Lewis Lorton,[1] *D.D.S., M.S.D. and William H. Langley,*[2] *M.S.*

Design and Use of a Computer-Assisted Postmortem Identification System

REFERENCE: Lorton, L. and Langley, W. H., **"Design and Use of a Computer-Assisted Postmortem Identification System,"** *Journal of Forensic Sciences*, JFSCA, Vol. 31, No. 3, July 1986, pp. 972-981.

ABSTRACT: The organization and rationale for the design of a computer-assisted postmortem identification system are discussed along with results of the use of this system in extensive simulation trials on a database of 578 records. The selectivity of dental characteristics is so great that any individual with 4 or more characteristics (either fillings or missing teeth), can be separated from a group of 578 people for final verification of the identity match. The effects of errors in the database are discussed and the actual effects of different error rates on identification are shown. Error rates of up to 30% have only small effects on the ability of the system to pick out correct identity matches. The system is presently implemented on a portable microcomputer, a representative desktop computer, and a large minicomputer. The present efforts include statistical analysis of an enlarged database and testing of a data acquisition system to allow the building of a large identification database (25 000 records) in a quick and economical manner.

KEYWORDS: odontology, human identification, computers, dentition, postmortem identification, identification, mass casualty, forensic dentistry, dental characteristics, teeth

In a previous paper [*1*], certain precepts were suggested for a computer-assisted system for postmortem identification (CAPMI). The system should be simple, based on clear logical decisions, should not require decisions on subjective matters, and should be inexpensive to implement and to use. Dental characteristics were tested as the primary comparison factors for the initial work because of the clarity of the decision process, the inherent selectivity of dental characteristics, and their ubiquity in postmortem identification.[3] This paper describes the design of the identification system, and presents the methods, results, and conclusions of the field trial ($n = 578$).

The system is designed to optimize the initial part of the identification process, designated the "comparison/selection" phase. In this phase the large database is searched for the most possible matches to the unknown in question and a smaller subset of most probable identities is formed for further examination in the second or "verification" phase.

While testing the reliability of dental characteristics to serve as comparison/selection factors, we also, simultaneously, tested the concepts of a computer-assisted system of identifica-

Received for publication 22 April, 1985; revised manuscript received 27 Nov. 1985; accepted for publication 3 Dec. 1985.

[1]Colonel, Dental Corps, USA; Chief of bioengineering branch, U.S. Army Institute of Dental Research, Presidio of San Francisco, CA.

[2]Chief, Applications Team, Information Sciences Branch, Letterman Army Institute of Research, Presidio of San Francisco, CA.

[3]Tadao Furue, personal communication, chief forensic anthropoligist, U.S. Army Central Identification Laboratory, Hawaii.

tion. The design proved to have a high degree of selectivity even when the database was intentionally contaminated with 10 to 40% errors.

Functional Description

An identification system was developed that will accept the dental and physical records of an unknown to be identified, compare this "key" record with a database made up of equivalent records from the population at risk (known as the "object" records), and return a list of the best match possibilities according to certain specific criteria. Records can be coded with only a specific set of possible codes that describe particular physical and dental characteristics. The characteristics are listed and described in Appendix A.

In this test, the physical characteristics, other than dental, were not used as selection parameters since the object of the trial was to evaluate the selectivity of the dental characteristics alone. Certain information that could be useful to the "verification" phase is not included in this comparison/selection process, although it is coded on the record. Previous research has shown that including more specific factors, such as the type of filling material, in the comparison routines greatly increases complexity of the logical processes with no increase in the power of the selection. By using only the information on the surfaces restored and teeth missing, the decision logic is greatly simplified and the process of comparison greatly hastened with *no loss of selectivity*.

The logical processes for the comparison routines are described in Appendix B.

System User Interaction

The computer comparison system was designed to be used by identification personnel with little or no computer knowledge. Programmed functions are chosen from menus displayed on the screen; all acceptable commands are displayed on the computer screen to be chosen. Most functions require only the selection of the appropriate item from a menu. All mistakes or unacceptable entries are signaled by plain-English error messages.

The database can be constructed at the site of a mass casualty from records collected at the time or from an antemortem database, which has been established before the need. The database can be downloaded from its storage site to a local computer.

Upon command each record from the "object" file is compared with the "key" file tooth by tooth, and the numbers of matches, mismatches, and possible matches are recorded for each record in the object database. (This assumes that the physical characteristics are not included in the parameters requested.) The "object" file is then reordered in descending order of numbers of match and possible match counts. For example, the first record(s) on the best matches list may have 32 matches, no possible matches; the next record(s) would have 31 matches, 1 possible match; the next 31 matches, 0 possible, 1 mismatch, and so forth. The user decides the number of records reported out on the best matches list and can, indeed, specify that the entire "object" database be listed in order of its probability.

When the physical parameters are included in the search, the comparisons for the dental characteristics are made only on those records which match the physical criteria specified. The criteria used, or the range where appropriate, can be changed after each search to allow a wider or narrower selection of records to be considered.

The generated record list of the most likely matches for the unknowns can be used, along with complete dental charts and any other pertinent information, to sort through the records to establish and verify the identity.

Computer Decisions

A seductive way to handle the task of selecting and ordering possible identity matches from a large database is to leaf through the records and *eliminate from immediate consider-*

ation those records which have disqualifying mismatches in dental characteristics. This is done rather than tallying and accumulating a running total of the count of matches, mismatches, and possibles for each comparison.

During the development of the record comparison algorithms, the records were reordered in two separate ways. For each identification problem the database was reordered (1) by maximum number of matches, and within each category of equal number of matches, by least number of mismatches (or most number of possible matches, which is the same); and (2) by minimum number of mismatches and, within each equal set of mismatches, the maximum number of matches.

Intuitively, one would say that these two methods should produce the same list. In fact, they did, *until errors were introduced into the database*. Where any errors existed in the database, the method of listing by least number of mismatches reduced the chance that the correct match was high on the most probables list in proportion to the percentage of errors in the database.

An exaggerated example may clarify this point. An "object" database consists of the dental records of 100 people. Of these, 99 have no missing teeth and no restorations. One person has every molar restored with a multisurface restoration (twelve teeth restored in all). In the rush of charting one tooth which has a three-surface filling was mischarted as having only two surfaces (MO[3] rather than MOD). If this particular person were found dead and his record, correctly charted was entered into the system for search, two outcomes would be possible. If the records in the object database are reordered by maximum matches, minimum mismatches, the correct match would be first on the list with 31 matches, 1 mismatch. Next would follow the other 99 records, each with 20 matches, 0 mismatches, and 12 possible matches. If the list were ordered by minimum mismatches, maximum matches, the correct match would be last. The order would be 99 records with 0 mismatches, 20 matches and 12 possibles, then the actual correct record with 1 mismatch and 31 matches.

This second method is essentially the same one a person uses when he defers consideration of records with mismatches as he sorts through a large database. The better method of ordering the results—keeping track of the number of correct matches for each record-record match—is arduous for a human investigator but trivial for a computer system.

Field Trial of the Computer-Assisted Postmortem Identification System (CAPMI)

A trial of the data-gathering methods and the comparison/selection algorithm was carried out at a U.S. Army post. The goals of this trial were to estimate the selectivity of the data from a simple dental examination when applied to a database of at least 500 individuals and to establish the feasibility of collecting a database simply for identification. A prime consideration was, "could the data be collected inexpensively in time and effort?"

Methods

Simple visual dental examinations were performed on 578 soldiers who were seen at troop dental clinics. Data were gathered by investigators who acted as dental clinic examination officers for the period of the trial. The data recorded included the soldier's serial number, (SSN) (only to ensure the elimination of duplicate records), age, sex, and a description of the restorative status of each of his teeth according to the method described in previous sections. After the records were checked to eliminate duplicates, the SSN was deleted and a coded record number was substituted. These examinations were then entered into "known" record files of the Computer Assisted Postmortem Identification System (CAPMI) in a Data General Eclipse MV8000 computer at the Letterman Army Institute of Research.

To test the ability of the system to match identities correctly under varying conditions, other "unknown" files were constructed with only certain groups of teeth included for use in the comparison/selection process:

all 32 teeth	only teeth 1 to 8 included
only teeth 1 to 16 included	only teeth 6 to 11 included
only teeth 17 to 32 included	only teeth 24 to 32 included

In addition, the effect of errors in the known database on the selectivity of the system was tested by forcing errors into the file output before sorting. Error rates of 10, 20, 30, and 40% were tested and the effects on the identification process were examined.

Results

A total of 578 records was used in the initial comparison runs. Although a maximum of 43 conditions could be recorded for any single tooth by using combinations of acceptable codes, the numbers of codes for any tooth never approached the maximum. In the original data collected, the maximum number of conditions (23) was recorded on the lower first and second molars, while the fewest conditions (10) were recorded on the maxillary centrals, laterals, and mandibular third molars. (See Table 1.)

See Table 1.

The most common conditions found on each tooth type are summarized in Table 2. The distribution of virgin, restored, and missing teeth, the complete frequency distribution, and

TABLE 1—*Possible codes.*

modfl	modf	mod	mo	m—mesial
	modl	mol	do	o—occlusal/incisal
	dofl	mof	of	d—distal
	mdfl	dol	ol	f—facial
	mofl	dof	md	l—lingual
		lof	fl	v—virgin
		mfl	ml	c—crown
		mld	df	x—missing
		dfl	dl	
		mfd	ml	

TABLE 2—*Most common tooth conditions recorded, %.*

	C	L	C	P_1	P_2	M_1	M_2	M_3
				MAXILLARY				
1	v/82.9	v/84	v/91.4	v/73.3	v/67.21	v/31.7	v/45.5	v/62.5
2	c/4.8	l/3.0	m/[a]	do/10.7	do/8.2	o/19.6	o/26.6	x/26.6
3	m/2.7	m/2.9	d/[a]	o/7.2	o/8.0	ol/14.6	ol/8.3	o/4.3
4	x/2.3	x/2.6	f/[a]	mod/2.6	mod/7.1	mo/6.8	x/3.4	...
	...	...	c/[a]	...	...	...	...	...
				MANDIBULAR				
1	v/96.7	v/97.5	v/96.7	v/83.6	v/71.2	v/32.1	v/39.4	v/65.4
2	c/0.7	m/0.6	f/0.9	o/7.2	o/10.1	o/17.6	o/28.4	x/26.1
3	x/0.7	c/0.3	c/0.5	do/4.1	do/9.2	of/13.3	of/8.6	o/4.5
4	m/0.5	m&d/0.3	m/0.3	m/1.6	mod/3.0	x/10.7	mo/5.7	of/1.8

[a]Tied at 1.3.

summary statistics are shown in Figs. 1, 2, and 3. Table 3 presents a summary of the incidence of virgin, restored, and missing teeth by tooth location. Cross-tabulations of certain combinations of teeth showed that many restorations considered as "common" provide amazingly selective identification points.

The rank distributions of the comparison run which included all 32 teeth from each unknown case are summarized in Fig. 4 as percentages of the total cases. There are 45 cases with no restorations or missing teeth (7.8%) and 8 cases with all 4 third molars removed but no other dental work (1.38%). If these cases are not considered, then in 94.47% (495/524) of the cases the correct identity match is listed first on the list of matches, and in 5.34% (28) the correct matches are second on the list.

The introduction of errors at different rates caused some decrease in the natural selectivity of the system. The resistance to errors is a function of the number of dental characteristics in

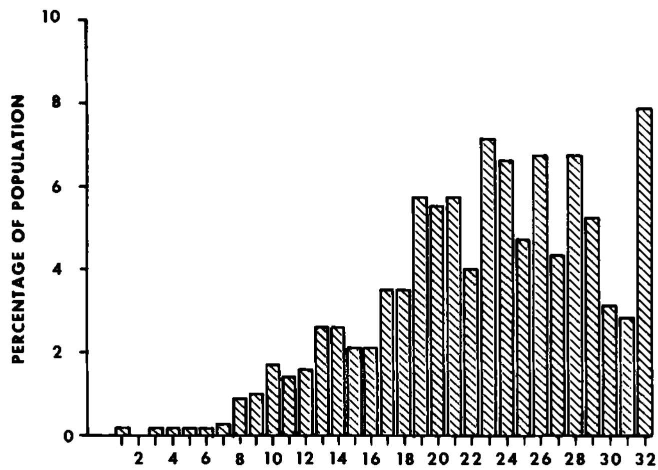

FIG. 1—*Number of virgin teeth in population of 578 soldiers.*

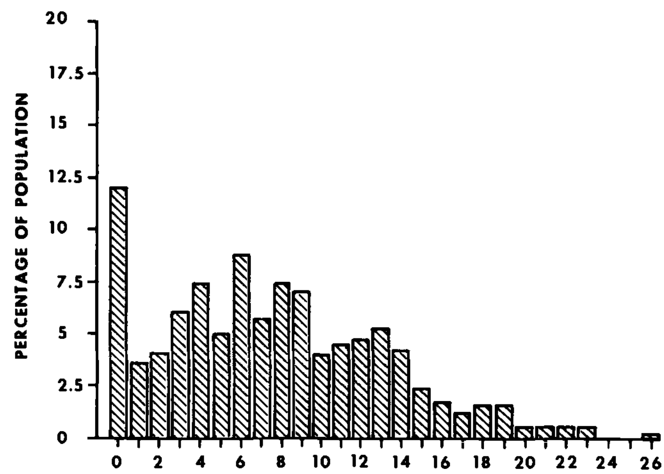

FIG. 2—*Number of restored teeth (n = 78). During comparison and selection routines, it is not necessary to record individual restorations; it is necessary to record only the surfaces restored.*

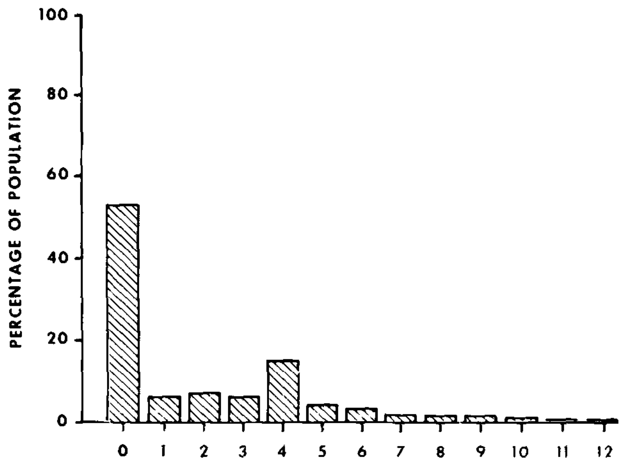

*FIG. 3—Number of missing teeth. This graph is skewed because of the non-normal distribution of the ages within the target population, age range 17 to 28 (*n *= 578).*

TABLE 3—*Percentage of virgin, restored, and missing teeth by tooth type:* n *(subjects) = 578.*

	C	L	C	P_1	P_2	M_1	M_3	M_3
				MAXILLARY				
Virgin	82.8	84	91.5	73.3	67.7	31.7	45.6	67.5
Restoration	14.9	13.4	7.2	24.2	29.8	62.8	51.8	5.9
Missing	2.3	2.6	1.2	2.5	2.4	5.5	3.4	26.6
				MANDIBULAR				
Virgin	96.7	97.5	97.0	83.6	71.2	32.1	39.4	65.4
Restoration	2.6	2.3	2.85	14.8	27.1	57.2	55.5	8.5
Missing	0.7	0.2	0.15	1.6	1.7	10.7	5.1	26.1

the known data record. The effects of various error rates upon the groups of records with different numbers of dental characteristics are shown in Fig. 5.

Each run was surveyed for the percentage of correct matches appearing in the first 10, 20, and 30 records. This percentage represents the chance of finding the correct match for an unknown file of a given number of dental characteristics if the error rate in the known database is estimable. If an unknown record had five or more dental characteristics, the chances of finding it in the top 5% of the sorted file were virtually 100% even with error rates up to 30% in the database.

Discussion

Three areas were evaluated in this project:

- Selectivity of dental characteristics—the performance of the dental characteristics as the primary selection mechanism. Did the simple dental records provide sufficient sensitivity and specificity?
- System efficiency—the efficiency and applicability of the algorithms used to perform

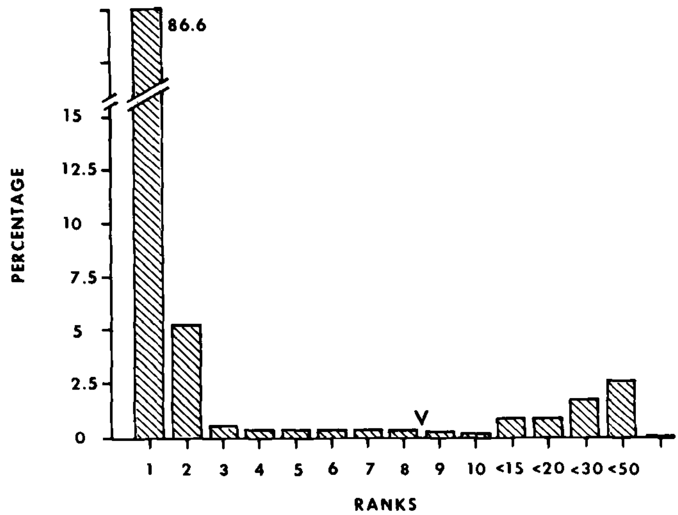

FIG. 4—*All 32 teeth used in the comparison/selection process. When all 32 teeth from the postmortem were unknown in the comparison/selection process, the correct identity match was selected first in 86.6% of the cases. The only cases whose correct matches occurred higher than tenth in the matches file had no restorations or missing teeth.*

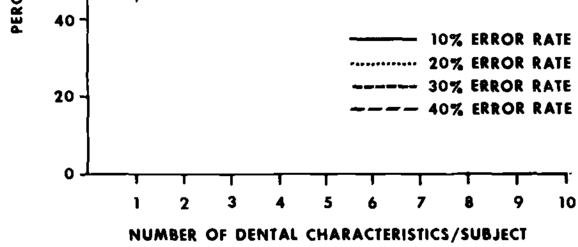

FIG. 5—*Errors were forced into the antemortem data to simulate recording mistakes in establishing the database. Error rates as high as 40% affect the identification of postmortem remains much less than intuition would expect. This graph can be understood as follows: If a database of 500 records had a recognized error rate in transcription of examinations of 20%, and the postmortem record had seven dental characteristics, the correct match for that record would appear in the top ten records (2%), about 96% of the time. Apparent inconsistencies in the behavior of the graph lines are a result of the randomized insertion of the errors in the database.*

the comparison and sorting processes within the computer programs. Did the program provide an efficient and effective aid to the verification process?

• System use—the "user friendliness" of the collection, comparison, and reporting mechanism including the impact of the required dental examination of the dental care system. Is the system, as conceived, feasible?

Selectivity of Dental Characteristics

This system was developed based on the hypothesis that dental characteristics provide enough information to select the best match identifications from a data set of any size. The system does allow the entry of a complete set of descriptors such as age and sex. For complete description see Appendix A. The primary intent of this research effort was to estimate the selectivity of dental characteristics alone. Selectivity can be understood in this context as the general ability to pick the correct identity match from a database. It follows that the selectivity is directly related to the number of dental characteristics in the unknown, the number of teeth available for comparison, and the size of the object database.

An estimate of the selectivity of dental characteristics can be derived from some simple examples. Approximately 92% of the population have one or more dental characteristics while about 82% have four or more. Thus the worst-case situation, where the unknown has no missing or restored teeth, will still eliminate 92% of the population from consideration.

Keiser-Neilsen [*2*] said that certain teeth such as the lower first molar are too commonly restored to be of any use for forensic science purposes. This statement may be accurate for verification purposes but not for the comparison/selection process. However, the data in Table 1 indicate that, although certain common conditions occur the majority of the time, many different types of restorations can occur on any teeth. When less common restorations occur, they are highly specific in their selective ability.

As an example of a worst-case situation: If a postmortem remains had only three usable teeth—the upper first premolar, and the upper first and second molars, and these three teeth had no fillings of any kind—this selection/sort would still eliminate 75% of the database from consideration. If any of these teeth had any type of restoration, the percentage of the population removed from consideration increases dramatically. Any information on the condition of the teeth reduces the considered population, because every tooth exhibits some variation in the manner of its restoration.

The more dental characteristics present in any dentition or dentition segment, the better the ability of the system to report the correct identity higher in the best-matches list.

The figures can be interpreted as follows. As shown in Fig. 4, when all 578 unknowns were compared to the known database, the correct match was first in 86% of the identification reports. Approximately 5% showed up second on their respective identification reports. As expected, the fewer the teeth used in the comparison, the lower the ability to report the correct match high on the list. In Fig. 4, the layer of high ranks is due to the proportion of subjects with no restorations. For example, where the comparison run was performed with records of only teeth 6 to 11, the low proportion of correct matches returned in the top 10 or 20 of the best matches list is due to the high proportion of cases with only virgin teeth and no other dental characteristics. If only cases with one or more dental characteristics are considered, then in all comparison/selection runs the specificity is vastly improved. If all 32 teeth are used and only cases with one or more dental characteristics are considered, then 97% are in the first or second places on the list and 99.4% are in the first ten records.

To assay the effect of errors in the "known" antemortem records or changes caused by dental work performed between antemortem and postmortem records, errors were introduced into the data files at rates of 10, 20, 30, and 40%. These errors were all seen by the program as "mismatches" and so were a worst-case situation. Figure 5 illustrates the effects of the successively higher error rates on "unknown" record matching. The erratic behavior of the lines at certain points is attributed to the randomness of the error application. The trends are obvious: Even with an error rate as high as 40%, the selectivity of the system is still great. More than 80% of the time the correct match is in the first ten records (approximately 2% of the database).

Conclusions Derived From the Clinical Trial

The CAPMI system is a quick and efficient method of selecting the most appropriate subset of records for use by a forensic science identification team. It has a high degree of selectivity even when the database is contaminated with 10 to 40% errors. The system can accommodate both situations—a mass disaster where the database must be constructed on site or the identification of an individual from a preexisting database.

The cost to collect the data or construct a database is small in comparison with costs to collect and sort records manually. Clinical trials have shown that an identification examination can be done at the same time that other procedures are performed, thus incurring no additional chair time.

The selectivity of the system is dependent on the number of dental characteristics in the unknown subject, the error content in the database of known records, and the size of the known database.

Implementation

The CAPMI system was initially developed on a Data General Eclipse MV8000 computer and has now been transferred to a 10-lb (4.5-kg) portable computer (GRiD Compass II 1129). This portable machine can handle and sort up to 900 records without the use of additional disk storage. The computer is run under the MS-DOS operating system; the application software is coded in Fortran 77 and 8088 assembler programming languages. It supports the system functions of data entry, comparison/selection, and reporting. The versatile power requirements (the computer can use 110 or 220 V, or any 12-V battery source) allow the portable to be used at the site of a mass disaster anywhere in the world. Comparison/selection can be made on data downloaded from a preexisting database, or a database can be constructed on site.

The system has been transferred to a desktop computer for use as a prototype local system in a network of data collection points.

APPENDIX A

Physical Characteristics

Physical characteristics that can be entered are height, weight, sex, eye color, hair color, race, and blood type. Appropriate codes are given either in on-screen menus or on the optical mark read coding sheets. An extensive array of distinguishing marks and physical characteristics is also represented by available codes.

Dental Characteristics

Dental characteristics that are used by the comparison mechanisms are limited to those certain codes describing the type and site of a dental restoration, and the presence or absence of the tooth.

Codes were limited in describing restored and missing teeth in order to adhere to certain rules for the classification, to simplify the decision logic, and eliminate, as much as possible, diagnostic *opinion*.

All of the diagnoses are to be made without benefit of radiographs.

To make the decision and coding process as simple and orderly as possible, the descriptors are limited to the description of the presence of the tooth and its restored status.

The codes—M, O, D, F, L, or I or any combination—were designated as natural identifiers on the Mesial, Occlusal, Distal, Facial, Lingual, and Incisal surfaces. No distinction was made between a tooth with one **MO** restoration and a tooth with an **M** restoration and an **O** restoration. Other codes were: C—any crown or crown preparation, X—missing and healed, V (or blank)—virgin, N—not chartable, and /—no information; this portion of the skull is not present.

APPENDIX B

Comparison of the Dental Records

There are three possible results for each tooth when the condition recorded in a "postmortem" record is compared with the condition recorded in an "antemortem" record. Possible results are:

Match—The condition of the tooth in the postmortem record is the same as that of the antemortem record.

Possible—The condition of the tooth in the postmortem record may have evolved from the condition recorded in the antemortem record.

Mismatch—The condition of the postmortem record is not the same as the antemortem record and the "possible" condition does not exist.

References

[*1*] Lorton, L. and Langley, W. H., "Postmortem Identification: A Computer-Assisted System. Research and System Design," U.S. Army Institute of Dental Research, WRAMC, Washington, DC.

[*2*] Keiser-Nielsen, S., *Person Identification by Means of the Teeth: A Practical Guide*, John Wright and Sons, Bristol, U.K., 1980.

Address requests for reprints or additional information to
Lewis Lorton, D.D.S., M.S.D.
U.S. Army Institute of Dental Research
USAIDR c/o KACH
Ft. Meade, MD 20755