Forensic Science International 340 (2022) 111463

Contents lists available at ScienceDirect

Forensic Science International

journal homepage: www.elsevier.com/locate/forsciint

Victim identification from the September 11, 2001 attack on the World Trade Center: Past trends and future projections

Bradley Adams[a,*], Julia Warnke-Sommer[b,c], Jennifer Odien[a], Angela Soler[a], Franklin Damann[c]

[a] New York City Office of Chief Medical Examiner, 520 First Avenue, New York, NY 10016, USA
[b] Illumina Inc., 5200 Illumina Way, San Diego, CA 92122, USA
[c] Defense POW/MIA Accounting Agency, 106 Peacekeeper Drive, Bldg. 301D, Offutt AFB, NE 68113, USA

ARTICLE INFO

Article history:
Received 19 July 2022
Accepted 8 September 2022
Available online 13 September 2022

Keywords:
World trade center
Victim identification
Terrorist attack
9/11
Good-toulmin estimator
Identification accumulation curve

ABSTRACT

Victim identification following mass fatality events is critically important. Extensive traumatic injuries and body fragmentation add complexity to this process. World Trade Center (WTC) identification efforts have been ongoing for over 20 years and this study tracks identification trends from the 2753 known WTC victims and the 21,905 recovered remains. For identified victims, data include the number of remains identified, date(s) of the identification(s), and identification modalities. Results show a heavy reliance on DNA due to body fragmentation. Other modalities played an important role initially, but DNA eventually became the singular identification modality. For large-scale disasters involving significant body fragmentation, aggressive DNA testing strategies are critical for victim identification. Over time, the number of linked remains (portions of previously identified individuals) will greatly outnumber the new identifications (first-time identifications). A novel approach using statistical modeling from ecology studies was applied to estimate future WTC identification rates using Identification Accumulation Curve extrapolation with the Good-Toulmin estimator. Projections indicate there will be 76 first-time identifications (95% CI: 49–117) through the successful DNA testing of 3404 unidentified, fragmentary remains. The remainder of the identifications would be additional portions of previously identified victims. These results may be instructional for management of other large-scale, protracted victim identification efforts.

1. Overview of the September 11, 2001 attack on the World Trade Center

On September 11, 2001, both of the 110-story World Trade Center (WTC) towers in New York City were attacked by terrorists. At 8:46 in the morning the North Tower (WTC 1) was struck by American Airlines (AA) Flight 11, which had departed earlier from Boston's Logan International Airport carrying 81 passengers (including five terrorists), two pilots, and nine flight attendants [1,2]. At 9:03 in the morning the South Tower (WTC 2) was struck by United Airlines (UA) Flight 175, which had also departed from Boston's Logan International Airport. It was carrying 56 passengers (including five terrorists), two pilots, and seven flight attendants [1,2]. The South Tower collapsed first at 9:59 and the North Tower collapsed shortly afterward at 10:28 [2].

The coordinated attack on the World Trade Center resulted in the death of 2753 known victims and the eventual recovery of 21,905 remains. As of the 20-year anniversary of the attack, 1647 known victims (59.8%) had been identified and 1106 were still unidentified (40.2%) (Table 1). From the recovered remains, 14,750 have been associated with a victim while 7155 have not yet been identified despite prior efforts (Table 1).

2. The role of the New York City Office of Chief Medical Examiner

The New York City Office of Chief Medical Examiner (NYC OCME) has been responsible for positive identification of the victims, issuance of death certificates, and repatriation of remains to victims' families [3,4]. The NYC OCME issued 2750 death certificates associated with the 9/11 attack on the World Trade Center. Three death certificates were issued outside of New York City to victims who died in health care facilities outside of city jurisdiction from injuries sustained that day. All victims received a death certificate regardless of their identification status. For victims who have never had remains identified, the death certificates were issued by the NYC OCME through judicial decree after case review by the Surrogate's Court in New York City [3]. All death certificates were issued with a

* Corresponding author.
E-mail address: badams@ocme.nyc.gov (B. Adams).

https://doi.org/10.1016/j.forsciint.2022.111463
0379-0738/

Table 1
Identification Statistics as of September 11, 2021[a].

Total Known Victims	2753
Identified Victims	1647 (59.8%)
Unidentified Victims	1106 (40.2%)
Total Remains Recovered	21,905
Remains Identified	14,750 (67.3%)
Remains Unidentified	7155 (32.7%)

[a] The reported values only reflect victims of 9/11 and do not include the terrorists.

manner of death of "homicide" and the cause of death for most victims was determined to be blunt force trauma [3]. For this disaster, it was determined that only external examinations would be conducted and no autopsies were performed [3,5].

Since 2001, the NYC OCME has been in direct communication with WTC families regarding victim identification developments, disposition of remains, and to answer any specific questions. Due to the protracted identification efforts, the NYC OCME developed a notification process to follow whenever remains are identified. For new identifications (i.e., first time remains have been identified to a specific victim), the family will usually be notified in person by their local law enforcement agency. If the newly identified remains are linked to a previously identified victim, the Notification Request Form is used to determine next steps. This form outlines family member requests to either be notified each time remains are linked or to never be notified again. For those families who choose to be notified of additional remains, notifications are usually handled by the NYC OCME over the phone. Families can change their notification wishes at any time, regardless of their original wishes.

3. Methods

The World Trade Center victim list used for this analysis includes individuals who died as a direct result of the impact by the planes and the subsequent collapse of the towers, excluding the terrorists. It does not include responders who may have died due to health complications associated with the protracted search and recovery efforts.

The data for this analysis were collected using two computerized databases, known as "WTC DataEase" and the "OCME WTC Integrated Case Management System." Additionally, every victim and all recovered human remains have their own separate hard-copy case folders. Information compiled for known victims includes identification status, employer, and last known location. For victims who have been identified, additional information was collected from the WTC DataEase program regarding the number of remains identified, date(s) of the identification(s), and identification modalities (e.g., DNA, dental, fingerprints).

The extracted data were subjected to additional statistical analyses to determine the rate of identifications over time and to generate an estimate of how many first-time identifications may be possible with continued DNA testing. Identification data for each tower and aircraft were evaluated using an Identification Accumulation Curve (IAC). The IAC is defined as the number of individuals identified $|I_{id}|$, $I_{id} \in I$ as a function of remains identified $|R_{id}|$, $R_{id} \in R$, where R is the remains set belonging to I individuals with $|R| \geq |I|$. It is analogous to the well-known Species Accumulation Curve (SAC) which was first applied in the ecology domain for biodiversity analysis [6] and more recently in large-scale biological experiments such as DNA sequencing applications [7,8]. The SAC is defined as the number of observed species as a function of sampling effort. The analysis of species accumulation curves can address questions important in ecology such as, "How many new species are likely to be discovered if the environment is continuously sampled?" and "What is the species richness, or the total number of distinct species for a given community?" In this study, the accumulation curve is applied to determine the rate of new WTC identifications (i.e., first time remains are identified to a victim) that can be expected through additional DNA testing and provides an estimate of the total number of individuals present in the set of recovered WTC remains.

To estimate the identification trends of the WTC remains, a rarefaction curve with a 95% confidence interval was created for the WTC site by randomly permuting the IAC 1000 times and taking the average of the generated curves. The Good-Toulmin estimator [9] was used to extrapolate the total number of new identifications that would occur through continued DNA testing. The classical non-parametric empirical Bayes Good-Toulmin estimator was implemented [9]. The Good-Toulmin estimator is given by:

$$U = -\sum_{i=1}^{\infty} \left(-\frac{m}{n}\right)^{i} \Phi_i$$

Where, for the purpose of this paper, m is the number of remains to be identified, n is the number of remains that have been identified, and Φ_i is the number of individuals that have i associated remains. The estimator estimates U, the number of new individuals identified as a function of m additional remains identified. The Good-Toulmin estimator is accurate for $\frac{m}{n} \leq 1$, approximating U within $\sqrt{n}\frac{m}{n}$ [10]. A 95% confidence interval was constructed for the Good-Toulmin estimate following the methods of Chao et al. [11] and by applying a

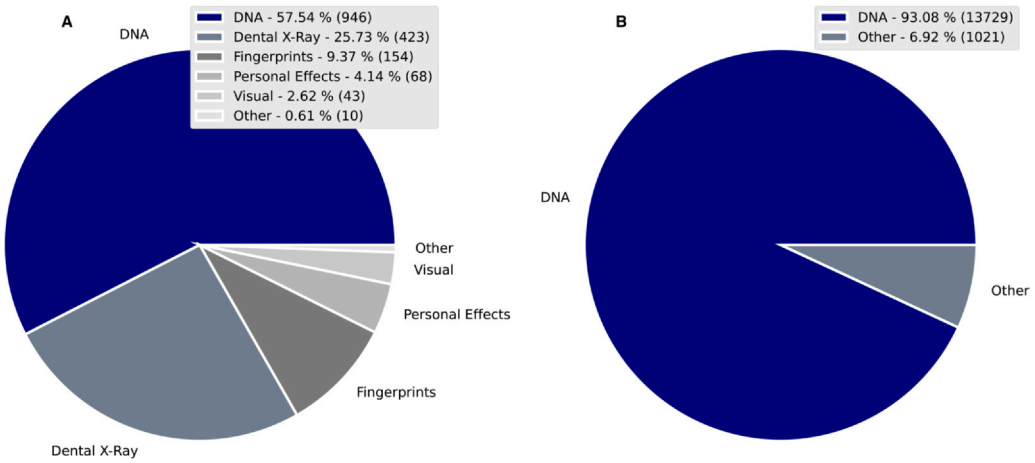

Fig. 1. Pie charts showing identification modalities. Chart (A) shows primary modalities used for the initial victim identifications (1644 individuals identified by NYC OCME). Chart (B) shows modalities used for all identified remains (14,750).

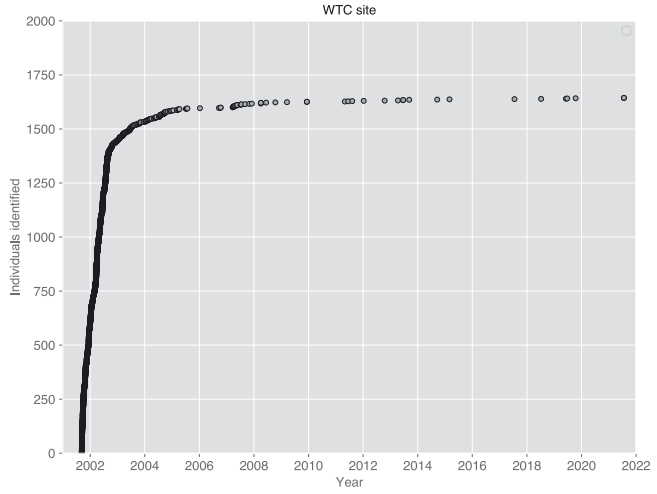

Fig. 2. Identification accumulation curve showing new identifications by year and all locations.

log-transformation procedure for better coverage [12]. These results can be used to determine the potential benefits of additional sampling and the likelihood of making new identifications. All methods for the IAC analysis and the Good-Toulmin estimator were implemented in Python (3.8) using Pandas [13], NumPy [14], and SciPy [15] packages.

4. Victim identification modalities

Soon after the attack, a collective policy decision was made to attempt identification on all recovered remains, regardless of size or condition. This type of mass fatality incident with nearly 22,000 recovered remains is then comparable to a scenario involving nearly 22,000 decedents since each fragment, regardless of size, must be considered to potentially come from a different individual and undergo identification efforts.

Different modalities have been utilized in the identification efforts of the WTC victims. Primary identification modalities include DNA, Dental, Fingerprints, Visual, and Personal Effects. Additionally, there is an "Other" category that includes identification methods that do not fit into the abovementioned modalities and may include anthropological reassociation of remains, contextual information, medical records, and distinctive tattoos.

Fig. 1 shows the modalities used for the initial (i.e., first-time) identification of each of the 1647 identified victims. It also shows the modalities used to identify each of the 14,750 remains, many of which are additional portions of previously identified victims. Although DNA has played a major role in the WTC victim identification, dental and fingerprints account for 35.10% of the initial victim identifications (Fig. 1A). In comparison, if the identification modalities are explored for every one of the 14,750 identified remains, DNA accounts for 93.08% of the identifications (Fig. 1B). Other modalities besides DNA were more prevalent during 2001 and 2002, but from 2003 onward 9018 out of the 9063 remains (99.50%) were identified based on DNA results. This clearly shows the importance of DNA in the identification of WTC remains where other identification modalities (e.g., fingerprints, dental, visual) are often not viable options. These findings are consistent with identification trends seen with other mass fatality incidents involving fragmentary remains [16].

Since it was acknowledged early on that the identification efforts would be prolonged, a decision was made by the NYC OCME to desiccate all the fragmentary remains. This drying process allowed for long-term storage of the remains without complex climate control measures or harsh chemicals. All remains were placed in individually labeled, vacuum-sealed packages after completion of the desiccation process. Through desiccation, the viability of DNA was also maintained which allows for continued resampling and testing.

DNA testing generally provides three possible identification outcomes: 1) A new identification (first time remains are identified to a victim); 2) Linked remains (additional remains linked to a previously identified victim); or 3) Insufficient DNA (reportable profile not

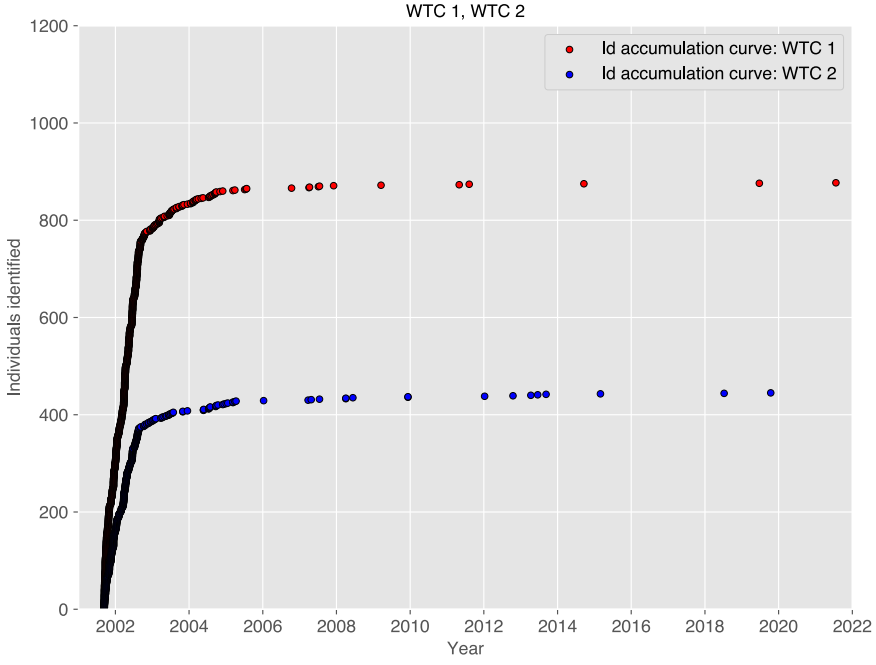

Fig. 3. Identification accumulation curves for the North (WTC 1) and South (WTC 2) Towers showing a similar identification trend between both locations.

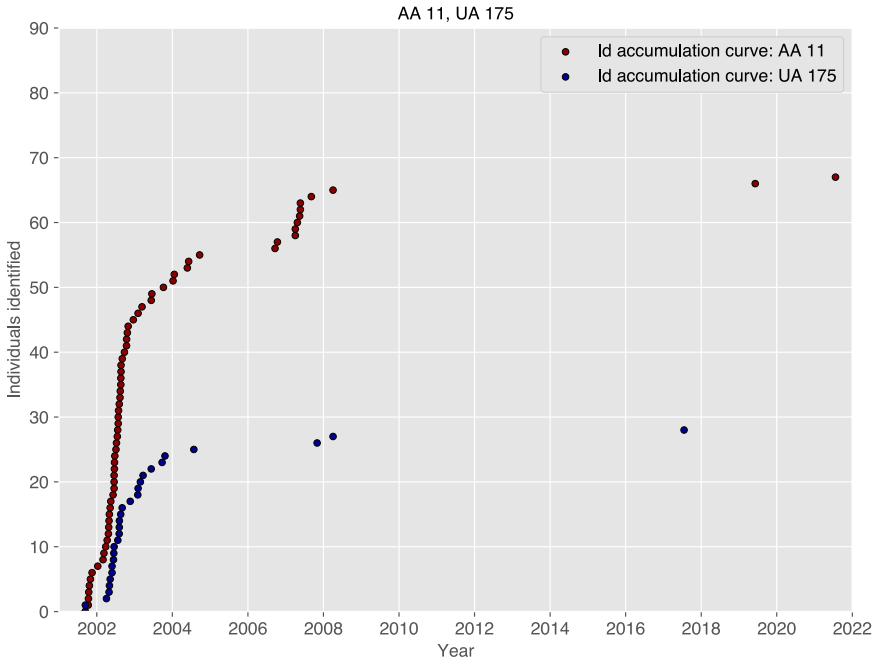

Fig. 4. Identification accumulation curve for AA Flight 11 and UA Flight 175. There is an increase of new identifications from AA Flight 11 in 2007 due to DNA testing associated with additional remains discovered around this time.

generated). The concepts of a "new identification" and "linked remains" are important distinctions as data are presented in this study.

5. Current identification trends

5.1. New victim identifications

As would be expected in any mass fatality incident, new identifications (i.e., first time remains are identified to a victim) occurred regularly in the days, weeks, and months immediately following the 9/11 attack. As identification efforts progressed over many years, new identifications became far less frequent and the subsequent identification of remains to previously identified individuals became the dominant trend. The number of new identifications was highest between 2001 and 2003 and by 2005 the frequency of new identifications had plateaued with most DNA results linking additional remains to previously identified individuals (Fig. 2). In the sixteen-month period from the day of the attack until the end of 2002 there were 1446 new identifications, representing 88% of all new identifications made during the entire 20-year study period. In

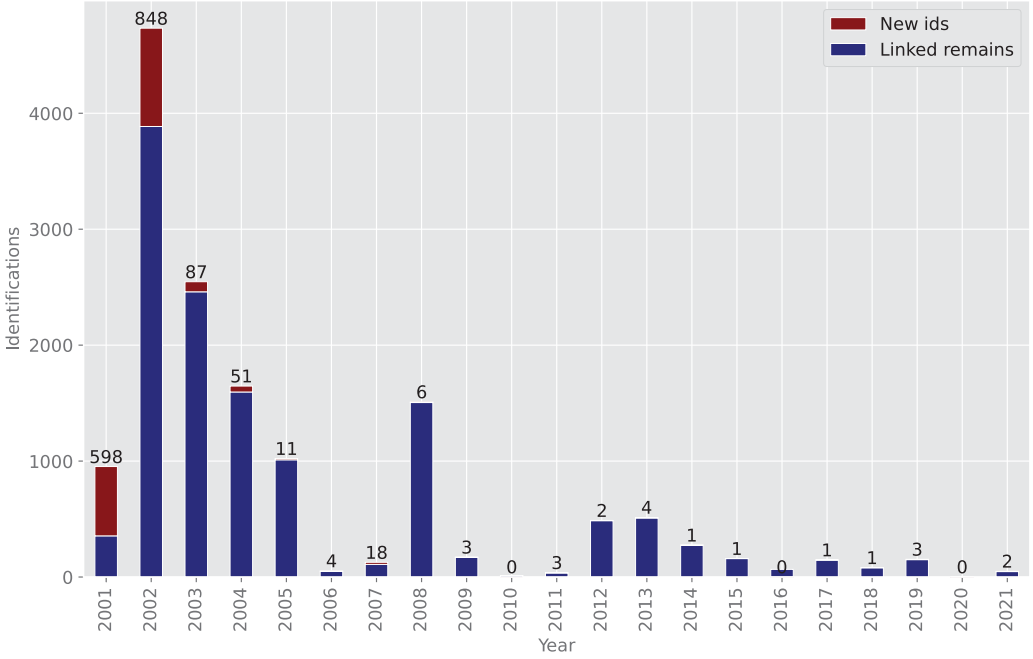

Fig. 5. Ratio of new identifications (in red and numbers at top of columns) and linked remains (blue columns). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

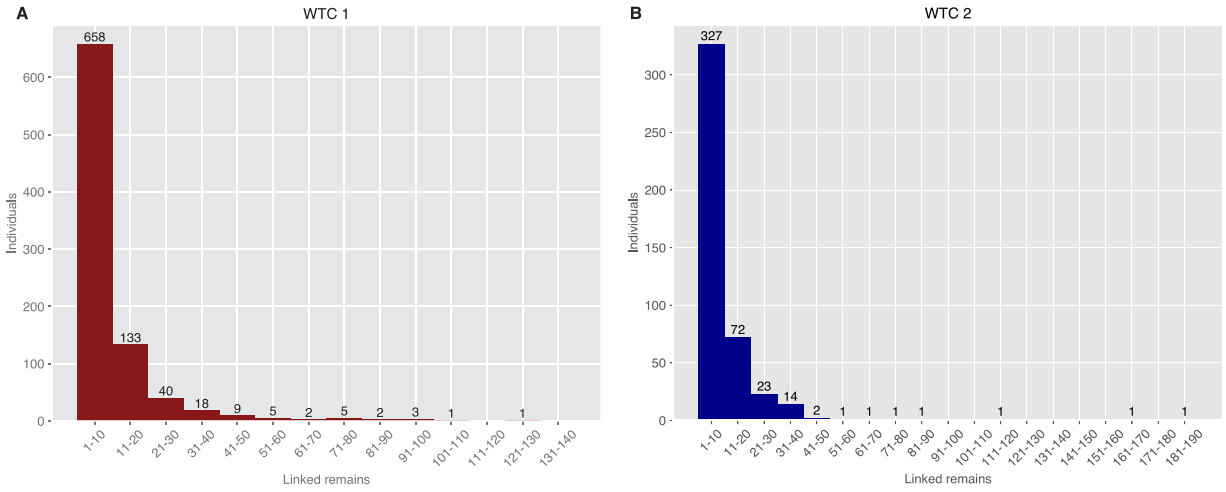

Fig. 6. Linked remains associated with North Tower (A) and South Tower (B) victims.

comparison, from January 1, 2006 to September 11, 2021, there were 3798 links to previously identified individuals and 49 new identifications.

Fig. 3 compares the rate of new victim identifications in the North and South Towers. Both towers exhibit similar trends with a steep initial climb in the early years and then a plateau. Comparison of the rate of new identifications between the airplane victims shows more variation (Fig. 4), with a spike occurring in the rate of new identifications associated with AA Flight 11 in 2007. This is related to the unexpected discovery around that time of hundreds of small bone fragments on the rooftop of a building, called Deutsche Bank, located immediately south of the twin towers [5,17]. Reportable DNA profiles were developed for 325 remains from the Deutsche Bank roof and all were associated with AA Flight 11 victims, resulting in 6 new victim identifications.

Fig. 5 shows that the number of linked remains has exceeded the number of new identifications in every year except 2001 (September to December). Again, based on the high fragmentation, this is an expected outcome, especially with an aggressive DNA testing strategy. The spike in new identifications and linked remains in 2006 and 2007 is associated with renewed search and recovery efforts that resulted in the discovery of additional remains on the Deutsche Bank roof and other locations around the WTC site during that period [5,17].

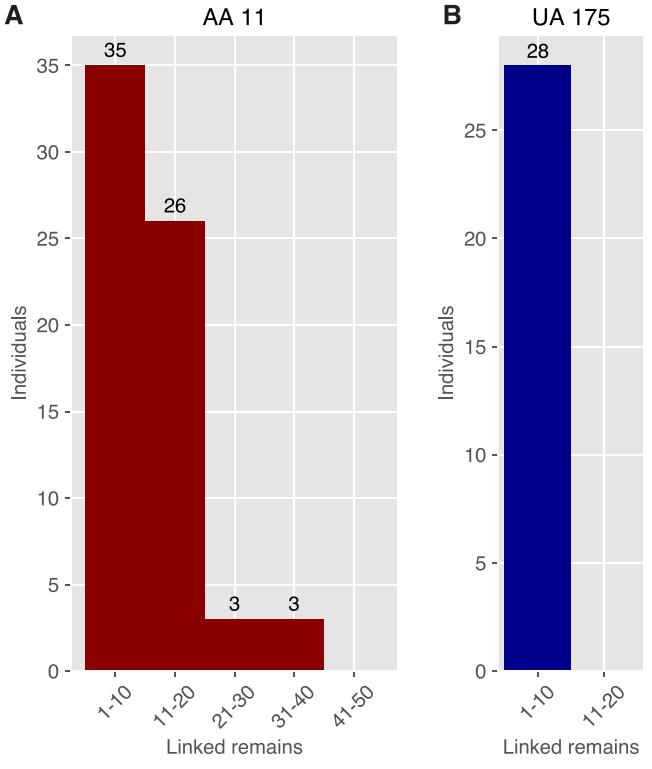

Fig. 7. Linked remains associated with AA Flight 11 (A) and UA Flight 175 (B).

5.2. Linked remains

The frequent identification of multiple remains to the same victim (i.e., linked remains) was explored to see if there are trends based on incident location (Figs. 6 and 7). For identified victims associated with the North Tower, the majority of individuals (75.03%) have between 1 and 10 linked remains (Fig. 6). There are 19 identified individuals (2.17%) associated with the North Tower who have over 50 linked remains. For identified victims associated with the South Tower, most individuals (73.48%) have between 1 and 10 remains linked to them (Fig. 6). There are 7 identified individuals (1.57%) associated with the South Tower with over 50 linked remains. The data show that the relationship between the number of linked remains and the number of victims is very similar between the two towers. Of note, the identification of a large number of fragmentary remains to a single victim does not necessarily equate to a higher degree of overall body completeness.

For identified victims associated with AA Flight 11, the majority of individuals (91.04%) have between 1 and 20 remains linked to them (Fig. 7). There are no identified individuals with over 50 linked remains associated with AA Flight 11. All of the identified victims (100%) associated with UA Flight 175 have between 1 and 10 linked remains (Fig. 7). The data show that the number of linked remains to victims from the two airlines follows a different pattern. A portion of this discrepancy can be attributed to the additional remains discovered on the rooftop of the Deutsche Bank building, since 325 bone fragments were associated by DNA to 54 different victims from AA Flight 11. If the Deutsche Bank remains were excluded from the counts, most of the AA Flight 11 victims would have between 1 and 10 remains and none would have more than 18 remains, showing the effect of the isolated discovery on the rooftop.

Table 2
Monte Carlo Cross-Validation.

	50% subsampling	60% subsampling	70% subsampling	80% subsampling	90% subsampling
Mean	1647.1	1646.6	1646.6	1647.1	1647.2
RMSE	40.1	20.4	14.4	10.2	6.6

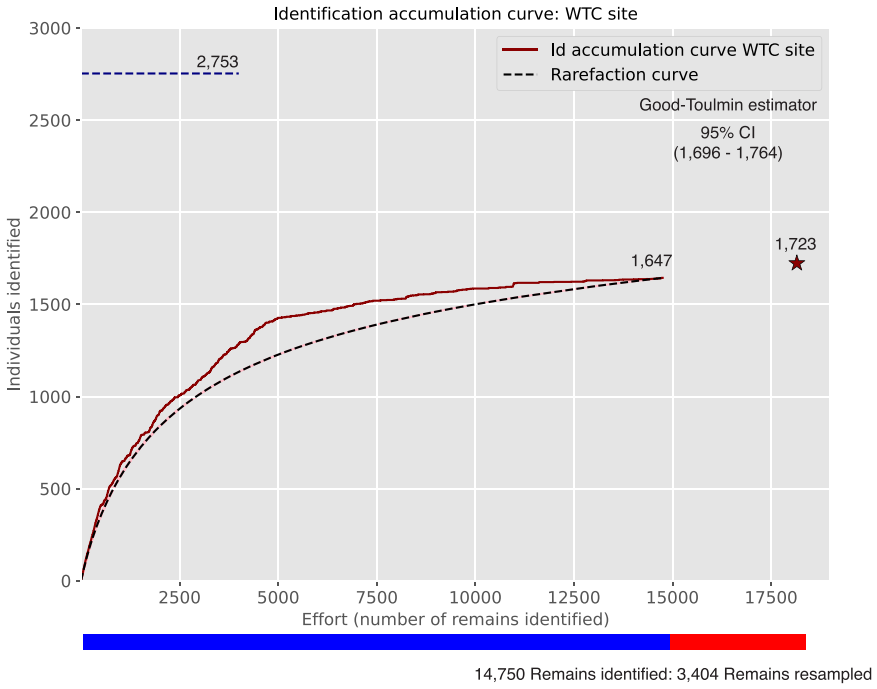

Fig. 8. Identification accumulation curve (*c.f.* species accumulation curve) and the Good-Toulmin Estimator showing projected new WTC identifications based on additional DNA testing,

6. Future identification projections

The objective of disaster victim identification is to identify the set of individuals *I* represented by a set of recovered remains *R*, where |R| ≥ |I|. The total number of individuals represented by a set of remains is challenging to determine due to the nature of disaster events. The set of recovered remains may not include all victims of a disaster event and/or the total number of disaster victims itself may be uncertain. The completion of an identification project could take years depending on the disaster scale, condition of the recovered remains, availability of DNA references, agency resources, and policy decisions. For large-scale identification projects, it is beneficial to establish criteria for determining end points early in the identification process [16,18]. An estimate of the number of individuals present in the recovered remains set is useful for guiding stakeholder decisions as well as for providing resolution for large-scale identification efforts spanning many years.

As DNA testing becomes more sensitive, WTC remains are periodically retested if they did not previously produce a viable profile. Due to contamination and degradation issues associated with soft tissue, only remains with bone are currently eligible for ongoing DNA testing efforts. As of September 11, 2021, there are 3404 remains deemed eligible for bone retesting efforts.

To estimate the number of new WTC identifications that could occur through continued DNA testing, the Good-Toulmin estimator was utilized. First, a cross-validation study was conducted to evaluate the estimator following the Monte-Carlo cross-validation method. The set of 14,750 identified remains representing 1647 individuals was subsampled 1000 times without replacement at 90%, 80%, 70%, 60%, and 50%. The Good-Toulmin estimator was applied to each subsample to generate an estimate of the total individuals for the identified remains set. Table 2 shows the mean estimate at each subsampling level along with the root mean square error (RMSE). The results show excellent correspondence with the actual number of individuals. Expectedly, the RMSE decreases as the subsampling coverage increases.

Utilizing the Good-Toulmin estimator with n=14,750 (identified remains) and n=3404 (unidentified remains suitable for DNA retesting), it is projected that there are a total of 1723 victims (95% CI: 1696 - 1764) represented within this WTC data set of 18,154 remains. The estimate and log-transformed CI were rounded to the nearest integer number. The IAC, rarefaction curve, and associated Good-Toulmin Estimate can be found in Fig. 8. The results of the statistical modeling suggest that successful DNA testing of the remaining 3404 unidentified remains would result in 76 new identifications (95% CI: 49–117). The remainder of the DNA associations are expected to be linked to previously identified individuals. These results suggest that while the NYC OCME will continue to make new identifications through ongoing DNA testing, over a thousand of the known WTC victims may never have remains identified to them. This is likely due to body fragmentation, burning, and the complex recovery operations.

7. Conclusions

In summary, the WTC data collected by the NYC OCME over a 20-year period provides valuable information about the victim identification process in a large-scale, highly fragmented, and complex mass fatality incident. The results show that there will be a heavy reliance on DNA for victim identification in this type of incident. As

expected, there is an initial spike in new identifications but as time progresses and DNA testing continues, it is far more frequent for fragmentary remains to be linked to previously identified individuals. Identification rates for victims of both towers follow very similar trends based on the identification accumulation curves. The identification accumulation curves for the airplanes showed different trends, primarily due to the unexpected discovery of a large number of AA Flight 11 remains on a rooftop.

Although it is infrequent that new identifications are still made, the NYC OCME continues to use new DNA technology to retest WTC samples that have not previously produced reportable DNA profiles. The NYC OCME has made a commitment to the families of the WTC victims to continue identification efforts in hopes of identifying as many of the 1106 known victims who, to date, have no remains identified to them. Statistical modeling using methods derived from species richness modeling in ecology studies suggests that an additional 76 WTC victims (95% CI: 49–117) may have first-time identifications through continued DNA testing efforts of the unidentified remains.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Acknowledgements

Jason Graham, Christopher Rainwater, Mary Fetchet, Amy Mundorff, Mark Desire, Carl Gajewski, Andrew Schweighardt, and Justin Goldstein provided input and assistance in the development of this manuscript.

References

[1] A.Z. Mundorff, B.J. Adams, J.E. Byrd (Eds.), *Anthropologist Directed Triage: Three Distinct Mass Fatality Events Involving Fragmentation and Commingling of Human Remains*, in *Commingled Human Remains: Methods in Recovery, Analysis, and Identification*, Academic Press, San Diego, 2014, pp. 365–388.

[2] National Commission on Terrorists Attacks Upon the United States, *The 9/11 Commission Report: Final Report of the* National Commission on Terrorist Attacks Upon the United States. 2004, Washington, DC: Government Printing Office.

[3] J.R. Gill, 9/11 and the New York City Office of Chief Medical Examiner, Forensic Sci. Med Pathol. 2 (1) (2006) 29–32.

[4] G. Mackinnon, A.Z. Mundorff, T.J.U. Thompson, S.M. Black (Eds.), *World Trade Center - September 11th, 2001*, in *Forensic Human Identification: An Introduction*, CRC Press, Boca Raton, 2006, pp. 485–499.

[5] J.R. Gill, et al., E.E. Turk (Ed.), *The 9/11 Attacks: The Medicolegal Investigation of the World Trade Center Fatalities*, in *Forensic Pathology Reviews*, Humana Press, New York, 2011, pp. 181–196.

[6] R.A. Fisher, A.S. Corbet, C.B. Williams, The relation between the number of species and the number of individuals in a random sample of an animal population, J. Anim. Ecol. 12 (1) (1943) 42–58.

[7] C. Deng, T. Daley, A.D. Smith, Applications of species accumulation curves in large-scale biological data analysis, Quant. Biol. 3 (3) (2015) 135–144.

[8] C. Deng, et al., Molecular heterogeneity in large-scale biological data: techniques and applications, Annu Rev. Biomed. Data Sci. 2 (2019) 39–67.

[9] I.J. Good, G.H. Toulmin, The number of new species, and the increase in population coverage, when a sample is increased, Biometrika 43 (1–2) (1956) 45–63.

[10] A. Orlitsky, A.T. Suresh, Y. Wu, Optimal prediction of the number of unseen species, Proc Natl Acad Sci U S A 113 (47) (2016) 13283–13288.

[11] Chao, A., et al., *Appendix G. A bootstrap method to construct an unconditional variance estimator for any interpolated or extrapolated estimator. Wiley. Dataset.* ⟨https://doi.org/10.6084/m9.figshare.3568407.v1⟩. 2016.

[12] A. Chao, Estimating the population size for capture-recapture data with unequal catchability, Biometrics 43 (4) (1987) 783–791.

[13] McKinney, W., *Data structures for statistical computing in python*, in Proceedings of the 9th Python in Science Conference. 2010. p. 56–61.

[14] C.R. Harris, et al., Array programming with NumPy, Nature 585 (7825) (2020) 357–362.

[15] P. Virtanen, et al., SciPy 1.0: fundamental algorithms for scientific computing in Python, Nat. Methods 17 (3) (2020) 261–272.

[16] E.J. Kontanis, P. Sledzik, *Resolving Commingling Issues During the Medicolegal Investigation of Mass Fatality Incidents*, in *Commingled Human Remains: Methods in Recovery*, in: B.J. Adams, J.E. Byrd (Eds.), Analysis, and Identification, Academic Press, San Diego, 2014, pp. 447–468.

[17] P.S. Sledzik, et al., D.W. Steadman (Ed.), *Disaster Victim Recover and Identification: Forensic Anthropology in the Aftermath of September 11*, in *Hard Evidence: Case Studies in Forensic Anthropology*, Prentice Hall, Upper Saddle River, NJ, 2009.

[18] L.G. Biesecker, et al., Epidemiology. DNA identifications after the 9/11 world trade center attack, Science 310 (5751) (2005) 1122–1123.