Using the Capture-Recapture Technique to Estimate the Ascertainment-Corrected Incidence of Musculoskeletal Injuries During Marine Corps Recruit Training

Mita Lovalekar, MBBS, PhD, MPH*; Karen A. Keenan, PhD, ATC†; Debora E. Cruz, MSc*; Bridget A. McFadden, PhD‡; Sidra Montgomery, PhD§; Shawn M. Arent, PhD, CSCS*D, FACSM, FISSN, FNAK©†; COL Bradley C. Nindl, MSC, USAR, PhD, FACSM*

ABSTRACT

Introduction: Injury epidemiology research with military populations typically utilizes data obtained through medical chart review (MCR) or injury self-reports (ISRs). MCR data will not capture musculoskeletal injury (MSKI) data for which medical care was not sought, which is common during military recruit training. Injury self-report is affected by issues with recall, especially for MSKIs perceived as less severe. U.S. Marine Corps (USMC) recruits participate in an intense 13-week recruit training program during which they are susceptible to MSKIs. The purpose of the current analysis was to utilize a novel statistical method, the capture-recapture (CRC) technique, to account for the undercounting inherent in MSKI data sources and estimate the ascertainment-corrected cumulative incidence of MSKIs during USMC recruit training.

Materials and Methods: Data for the current study were derived from a larger study, the USMC Gender-Integrated Recruit Training Study, which was initiated to provide data-driven recommendations to increase gender integration in USMC recruit training. The estimated cumulative incidence of MSKIs during 13-weeks of USMC recruit training was calculated from the 2 sources of MSKI data (MCR, ISR) and using CRC analysis. Medical charts were reviewed to extract data about MSKIs that occurred during recruit training. Self-reported MSKI data for the same period were obtained from recruits at the end of recruit training. MSKIs were classified according to their anatomical location and type. The Chapman modification of the Lincoln-Peterson estimator was utilized to conduct the CRC analysis.

Results: Medical chart review and ISR MSKI data were available for 464 USMC recruits (age: 19.1 ± 1.9 years; gender: men 70.0%). The observed 13-week cumulative incidence of MSKI in the sample was 21.8% in the MCR and 28.4% in the ISR, while the CRC incidence was much higher (62.0%). The MCR and ISR ascertainment were 35.1% and 45.9%, respectively, while the overall ascertainment or completeness of MSKI data when 2 sources were used was moderate (65.0%). When stratified by MSKI anatomical location, the overall ascertainment varied by anatomical location of the MSKI. It was highest for lower extremity MSKIs (64.8%), but lower for upper extremity (38.9%) and spine (33.3%) MSKIs. The overall ascertainment also varied by MSKI type; it was highest for sprain (55.1%), followed by strain (54.8%), and the pain/spasm/ache (43.3%).

Conclusions: This was the first study to utilize the CRC technique to access the ascertainment-corrected incidence of MSKIs among USMC recruits. There was significant undercounting in both sources of the data analyzed, and the extent of undercounting varied by both MSKI anatomical location and type. When 2 sources of data were utilized simultaneously, the percent of CRC-estimated MSKIs observed from 2 sources of data was more complete. There is a need for further application of the CRC technique to MSKI data in military populations to provide a more complete assessment of MSKIs. Identification of modifiable factors that influence completeness of MSKI data obtained during military recruit training is also warranted.

INTRODUCTION

Musculoskeletal injuries (MSKIs) occur frequently in military populations, especially during training. Musculoskeletal injuries are a common cause of pain, disability, morbidity, high financial cost, and loss of tactical readiness in military populations.1–3 Epidemiological studies of MSKIs require the...
availability of valid data to describe the burden of MSKIs, identify risk factors, and design MSKI prevention programs. Studies of MSKIs in military populations utilize various sources of data, including injury self-reports (ISRs), medical chart review (MCR), or data from centralized databases. Each of these sources of data has its own advantages and disadvantages.

Injury self-reports are cost-efficient to obtain, are expected to contain information about MSKIs for which medical care was not sought, and may contain information that might not be collected as a part of routine medical care (e.g., contact vs. non-contact mechanism of injury, lost duty days). Even though ISR data are easy to obtain, their validity may be adversely impacted by issues with recall, especially for older and less severe MSKIs. This can lead to incompleteness or “under-ascertainment” of MSKI data collected via ISRs. Medical charts are expected to contain valid data since they are written by health professionals; however, MCR data may also be incomplete as there may be a reluctance to seek medical care, especially in military populations, to avoid profile and/or due to fear of reprisal for seeking services.

To adjust for the inherent under-ascertainment in these individual sources, MSKI data from multiple sources can be combined using the process of capture-recapture (CRC) to estimate an ascertainment-corrected incidence of MSKIs as well as to assess and compare the amount of underreporting from individual sources of MSKI data. The CRC technique was first utilized in ecology to estimate the population size of animals by a process of repeated sampling of tagged animals. This method was subsequently adopted by epidemiologists to study the validity and completeness of diverse health-related data. To date, only 1 study has utilized the CRC technique for MSKI data in a military population, specifically active duty Army soldiers; however, the CRC technique has not yet been utilized in military recruits. Recruit training occurs in a unique milieu, and there may be a greater reluctance to seek medical care among recruits to avoid profile and to graduate on time. Having a more accurate and complete assessment of MSKI data during recruit training is essential. An accurate estimate of the occurrence of MSKIs is needed to understand the true burden, assess causes, and design targeted interventions aimed at mitigating MSKIs that are associated with high incidence and cost.

U.S. Marine Corps (USMC) recruits undergo an intensive 13-week recruit training program designed to transform civilians into basically trained Marines, during which they are susceptible to MSKIs. The purpose of this analysis was to utilize 2 sources of MSKI data (MCR and ISR) to analyze the completeness of data collected from each of the 2 sources, and to estimate the ascertainment-corrected cumulative incidence of MSKIs during USMC recruit training. It was hypothesized that both sources of MSKI data would lead to undercounting of MSKIs and the CRC estimate of MSKI cumulative incidence would be much higher than the observed MSKI incidence.

METHODS
Musculoskeletal injury data included in the CRC analysis were derived from a larger research study, the USMC Gender-Integrated Recruit Training Study, which was initiated to provide data-driven recommendations for alternate training models to increase gender integration during USMC recruit training.

Study Participants and Human Subjects’ Approval
The research team provided an overview of the study to USMC recruits during group briefs and invited them to participate. Recruitment of study participants occurred in the presence of an ombudsman to minimize risk of coercion and all the participants provided written informed consent before data collection. Higher-ranking Marines were not present during the briefing and consent process to prevent undue influence. All recruits who were medically cleared to participate in recruit training were eligible to participate in this study. Human subjects’ approval for the study was obtained from the University of Pittsburgh Institutional Review Board (STUDY20120069: sIRB USMC Gender-Integrated Recruit Training Study) and was endorsed by the USMC Human Research Protections Official.

Each participant was assigned a unique identification number, and MCR and ISR MSKI data for each participant were linked using each participant’s identification number. The MCR data were available for 584 recruits, and ISR data were available for 467 recruits. A total of 464 recruits had both MCR and ISR data and were included in the CRC analysis.

MSKI Operational Definition
An MSKI was defined as an injury to the musculoskeletal system (bones, muscles, tendons, ligaments, etc.). This included conditions such as sprains, strains, fractures, pain/spasm/ache, and tendonitis, but not superficial injuries such as contusions and lacerations. The MSKI also were classified by anatomical location: lower extremity, upper extremity, spine, torso, or head/face.

The survey that was utilized to collect ISR data was administered to study participants on a tablet. The questions in the ISR survey were hierarchical, and the next question presented to a participant depended on their response to the previous question. This hierarchy in the ISR survey was designed to increase the accuracy of data entry by the study participants, as well as ease of use.

MSKI Data Collection
Both MCR and ISR data were collected from USMC recruits who volunteered to participate in the study. Medical Chart Review data included in this analysis were collected prospectively by clinicians at the Marine Corps Recruit Depots (MCRDs) during recruit training. The operational definition of MSKIs was relayed to the clinicians at the MCRDs. The study team had discussions with USMC clinicians about the
MSKI data, but the MSKI data were not International Classification of Diseases (ICD)-coded. A certified athletic trainer at MCRD Parris Island reviewed consented study participants’ Armed Forces Health Longitudinal Technology Application records and provided information about MSKIs to a study team member, who then transcribed the information. MSKI data from MCRD San Diego were derived from Military Health System Genesis via file transfer from base clinicians.

ISR data were collected at the end of recruit training via a customized survey through the University of Pittsburgh licensed REDCap software (National Institutes of Health support through Clinical and Translational Sciences Institute at the University of Pittsburgh Grant Number UL1-TR-001857).

Study participants were asked to provide details about all the MSKIs sustained during recruit training. They were provided information about the study team’s operational definition of MSKI, given multiple examples, and were encouraged to ask any questions they had. The questions in the ISR survey included MSKI date, anatomical location, sublocation, and MSKI type. The questions were hierarchical and the questions presented to the recruit depended on their response to the previous question. Based on the anatomical location and type of MSKI, the number of questions in the ISR ranged from 8 to 15. ISR data were collected during August 2021 at MCRD Parris Island and November 2021 at MCRD San Diego.

**Statistical Analysis**

The CRC analysis is a statistical technique to adjust for the underreporting inherent in many sources of MSKI data and estimate the “ascertainment-corrected” or CRC estimate of MSKI. The CRC estimate is interpreted as MSKI incidence if data were complete, and no MSKIs were “hidden” or missing in the counting process (Fig. 1).

The Chapman modification of the Lincoln-Peterson estimator was utilized to conduct the CRC analysis in the current study. This allowed an estimate of “hidden” or uncounted MSKIs (Fig. 1) when 2 sources of data were used, and an ascertainment-corrected cumulative incidence of MSKIs during recruit training was calculated.

The CRC analysis was repeated after stratification by the leading MSKI locations (lower extremity, upper extremity, and spine) and types (i.e., strain, sprain, and pain/spasm/ache) in the dataset.

**RESULTS**

Medical chart review and ISR were available for 464 USMC recruits (age: 19.1 ± 1.9 years; gender: men 70.0%; BMI: 24.5 ± 3.0 kg m$^{-2}$) and were included in the CRC analysis. Table II includes data about the observed number of injured recruits. A total of 101 recruits were identified as injured during the recruit training using MCR, 132 recruits were identified as injured using ISR, and 46 recruits were identified as injured from both sources of data. The observed number of injured recruits was 187 (formula in Table I, part a).

The results from the CRC analysis are also included in Table II. The CRC estimate of the number of injured recruits was 287.6 (95% CI, 239.5-335.8) (formulae in Table I, part b). This CRC estimate was much higher than the observed number of injured recruits (187). The overall ascertainment was moderate (65.0%), while the MCR and ISR ascertainment were 35.1% and 45.9%, respectively (formulae in Table I, part d).

The CRC analysis was repeated after stratification by MSKI locations and types. Overall ascertainment varied by anatomical location of the MSKI. It was highest for lower extremity MSKI (64.8%), but relatively lower for upper extremity (38.9%) and spine (33.3%) MSKI. For spine MSKI, the MCR ascertainment (19.0%) as well as the ISR ascertainment (16.7%) were low. ISR ascertainment was higher than MCR ascertainment for lower extremity MSKI (ISR: 44.5%, MCR: 36.3%). In contrast, MCR ascertainment was higher than ISR ascertainment for upper extremity (MCR: 22.6%, ISR: 20.1%) and spine MSKI.
TABLE I. Formulæ Utilized for Conducting 2-source CRC Analysis among USMC Recruits

a Observed number of injured recruits during recruit training
Observed number of injured recruits = (number of recruits identified as injured in medical charts + number of recruits identified as injured in self-reports) − (number of recruits identified as injured in both medical charts and self-reports).

b Estimated number of injured recruits during recruit training: CRC utilizing the Chapman modification of the Lincoln-Peterson estimator

\[ N = \frac{(M+1)(S+1)}{(O+1)} - 1 \]

\[ \text{variance}(N) = \frac{(M+1)(S+1)(M-O)(S-O)}{(O+1)(O+2)} - 1 \]

95% confidence interval for \( N = N \pm 1.96 \sqrt{\text{variance}(N)} \)

Where:
- \( N \) = CRC estimate of recruits injured.
- \( M \) = Number of recruits identified as injured in medical charts.
- \( S \) = Number of recruits identified as injured in self-reports.
- \( O \) = Overlap (number of recruits identified as injured in both sources of data).

c Cumulative incidence of injuries during recruit training

Capture – recapture cumulative incidence (\( \% \)) = \( \frac{\text{Estimated number of injured recruits in the capture – recapture analysis} \times 100}{\text{Number of recruits included in the analysis}} \)

Medical chart review cumulative incidence (\( \% \)) = \( \frac{\text{Observed number of injured recruits in the medical charts} \times 100}{\text{Number of recruits included in the analysis}} \)

Injury self – report cumulative incidence (\( \% \)) = \( \frac{\text{Observed number of injured recruits in the injury self-report} \times 100}{\text{Number of recruits included in the analysis}} \)

d Calculation of ascertainment percent

Overall ascertainment (\( \% \)) = \( \frac{\text{Observed number of injured recruits} \times 100}{\text{Capture – recapture estimate of recruits injured}} \)

Medical chart ascertainment (\( \% \)) = \( \frac{\text{Observed number of injured recruits in the medical charts}}{\text{Capture – recapture estimate of recruits injured} \times 100} \)

Self – report ascertainment (\( \% \)) = \( \frac{\text{Observed number of injured recruits in the injury self-report}}{\text{Capture – recapture estimate of recruits injured} \times 100} \)

Overall ascertainment also varied by MSKI type. Overall ascertainment was highest for sprain (55.1%), followed by strain (54.8%) and the pain/spasm/ache (43.3%). For injuries identified as pain/spasm/ache, MCR ascertainment (27.1%) was higher than ISR ascertainment (21.9%).

Table III includes data about the cumulative incidence of MSKIs during recruit training (formulæ in Table I, part c). The MCR and ISR cumulative incidences were both low (MCR: 21.8%, ISR: 28.4%). The ascertainment-corrected CRC cumulative incidence (62.0%, 95% CI, 51.6-71.4) was much higher than the MCR as well as the ISR cumulative incidences. MSKIs were stratified by anatomical location and by MSKI type and the analysis was repeated within each stratum. In each of the stratified analyses, the CRC cumulative incidence was much higher than both the MCR and the ISR cumulative incidence.

The MSKI cumulative incidence during recruit training was higher in the ISR data (132/464 = 28.4%) compared to MCR data (101/464 = 21.8%). A higher percentage of recruits were also identified as injured in the ISR as compared to the MCR for lower extremity MSKI (ISR: 22.2%, MCR: 18.1%) and sprain (ISR: 3.9%, MCR: 1.7%). In contrast, a higher percentage of recruits were identified as injured in the MSR as compared to ISR for upper extremity MSKI (ISR: 22.2%, MCR: 18.1%) and spine MSKI (MCR: 3.4%, ISR: 3.0%), strain (MSR: 3.0%, ISR: 1.9%), and pain/spasm/ache (MCR: 13.4%, 10.8%).

TABLE II. CRC Estimates and Percent Ascertainment of MSKIs among 464 Recruits during USMC Recruit Training

<table>
<thead>
<tr>
<th>MSKI type</th>
<th>Number of subjects identified as injured in medical charts</th>
<th>Number of subjects identified as injured in self-reports</th>
<th>Number of subjects identified as injured in both sources</th>
<th>Total number of subjects identified as injured</th>
<th>CRC estimate* of number of injured subjects (95% CI)</th>
<th>Overall</th>
<th>Medical chart</th>
<th>Self-report</th>
</tr>
</thead>
<tbody>
<tr>
<td>Any/all MSKI</td>
<td>101</td>
<td>132</td>
<td>46</td>
<td>187</td>
<td>288 (239-336)</td>
<td>65.0</td>
<td>35.1</td>
<td>45.9</td>
</tr>
<tr>
<td>Lower extremity</td>
<td>84</td>
<td>103</td>
<td>37</td>
<td>150</td>
<td>232 (188-275)</td>
<td>64.8</td>
<td>36.3</td>
<td>44.5</td>
</tr>
<tr>
<td>Upper extremity</td>
<td>18</td>
<td>16</td>
<td>3</td>
<td>31</td>
<td>80 (25-135)</td>
<td>38.9</td>
<td>22.6</td>
<td>20.1</td>
</tr>
<tr>
<td>Spine</td>
<td>16</td>
<td>14</td>
<td>2</td>
<td>28</td>
<td>84 (16-152)</td>
<td>33.3</td>
<td>19.0</td>
<td>16.7</td>
</tr>
<tr>
<td>Strain</td>
<td>14</td>
<td>9</td>
<td>3</td>
<td>20</td>
<td>37 (15-58)</td>
<td>54.8</td>
<td>38.4</td>
<td>24.7</td>
</tr>
<tr>
<td>Sprain</td>
<td>8</td>
<td>18</td>
<td>3</td>
<td>23</td>
<td>42 (17-66)</td>
<td>55.1</td>
<td>19.2</td>
<td>43.1</td>
</tr>
<tr>
<td>Pain/spasm/ache</td>
<td>62</td>
<td>50</td>
<td>13</td>
<td>99</td>
<td>229 (141-316)</td>
<td>43.3</td>
<td>27.1</td>
<td>21.9</td>
</tr>
</tbody>
</table>

*Number has been rounded to the nearest whole number for practical interpretation. The CRC estimates for anatomical locations and injury types will not sum to the Any/All total. Some recruits experienced more than 1 injury.

Overall ascertainment also varied by MSKI type. Overall ascertainment was highest for sprain (55.1%), followed by strain (54.8%) and the pain/spasm/ache (43.3%). For injuries identified as pain/spasm/ache, MCR ascertainment (27.1%) was higher than ISR ascertainment (21.9%).
DISCUSSION

The purpose of this 2-source CRC analysis was to estimate the ascertainment-corrected cumulative incidence of MSKIs during USMC recruit training. In agreement with the hypothesis, the CRC estimate of the incidence (62.0%) was much higher than the observed incidence in each of the 2 individual sources of MSKI data—MCR (21.8%) and ISR (28.4%). In addition, the overall ascertainment in the CRC analysis, which can be interpreted as the completeness of data, was lowest for spine MSKIs (33.3%), and varied by MSKI anatomical location and type.

MCR, ISR, and CRC Cumulative Incidence

The observed cumulative incidence of MSKIs during recruit training was low (MCR: 21.8%, ISR: 28.4%), whereas the estimated CRC incidence, after combining information in the 2 sources of data, was much higher (62.0%). This finding highlights the inherent incompleteness of MSKI data from each source and agrees with the findings from previous studies in civilian as well as military populations. A CRC analysis was utilized to assess the completeness of MSKI data from bicycle crashes in New Zealand that were attended to by medical personnel and from the police data, obtained through record linkage. The quality of the record linkage data was assessed by comparison to ISR data. There was a moderate agreement (kappa = 0.55) between the linked and ISR data. The record linkage data were incomplete, and only 63.1% of self-reported injuries were retrieved through linkage from insurance claims, hospital discharges, mortality records, and police reports. The ISR data were also incomplete; 1 or more crashes were recorded in the linked record data for 5.6% of participants who did not self-report a crash. The findings of the current study agree with this previous study of bicycle crashes. A similar CRC study to estimate the 1-year cumulative incidence of MSKIs among active duty U.S. Army soldiers showed that there was undercounting from both sources of data (MCR and ISR). The percentage of CRC-estimated MSKIs observed in the previous analysis, or the estimate of the completeness of data, was low (MCR: 32.9%, ISR: 36.1%).

The overall ascertainment in the current study, after utilizing 2 sources of MSKI data, was moderate-to-good (65.0%). An overall ascertainment of 65% means that nearly two-thirds of all MSKIs were counted when information obtained from MCR and ISR were combined. When each individual source of data was used in isolation, the ascertainment was poor—only 35.1% for MCR and slightly higher (45.9%) for ISR. The current analysis utilized 2 available sources of data (MCR and ISR). Previous research has demonstrated that MSKIs are underreported even when more than 2 sources of data are utilized. It is expected that as more independent sources of MSKI data are included in the CRC analysis, the completeness or ascertainment will improve. A CRC analysis of fatal occupational injuries in Nicaragua utilized a very large number of data sources (n = 10). Despite the unusually large number of sources as well as the severe nature of the injury, the observed rate (8.3 per 100,000 employed) was much lower than the CRC estimated rate (11.6 per 100,000). These findings from various CRC studies underscore the importance of improving the quality of each individual data source, as well as creating better record linkage between sources.

Validity and Completeness of MSKI Data Obtained by MCR and ISR

In the current study, the overall ascertainment was higher for ISR (45.9%) compared to MCR (35.1%). Military recruits are in a unique situation during recruit training. To avoid profile and to increase the likelihood that they graduate with their cohort, recruits may not seek medical care for MSKIs. This may result in recruits attempting to perceive MSKIs as less severe than what they truly are and/or to self-manage their MSKIs.

It is more cost-effective and time-efficient to collect ISR data. Also, ISR data are expected to contain information about all injuries that have affected study participants, including those for which medical care was not sought or was sought outside a system from which data are obtained. Not seeking medical care is especially problematic in military populations, which makes collection of ISR data especially important in this population. In addition, ISR data often contain information that is critical to MSKI prevention research, such as cause and mechanism of injuries, which are often absent in military outpatient medical records.

ISR data have their limitations as well because humans are inherently limited in their ability to remember information. This can lead to differential reporting based on demographic groups and undercounting of self-perceived minor MSKI, especially as time-since-injury increases. A study among Naval Special Warfare personnel, a military population especially prone to injuries, compared MSR and ISR MSKIs that occurred during a 1-year period. Recall for less severe, non-fracture MSKIs was higher for more recent compared to older MSKIs. This effect of time-since-MSKI on recall was not observed for fractures.

Medical records can be expected to contain accurate information about injuries, since medical records are created by

<table>
<thead>
<tr>
<th>Injury type</th>
<th>CRC incidence (95% CI)</th>
<th>Medical chart incidence</th>
<th>Self-report incidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Any/all injuries</td>
<td>62.0 (51.6-72.4)</td>
<td>21.8</td>
<td>28.4</td>
</tr>
<tr>
<td>Lower extremity</td>
<td>49.9 (40.6-59.2)</td>
<td>18.1</td>
<td>22.2</td>
</tr>
<tr>
<td>Upper extremity</td>
<td>17.2 (5.3-29.0)</td>
<td>3.9</td>
<td>3.4</td>
</tr>
<tr>
<td>Spine</td>
<td>18.1 (3.5-32.7)</td>
<td>3.4</td>
<td>3.0</td>
</tr>
<tr>
<td>Strain</td>
<td>7.9 (3.2-12.5)</td>
<td>3.0</td>
<td>1.9</td>
</tr>
<tr>
<td>Sprain</td>
<td>9.0 (3.7-14.3)</td>
<td>1.7</td>
<td>3.9</td>
</tr>
<tr>
<td>Pain/spasm/ache</td>
<td>49.2 (30.4-68.0)</td>
<td>13.4</td>
<td>10.8</td>
</tr>
</tbody>
</table>

CAPTURE-RECAPTURE ANALYSIS
health care providers. Medical records, however, may be prone to low concurrent validity due to coding and keypunch errors.\textsuperscript{25,26} Additionally, these records often do not contain information about the cause of injury, since the primary purpose of medical records is documenting medical care and not research. Information about the cause of injury is essential for optimal surveillance, evaluation of interventions, and creating injury prevention policies,\textsuperscript{27} which are essential sequential steps when applying the public health model to injury prevention.\textsuperscript{23}

**Capture-Recapture for Assessing Completeness of MSKI Data**

Capture-recapture is a statistical technique that can be utilized to compensate for the undercounting inherent in various sources of data and to calculate an ascertainment-corrected estimate of the cumulative incidence or the prevalence.\textsuperscript{28} The CRC technique has been utilized extensively in disease and injury surveillance for diverse conditions such as: road traffic mortality,\textsuperscript{29}; mortality before and during the COVID-19 pandemic,\textsuperscript{30}; prevalence of opioid use disorder,\textsuperscript{31}; and the incidence of type-1 diabetes in children.\textsuperscript{32} The application of the CRC technique to injury data in military populations, however, has been limited. A study among active duty U.S. Army soldiers utilized a 2-source CRC analysis to estimate the 1-year cumulative incidence of MSKIs.\textsuperscript{17} To the best of our knowledge, no research study has utilized the CRC technique in a military recruit population, where the risk of MSKIs is much higher than among active duty military populations and the duration of training is short.

**Effect of MSKI Severity and Time since MSKI on Completeness of Data**

Previous research among Army soldiers has shown that the overall ascertainment of severe MSKIs, such as fractures, is typically high.\textsuperscript{17} A CRC analysis of the 1-year cumulative incidence of MSKIs among active duty Army soldiers\textsuperscript{17} showed that the percent of undercounting was comparatively low for severe MSKIs such as fractures (overall ascertainment: 75.0%), as compared to less severe MSKIs such as strain (43.8%) or sprain (53.8%).\textsuperscript{17} This strong impact of MSKI severity on completeness of MSKI data was not seen in the current study. In contrast to the study among Army soldiers\textsuperscript{17} that was conducted over a 1-year period, participants in the current study were in a 13-week recruit training program. Recruits with severe MSKIs, such as traumatic fractures, would have attrited from their recruit training cohort; therefore, their data were not available for inclusion in the analyses. Further, the shorter time between time-of-injury and collection of ISR data may reduce risk of lack of recall.

Previous research has investigated the combined effect of time-since-injury and injury severity of recall on injuries and poisonings reported in the National Health Interview Survey.\textsuperscript{8} As the time-since-injury increased, there was a decrease in the number of less severe incidents reported; however, this trend was not observed for severe injuries involving fractures and hospitalization.\textsuperscript{8} This effect of time-since-injury on recall could not be assessed in the current study, as the duration of the current study was relatively short—13 week of USMC recruit training. Also, likely due to the short duration of the study, the overall ascertainment or completeness of data in the current study (65.0%) was higher than in a similar CRC study conducted among active duty Army soldiers over a 1-year period (57.4%).\textsuperscript{17} These findings are in agreement with the previous research studies that have recommended relatively short recall periods of 2 months\textsuperscript{6} or as short as 5 weeks\textsuperscript{8} to improve the accuracy of recall. It is also possible that recruits in the current study may have had better on-site access to medical care compared to the active duty Army soldiers analyzed in previous research.\textsuperscript{17}

**Effect of Anatomical Location and MSKI Type on Completeness of Data**

In the current study, overall ascertainment was lower for spine (33.3%) and upper extremity (38.9%) compared to lower extremity (64.8%) MSKIs. This lower ascertainment could be due to underreporting in the ISRs or not seeking medical care, or a combination of both factors. The findings from the current study agree with similar observations among other military populations (i.e., active duty U.S. Army soldiers,\textsuperscript{17} Special Operations reserve soldiers\textsuperscript{33}) and workers’ compensation reports.\textsuperscript{34} A study of the validity of ISR of no previous back pain during a 3-year consecutive period among a group of Special Operations reserve soldiers showed that more than 80% of the study participants reported at least 1 back pain episode during monthly interval testing. Despite the high incidence of back pain in this study sample, most (97%) continued to self-report no history of back pain in the annual medical certificate. The authors theorized that the soldiers included in the study considered even severe back pain to be “par for the course,” and not a medical problem.\textsuperscript{33} A study among active duty Army soldiers showed that overall ascertainment was lower for spine injuries (31.9%) as compared to other anatomical locations analyzed.\textsuperscript{17} The findings from the current study also agree with an observed low rate of reporting upper extremity injuries in workers’ compensation reports.\textsuperscript{34} These observations suggest the possibility of differential lack of recall for MSKIs by anatomical locations, especially among those who have a good self-perception of health, which needs to be considered when assessing self-reported MSKI data.

**Study Limitations**

The participants in this study were a unique sample of USMC recruits assessed during recruit training and were highly susceptible to the risk of MSKI. The findings of this study may not be generalizable to other populations, including other military populations. An important assumption of the CRC
analysis is independence of the sources of data, which usually is not practical in epidemiology data; recruits who had an MSKI in their MCR are more likely to report an MSKI in their ISR. Positive dependence between sources of data utilized in the CRC analysis can lead to conservative underestimation of injured participants. Future research to identify methods to correct this underestimation in military populations is recommended. Negative dependence can lead to a serious problem with overestimation of MSKI incidence, but this is unusual in multiple sources of epidemiology data[11] and was not expected in this study. Recruits with severe MSKIs would have attrited from their recruit training cohort; therefore, their data were not available for inclusion in the analyses.

It is plausible that MSKIs are underreported in both sources of data (MCR and ISR) as some minor MSKIs that do not impact physical performance may not be recognized as “injuries” by the study participants. MCR MSKI data were obtained separately from MCRDs Parris Island and San Diego, from USMC clinicians. Since these data were not ICD-coded, it was not possible for the study team to confirm that the same codes were pulled from both MCRDs. Also, if the study team were able to extract ICD-10 Clinical Modification-coded data from other repositories such as the Military Health System Data Repository or Defense Medical Surveillance System, we may have captured a greater number of MSKIs in the MCR. Due to the relatively short duration of USMC recruit training (13 weeks) as well as the absence of severe MSKIs such as traumatic fractures in the dataset, it was not possible to assess the effect of MSKI severity and time-since-MSKI on the completeness of MSKI data obtained by ISR.

CONCLUSIONS

Availability of valid and reliable MSKI data is essential in injury epidemiology research in the military and in other populations. The results of this study suggest that, in military recruit populations, MSKI incidence estimated using only 1 source of data may suffer from issues with undercounting. The CRC method has the potential utility to correct for the under-ascertainment when measuring MSKI incidence, as well as to assess and compare the amount of underreporting from various sources of MSKI data. Since various sources of MSKI data in military populations are prone to issues with under-ascertainment, it is important to weigh the advantages and disadvantages of MSKI data obtained by MCR and by ISR before utilizing them in injury epidemiology research in military populations. The findings from this and other CRC studies underscore the importance of improving the quality of the individual data sources, as well as creating better record linkage between sources. There is a need for future research to identify modifiable factors that influence recall of data in the ISR, and application of the CRC technique to the assessment of completeness of various MSKI attributes such as MSKI causes and mechanism.

FUNDING

U.S. Marine Corps AWD00030104 M95494-20-C-0021; no COI.

SUPPLEMENT SPONSORSHIP

This article appears as part of the supplement “United States Marine Corps Gender-Integrated Recruit Training Study,” sponsored by the University of Pittsburgh.

CONFLICT OF INTEREST STATEMENT

None declared.

DATA AVAILABILITY

The original data underlying this article cannot be shared publicly due to ethical reasons and the sensitive nature of the data.

REFERENCES


