# Examining Statistical Differences Between Standard Osteological Measurements Taken in Situ versus a Laboratory Setting: Preliminary Results

Johanna Kelly, M.Sc., Crystal Forrest, PhD., & Alexis Dunlop, M.Sc

## Abstract

Legislative parameters governing bioarchaeological projects undertaken by cultural resource management (CRM) companies often dictate the type of analysis conducted. In situations where analysis cannot be executed in a laboratory setting due to policy restrictions or reasons of expediency, researchers turn to conducting analysis in the field. This study aims to determine if there is a statistically significant rate of interobserver error between lab and *in situ* measurements.

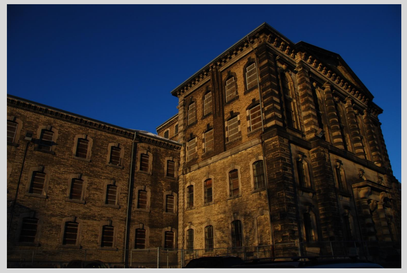
Standard osteological measurements from 15 individuals from the Old Don Jail in Toronto, Ontario, Canada, were taken by three researchers. The remains were observed *in situ* 12 months prior to being analyzed in a laboratory setting. A paired t-test was performed to determine if there was significant difference between the two sets of measurements. The mean difference (M=0.7182, SD=4.9712, N=719) was statistically significant (t(718)=4.2108, two tail p=0.000028), suggesting that there are significant differences between measurements taken in the field and those taken in the laboratory. The results of this study are important given that bioarchaeologists strive for high accuracy, precision and replicability in the field and in the lab.

## Introduction

The Toronto Jail (known locally as the Old Don Jail) is located in the east end of Toronto, Ontario. In active use for over one hundred years starting in 1862, the jail was the site of 34 executions by hanging. Of those individuals, 15 men were buried on jail property. After the closing of the old section of the jail in 1977 and the demolition of the yard walls, the locations of the 15 burials were forgotten. With the aid of a newly discovered map dating from the 1950s identifying the East Exercise Yard as the prison cemetery, an exploratory excavation was undertaken in 2007 by Archaeological Services Inc. (ASI) to re-locate the burials and define the extent of the cemetery. Historic period documents also provided information about the identities of the individuals who has been buried at the jail and not in family plots at other cemeteries.

The burials of all 15 individuals were investigated in accordance with the Ontario Cemeteries Act, and measurements were taken in situ in order to match physical characteristics with historic descriptions of the hanged men.

All 15 individuals were exhumed in 2008 under the guidance and permission of the Ontario Registrar of Cemeteries. The remains were then subjected to analysis by ASI staff in a laboratory in conjunction with researchers from the University of Toronto. This work duplicated the measurements collected in the 2007 field season.

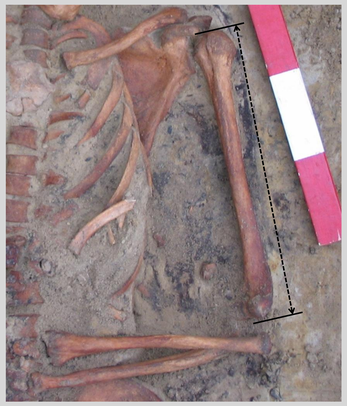


The Old Don Jail, located in Toronto, Ontario, Canada.

## Methods

Each skeleton was measured twice, both in field and laboratory contexts. Elements were measured according to standards provided in Buikstra and Ubelaker (1994). The first set of measurements done in the field was completed under the stipulation that the remains be minimally moved, if at all. This was done because decisions had not yet been made as to whether the burials would be allowed to remain in their original location, or if they would be moved to an alternate final resting place.

Since little movement of remains was optimal, it was not possible to use an osteometric board to measure remains, although sliding and spreading calipers as well as soft tape were used. In order to obtain measurements that were equivalent to using an osteometric board, it was necessary to "pedestal" the individual elements so that measurements could be taken. After detailed information regarding the archaeological features of the burials was collected, the soil surrounding each element was removed so that the element remained on a small, raised pedestal allowing for maximum exposure of each element for measurement. For measurements that are typically done using an osteometric board, the ends of the elements were delimited using conventional rulers, which were arranged in a parallel fashion. Measurements between the ends were obtained by running a soft tape between these rulers.



Example of pedestalled long bones of the arm; dirt was cleared from proximal and distal end to allow rulers to be placed in proper position (solid black line), the anatomical length was then recorded from ruler to ruler (dotted black line)

## Methods continued

In order to facilitate the collection of measurements of circumference of long bones, small tunnels were dug under the remains so that instruments could be used to obtain these types of measurements.

Measuring circumference in the field with a soft tape

All remains at the site were interred in a supine position, which meant that certain measurements (especially remains located dorsally) were impossible without moving the remains. In these cases, these measurements were simply not taken.

Laboratory measurements were completed using conventional tools, and were executed according to instructions provided in Buikstra and Ubelaker (1994).

Examples of equipment used during lab analysis;
Left: digital calipers, above: osteometric board

## Analysis

In the interests of accuracy and replicability, measurements were excluded if they failed to satisfy either of the following two criteria:

1. for any one standard measurement there must be at least two values: one recorded during the field analysis and one recorded during the laboratory analysis;
2. and each value must be a precise value recorded to two decimal points.

In certain instances (problematic positioning of an element, damage, or pathology) measurements were marked with an asterisk to indicate an estimated value. Such values were excluded.

A paired t-test was used to compare the sets of field and lab measurements, where the null hypothesis postulated that the difference between measurements would be zero.

## Results

The results of the paired t-test showed that there were statistically significant differences between measurements taken in the field *in situ* compared to those taken in a laboratory setting (p<0.001) (Tables 1 and 2).

| | *Variable 1* | *Variable 2* |
|---|---|---|
| Mean | 76.14030598 | 75.35909597 |
| Variance | 8354.000247 | 8386.837503 |
| Observations | 719 | 719 |
| Pearson Correlation | 0.998523663 | |
| Hypothesized Mean Difference | 0 | |
| df | 718 | |
| t Stat | 4.210838205 | |
| P(T<=t) one-tail | 1.43383E-05 | |
| t Critical one-tail | 1.646978626 | |
| P(T<=t) two-tail | 2.86765E-05 | |
| t Critical two-tail | 1.963273425 | |

*Table 1. Results of t-test*

| | |
|---|---|
| Mean | 0.781210014 |
| Standard Error | 0.185265785 |
| Median | 0.5 |
| Mode | 2 |
| Standard Deviation | 4.971202669 |
| Sample Variance | 24.71285598 |
| Kurtosis | 32.31253858 |
| Skewness | -2.707694377 |
| Range | 81.7 |
| Minimum | -58 |
| Maximum | 23.7 |
| Sum | 562.47121 |
| Count | 720 |
| Confidence Level(95.0%) | 0.363726538 |

*Table 2. Summary statistics*

The mean difference between paired measurements (M=0.78, SD=4.97, N=719) is significantly greater than zero, suggesting that a substantial degree of error is introduced between the two settings. The probability of this result being due to chance can be stated as 0.000028, which is a statistically significant result as it is less than the commonly accepted cutoff of 0.05. A 95% confidence interval about the mean for difference in measurements is (0.42, 1.05).

## Discussion and Conclusions

The observed statistically significant differences in field and laboratory measurements is in many ways unsurprising. In the field, the inability to manipulate remains into optimal positioning for measuring, the adherence of dirt and debris to the bone, and the inability to use conventional tools for some measurements introduces a substantial degree of error. Despite the best efforts to ameliorate these problems, they had a negative effect on the accuracy of the results, and certainly affected replicability in this instance.

Documentation of archaeological human remains in a field context is becoming more common in Ontario as policy is shaped by the increasing desire to respect the wishes of descendant groups. This is particularly true in instances in which Aboriginal remains are encountered. Bioarchaeologists attempt to balance scientific credibility with cultural relativist approaches to the investigation of human remains, and at times this is difficult to accomplish. The results of this study suggest that in cases in which remains are studied *in situ*, the precision, accuracy, and replicability of the data that are collected may be compromised. Since the collection of osteological data in the field is increasingly common, innovative methods are needed in order to assure the quality of the data.