP abstracted the wrong points (2) ESG or AOP abstracted N points or X points that the other did not extract. (These points should not be considered for comparison when AOP is establishing a centroid. Only the line companies (A-E) grid points should be compared), (3) ESG inputted a grid for a village in which AOP did not. It is apparent that for each day in which there is a major difference in the locations, AOP and ESG should analyze these data in order to identify the reason for the discrepancy. This has not been done; however, we have requested AOP's reabstracted grid coordinate points in order to perform this analysis. ESG will not make assumptions or initiate reports until we verify what the data reflects. This can only be accomplished by checking the source

documents.

k. AOP Statement (paragraph 1, page 15). "We have also discussed the reabstraction of other units, but ESG has been unwilling to commit themselves to reabstracting additional data."

ESG Comment: ESG policy has been stated at many meetings over the last year that we will be happy to reabstract any battalion when asked to do so. The aforementioned statement is incorrect. ESG has only been asked to reabstract battalion #4. We agreed to reabstract this battalion because it was one of the early battalions tracked when using the CDC KAYPROS (word processors). We later discarded the KAYPROS because the researcher could not edit the data once it was keypunched. AOP recommended that several units be reabstracted but AOP failed to demonstrate ESG expertise in the battalion tracking arena. ESG report dated 10 September 1985 (see TAB B) documents the problems ESG found with AOP's latest attempt at quality control. AOP wanted ESG to reabstract 3 or 4 battalions as a result of their Quality Control Report dated 1 July 1985. ESG demonstrated to AOP that their report was less than acceptable and AOP dropped the notion to reabstract the battalions. We are currently working closely with Dr. Riduan Joesoef, AOP to develop a quality control scheme in which the quality of data can be accurately evaluated. Dr. Joesoef has worked well with members of the staff to solve these problems and we appreciate his interest and efforts.

l. AOP Statement (paragraph 2, page 15). "The approach used by ESG to abstract military documents is mechanical."

ESG Comment: We agree with AOP that ESG developed the current method for abstracting company location data. This was not the case for the early abstraction phase when detailed instructions were given to ESG (which turned out to be poor decisions confirmed at the last Science Panel meeting). However, AOP has continuously provided ESG with rules which has limited ESG on the decisions it could and should infer. There are numerous decisions or determinations ESG could initiate if we were given a free hand to utilize our expertise of U.S. Army combat unit data. We are convinced ESG's current methodology is the best possible method for tracking combat battalions. Clearly, the disagreements, technical as they may be,

point up the need for the separation of responsibilities to make the study believable and that it be done by military analyst who have worked on this project for the last three years. In the interest of getting on with the project, the Science Panel might consider assigning the total mission of selecting the study subjects to ESG, thus allowing AOP to conduct the interviews, physicals and statistical analysis. During the protocol development of this study ESG planned to conduct the research and accomplish the automation process, thus providing the principal investigators blind lists of study subjects. AOP wanted the automation portion of the recording of data and so it was agreed. This concept of the provision of blind lists was also discussed and agreed upon by the VA when it was thought that the VA would be doing the study.

m. AOP Statement (paragraph 2, page 15) "AOP has requested that ESG involve other military experts in decisions concerning the best methods for locating units, but ESG considered this unnecessary."

ESG Comment: These techniques for abstraction were dictated by AOP before they understood very much about the battalion records. The AOP criteria as bad as it was, was followed to the letter by ESG. The reason ESG has not called in other so called "experts" to locate military units is that ESG already has cornered the market in these specialists with Vietnam combat experience and records management expertise.

n. AOP Statement (paragraph 3, page 15). " All parties then agreed to the establishment of a group consisting of persons outside the Army."

ESG Comment: ESG rejected such a proposal because the best current expertise in records extractions is in ESG. These persons suggested by AOP as reviewers were less competent than the personnel now doing the extraction process. We never agreed to a final peer review selection process hence the last sentence is not true. Not stated in this paragraph is ESG's suggestion to let the 4 major Veterans organizations each provide 2 former combat company or battalion commanders to come in and review the record extraction process. This was summarily rejected by AOP but it is never mentioned. AOP tends to write down only



their own proposals and omit any reference to counter proposals from ESG particularly if AOP does not care for the idea or suggestion. Another point is that this peer review idea from AOP was never presented to the WHAOWG Science Panel for consideration. Rather it was an arbitrary unilateral decision by AOP that such a review was called for. AOP should clearly be made to understand that ESG is in no manner a sub-contractor to AOP. We are very willing to serve in the role of an equal research partner.

o. AOP Statement (paragraph 1, page 20). "We asked Mr. Tavia Gordon, a consultant with AOP, to establish the review group."

ESG Comment: Mr. Gordon is not known to us. We have never met nor communicated with him.

p. AOP Statement (paragraph 2, page 20). "Mr. Gordon did contact Mr. Stanton, and after further problems were discussed with the location data supplied by ESG, AOP employed Mr. Stanton as an AOP staff member."

ESG Comment: ESG is well aware of Mr. Stanton's background as a historian. If Mr. Stanton should discover any potential problems with the data we would be happy to meet with him.

q. AOP Statement (paragraph 4, page 20). "A meeting took place between ESG and AOP on November 12, 1985. We discussed the information available for battalion #14 for April 1967. ESG had reabstracted the information on the unit for presentation at this meeting. In this reabstraction, locations were found where none had been supplied to AOP by ESG from their original abstraction. In the data originally sent, AOP had location information for only 4 of 30 days, while after reabstraction, information was obtained for 27 of 30 days. It also appears that there were numerous abstraction errors in the data AOP originally received, some of which resulted in placing units almost 100 kilometers from where they actually were."

ESG Comment: AOP had established the criteria which resulted in location data for only 4 days out of 30 for battalion #14. When ESG was freed of the AOP imposed criteria then ESG was able to account for 27 of the 30 days with company locations. This merely proved the fal-

lacious nature of the original AOP imposed location criteria.

r. AOP Statement (paragraph 5, page 20). "Based on the reabstraction of military records for four different months involving three different battalions mentioned above, we suspected that we were not receiving all possible information on a company's location on all days during 1967 and 1968. While we had concerns about the quality of location data being received from ESG prior to this meeting, the data presented at the meeting indicated that they may be of such poor quality that they may compromise the scientific credibility of the study. We also believe, based on our own research and the reabstraction of data presented by ESG, that it is possible to obtain a data set of reasonably high quality if the abstraction process is changed to collect more of the available information from the military records. Moreover, we believe that unit location information of sufficient completeness, accuracy, and reproducibility to withstand the scientific scrutiny which will occur at the completion of the study is more likely to be obtained if the abstraction process is directed by AOP scientists, rather than ESG."

ESG Comment: We agree that because of the AOP imposed criteria, that the data on daily company locations may be faulty. Why doesn't AOP admit this was of their doing and that if ESG is permitted to freely abstract and interpret daily locations the problem will be readily solved with an appreciable improvement in accuracy. AOP gives no information whatsoever as to how they ensured quality control of daily location data inputs to the AOP computers. In no way will ESG agree to any transfer of the abstraction process to AOP scientists who have little or no background experience with Army operational records. For further reference see TABs A and B.

s. AOP Statement (paragraph 2, page 21). "AOP generally would not describe a meeting in a scientific report, but we believe this information necessary to justify our recommendation that AOP needs more direct control over the abstraction process. In this way AOP can be responsible for the validity of the exposure indices and can enlist the services of additional experts in the areas of military operations and records. These

comments are not intended to minimize the contribution that ESG has made to the study through the abstraction of information on individual study participants from military personnel files, and we trust that this contribution will continue throughout the remainder of the study."

ESG Comment: A report of any kind should be presented fairly and supportable. While we can understand it is much more convenient not to be subject to peer review, AOP does not wish a review from above comments. We have gone the extra mile all along and because we want to see the study through for the sake of our Nation's Vietnam veterans we will continue to complete our mission.

t. AOP Statement B. AOP recommendations for obtaining valid location information, page 21 & 22, "Transfer responsibility for the collection of the location data to AOP. We believe it would benefit the Agent Orange Study to transfer the responsibility for the collection of data on unit locations to AOP. It is our judgment based on past experience and statistical analysis that with current resources ESG cannot, on a timely basis, abstract information with sufficient accuracy and completeness to withstand scientific scrutiny the completion of the study. AOP would hire a contractor to abstract all the unit location data needed for the study according to uniform procedures specified by AOP. Performance standards would be written into the contract and payment would depend on the quality of service. AOP does not foresee the need to request an additional Congressional appropriation for this work. We understand that the records being used for unit locations are controlled by the National Archives and are available to the public. Therefore, we do not foresee any difficulties with continuing to gain access to military information. However, we would ask DOD to help expedite the process. We also would renew our request to DOD for short term consultative services from members of the Military History Institute and the War College. AOP would develop a data collection procedure on knowledge already possessed and information on the design from the experts on the available military documents and the conduct of the Vietnam War."

ESG Comment: We totally disagree with AOP's

recommendations. There are many issues throughout this report that are presented from one single point of view in AOP's favor. This attempt by AOP to take over this portion of the study negates the requirements for checks and balances to make a study creditable and acceptable to the veterans who have repeatedly expressed confidence in ESG. They deserve the best that we can give them. Do not take that away from them. Vietnam was an honorable war. We all did our best. All AOP has to do is ask ESG directly for what they want as has been the case in the past and we will produce the information. In summary: Non-Concur by ESG. We do not believe that AOP has the breadth of experience in military records abstraction so that they can supervise any contractor (nor can a contractor be found with expertise in personnel such as already available in ESG).

u. AOP Statement (paragraph 3, page 22). "This approach may have some merit, but we do not yet know how much of the location data would have to be reabstracted. If this approach necessitates a substantial amount of reabstraction, it would be preferable to reabstract all of the information as outlined in option 1. AOP does not recommend this approach since everyone agrees (ESG, AOP, and others involved in recent discussions) that more location information exists in the records than is being collected presently and this new approach ignores those data. Furthermore, since this approach requires that we reabstract only a subset of the data, we are left with the inaccuracies in the remainder. If AOP adopts this approach we would increase the quality control on the data being collected and have the reabstraction done through a contractor supervised by AOP as with the first alternative. The scope of work, however, would be more limited than that described in option 1."

ESG Comment: AOP does not have an understanding of the problem. After ESG has already qualified the study subjects (163 data elements per individual) they are later disqualified by AOP because of future new qualifications. The job is manageable and can be accomplished with a few concessions from AOP and AOP's admission that their location criteria is faulty.

v. AOP Statement: (page 22, Cancel the Agent Orange Study).

"While we believe that there is sufficient justification for continuing the study and that the location of companies in Vietnam can be determined, we also believe that the study must be based on the most complete and accurate information available. AOP cannot suggest that the study be conducted with less than good quality unit location data since these data are critical to accurate exposure assessment and since so much money and effort are going into assuring that all other aspects of the study result in valid data."

ESG Comment: We concur - let us get on with a more accurate daily location process as conceived by ESG and accomplished by the highly trained and competent ESG staff. Let ESG select the study cohorts (names of individuals) without AOP interference and unreasonable constraints which have been proven wrong by AOP themselves, as exemplified by their report.

w. AOP Statement (Section IV - Conclusions, page 23). "Our major concern is the completeness, accuracy, and reproducibility of the unit location information supplied to us by ESG. Therefore, AOP would like to discuss the assumption of direct responsibility for the collection of unit location information."

ESG Comment: Non-Concur by ESG as to any AOP assumption of direct responsibility for abstraction of unit location information.

x. AOP Statement: (Page 24, Appendix I. Obtaining Unit Location Information).

"We discussed the accuracy and completeness of the data supplied by ESG in Section III. In this appendix we describe documents and methods presently employed by ESG and AOP to obtain company-level location information.

The methods used to abstract the data do not provide a location for every company on each day of the period of the study. Consequently, gaps exist in our knowledge of company locations for approximately 50 percent of the days. In part A of this appendix we describe data available from

battalion, brigade, and division level documents. In part B we evaluate ESG's suggestion that company-level morning reports be used to fill gaps in unit location information."

ESG Comment: The lack of data for 50% of the days in AOP's opinion comes from use of incomplete data and the lack of authority to do interpretation by ESG to fill in these gaps. It is a difficult process and AOP does not understand how such a process can be made to work by ESG.

y. AOP Statement (paragraph 3, page 24). "AOP has received data from 37 of these battalions."

ESG Comment: ESG has already provided AOP data on 40 battalions. By the end of the year 55 battalions will have been reviewed.

z. AOP Statement (paragraph 4, page 24). "The map coordinates gleaned from these records are in the Universal Transverse Mercator (UTM) system which uses a two letter and six digit designation of location."

ESG Comment: ESG also abstracts two letter and six digit grid coordinate locations. Early in the study we were told by AOP if the Alpha letters were not recorded we could not use them. (See letter TABs D & E).

a.1. AOP Statement (paragraph 4, page 27). "Unfortunately AOP finds that it is not always clear as to what the location information listed in the morning report refers, particularly when only an APO is available."

We believe, based on our discussions with military experts and the analysis of over 40,000 morning reports, that the location refers to the point of mail distribution which was the base camp of the company's brigade or division. If this information is presently available, it is more easily retrieved from USARV station lists."

ESG Comment: Morning Reports are the document which would record the actual day a unit moved it's base camp. USARV Station Lists would not provide this information. Also the USARV Station Lists are not updated to reflect a unit's move until many months later. Morning Reports are definitely the best source document to use when verifying a unit's major base camp.

b.1. AOP Statement (paragraph 2, page 28)"....morning reports refer to either the brigade or division base camps, without further documents, we cannot determine which."

ESG Comment: ESG could easily furnish additional documents to verify the morning reports. This method is done all the time in connection with the many other ongoing studies where detailed tracking is required. AOP acts as though one could use Morning Reports exclusively. Such is not the case. ESG has maintained that the MR's must be used in conjunction with other reports and daily journals, then they make sense. This paragraph points out the naivete of the AOP approach.

c.1. AOP Statement (paragraph 1, page 32). "One major problem with obtaining locations from morning reports is that we have been unable to locate these reports for 40 percent of the days being studied. When the reports are available, only an APO number occurs 48 percent of the time."

ESG Comment: There are morning reports for every month which would verify the major base camp location for each and every day. An experienced researcher can interpret the APO number and locate the unit's main base camp location. There are many reasons why morning reports are missing. No change in status is one example. Our trained researchers have the ability to track the unit using other data source materials.

d.1. AOP Statement (paragraph 1, page 32). "Even when coordinates are listed the coordinates often occur again and again for the entire time period which indicates that the coordinates refer to a fixed location such as that of a division base camp and not the actual location of the company."

ESG Comment: This is to be expected as major base camps do not move. They were constructed and protected with chain link fences, barbed wire, bunkers and open fields of fire. However, combat companies do move in and out of these base camps.

e.1. AOP Statement (paragraph 3, page 32). "If the locations obtained from morning reports represent these locations, they should, on average, be reasonably close to the field locations noted for the companies in other documents. This analysis, however, indicated that the locations obtained from the morning reports are not truly representative of the physical location of the company."

ESG Comment: This is an apparent misunderstanding on the part of AOP. There is a significant difference between a major base camp location and a field location. Being in the field to an infantryman can mean many things. Out on search-and-destroy missions, ambushes, fire support bases, the taking of an objective, retrograde movement and a host of other battle situations such as serving as a screening force, roads, bridges, protection and the like. Some of these missions may take the company far afield from the Division base location.

Sincerely,

  
Richard S. Christian, C.R.M.  
Director

Enclosures

1. AOP Quality Control Report
2. CDC Quality Control Report
3. Chair, Science Panel Memo, 15 November 1985
4. Epidemiologist Letter, 18 November 1985
5. AOP Letter, 10 February 1984



March 21, 1985

Statistician, Agent Orange Projects

## Battalion Tracking Procedures

Dan McGee, Senior Statistician, Agent Orange Projects

This memorandum reports the results of a review of the battalion abstraction process for the Agent Orange Study by Centers for Disease Control (CDC). A previous report by me gives details of the battalion abstraction process, including a review of army documents and terminology, a description of each data field on the battalion abstraction form, and a section on how to interpret text in the army documents. This report was sent to Mr. Don Hakenson of The U.S. Army and Joint Services Environmental Support Group (ESG) on February 6, 1985 for review.

The primary purpose of our review is to see if a group of researchers independent of the ESG can extract the same information on a battalion's locations using the same army documents. Specifically, can two independent groups arrive at the same overall battalion location for each day from their respective abstractions? The secondary purpose of the review is to numerate disagreements in abstractions. Characterizing disagreements will help describe the data and perhaps be useful in editing the battalion data sets for analysis.

CDC selected randomly four time periods to examine: May and June, 1967 of Battalion 4; April, 1967 of Battalion 12; and June, 1967 of Battalion 21. Drew Baughman and Christie Ernst researched the records for these battalions without knowledge of the data the ESG had abstracted. Appendix A contains all documents researched for this report. We used the procedures outlined in the report mentioned above; from all accounts, these procedures appear to be those that the ESG uses for their abstraction.

The first part of this report estimates the similarity and looks at differences between ESG and CDC abstractions. The second part compares the ESG set of locations before and after editing disagreements against the set of procedures for abstraction documented by CDC.

### I. ESG-CDC Comparison

The left side of Table 1 shows the distance in kilometers between the ESG and CDC overall battalion locations and the number of abstracted locations by battalion, date, and research group. For reasons not discussed here, cluster analysis is used to compute an overall battalion location for each day, for the ESG and CDC data sets separately. Appendix B outlines the algorithm to compute an overall battalion location or centroid.

TAB A

Table 1

Between-centroid distance and number of locations for initial ESG and CDC abstractions and for ESG abstraction before and after editing by CDC, by battalion and date.

Battalion 4:		(ESG, CDC)		(ESG, ESG/CDC)	
Date	Distance (km)	Number of Locations		Distance (km)	Number of Locations
May 1	1.1	16	16	0.04	16 14
May 2	0.6	12	15	0.1	12 14
May 3	-	0	24	-	0 23
May 4	0.04	14	20	0.02	14 18
May 5	0.3	20	24	2.4	20 21
May 6	0.03	22	32	0.02	22 26
May 7	0.0	7	7	0.0	7 7
May 8	4.1	11	4	5.2	11 4
May 9	22.0	2	1	11.0	2 3
May 10	-	0	0	-	0 0
May 11	-	1	0	-	1 0
May 12	-	1	0	-	1 0
May 13	30.3	10	9	22.0	10 8
May 14	3.1	17	16	2.9	17 16
May 15	24.5	13	8	24.9	13 8
May 16	1.6	5	9	2.0	5 5
May 17	0.3	35	15	0.3	35 18
May 18	-	0	25	-	0 18
May 19	1.1	17	22	0.5	17 14
May 20	0.4	15	6	3.3	15 10
May 21	0.04	22	18	0.5	22 19
May 22	3.7	12	5	1.0	12 12
May 23	6.0	24	5	8.1	24 20
May 24	17.5	25	10	16.6	25 24
May 25	3.0	16	7	0.2	16 17
May 26	2.7	23	10	0.1	23 19
May 27	2.3	13	3	0.04	13 13
May 28	18.5	20	7	0.0	20 20
May 29	4.5	13	1	0.2	13 7
May 30	1.4	14	7	1.0	14 15
May 31	1.8	26	11	0.2	26 24
Subtotal		426	337		426 417

Table 1 (Continued)

Between-centroid distance and number of locations for initial ESG and CDC abstractions and for ESG abstraction before and after editing by CDC, by battalion and date.

Battalion 4:		(ESG, CDC)		(ESG, ESG/CDC)	
Date	Distance (km)	Number of Locations		Distance (km)	Number of Locations
June 1	0.7	14	3	0.5	14
June 2	1.9	38	13	0.005	38
June 3	0.5	17	12	0.04	17
June 4	19.7	27	23	0.1	27
June 5	1.1	46	12	19.3	46
June 6	17.1	41	8	24.6	41
June 7	19.3	25	6	19.5	25
June 8	3.1	30	10	0.1	30
June 9	42.8	25	2	37.0	25
June 10	21.7	23	6	18.0	23
June 11	22.1	21	8	19.1	21
June 12	3.6	20	6	0.2	20
June 13	24.6	13	4	19.6	13
June 14	19.5	32	12	19.1	32
June 15	22.4	17	7	19.8	17
June 16	23.5	29	15	20.1	29
June 17	24.1	18	10	19.5	18
June 18	21.6	25	17	18.9	25
June 19	24.2	30	4	19.4	30
June 20	36.4	39	29	19.4	39
June 21	28.3	26	14	19.9	26
June 22	19.7	48	30	19.2	48
June 23	21.8	41	27	19.7	41
June 24	23.2	37	19	0.2	37
June 25	1.5	30	12	19.7	30
June 26	46.8	20	20	19.9	20
June 27	24.0	35	9	20.1	35
June 28	4.3	50	33	1.2	50
June 29	3.1	37	28	19.7	37
June 30	22.7	31	22	20.0	31
Subtotal		885	421	885	796

Table 1 (Continued)

Between-centroid distance and number of locations for initial ESG and CDC abstractions and for ESG abstraction before and after editing by CDC, by battalion and date.

Battalion 12:		(ESG, CDC)		(ESG, ESG/CDC)	
Date	Distance (km)	Number of Locations		Distance (km)	Number of Locations
April 1	7.6	5	10	12.6	5 10
April 2	0.07	15	22	0.1	15 20
April 3	3.6	8	7	0.2	8 11
April 4	-	0	0	-	0 0
April 5	7.9	5	5	0.0	5 5
April 6	-	0	0	-	0 0
April 7	0.3	4	5	0.0	4 4
April 8	1.5	10	13	0.7	10 14
April 9	1.9	8	8	1.9	8 8
April 10	0.1	8	9	0.1	8 9
April 11	0.02	18	24	0.01	18 25
April 12	49.1	6	11	2.9	6 10
April 13	0.04	18	21	0.01	18 22
April 14	2.3	16	20	3.5	16 20
April 15	5.3	19	16	5.3	19 18
April 16	0.2	13	18	0.04	13 22
April 17	9.9	7	13	9.9	7 12
April 18	12.3	5	2	0.5	5 6
April 19	-	1	0	-	1 0
April 20	1.7	2	1	1.7	2 1
April 21	-	0	0	-	0 0
April 22	-	0	0	-	0 0
April 23	4.1	1	11	4.0	1 8
April 24	3.7	8	15	3.7	8 13
April 25	5.6	11	22	7.1	11 19
April 26	3.2	11	16	4.5	11 14
April 27	15.3	7	15	4.1	7 12
April 28	0.1	10	14	0.1	10 14
April 29	0.2	9	14	0.2	9 10
April 30	5.1	10	19	5.1	10 19
Subtotal		235	331		235 326

Table 1 (Continued)

Between-centroid distance and number of locations for initial ESG and CDC abstractions and for ESG abstraction before and after editing by CDC, by battalion and date.

Battalion 21:		(ESG, CDC)		(ESG, ESG/CDC)	
Date	Distance (km)	Number of Locations		Distance (km)	Number of Locations
June 1	1.5	2	1	0.0	2 2
June 2	1.1	16	19	1.1	16 20
June 3	0.1	12	12	0.01	12 13
June 4	0.1	17	18	0.1	17 19
June 5	3.1	1	9	3.1	1 8
June 6	7.7	12	18	2.0	12 13
June 7	0.1	27	35	0.1	27 35
June 8	1.6	18	32	0.4	18 33
June 9	0.6	2	5	0.5	2 5
June 10	0.1	21	28	0.1	21 29
June 11	3.2	16	20	0.1	16 22
June 12	2.3	1	7	2.3	1 7
June 13	0.4	25	20	0.03	25 33
June 14	1.3	18	10	0.2	18 28
June 15	2.5	7	10	0.9	7 11
June 16	0.2	7	14	0.2	7 11
June 17	0.3	6	7	0.9	6 7
June 18	1.6	4	9	0.3	4 9
June 19	2.6	4	4	0.2	4 7
June 20	3.3	5	11	3.0	5 10
June 21	1.1	11	15	0.03	11 13
June 22	0.1	16	27	0.01	16 19
June 23	2.1	12	17	3.2	12 15
June 24	1.8	8	16	0.5	8 11
June 25	0.5	13	20	0.1	13 15
June 26	2.1	1	4	2.6	1 2
June 27	9.2	18	16	9.4	18 22
June 28	1.4	5	15	0.2	5 16
June 29	0.006	19	11	0.004	19 21
June 30	5.7	2	3	5.3	2 4
Subtotal		326	433		326 460
Total		1,872	1,522		1,872 1,999

Below is a summary of the (ESG, CDC) comparison in Table 1.

Battalion	Extract Period	Number of Days with Between-Centroid Distance			Number of Locations		ESG/CDC
		< 5 km	> 5 km	Total	ESG	CDC	
4 May67	Apr84	20	6	26	426	337	417
4 Jun67	Apr84	9	21	30	885	421	796
12 Apr67	Aug84	16	9	25	235	331	326
21 Jun67	Oct84	27	3	30	326	433	460

The ESG abstracted more locations for May of Battalion 4 (426 versus 337) and for June of Battalion 4 (885 versus 421), and CDC extracted more locations for April of Battalion 12 (331 versus 235) and for June of Battalion 21 (433 versus 326). CDC's abstraction places the battalion at least five kilometers from ESG's 27 of 56 days for Battalion 4, nine of 25 days for Battalion 12, and three of 30 days for Battalion 21. Assuming standard procedures for abstraction, these differences are due to either error in ESG's abstraction, error in CDC's, or both. I scrutinized the documents, the same for ESG and CDC, and found the latter to be the case. Since the purpose of this report is to review ESG's abstraction of troop locations from army documents, the discussion that follows is limited to ESG's errors only. Please note that "errors" subsumes a variety of possible mistakes and oversights and together they refer only to the standard of abstraction outlined by CDC.

Table 2 shows the errors in abstraction by the ESG for each battalion by date. Five categories of errors are listed there: transcription errors, projected or planned locations, coordinates imputed from a master list, a miscellaneous group, and omissions. The following is a summary of these disagreements.

Battalion	Extract Period	Number of Locations	Errors				
			Xpos.	Proj.	Imp.	Misc.	Omis.
4 May67	Apr84	427	11	39	120	16	56
4 Jun67	Apr84	885	78	151	364	11	95
12 Apr67	Aug84	235	0	3	8	9	124
21 Jun67	Oct84	345	0	5	2	3	175

The transcription errors in Battalion 4 are mostly wrong translations for the grid. For example, an entry without a grid attached such as 556998 was taken as XT556998 instead of XS556998. A few transcription errors are wrong digits such as taking 1 for 6, or transposition of digits such as 96 for 69.

Many locations in Battalion 4 were simply lists of patrols, check points, and command posts, only some of which are later reported for some group's arrival at these locations by stating "in position" or "closed." Most of the projection errors are taking every location in the list, not just the ones later confirmed, and some are taking planned locations from a paragraph explaining the next day's operations or from the plans summary at the end of the journal.

The documents for Battalion 4 also contained many references to arrivals to and departures from Fire Support Base Nickel (Nickel), Cu Chi base camp, and Bao Trai village or airstrip. Of the 529 such references, 312 are for Nickel (all XT571046), 145 are for Cu Chi base camp (137 are XT6412 and 8 are XT6415), and 52 are for Bao Trai village or airstrip (four unique locations). Table 3 shows a check CDC made to see if and when the location for Nickel changes.

Table 3

The location of Fire Support Base Nickel recorded by the ESG and found in the day's Brigade Sitrep by CDC, by date for Battalion 4.\*

Date	ESG	CDC
May 11	-	-
May 12	-	-
May 13	XT571046	-
May 14	XT571046	XT571046
May 15	-	-
May 16	XT571046	XT571046
May 17	XT571046	XT527043
May 18	-	-
May 19	XT571046	XT565046
May 20	XT571046	-
May 21	XT571046	XT569048
May 22	XT571046	-
May 23	XT571046	-
May 24	XT571046	-
May 25	XT571046	-
May 26	XT571046	-
May 27	XT571046	XT568042
May 28	XT571046	XT568042
May 29	XT571046	XT568042
May 30	XT571046	XT568042

\*CDC did not check the first ten days in May and June. A "-" means no reference to Nickel in the document.

Table 3 (Continued)

The location of Fire Support Base Nickel recorded by the ESG and found in the day's Brigade Sitrep by CDC, by date for Battalion 4.\*

Date	ESG	CDC
June 11	XT571046	XT648042
June 12	XT571046	XT568044
June 13	XT571046	XT568044
June 14	XT571046	XT568044
June 15	XT571046	-
June 16	XT571046	XT568044
June 17	XT571046	XT568044
June 18	XT571046	XT568044
June 19	XT571046	XT568044
June 20	XT571046	XT568044
June 21	XT571046	XT568044
June 22	XT571046	XT568044
June 23	XT571046	-
June 24	XT571046	-
June 25	XT571046	-
June 26	XT571046	XT568044
June 27	XT571046	XT568044
June 28	XT571046	XT569044
June 29	XT571046	-
June 30	XT571046	-

\*CDC did not check the first ten days in May and June. A "-" means no reference to Nickel in the document.

All references to Nickel for May and June of Battalion 4 were recorded as XT571046 and coded as inferred from the day's Brigade Sitrep by the ESG, but CDC found different coordinates for Nickel in the Brigade Sitreps, indicating the location of Nickel changed.

The miscellaneous group of errors include taking coordinates for the wrong battalion, an airstrike, artillery or small arms firing, five-digit coordinates, and keypunch error.

Omissions were spotted in all three battalions reviewed. These were mostly "patrol in position," "dustoff" or "dustoff complete" (medical evacuation of men wounded in action), and "n/c" (no change) or "same position as last hour."

In addition to errors committed in abstracting coordinates, the ESG took the TIME from the text and not the "TIME IN" column for all documents reviewed. The ESG and CDC had good agreement for the first digit of the UNIT field, but there were some differences in the second digit.



Table 2

Number of locations abstracted by the ESG and number of errors by battalion and date.

## Battalion 4:

Date	Number of Locations	Errors				
		Te	Pr	Ri	Ot	Om
May 1	16	0	2	0	0	0
May 2	12	0	0	0	0	2
May 3	0	0	0	0	0	0
May 4	14	0	0	0	0	6
May 5	20	0	0	0	0	4
May 6	22	0	0	0	0	8
May 7	7	0	0	0	0	0
May 8	11	1	9	0	0	0
May 9	2	0	0	2	0	2
May 10	0	0	0	0	0	0
May 11	1	0	0	1	1	0
May 12	1	0	0	1	1	0
May 13	10	0	0	2	1	3
May 14	17	2	0	2	1	2
May 15	13	4	0	0	1	0
May 16	5	0	0	1	1	6
May 17	35	0	0	1	1	1
May 18	0	1	0	0	0	1
May 19	17	0	2	2	1	0
May 20	15	0	4	5	0	1
May 21	22	0	3	1	0	0
May 22	12	0	1	5	3	3
May 23	25	1	3	14	1	0
May 24	25	1	1	14	0	0
May 25	16	0	1	10	0	2
May 26	23	0	5	9	0	3
May 27	13	1	0	10	0	1
May 28	20	0	0	11	2	1
May 29	13	0	4	7	1	0
May 30	14	0	0	9	1	4
May 31	26	0	4	13	0	6
Subtotal	427*	11	39	120	16	56

## Legend:

Te = transcription error

Pr = projected or planned

Ri = researcher imputed coordinates from master list

Ot = wrong battalion, airstrike, artillery or small arms firing, five-digit coordinate, keypunch error

Om = omission

\* = includes namecode-only entries

Table 2 (Continued)

Number of locations abstracted by the ESG and number of errors by battalion and date.

## Battalion 4:

Date	Number of Locations	Errors				
		Te	Pr	Ri	Ot	Om
June 1	14	0	1	10	0	1
June 2	38	1	1	20	0	0
June 3	17	2	1	9	0	3
June 4	27	3	0	15	0	1
June 5	46	3	13	21	0	2
June 6	41	10	13	17	0	1
June 7	25	1	4	15	3	1
June 8	30	0	3	19	1	2
June 9	25	2	0	16	0	0
June 10	23	1	13	5	0	2
June 11	21	2	2	10	0	2
June 12	20	0	1	13	2	2
June 13	13	1	0	9	0	0
June 14	32	3	8	14	1	4
June 15	17	1	4	6	1	1
June 16	29	1	7	7	0	0
June 17	18	0	0	10	0	0
June 18	25	0	1	4	0	0
June 19	30	0	9	13	1	1
June 20	39	4	4	12	1	8
June 21	26	5	8	10	0	6
June 22	48	3	3	12	1	5
June 23	41	6	3	16	0	7
June 24	37	6	2	18	0	5
June 25	30	4	5	12	0	3
June 26	20	5	3	6	0	11
June 27	35	4	15	14	0	4
June 28	50	5	10	7	0	4
June 29	37	1	12	15	0	8
June 30	31	4	5	9	0	11
Subtotal	885	78	151	364	11	95

## Legend:

Te = transcription error

Pr = projected or planned

Ri = researcher imputed coordinates from master list

Ot = wrong battalion, airstrike, artillery or small arms firing, five-digit coordinate, keypunch error

Om = omission

Table 2 (Continued)

Number of locations abstracted by the ESG and number of errors by battalion and date.

## Battalion 12:

Date	Number of Locations	Errors				
		Te	Pr	Ri	Ot	Om
April 1	5	0	2	1	0	10
April 2	15	0	0	0	0	5
April 3	8	0	0	0	0	3
April 4	0	0	0	0	0	0
April 5	5	0	0	2	0	2
April 6	0	0	0	0	0	0
April 7	4	0	0	0	1	3
April 8	10	0	0	0	2	6
April 9	8	0	1	0	0	2
April 10	8	0	0	0	1	2
April 11	18	0	0	0	0	7
April 12	6	0	0	0	0	5
April 13	18	0	0	0	0	3
April 14	16	0	0	0	1	5
April 15	19	0	0	0	2	0
April 16	13	0	0	0	0	9
April 17	7	0	0	0	1	7
April 18	5	0	0	3	1	2
April 19	1	0	0	1	0	0
April 20	2	0	0	1	0	0
April 21	0	0	0	0	0	0
April 22	0	0	0	0	0	0
April 23	1	0	0	0	0	9
April 24	8	0	0	0	0	8
April 25	11	0	0	0	0	8
April 26	11	0	0	0	0	5
April 27	7	0	0	0	0	8
April 28	10	0	0	0	0	4
April 29	9	0	0	0	0	2
April 30	10	0	0	0	0	9
Subtotal	235	0	3	8	9	124

## Legend:

Te = transcription error

Pr = projected or planned

Ri = researcher imputed coordinates from master list

Ot = wrong battalion, airstrike, artillery or small arms firing, five-digit coordinate, keypunch error

Om = omission

Table 2 (Continued)

Number of locations abstracted by the ESG and number of errors by battalion and date.

## Battalion 21:

Date	Number of Locations	Errors				
		Te	Pr	Ri	Ot	Om
June 1	6	0	1	0	0	0
June 2	17	0	0	0	0	1
June 3	12	0	0	1	0	11
June 4	17	0	0	0	1	3
June 5	1	0	0	0	0	8
June 6	14	0	0	1	1	9
June 7	27	0	0	0	0	9
June 8	18	0	0	0	0	15
June 9	3	0	0	0	0	4
June 10	23	0	0	0	0	9
June 11	16	0	0	0	0	5
June 12	2	0	0	0	0	8
June 13	25	0	0	0	0	1
June 14	18	0	1	0	0	5
June 15	10	0	1	0	0	5
June 16	8	0	0	0	0	5
June 17	6	0	0	0	0	4
June 18	5	0	0	0	0	5
June 19	4	0	0	0	0	3
June 20	5	0	0	0	1	10
June 21	11	0	0	0	0	4
June 22	16	0	0	0	0	7
June 23	12	0	1	0	0	5
June 24	8	0	0	0	0	9
June 25	14	0	1	0	0	8
June 26	1	0	0	0	0	4
June 27	18	0	0	0	0	4
June 28	5	0	0	0	0	10
June 29	19	0	0	0	0	1
June 30	4	0	0	0	0	3
Subtotal	345*	0	5	2	3	175
Total	1,865	89	198	494	39	450

## Legend:

Te = transcription error

Pr = projected or planned

Ri = researcher imputed coordinates from master list

Ot = wrong battalion, airstrike, artillery or small arms firing, five-digit coordinate, keypunch error

Om = omission

## II. ESG-ESG/CDC Comparison

13

To estimate the affect, if any, of errors in abstraction on the daily overall battalion location, I compute the distance between the daily centroids derived from the ESG set of data before and after editing the errors listed in Table 2. All edits to the ESG data set are based on the rules of abstraction detailed by CDC in a previous report. The right side of Table 1 tabulates this distance by date for each battalion. The following summary of these distances shows a difference of at least five kilometers 28 of 56 days for Battalion 4, five of 25 days for Battalion 12, and two of 30 days for Battalion 21.

Battalion	Extract Period	Number of Days with Between-Centroid Distance			Number of Locations	
		< 5 km	> 5 km	Total	ESG	ESG/CDC
4 May67	Apr84	20	6	26	426	417
4 Jun67	Apr84	8	22	30	885	796
12 Apr67	Aug84	20	5	25	235	326
21 Jun67	Oct84	28	2	30	326	460

Extrapolating these figures to the rest of the battalion tracking data, I recommend that Battalion 4 be re-abstracted using the current set of tracking procedures, and one randomly selected week from each of Battalions 1-3, 5-7 be reviewed. This statement is based on: (1) the poor reproducibility of the daily overall battalion location due to errors in abstraction for about 50% of the days CDC reviewed for Battalion 4; (2) the present Battalion Abstraction Form was not implemented until Battalion 8, and therefore a different sort of abstraction may have occurred before Battalion 8; and (3) the current set of tracking procedures were not formalized for the early battalions.

Drew Baughman

CDC:CEH:CDD:AOP:DBaughman:pfh 3/21/85  
DOC#96 SEC12

## Appendix B

I used PROC FASTCLUS in SAS to cluster points for each day, for each data set separately. This procedure uses an agglomerative, nearest neighbor method of clustering. It uses the first nonmissing observation (point) as the initial cluster seed, defines it as the first cluster, and proceeds sequentially through the data set, computing the Euclidean distance from the present observation to each of the cluster seeds. An observation is considered a new seed if its minimum distance to previous seeds is greater than 3 km (chosen by user).

Each observation is assigned to the cluster with the nearest seed, and after an observation is processed, that cluster's seed is recalculated as the mean of the observations currently assigned to the cluster. Cluster seeds are iteratively recomputed up to five times (chosen by user).

I allowed for a maximum of five clusters on any day and computed the centroid of all cluster centroids by day for the ESG and CDC sets of data, separately.



DEPARTMENT OF THE ARMY  
US ARMY & JOINT SERVICES ENVIRONMENTAL SUPPORT GROUP  
1730 K STREET N.W. ROOM 210  
WASHINGTON, DC 20006-3868

3-1

CHY SSD

REPLY TO  
ATTENTION OF

DAAG-ESG

10 September 1985

*J. McGee*  
*10 Sep 85*

MEMORANDUM FOR DIRECTOR, ESG

SUBJECT: CDC Quality Control Report, Dated 1 July 1985

1. On 16 July 1985, Bob Delaney, Drew Baughman and Riduan Joesoef from CDC in Atlanta, Georgia, visited our organization to discuss battalion abstraction procedures and quality control.
2. During the visit Mr. Joesoef provided Mr. Hakenson of our staff a copy of a quality control report he had sent to Mr. McGee concerning a quality control abstraction comparison between CDC and ESG. He reported that the CDC abstraction produced more points of locations than ESG had abstracted.
3. CDC based their analysis on a one week reabstraction on battalions #1, 2, 5, and 7. Their recommendation was that battalion #2, 5 and 7 be reabstracted. This was based on CDC's reabstraction of the battalion data as the "gold standard". I would like to point out that the procedures used by CDC abstractors were obtained from a memo dated March 21, 1985, written by Mr. Drew Baugman, after spending 3 days at our organization attempting to learn the abstraction process in January 1985.
4. Upon further reexamination and analysis of the CDC and ESG abstraction process we were able to uncover CDC abstraction errors and key punch errors by our staff. On battalion #2 there were two days in which CDC and ESG did not agree. This was due to a key punch error in which two days of grid coordinates were abstracted for the same day. Except for the key punch error, all data was exact. In our opinion this did not warrant a total reabstraction.
5. On our examination of battalion #7 we found numerous grid coordinate errors by the CDC abstractor. CDC showed they found 20 grid coordinate points. Our examination reduced the number of points to 9. ESG's original abstraction produced 15 good grid coordinate locations. Listed below are the types of recurring grid coordinate locations that CDC abstracted.

- a. VC Sighting - When the tracked unit sighted VC and provided a grid location the CDC abstractor recorded this grid. This is the VC's location, not the tracked unit location.

TAB B

10 September 1985

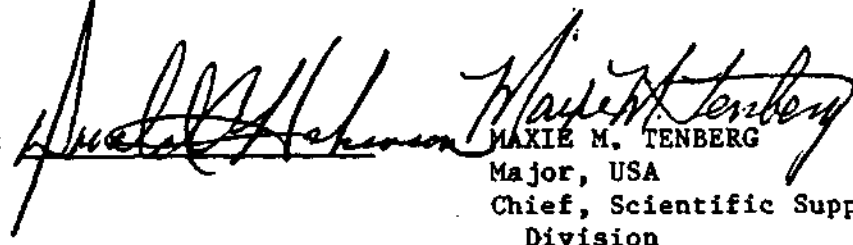
SUBJECT: CDC Quality Control Report, Dated 1 July 1985

b. When an artillery unit fired in an area, CDC recorded this location. This is not the units location but where that unit was firing.

c. One entry that was recorded was that the tracked artillery unit was supporting another unit in a specific location. This was not the tracked units location but another unit that was requesting artillery support (fire support).

6. Due to the discrepancies we found in the CDC abstraction process, we do not feel the battalions warrant reabstraction. Our new researchers have made the same sort of errors, but are taught through constant training to find and identify these types of incidents and not to record them. The only conclusion I can draw from this report is that CDC is not proficient enough at this stage of the battalion tracking process to produce reports criticizing ESG. However, this did not stop CDC from producing and circulating a highly critical quality control report. I would also like to point out that Mr. Joesoef had never extracted grid coordinates from battalion records before. I would strongly suggest that CDC follow the ESG SOP Abstraction Procedures dated May 7, 1985 when abstracting grid coordinate locations.

Concur:

  
MAXIE M. TENBERG  
Major, USA  
Chief, Scientific Support  
Division



Date: July 1, 1965.

From: Riduan Joesoef

Subject: Results of quality control battalion 1,2,5, and 7.

To: Dan McGee

This report summarizes the results<sup>2</sup> of quality control for battalion tracking 1,2,5, and 7. For battalion 3 CDC received brigade daily journal instead of battalion daily journal, thus, a comparison to the ESG abstraction was not possible. However, a comparison of battalion daily journal and brigade daily journal was conducted to give a better understanding of the journals.

#### Quality Control Procedures

The same quality control procedures as described in Drew Baughman's memo dated March 21, 1965 were applied, except for data entry, edit, and verification work. This extra work was conducted locally by Agent Orange personnel to reduce keypunching error to a minimum. In addition, only one week period (7 days) of abstraction for each battalion was abstracted.

#### ESG - CDC comparison

A comparison between ESG and CDC abstractions indicated that the CDC abstraction produced more points of locations (Table 1). In regard to the between centroid-distance only battalion 5 and 7 show significant distances (Table 1).

To assess further the quality of ESG abstraction, an evaluation of the number of locations or days which were in the source documents but omitted by ESG abstraction was conducted. In this report, this type of error was defined as omission error which consisted of locations and days omission.

Table 1  
Between centroid-distance and number of locations  
for ESS and CDC abstraction

Battalion	Date	Distance (km)	Number of Locations	
			ESS	CDC
1	Oct. 15, 1967	0	2	2
	Oct. 16, 1967	0	6	6
	Oct. 17, 1967	2.23	5	26
	Oct. 20, 1967	-	4*	0
	Oct. 21, 1967	-	0	10
			17	15
2	Mar. 18, 1968	0	1	1
	Mar. 20, 1968	0	1	1
	Mar. 21, 1968	0	1	1
	Mar. 22, 1968	3.65	15	20
	Mar. 23, 1968	-	0	11
			19	23
5	Mar. 19, 1968	-	1*	3
	Mar. 20, 1968	26.03	3	3
	Mar. 21, 1968	0.45	1	2
	Mar. 22, 1968	79.10	2	1
	Mar. 23, 1968	-	0	1
	Mar. 25, 1968	38.99	5	3
			12	13
7	Nov. 17, 1968	-	0	20
	Nov. 18, 1968	-	1	8 →
	Nov. 19, 1968	<del>12.30</del>	1	21 error CDC
	Nov. 20, 1968	19.41	4(5)	21
	Nov. 21, 1968	3.53	3(2)	22(1) error
	Nov. 22, 1968	-	3	52
	Nov. 23, 1968	-	3	22
			15	20

\*Locations include name-code only.

*This data needs  
to be rechecked*

*wrong date*

These errors were critical because they reduced the number of location points and unique days obtained from the abstraction. Another type of error was inclusion errors. These errors were defined as the number of locations or days which were not in the source document but included by ESG abstraction.

To evaluate the omission and inclusion errors, the ESG and CDC abstractions were combined and edited to produce a compromised, optimal abstraction. This optimal abstraction was named the ESG and CDC abstraction and used as a "gold standard" to evaluate the ESG abstraction. Two summary tables-- Table 2 and 3 --were created for these purposes.

Table 2 shows the number of locations and days for omission and inclusion errors by battalion. For example ESG omitted 14 location points for battalion 2 which constituted a 64% (14/22) omission error. This means that ESG abstraction missed 64% of the location points that could be abstracted from that battalion. The number of days that ESG omitted for battalion 2 was 1 day which constituted a 20% (1/5) omission error. On the other hand, ESG abstraction included 11 more location points (these 11 points were not in the source document).

Table 2  
Number of locations and days errors  
abstracted by ESG by battalion

Batt.	Number of Locations by ESG&CDC	Location-Errors				Number of Days by ESG&CDC	Day-Errors			
		Omission N	P	Inclusion N	P		Omission N	P	Inclusion N	P
1	21	3	7%	0	0%	5	1	20%	0	0%
2	22	14	64%	11	50%	5	1	20%	0	0%
5	13	8	62%	4	31%	0	2	33%	0	0%
7	20	13	65%	10	50%	6	1	17%	1	17%

Batt.=Battalion

N= Number of omission or inclusion

P= Percent of omission or inclusion

*This is more  
bullshit!*

In summary, battalion 2, 5, and 7 have the most errors for either locations or days.

To investigate the two types of errors at company level, Table 2 was broken down by company to produce Table 3. This table shows the number of locations and days errors at company level for each battalion. A similar pattern was noticed. The omission errors for companies in battalion 2, 5, and 7 were serious. For battalion 1 only company B has significant errors. It is worth noting that only division daily journals were available for battalion 5. Consequently, only two companies C and D were produced from the division journal abstraction. Unfortunately, company C has a 100% omission error for both locations and days.

In conclusion, reabstraction of battalion 2, 5, and 7 are necessary to increase coverage of locations and unique days for either company or battalion. If time permitted battalion 1 should also be reabstracted.

Table 3  
Number of locations and days errors  
abstracted by ESG by battalion and company

*This is now  
bull shit.*

Batt.	Comp.	Number of Location by ESG&CDC	Location-Errors				Number of Days by ESG&CDC	Day-Errors			
			Omission		Inclusion			Omission		Inclusion	
			N	P	N	P		N	P	N	P
-----											
1	A	3	0	0%	0	0%	1	0	0%	0	0%
	B	9	3	33%	0	0%	3	1	33%	0	0%
	C	4	0	0%	0	0%	3	0	0%	0	0%
	D	1	0	0%	0	0%	1	0	0%	0	0%
	E	0	-	-	-	-	0	-	-	-	-
	X	4	0	0%	0	0%	2	0	0%	0	0%
2	A	3	1	33%	1	33%	2	1	50%	0	0%
	B	3	2	66%	1	33%	2	1	50%	0	0%
	C	3	0	0%	0	0%	3	0	0%	0	0%
	D	0	-	-	-	-	0	-	-	-	-
	E	0	-	-	-	-	0	-	-	-	-
	N	3	2	66%	0	0%	2	1	50%	0	0%
	X	10	9	90%	8	80%	1	0	0%	0	0%
5	A	0	-	-	-	-	0	-	-	-	-
	B	0	-	-	-	-	0	-	-	-	-
	C	4	4	100%	0	0%	3	3	100%	0	0%
	D	0	-	-	-	-	0	-	-	-	-
	E	0	-	-	-	-	0	-	-	-	-
	N	9	4	44%	0	0%	5	3	60%	0	0%
7	A	6	3	50%	0	0%	4	3	75%	0	0%
	B	4	4	100%	2	100%	2	2	100%	0	0%
	C	0	-	-	-	-	0	-	-	-	-
	D	6	2	33%	2	33%	3	1	33%	0	0%
	E	4	4	100%	0	0%	4	4	100%	0	0%

Batt.=Battalion, Comp.=Company  
N= Number of omission or inclusion  
P= Percent of omission or inclusion

# Battalion and Brigade Daily Journals Comparison

Table 4 indicates that battalion daily journal has more location points and days than brigade daily journal. However, only 29% of the location points in the brigade journal matches with the points in the battalion journal.

Table 4

Number of Locations and Days for  
Battalion and Brigade Daily Journals

Daily Journal	NUMBER OF Locations	NUMBER OF Days
Brigade	35	5
Battalion	114	7

  
Riduan Joescher

CC. Drew Baughman

Dennis Smith

Debbi Kotlovker

18 November 1985

Mr. Richard Christian  
Director  
U. S. Army and Joint Services Environmental Support Group  
1730 'K' Street, Room 210  
Washington, D. C. 20006-3868

*Duck*  
Dear Mr. Christian,

This letter is in follow-up to our telephone conversation of 18 November 1985, concerning the abstraction of grid coordinates from various military documents.

Beginning in the early planning stages of the Agent Orange Study, the Centers for Disease Control (CDC) provided very explicit instructions regarding the abstraction of grid coordinates from military documents. In numerous planning meetings, Dr. Lee Anest stated which coordinates should/should not be taken. He stated that coordinates missing letters or digits (Note: an exception being a four [4] digit coordinate) should not be taken as the true location could not be verified. Moreover, this point of interest was discussed numerous times with several other individuals from CDC and at no time was ESG given any instructions to the contrary. CDC stressed time and again the importance of consistency in data abstraction and not extrapolating on any incomplete data. Unless the document clearly stated that the specific unit of interest was at a particular location, coordinates were not taken. For example, enemy locations were not taken unless U. S. troops were also at that same location.

Finally, as we are both fully aware, the military documents are not perfect and they were never intended to be used for any such scientific endeavor. Never-the-less, despite their limitation, I firmly believe that they do contain a large quantity of excellent information.

Should you require any further clarification, don't hesitate to contact me.

*Rob Lipnick*  
Robert J. Lipnick, DSc  
Epidemiologist

TAB D



DEPARTMENT OF HEALTH & HUMAN SERVICES

Public Health Service

Centers for Disease Control

February 10, 1984

Rob Lipnick, Ph.D.  
c/o Mr. Dick Christian  
Environmental Support Group  
Suite 210  
1730 K Street, N. W.  
Washington, D. C. 20006

Dear Rob:

The AOP staff express our appreciation to you and your staff for taking the time to explain the methods and materials used in the Battalion search for the Agent Orange Study Pretest. The Agent Orange Project staff have reviewed your questions concerning the battalion search and would like to provide necessary guidelines to standardize the abstraction procedure for data collection in subsequent Battalion searches. In order to adequately address these issues, we would like ESG to provide us with some detailed information concerning the locations and movements of infantry and artillery units selected for the Agent Orange Study.

CDC would like ESG to initially track the companies/batteries, excluding the headquarters company/battery, in each of the following four battalions:

1. 1st Battalion/2nd Inf. Regiment/1st Brigade/1st Inf. Division
2. 2nd Battalion/2nd Inf. Regiment/3rd Brigade/1st Inf. Division
3. 5th Battalion/2nd Artillery Regiment/II FForce.
4. 1st Battalion/5th Artillery Regiment/DivArty/1st Inf. Division

(Note that on the original CDC list of 48 Battalions, these units were #1, #2, #6, and #7.) These battalion searches will include tracking at the battalion, company/battery, platoon and squad unit levels.

Currently, Paul Simpson, a computer programmer with the Project, and I are working on a data entry and management system using the computer software "Infostar" for your use in abstracting data in a standardized way from the military operational and intelligent reports onto the KAYPRO II personal computer. He and I plan to be in Washington on February 13 through 15 to train the ESG researchers how to use the computer program. I have been corresponding with Joan Wilson of your staff about the format of the data entry form. A sample of the proposed data entry form is attached. Most of the fields are self-explanatory. The "time" field refers only to the logging in or out of coordinates specified in the Brigade and Battalion Daily Journals. The form will be pretested here using some data sent down to us by Joan. The remaining "bugs" in the system can be worked out during our stay in Washington. Other details about transferring the data to CDC will also be decided at that time.

TAB E



Page 2 - Dr. Lipnick

As a measure of accuracy of data abstracted from the Daily Journals and other military documents, CDC will conduct a replicate study of these four battalions. Essentially, this will entail designating two CDC staffers with the Agent Orange Project to replicate at least a portion of the abstraction process conducted initially by an ESG team. The design and content of the replicate study is being discussed and developed. This study will not interrupt ESG's progress on the Battalion search. CDC will provide ESG with details within the next few weeks.

The replicate study is not in lieu of the quality control program ESG will perform for the Battalion search. We would like to have in writing a description of your quality control procedures including (1) training of new researchers, (2) review of data abstracted from the military records, and (3) documentation of methods and materials. The CDC staff would, of course, like to review and comment on the quality control program.

A double-blind quality control study using CDC staff would not be feasible. However, the replicate study can be conducted independently of the ESG research team. The goal in this exercise would be to compute the concordance or discordance of location information abstracted by the two groups. If such a study were not conducted, accuracy of the abstraction method could be open for criticism as not adequately evaluated.

For the initial searches on the four battalions, it would be best if a team of two ESG researchers searches the entire two years of a given battalion. In this way, inter-team variability within a battalion can be avoided, thus the statistical evaluation of the replicate study will be more straight forward.

Thank you for your time and cooperation.

Sincerely yours,



Lee Annest, Ph.D.  
Geneticist and Statistician  
Agent Orange Projects, CEH, CDC

Enclosure

## Memorandum

Date November 15, 1988

From Carl Keller *CK*  
Chair, Science Panel, Agent Orange Working Group

Subject Progress of Cohort Selection for Agent Orange Epidemiological Study

To Dixon Arnett  
Acting Chair, Agent Orange Working Group

At a meeting of the Science Panel held after the regular meeting of the AOWG on November 6, 1988, the progress of the cohort selection process for the Agent Orange Epidemiological Study being conducted by CDC was discussed. Following the suggestion by congressional staff that meetings be held between CDC Agent Orange Projects staff and the Environmental Support Group in the presence of representatives from the Science Panel, AOWG and OTA, and that the Congressional staff be notified when such meetings were to be held, the Science Panel selected me, as a subcommittee of one, to work with the other participants. I agreed to participate on those terms, and Ms. Helen Gelband as an observer to the AOWG and as the responsible person at OTA also agreed to attend such meetings. We scheduled a meeting between the principals for November 12 to be held in Washington at the offices of the ESG and I requested that both CDC and ESG select the same unit and time period and determine its daily location by grid coordinates in Vietnam using their respective methods so that we could compare the two methods in concrete rather than theoretical terms. It was hoped that we would be able to judge which method would be more suitable for estimating possibilities for exposure to Agent Orange among Army combat troops in Vietnam, and thus enhance the quality of inferences concerning the health effects which might be due to such exposure if the results of the study so indicated. In order to clarify the purpose of the meeting, I prepared an agenda, attached, which outlined two proposed methods for locating company sized units on those days for which no precise grid coordinate for the given company was available in battalion daily journals or higher level reports. In addition to the principals listed on the agenda and their colleagues, the meeting was attended by Mr. Victor Raymond, Majority Staff Member for the Hospitals and Health Care Subcommittee of the House Veterans Affairs Committee.

At the beginning of the meeting, CDC indicated that the procedure outlined in method "b" of the agenda was clearly the more accurate but was probably not feasible to do for all units which would have to be tracked for the study. The method is far too hand-intensive, relies on information which has not and possibly can not be abstracted and coded from the source documents and requires a level of sophistication for interpreting records which is not easily acquired on a large enough scale to avoid excessive delay in the completion of the study. During the comparison of daily grid locations identified by the two methods on a day-by-day basis, however, it became apparent that there were significant gains in the accuracy of estimating the daily location of companies to be had by the contextual approach. The remainder of

TAB C

the meeting was devoted to discussions of how to make method "b" more feasible for use in the ongoing study. ESB already has sufficient depth of expertise in records interpretation to undertake the task for a substantial number of units. It was suggested by OTA that a good way to screen out units for which it would not benefit the study to accurately locate is to use method "a" to identify units which are more than 25 kilometers from any herbicide application during one-month time periods, and to consider them as unexposed during that month. Computerized data already on file at CDC should be accurate enough to make this an efficient and feasible screening process and the anticipated clustering of herbicide applications should yield a substantial number of battalion-months that will not require a hand search for locational information. As an example, the battalion containing the company which was selected at random to prepare for the meeting was never within 25 kilometers of a herbicide application during the month selected, and it would have been unnecessary to more accurately locate any of the companies of that battalion during that period. All participants in the meeting agreed that the procedures suggested and discussed would benefit the timely progression and usefulness of the Agent Orange Epidemiological study.

Agenda for meeting  
1 PM November 12, 1965 1730 E St. Washington, D.C.

Scheduled Participants	
Don McQuee	ADP/CNC
Dick Christian	ESG/DOB
Edna Gelband	OTA/US Congress
Carl Keller	ADWG/ENS

The main order of business will be to attempt to describe an appropriate method for inputting the specific UTM grid locations of Infantry Companies on those days for which a precise location is not given in the Battalion Daily Journal or other battalion and higher level reports. At least two methods have been proposed as follows:

- a) A method developed by the ADP at CNC would put a Company next to other companies or elements of the same Battalion which have a known precise location given for the same day on those days for which there is no information concerning its location in the Battalion or higher level reports. This method has the advantage of employing a computational algorithm and can thus be efficiently accomplished with a computerized data base.
- b) A method proposed by ESG of DOB would use all available information, including Company Morning Reports, to estimate where each company is on those days for which no precise location exists among Battalion and higher level records. It is not clear whether this approach constitutes an "informed guess", how much information would be routinely used to assign a location to all companies which might be selected for the Agent Orange Study, nor how much time would be required to accomplish this task.

So far, communication among the principals and their agents has not resolved the issue of whether one method is more accurate, or more reliable, than the other. ESG maintains that the Morning Reports contain useful, and even essential, information on the location of company-level units for the purposes of the Study (i.e., watching company locations with herbicide application locations contained in the Services Herbs Tape). The ADP has compared the location indicated on the Morning Reports with that obtained from their own method for days for which the location of a number of companies was known from other sources, and found their own method to be more accurate. It was suggested by OTA and ADWG that such a test was inappropriate since it may well be the case that on those days for which a given company is listed in battalion records it is also more likely to be operating in the field and near other companies of the same battalion, while on those days that it is not listed among battalion records it is more likely to be in its base camp or other inactive position that may be best indicated on the Morning Report.

In order to help resolve the issue, it was proposed on October 30 to compare the results of using each of these methods on the same units over the same time period. Since the problem almost certainly involves how records are used as well as which records are used, it is essential that the results to be compared are prepared by the investigators at ESG and ADP according to their respective methods. Congressional Staff present on October 30 endorsed this plan and have even requested that it be formalized in writing. For the purposes of the present meeting, I am requesting that at least all of the companies in one or two battalions be "located" for a specified one or two month period using both methods so that we can compare the results today. In order to compare results, it will be necessary to go through each bit of information and decision in some detail since there is no "gold standard" for comparison. It will also be useful to consider the feasibility of both methods for the full study and the time and effort required to implement them as well as the effects of any misclassification introduced in the assignment to high and low exposure cohorts.