Identifying Determinants of Performance for Females Completing a Paramedic Physical Employment Test

by

Alexander Lawrence Malone

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Master of Science
in
Kinesiology and Health Sciences

Waterloo, Ontario, Canada, 2023

Author’s Declaration

This thesis consists of material all of which I authored or co-authored: see Statement of Contributions included in the thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Statement of Contributions

Daniel Armstrong, Renée MacPhee and Steven Fischer were contributing authors for section 4 (Research Question 1), which was submitted for publication, and section 5 (Research Question 2).

Abstract

Background: Sex disparities exist in employment and injury rates in the paramedic sector. Low success rates among females attempting physical employment standards could explain the elevated injury risk among female paramedics. Identifying factors that underpin successful work-related performance can inform pre-hire and return-to-work based physical training programs to address these disparities.

Purpose: The purpose of this thesis was to identify the determinants of successful physical performance for females engaged in paramedic tasks.

Research Question 1: Participant demographics, college type, employment status and heart rate were obtained from female participants who completed the Ottawa Paramedic Physical Abilities Test (OPPAT™), a physical employment standard for paramedics. These data were used in a logistic regression model to determine which factors could predict the likelihood of successfully completing the OPPAT™. Females who were actively employed, who were educated in a public paramedic college, who had higher body mass, or those who had lower BMI were more likely to successfully complete the OPPAT™.

Research Question 2: Lift duration and the time between peak knee and hip joint angular velocity during the Scoop and Barbell lift were compared between females who passed and failed the Ottawa Paramedic Physical Abilities Test. Four ANCOVAs were used for these comparisons where college type (public or private) and employment status (employed or unemployed) were used as categorical factors and body mass and BMI were used as covariates. No significant differences were found between passing and failing females.

Discussion: Modulating demographic factors that increase the likelihood of success could lead to improved performance outcomes, but other determinants should be explored to improve the predictive ability of the current model. Future research should continue to leverage emerging technology, such as markerless motion capture and unsupervised machine learning, to identify determinants of success for females in paramedic tasks.

Acknowledgements

I would like to acknowledge all of the active duty and incoming paramedics for consenting to have their demographic and performance information from the OPPAT^{TM} collected for research purposes.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

ANCOVA	Analysis of Covariance
BMI	Body Mass Index
FCE	Functional Capacity Evaluation
HJA	Hip Joint Angle
HJAV	Hip Joint Angular Velocity
HR	Hear Rate
HvK	Hip vs Knee Peak Joint Angular Velocity Timing
KJA	Knee Joint Angle
KJAV	Knee Joint Angular Velocity
MP	MediaPipe
OPPAT	Ottawa Paramedic Physical Abilities Test
OR	Odds Ratio
PES	Physical Employment Standard
RMSD	Root Mean Squared Difference

1 Introduction

Paramedics provide vital medical services to the public but continue to have an unbalanced sex representation in the workforce and experience elevated injury rates compared to the general population. Female underrepresentation in the workplace has remained consistent, with females comprising less than a quarter of working paramedics in the last decade (Crowe et al., 2020). This could be connected to the elevated injury risk associated with this occupation. In Australia, injury risk for active paramedics is seven times greater than the national average (Maguire et al., 2014). While this discrepancy in injuries with the general population is substantial, the risk is further compounded among female paramedics. Despite comprising such a small portion of the workforce, female paramedics are at greater risk of injury compared to males (Kearney, Muir & Smith, 2022). The elevated risk for injury among female workers should be addressed, in addition to the injury risk in the broader paramedic community. Previous research indicates that a disconnect exists between paramedic workforce capacity and public demand, particularly with the ability to safely complete physical tasks. Systematic reviews have linked elevated call volumes with injury rates in paramedics, with the majority of injury claims and time off resulting from musculoskeletal injury manual handling duties (Kearney et al., 2022). The relationship between volume and injury rates could signal a surge of work-related injuries with recent increases in demand for paramedic services. From 2016 to 2018, Ontario experienced over a 10% increase in call volume despite only a 3% increase in population (Pasma, 2020). When discussing similar trends in Australia, MacQuarrie (2019) stated that recent increases in call volume show that paramedics are busier and under more pressure to deliver care to the public. The associated risks and rising trends in call volume highlight an urgent call to action in support of the paramedic workforce. With physical duties being the leading cause of workplace injury, researchers need to better understand how individuals, but females in particular, can meet the physical demands of the paramedic tasks.

Paramedic candidates must pass a series of written, practical and physical tests before they are eligible to be hired. The practical portion is a physical employment standard (PES) that is also implemented as a return-to-work evaluation that suggests individuals have the physical capacity to meet the demands of the job. The Ottawa Paramedic Physical Ability Test (OPPAT™) exists as the current physical employment standard in Ontario to prove a candidate's readiness for duty. Maguire and colleagues (2005) have noted that injury rates may be directly associated with employee turnover rate, career length and quality of service provided. Considering the proportion of injuries related to physical labour, focusing on improving physical readiness for duty among pre-hire candidates and current employees will inform strategies to improve both the viability of the employment pool and may also strengthen worker retention. This also provides an opportunity to add to the limited research on the determinants of success for paramedics. Like other PESs, the outcome of the OPPAT™ is a binary pass or fail, but several reasons might underpin why an individual does not pass. Did the candidate lack sufficient strength to perform a demanding lifting task? Did the candidate select a movement strategy that increased the difficulty of the task? Presently, this is a gap in our understanding of paramedic PES performance which can be explored using data collected within the test itself.

We do not know which factors influence the likelihood of passing or failing the OPPAT™, which in turn, limits the ability to develop appropriate interventions to maximize an individuals' ability to meet to demands of the job. Research on other PES performances have found connections between certain anthropometric data and success rates (Jamnik, Thomas & Gledhill, 2010; Reilly et al., 2016). If found to be influential of physical performance, strategies can be formed to modulate these determinants when possible, or plans can be established for alternative solution when determinants are more resistant to change. This is not to say that those with unfavourable characteristics are doomed to perform poorly in occupational task. Sex has been linked to poor performance in other PES performance (Harbin & Olson, 2005; Jamnik et al., 2010; Reilly et al., 2016) and elevated injury rates in manual occupational work (Gagnon Plamondon, & Larivière, 2018; Kearney et al., 2022). While many physiological differences exist between the average males and females that limit PES performance (Roberts et al., 2016), there are females who can meet the physical demands required to complete manual occupational tasks. It has been suggested that a performers' skill or movement competency can allow them to leverage their available degrees of freedom to be successful in a task despite certain limitations (Armstrong et al., 2019a). Therefore, the inclusion of a movement comparison will help identify underlying gaps in performance between successful and unsuccessful females in the OPPAT™.

In addition to anthropometrics, distinct movement differences when performing manual materials handling tasks have been identified between males and females (Gagnon et al., 2018; Kranz et al., 2021; Makhoul et al., 2017; Plamondon et al., 2014). Specific to paramedics, leg power has been linked to injury risk and linked to spine moments during lifting tasks (Lentz et al., 2019; Makhoul et al., 2017). Additionally, lower body joint coordination was found to be distinctly different between not only males and females, but also novice and elite employees (Plamondon et al., 2014). From this previous research, we expected differences in movement coordination and a speed-based metric between successful and unsuccessful female candidates. The use of 2D markerless motion capture to extract these outputs was proposed to maximize the external validity of the movement comparisons by limiting the interreference with performers in the OPPAT^{TM}. Through evaluating movement and demographic differences linked to successful performance, interventions can be informed to increases the number of female candidates physically ready for duty and maintain this physical ability in current paramedics.

The purpose of this thesis project was to identify factors, or combination of factors, that determine pass/fail results from females performing the OPPAT^{TM}. By first identifying determinants of performance, future research can guide precise interventions that improve occupational performance and ultimately to bolster the paramedic workforce.

2 Literature Review

2.1 Paramedics

Paramedic services are in high demand but are physically demanding and strain both the capacity of the individual workers and the workforce as a whole. After conducting a systematic review of studies from 2004 to 2019, Kearney and colleagues (2022) found that paramedics experience 33 injuries per 100 full-time equivalent workers, with 64% of injuries being musculoskeletal (Kearney et al., 2022). This contrasts the U.S. Bureau of Labor Statistics (2022), which reported a national average of < 3 injuries per 100 full-time equivalent workers in 2021. With paramedics experiencing injuries at a rate 10 times that of the national average, an intervention is needed to support paramedics. There is not only a moral motivation to aid these emergency responders, but a financial incentive as well. While Ontario experienced >10% increase in call volume, annual paramedic overtime hours increased by 14.8% and Workplace Safety and Insurance Board claims increased by 63.5% (Pasma, 2020). The combined annual cost of overtime hours and injury claims totaled $47.7 million in 2018 (Pasma, 2020), showing not only a physical toll on employees but a financial burden to the community. It has also been suggested that these consistent trends, particularly regarding injury claims, could mean that paramedics working while injured and/or are not providing patient care to the best possible standard. If true, this could mean that paramedics are not only struggling to service the public, but the quality of service they provide may be declining (Maguire & Smith, 2013). It is clear that paramedics require additional support to meet to growing demand of the public, but there are different avenues through which assistance can be provided. Additional recruiting could add to the current work forces raw numbers, but does not provide support to the capacity of the individual worker. Rather than spending upwards of $7.7 thousand per individual to expand paramedic recruitment (Avesta Systems, Inc & the American Ambulance Association, 2019), focusing efforts on physical performance within the current candidate pool and incumbent paramedics may be an effective support option.

2.2 PES

With physical duties being the greatest cause of workplace injury (Kearney et al., 2022), an evaluation involving these physical tasks can be used to explore predictors of successful performance. In some work environments, the occupational demands of a job can be readily modified to alleviate the physical stress placed on the worker. However, if these occupational requirements cannot be modified due to the essential nature of the work being done, employees need to be capable of meeting the demands of the job (Harbin & Olson, 2005). Jamnik et al. (2010) posits that there is a responsibility to employees and the general public to ensure that all applicants are capable of meeting the physical demands that they will encounter on the job, therefore, Physical Employment Standards (PES) requirements cannot be lowered to accommodate applicants who are unable to meet the established standards. PES are an evaluation of an individual's ability to perform job-specific tasks that are essential to the occupation, and can be used for employee placement and/or as a pre-hire physical requirement. Focusing on one's ability to complete job-specific tasks has also been linked to work injury incidence, more so than physical testing alone (Harbin and Olson, 2005). Using Functional Capacity Evaluation (FCE) performance (where an FCE is a more medicalized form of physical testing) to determine if individuals had the proven physical capacity to perform essential functions of job, Harbin and Olson (2005) found fewer injuries (3%) among those whose FCE results indicated physical capacity that matched or exceeded the demands of the job compared to individuals whose FCE performance did not meet the demands of their job (33%). Improving job specific performance could reduce the incidence of injury, thereby increasing both the number of paramedics in the field and overall length of employment (Maguire et al., 2005). To be an effective test of physical requirements, evaluations need to be both specific to and representative of occupational demands (Harbin & Olson, 2005), so an evaluation exclusive to paramedic duties is required. The paramedic sector has recently introduced a PES, the Ottawa Paramedic Physical Ability Test (OPPAT™), designed in accordance with best practice guidelines as outlined by Jamnik, Gumienak and Gledhill (2013). In the establishment of this test, a development team was formed and conducted a physical demands description of paramedic duties, then characterized the most frequent and demanding tasks. The team then developed, piloted and refined a test before establishing scientific validity. The OPPAT™ exists as a legally defensible physical employment standard through which physical readiness for paramedic employments can be tested. Identifying determinants of success in the OPPAT™ will inform which strategies should be adopted to maximize performance for all candidates in the OPPAT™, particularly female participants. While there are research gaps regarding the specific determinants of performance for paramedic tasks, similar occupational performance data can inform the starting point of investigations with this population.

2.3 Demographic Determinants of Performance

Biological sex is a factor that has been linked to occupational performance and injury risk in various work sectors. Jamnik, Thomas & Gledhill (2010) reported pass rates of 29% for females and 73% for males in a fitness test for correctional officer applicants (FITCO). There have been similar performance findings from OPPAT™ investigations as well. When first attempting the OPPAT™, 29% of the group failed, all of which were female (Armstrong et al., 2019c). The alignment of findings between these studies shows that female physical performance in the paramedic sector is no exception to trends seen in other laborious occupations. Specific to injuries among emergency responders, a systematic review found that females were also more likely to be injured on the job compared to males (Lentz et al., 2019). This is not to say that all females have the same increased risk of injury. Nor does this mean that the one's sex is the ultimate determinant of how they perform occupational tasks. Several underlying determinants may be contributing to these associations we see between sex and both occupational performance and injury. On average, women have less muscle mass, strength, power, and endurance compared to men (Roberts et al., 2016). In an evaluation of fatigue during paramedic tasks, more than half of the females (53%) reached max heart rate, compared to only 10% of the male participants, indicating a higher level of fatigue with the same tasks (Barnekow-Bergkvist et al., 2004). Based on the underlying physiological differences between sexes, it has been suggested that the average female likely requires a greater percentage of her capacity to perform the same lifting task as the average male (Barnekow-Bergkvist et al., 2004). Accounting for these findings and hypotheses, perhaps mitigating these underlying differences may bridge the gap in performance outcomes between males and females.

Investigations have been conducted on the determinants of performance in the military and evaluations to inform interventions that can bolster the workforce and reduce injury. Reilly and colleagues (2016) measured body mass, fat mass, lean body mass, body mass index (BMI), body fat percentage, segmental lean mass, height, hip and waist circumference, sex, age, and dead mass (fat mass and load being lifted) as possible predictors of military performance on a PES. They found lean body mass to be one of the strongest predictors of performance, and suggest the use of fitness programs designed around these findings to improve performance. Expanding beyond anthropometric measures of organism constraints, Hydren and colleagues (2017) conducted a meta-analysis involving 9 reports on prediction of lifting capacity in military personnel. Multiple fitness domains were evaluated, including: body mass, aerobic capacity, dynamic strength, power, isometric strength, strength endurance, speed, isokinetic strength, flexibility and age. In addition to lean body mass, the results of their meta-analysis showed that power and dynamic strength were strong predictors of performance (Hydren et al., 2017). Outside of physiological determinants, work experience has also been associated with performance differences (Gagnon et al., 2018; Plamondon et al., 2010; Plamondon et al., 2014) and injury risk (Lentz et al., 2019; Maguire, 2011) between males and females. A similar investigation on performance in the OPPAT™ has yet to be conducted. To begin such an investigation, a wide range of determinants, inclusive of physiological and experiential features, were considered but informed from previous performance research.

2.3.1 Heart Rate Response

Cardiovascular fitness has been linked to injury risk and performance that is readily accessible during the OPPAT™ testing protocol. VO_2 max (maximal oxygen consumption) is often used as a reliable measure for cardiovascular fitness (Barnekow-Bergkvist et al., 2004; MacQuarrie, 2019). Research done by Poplin et al. (2014) found an association between VO_2 max and injuries in firefighters, where workers with lower VO_2 max (< 43 mL/kg/min) were 2.2 times more likely to be injured than those with a higher VO_2 max (> 48 mL/kg/min). While VO_2 max is commonly used, it is not the only measure that can be associated with cardiovascular fitness. In an evaluation of active paramedics, MacQuarrie (2019) found that paramedics with higher VO_2 max had lower heart rate (HR) and percentage of maximum heart rate (%HR max) whilst on calls compared to those with lower VO_2 max measurements. This finding provides confidence that a heart rate response to tasks can act as a substitute for VO_2 max values to represent a cardiovascular fitness. Morgan, Allison and Duhon (2012) compared heart rate response between different performance groups in a functional capacity evaluation. Performance groups were dictated by an individual's ability to repeat a task to volitional fatigue before or after being stopped by an evaluator based on biomechanical exposures. The higher performance group performed FCE tasks with significantly lower changes in heart rate compared to the lower performance group (Morgan et al., 2012). A real-time heart rate response measure should be included in the determinants of interest, in addition to other physiological characteristics associated with physical performance.

2.3.2 Body Mass and BMI

Similar to heart rate response, body mass and BMI can be modified over time and have been linked to performance. In paramedic specific tasks, heavier personnel were able to perform prolonged carrying tasks with less fatigue response compared to their lighter counterparts (Barnekow-Bergkvist et al., 2004). However, a participants size may encompass other underlying effects associated with their general stature. For example, the military PES performance investigation conducted by Reilly et al. (2016) found that greater lean body mass was one of the most positively associated predictor of performance and accounted for a portion of the sex differences in PES outcomes. Based on findings from Reilly et al. (2016) and Barnekow-Bergkvist et al. (2004), it is unclear if the overall size of the individual or both size and body composition are the better predictor of performance. To explore this relationship further, body mass will be included in tandem with BMI, which can be used as an analogous measure for lean body mass, and has already been noted as a measure of interest for paramedics. Studnek et al. (2010) evaluated the health status of EMS professionals and found that 27.7% were classed as “overweight” and 26% were classified as “obese” based on their BMI values. Sheridan (2019) identified 6 cross-sectional, international studies where excess weight was common in paramedics and linked to a decrease in both the quality of work and ability to perform physical activities. MacQuarrie (2019) suggests that elevated BMI levels in the paramedics may be associated with high injury rates as well. Considering the influence of BMI, using height, as a metric involved in the calculation of BMI, should also be explored.

2.3.3 Height

Height, being fairly consistent at working age, presents the most unique challenge of the underlying differences between sexes. Roberts and colleagues (2016) noted that height differences between males and females could be a significant contributor to the discrepancies in performance. Height was also found to explain a significant amount of variance in development of fatigue among females when performing a simulated stretcher-carrying task, even more so than body weight differences (Barnekow-Bergkvist et al., 2004). While it was not a factor compared in OPPAT^{TM} performance before, demographics from Armstrong et al. (2019c) showed that female candidates were 11.9 cm shorter on average compared to the male candidates. This same investigation found that those who failed from their sample were all female. Beyond paramedic tasks, Schenk and colleagues (2006) found that the relative height of a load in a functional capacity evaluation was inversely related to performance. The authors noted that, although height is a significant determinant influencing lifting capacity, the specific causes are unclear. Accounting for the limited research on the exact impact on performance, it is important to investigate height as a determinant of performance for future and current female paramedics.

2.3.4 Experience

Expanding beyond direct influence of anthropometrics, experience has also been connected with injury risk and performance. Specific to paramedics, Maguire (2011) conducted a review of transportation injuries in EMTs and paramedics over 3 years and found that 52% of injuries occurred in employees with 5 years of experience or less. Employees with less than a year of experience account for 30% of the total injuries in this study. Elevated injury risk with the less experienced employees could be explained by movement differences between novices and experts. In manual materials handling tasks, Plamondon and colleagues (2014) identified unique movement outcomes in novice employees, where novice employees exhibited less synchronous lower body movement compared to experts. In a later investigation, novices were found to differ from experts again, with a greater reliance on passive resistance and less active muscle force when competing a box transfer task (Gagnon et al., 2018). Specific to paramedics, Armstrong et al. (2019c) showed that the initial pass rate of participants engaging the OPPAT^{TM} increased by 19% after repeated exposures to the PES. Throughout the duration of this study, active paramedics had a pass rate of 100% (Armstrong et al., 2019c). Each of these findings show the performance benefit of practical experience in occupational performance. However, employees all start with no work experience, so can we account for experience differences even among new recruits?

Leveraging the findings from Armstrong and colleagues (2019c) and anecdotal observations from OPPAT^{TM} facilitators, exposure to paramedic specific tasks, rather than general work experience, is an analogous area of interests. It is suspected that, among paramedic recruits, those educated from a public paramedic college are exposed to more physical paramedic duties compared to those coming from a private paramedic college. If true, the additional exposure to paramedic duties may better prepare public college graduates in contrast to their private college counterparts. As both the college of study and employment status of candidates are readily accessible from OPPAT™, this information should be incorporated as a demographic of interest in relation to physical paramedic performance.

2.4 Kinematic Comparison

There are various ways an individual might coordinate their movements when performing an occupational task. Research by Armstrong and colleagues (2019a, 2020a) indicated a need to consider movement strategy when evaluating physical capacity. Measuring segmental kinematics to evaluate techniques was noted by Makhoul et al. (2017) as a potentially useful evaluation in aiding paramedics. One benefit of a kinematic comparison is that it can inform immediate interventions, in the form of specific movement suggestions, in contrast to training interventions which require time for physiological adaptations as well as life-style changes. Courtney, Francis & Paxton (2010) noted that high BMI is typically reported in paramedics but also discussed contributing personal factors such as elevated levels of fatigue and stress, as well as 14% lower physical activity than the general public. Additionally, Studnek et al. (2010) reported that 75% of paramedics did not meet the Centers for Disease Control and Prevention recommendations for physical activity. Furthermore, an evaluation of regular physical activity in paramedic schools based on the transtheoretical model showed that only 10.5% of these individuals were in either the action or maintenance stage of physical activity (Charkazi et al., 2014). Willingness to engage in long-term physical training interventions may be low, therefore, rather than relying only on long-term interventions that require significant biological, psychological and social changes to daily living, informing changes in movement that typically lead to successful performance may be more effective. Even when significant determinants of performance can be successfully modified, this does not always guarantee a change in movement behaviour or capacity. Sheridan (2019) indicated that even after a reduction in BMI from a training intervention, paramedics still had difficulties completing task specific work. Furthermore, in a situation where the determinants are more resistant to change, such as height, modifications to coordination may be the only viable option to directly influence exposure and performance. Perhaps the movement differences can explain both the performance and injury differences between males and females. Gagnon and colleagues (2018) found that females had larger joint forces, larger compression and anterior-posterior shear forces compared to males when performing manual materials handling tasks. Although experience has been linked to injury rates, Gagnon and colleagues (2018) did not find these same compression and shear force differences when comparing expert and novice males. The fact that females exhibited distinct movement outcomes compared to males performing the same tasks in other studies (Makhoul et al., 2017; Plamondon et al., 2014) provides additional incentive to include kinematic comparisons when identifying determinants of paramedic performance for females. To better inform a movement evaluation, successful factors related to performance should still be considered and adapted in a kinematic comparison.

Movement comparisons should still account for other interacting factors when investigated performance differences. In Newell's model of constraints (1986), biological, task and environmental constraints are stated restrict available movement options rather than initiate movement. Those with different constraints, or determinants, will have access to degrees of freedom in their movements. Therefore, a successful strategy implemented by a tall and heavy-weight candidate may not be within the realm of accessible movements for a short and light-weight candidate. With movement decisions being identified as an additional key factor in performance, an evaluation of kinematics must account for determinants of performance that constrain movement ability. Comparing performances between those with specific constraints will dimensionally compress the potential coordination options, leaving the resulting movement behaviours more predictable (Kay, 1988). Through this dimensional compression, the strategies from the one group are more likely to overlap with the strategies that are possible for the other group. Comparing the kinematics of successful and unsuccessful performances, while being open to the interaction of other significant determinants, will not only help identify successful strategies but also aid in interpreting the transferability of these movements between groups. For example, if passing and failing female participants exhibit different movement strategies, but only when there are also differences in body mass and experience, the movement outcomes from the successful group may be a result of these demographic differences and limit the transfer of this movement outcome to the failing group. In contrast, if passing and failing participants exhibit different movement strategies, even when matching for other key demographics, this successful movement strategy may be readily accessible for failing group. However, before kinematic comparisons could be made, we must identify specific tasks and kinematic outcomes to compare, and a collection method that fits the research goal.

2.4.1 Tasks

Similar to the identification of key demographics to that contribute to performance, specific paramedic tasks need to be selected for a kinematic comparison. In a survey of active paramedics, Fischer, Sinden and MacPhee (2017) noted that patient transfer and stretcher lifting were ranked as the most difficult tasks during a shift. The ranking of task difficulty matches paramedic injury reports from Maguire et al. (2014), where lifting accounted for the greatest proportion of injuries. In a systematic review, the majority (64%) of injury were musculoskeletal (Kearney et al., 2022). More precisely, the muscular stress while lifting and/or carrying was found to be a cause of injury, with 37-44% of workers' compensation claims being due to muscular stress while performing lifts specifically (Maguire et al., 2014). Armstrong et al. (2020a) ranked biomechanical exposures of paramedic specific tasks, noting the Scoop board lift as eliciting the greatest peak compression and anterior-posterior shear about the low-back (Armstrong et al., 2020a). Despite involving a lighter load compared to stretcher lifts, the Scoop board lift was found to be a highly demanding task, likely because the lift originated from the floor. The combination of subjective ratings from paramedics, injury statistics, and elevated biomechanical exposures associated with these lifting tasks indicate that they are extremely demanding for paramedics. Within the OPPAT™, the lifting tasks appear to be an event of interest to evaluate kinematic performance. Additionally, the elevated demands of these lifts provides a significant task constraint on the performer, further limiting coordination options afforded to them. This limitation on the potential degrees of freedom aids in the identification and transfer of specific kinematic outcomes from successful performers to unsuccessful performers. In addition to tasks of interest, specific kinematic measures need to be identified for performance comparisons.

2.4.2 Outcomes

There are many ways an individual can coordinate their own movements to achieve a specific goal. Referencing previous investigations on paramedic and general lifting performance guided the selection of specific outcomes. Lentz et al. (2019) found that leg power was lower in emergency responders who became injured, notably those with lower jump heights were three times more likely to have an injury than those with the highest jump height. Makhoul and colleagues (2017) also noted that increased lower body joint power decreased the low-back spine moments. Influence on low-back moments is of particular interest since the systematic review from Kearny and colleagues (2022) revealed that 40% of the paramedic injuries were to the trunk. Considering this, lift duration as an analogous time-based measure instead of lower body power, was considered a useful dependent variable in comparing passing and failing performances.

Another difference in movement that has been noted between males and females is lower body joint coordination. Plamondon et al. (2014) found that females performing lifting and loading tasks had greater and positive mean relative phase angles between the knee and hip. This indicated more out-of-phase i.e. asynchronous joint movements, with a more knee-leading lift strategy. This research group also noted that this movement strategy is similar to that employed by novice lifters (Plamondon et al., 2014). This could be associated with the findings from Gagnon, Plamondon and Larivière (2018) that both female and novice workers experience greater spine loading while lifting. It has been suggested that, in addition to body size and strength deficits in females, the lack of synchronous technique commonly exhibited could explain the elevated injury risk (Gagnon et al., 2018). This prompted the incorporation of a measure of lower body coordination for the kinematic comparison between passing and failing females in the OPPAT™. With the tasks and outcomes of interest being identified, the method of capturing kinematics should also be evaluated.

2.4.3 Collection

The nature of this project evoked considerations regarding markerless motion capture as the method of kinematic data collection. Armstrong and colleagues (2020a) expressed concern over the lack of external validity from lab-based kinematic evaluations with paramedic tasks. Evaluating kinematics via traditional 3D motion capture can be challenging as it is often cumbersome, time consuming and expensive (Colyer et al., 2018; Ota et al., 2020). Markered motion capture requires a change in the performance environment and tasks, as well as adding encumbrances to the individual, all of which could elicit changes in their movement decisions. Markerless motion capture has been considered a more suitable method to capture movement patterns in contrast to a traditional markered approach due to the minimal interference imposed on the individual's movement (Colyer et al., 2018). Obtaining real-time kinematics through markerless motion capture provides insight into the movement decisions during the OPPAT™ without modifying the constraints that make the OPPAT™ a defensible physical employment standard. Markerless motion capture makes use of pose-estimation, which is an automated machine learning method used to predict human motion data in 2D or 3D space (McKinnon, Sonne, & Keir, 2020; Remedios & Fischer, 2021). *Google MediaPipe* has been proven to be feasible in measuring joint angles for biomechanical evaluation by recent research by Lafayette and colleagues (2022), where *MediaPipe* outputs showed stronger agreement with Qualysis Tracking Manager markered motion capture to other establish pose estimation tools such as Kinect V2, Intel and Astra. Compared to other pose estimators, *MediaPipe* had the lowest root mean squared error and highest Pearson correlation with the Qualysis data when comparing joint angles. These findings indicate superior agreement with traditional markered motion capture compared to other pose estimation systems. Concerns have been raised about the reliability of z-axis coordinates from not only *MediaPipe* (Lafayette et al., 2022) but other pose estimation tools as well (Remedios & Fischer, 2021). However, 3D data of movement were not considered mandatory for this novel investigation based on other kinematic comparison studies. As an example, despite the other kinematic differences between expert and novice lifters identified by Plamondon et al. (2010), no significant differences were found in the frontal plane. This provided confidence that using only 2D outputs from *MediaPipe* particularly for lifting tasks with greatest movement in the sagittal plane, would still allow comparisons to be made between pass and fail groups.

3 Research Questions and Hypotheses

The review of current literature revealed gaps in research surrounding female determinants of performance in paramedic duties. Females appear to be less likely to meet the demands of the job in certain physically demanding occupations compared to males. However, among females, we do not know which factors are differentiate with those who can and cannot meet the physical requirements of the job. Certain demographics have been linked to PES performance in other physically demanding occupations, but we do not know which factors are associated with success in a paramedic PES. Movement differences have been seen between females and males in paramedic tasks, but similar comparisons have not be conducted between groups of females. Additionally, a movement comparison between those who are successful and unsuccessful in a PES has yet to be completed. The described gaps in the literature informed the research questions and hypotheses below:

Research Question 1: Do heart rate response, height, body mass, BMI, employment status and/or paramedic college type predict the likelihood of success in the OPPAT™ among female candidates?

Hypothesis 1: It is hypothesized that height and employment status will remain predictors of success likelihood.

Research Question 2: How do lift duration and hip vs knee peak joint velocity timing differ when performing the Barbell Lift and Scoop Lift tasks between passing and failing females in the OPPAT™, while accounting for the interaction of significant determinants from Research Question 1.

Hypothesis 2: It is hypothesized that there will be a significant interaction between determinants from Research Question 1 and pass/fail groups, and that the pass group will have shorter lift duration and less time differences between hip and knee peak joint angular velocities.

4 Predictors of Test Outcomes for Females on the Ottawa Paramedic Physical Ability Test

Authors: Alexander. L. Malone[1], Daniel. P. Armstrong[1], Renée. S. MacPhee[2] and Steven. L. Fischer[1]

[1]Department of Kinesiology, Faculty of Health Sciences, University of Waterloo, Waterloo, Canada
[2]Kinesiology & Physical Education and Health Sciences, Wilfrid Laurier University, Waterloo, Canada

Highlights

- Employment status, college type, body mass and BMI were significant predictors
- The model only accounted for 11% of the variance in performance outcomes
- Increased work-related experience and effective mass may improve female performance
- Additional factors, such as fitness and technique, should be included in future models

4.1 Introduction:

Sex disparities exist in both employment demographics and injury prevalence data within the paramedic sector. Over the last decade, females comprised 21-23% of the paramedic workforce, with minimal year-over-year growth (Crowe et al., 2020). Despite accounting for less than a quarter the workforce, more than 50% of physical injuries claims are reported by female employees (Maguire, 2011). Elevated injury rates are already common in the paramedic workforce, with injury risk seven times greater than the national average based on an Australian study (Maguire et al., 2014), but there appears to be an even greater risk for female paramedics. While there are benefits to increasing the presence of underrepresented groups in paramedic services (Crowe et al., 2020), ongoing issues with injury risk must also be addressed. Considering that females are not well represented and are at a higher risk for injury, generating strategies to support females entering the paramedic profession will improve both the capacity and diversity of the paramedic workforce.

With the growth in paramedic call volumes exceeding the growth in population (Pasma, 2020), it is imperative that the paramedic workforce is fortified with more paramedics, as well as a workforce that is more representative and able to work injury free. Paramedics perform a variety of duties that are integral to providing pre-hospital care, but also impose elevated injury risk for these service providers. Most serious workers compensation cases in paramedics (37-44%) are due to muscular stress while lifting during occupational tasks (Maguire et al., 2014). One method to reduce the incidence of injury is to improve physical occupational performance (Harbin & Olson, 2005; Legge, Burgess-Limerick & Peeters, 2013). Work-related physical performance has been used to predict injury outcomes, where those with capacity that matched or exceeded the physical demands of a job were 30% less likely to experience injury (Harbin & Olson, 2005). In a seven-year longitudinal investigation of injury incidence and physical occupational performance, those whose job-related functional capacity did not meet the demands of their job were 2.3 times more likely to experience injury compared to their coworkers (Legge et al., 2013). Measuring job-related physical capacity and understanding which factors underpin physical performance in paramedic tasks can help inform exercise-based strategies to help optimize performance.

To adequately test an individual’s physical capacity to meet the demands of an occupation, evaluations should be specific to the requirements of the job (Harbin & Olson, 2005). Physical

Employment Standards (PES) are implemented by employers to ensure that candidates, and in some cases even current employees, possess the physical capacity required to meet the demands of the job (Reilly et al., 2016). For paramedics, a PES has already been established. The Ottawa Paramedic Physical Ability Test (OPPAT™) exists as a legally defensible PES through which physical readiness for duty can be evaluated for prospective and current paramedics (Sinden, MacPhee, & Fischer, 2015). By using performance on the OPPAT™ as an evaluation of physical readiness for paramedic work, and by extension, risk of occupational injury, we can identify which performers have the necessary physical capacity to meet the demands of the job, and which factors likely underpin their abilities. Once identified, learnings can be used to inform strategies that can be implemented to help individuals who cannot meet the physical demands of paramedic work.

The OPPAT™ reflects the specific physical demands of paramedic work, but females are less likely to pass than males (Armstrong et al., 2019c). While females make up the majority of injury cases (Maguire, 2011) and typically perform poorly on physical employment standards relative to males (Gumieniak et al., 2018; Jamnik et al., 2010), not every female becomes injured while performing paramedic duties. Therefore, it is important to investigate how modifiable and non-modifiable personal factors influence OPPAT™ performance within a female cohort. For instance, among paramedics, higher cardiovascular capacity was found to be associated with a decreased likelihood of injury (Poplin et al., 2014), and a lower cardiovascular response was associated with improved occupational performance (Barnekow-Bergkvist et al., 2004; Morgan, Allison & Duhon, 2012). Therefore, heart rate as a measure of cardiovascular function, may help explain pass/fail outcomes on the OPPAT™. However, structural physical features, such as lean body mass (Hydren, Borges & Sharpe, 2017) and/or height (Schenk et al., 2006) may also be strong predictors of success. When evaluating female paramedics performing simulated ambulance work, Barnekow-Bergkvist and colleagues (2004) found that including participant height significantly improved the ability a model to predict the onset of fatigue. Considering the evidence from previous investigations, a measure of cardiovascular fitness, height and body mass may prove to be important predictors of pass/fail outcomes among females completing the OPPAT™.

The aim of this study was to identify factors that predict pass outcomes among females performing the OPPAT™. Based on prior research, we hypothesized that stature and mass would predict successful OPPAT™ performance. Additionally, given that some paramedic candidates are educated within the public college system and others within the private (for-profit) system, we also explored if their educational program was an additional predictor of pass/fail outcomes. By increasing our understanding about how modifiable (i.e., heart rate, lean body mass, educational program) and non-modifiable factors (i.e., height) influence performance on a job-specific performance test, we can better develop targeted screening and coaching methods to maximize every candidate's potential to successfully complete the OPPAT™.

4.2 Methods:

4.2.1 Study Design, Participants and Data Collection

This study employed a retrospective analysis of data collected from female paramedic recruits and active duty paramedics completing OPPAT™ between January 2021 and February 2022. Before attempting the OPPAT™, incoming and incumbent paramedics were required to complete a PAR-Q+ and receive medical clearance from a doctor if required. Those deemed ready to safely attempt the OPPAT™ were asked to consent to have their demographic information, heart rate (HR) data and pass/fail outcomes recorded for research purposes. The collection and secondary analysis of OPPAT™ participant data was approved by the ethics review board from the University of Waterloo and Wilfrid Laurier University.

Self-report questionnaires were administered prior to OPPAT™ completion and were used to gather data including height (m), body mass (kg), college of study, and employment status, where height and weight were also used to calculate body mass index (BMI). The college of study and employment status of each participant were dichotomised as public or private, and employed or not yet employed, respectively. Employed included active duty in a paramedic service or patient transfer company. HR and performance outcomes were recorded during OPPAT™ completion, where HR was measured by using a Polar chest-strap sensor (Kempele, Finland). HR values were recorded after successful lifting events during the test (i.e., Barbell Lift, Scoop Lift and Carry). Note that the Barbell Lift was the first challenging physical lift in the OPPAT™ and, as a result, only the post-Barbell Lift HR was retained for statistical analysis. A portion of participants failed the OPPAT™ at the Barbell Lift stage so HR from later lifts were not considered as a result of missing data for those who did not reach those points in the test.

The OPPAT™ was designed as a series of three circuits of increasing physical demand that are completed successively. Circuits were designed to replicate call types (no patient transfer required, stair chair extrication, and scoop board extrication), where successful performance required participants to complete all three circuits within 10 and 17 minutes, without dropping or unsafely handling equipment, or stopping in the middle of an activity. More details about the OPPAT™ are reported in Armstrong et al., 2019c, where the test was based on physical demands as reported on in Fischer et al., 2017.

4.2.2 Statistical Analysis

A logistic regression was used to determine if height, body mass, BMI, post-Barbell Lift HR, college type and/or employment status could predict the likelihood of females to pass the OPPAT™. A backwards stepwise approach (alpha <0.05 to retain, alpha <0.1 to remove) was applied by using the log likelihood test ($p < 0.05$) to identify significant predictors. All statistical analyses were competed using R (The R Foundation for Statistical Computing, Vienna Austria).

4.3 Results:

Of the 211 eligible participants, 135 passed and 76 failed the OPPAT™. Descriptive statistics stratified by height, body mass, BMI, post-Barbell HR, college type, and employment status are displayed in Table 1.

TABLE 1 – Descriptive statistics (± standard deviation) and frequency counts for demographics and heart rate data stratified by pass/fail outcomes

	Grand Mean (± SD)	Pass Mean (± SD)	Fail Mean (± SD)	Pass – Fail Difference
Height (m)	1.65 ± 0.07	1.66 ± 0.07	1.63 ± 0.06	0.03
Body mass (kg)	70.0 ± 14.2	72.6 ± 14.9	65.4 ± 11.5	6.9
BMI (kg/m^2)	25.7 ± 5.0	26.3 ± 5.0	24.8 ± 4.7	1.5
Barbell HR (bpm)	173.4 ± 12.4	173.0 ± 12.5	174.1 ± 12.2	-1.2
	Total (count)	Pass (count)	Fail (count)	% pass
Public college	143	99	44	69%
Private college	68	36	32	52%
Employed	83	60	24	71%
Not yet employed	128	76	52	59%

The results of the logistic regression indicated that a model including college type, employment status, weight, and BMI best predicted successful test outcomes in the OPPAT™ ($p < 0.01$, $df = 3$, *McFadden's pseudo* $R^2 = 0.11$) (Table 2). The likelihood of passing the OPPAT™ was increased for those training in a public college, those currently employed and for every unit increase in body mass. However, conversely, the odds of passing the OPPAT™ were decreased for every unit increase in BMI.

TABLE 2: Logistic regression results to predict the odds of passing the OPPAT™ using significant predictors including school type, body mass, BMI and employment status.

	p-value	OR	CI 5%	CI 95%
College Type (public)	0.009*	2.36	1.38	4.10
Body Mass (kg)	0.002*	1.12	1.05	1.18
BMI (kg/m^2)	0.010*	0.82	0.71	0.93
Employment Status (employed)	0.043*	1.94	1.14	3.36

Note: *indicates significant difference ($p<0.05$) between means of pass and fail groups.

4.4 Discussion:

The objective of this study was to identify factors that predict successful performance of the OPPAT™ among female participants. It is well established that females have a lower likelihood of success in physical employment testing, hence our focus on a population of female candidates. College type, employment status, body mass, and BMI were retained as significant predictors of OPPAT™ performance, where active employment, public college program education and higher body mass were positively associated with success and higher BMI was negatively associated with success.

From the final model, college type and employment status had the greatest impacts on the likelihood of success, where those employed (OR 1.94) or those training in a public college (OR 2.36) were twice as likely to pass the OPPAT™ than those who were unemployed or were trained in a private college paramedic program. The impact of on-the-job experience (i.e., active employment) was expected considering working paramedics should have proven their ability to complete paramedic tasks to become employed, in addition to continuous task-specific exposure during the workday to maintain employment. Maguire (2011) found that most of the injuries in

EMTs and paramedics occurred in employees within their first 5 years of employment, and 30% occurring within the first year. Lack of exposure to job related demands could be a contributing factor. This could imply that more experience increases one's ability to perform tasks. However, the finding that college type was a significant predictor is novel and may have important implications regarding the availability of practical experiences with patient handling during paramedic training. While limited information exists to understand curricular differences between publicly funded colleges offering paramedic programs relative to private (i.e., for profit) college offerings, anecdotally, publicly funded programs have been noted as offering more hands-on training with actual equipment. In fact, several programs report that candidates are required to demonstrate lifting ability as part of the program requirements. In contrast, participants who trained in private college often noted that the OPPAT™ was the first time where they had operated a stretcher or lifted a scoop board. Considering the role of employment status and college type as predictors, it seems that strategies to improve hands-on experience which in turn can help to develop tasks specific strength could be useful during the training phase, prior to employment. Accounting for the performance benefits from increased hands-on exposures, paramedic college curricula could be reviewed to ensure candidates are accruing hands-on experience before entering the workforce, perhaps coupled with appropriately scoped strength and conditioning to help candidates improve their overall physical capabilities to perform high-demand paramedic tasks (Armstrong et al. 2019c).

Body mass was also identified as a significant predictor of OPPAT™ performance. Based the model, the likelihood of passing was increased by 12% for every kilogram increase in body mass (OR 1.12). While some caution may be warranted here, our findings are consistent with Barnekow-Bergkvist et al. (2004) who demonstrated that heavier paramedics were able to perform prolonged carrying tasks with less fatigue compared to lighter paramedics. At face value, this finding may suggest that increasing mass is a viable strategy to improve the likelihood of passing the OPPAT™, thus demonstrating the ability to work as a paramedic; however, the type of mass is likely critical. Roberts and colleagues (2016) note that the smaller stature and muscle mass of the average female compared to the average male are linked to the strength and power differences between sexes. Providing context based on physical abilities testing in the military, Reilly and colleagues (2016) found that the second best predictor of performance on various military tasks was lean body mass, with the best predictor being a metric which included the maximum load carried. Considering this, it is possible that the relationship between body mass and performance is observed because heavier individuals are stronger on average (i.e., more lean body mass and great carrying capacity). While our results show that increased mass may improve performance, we assert that increased lean mass, which is related to strength, is likely the underlying causes of improved performance in heavier participants. This hypothesis requires further testing in future research.

In contrast to the findings with body mass, BMI was found to be negatively associated with OPPAT™ success. For every unit increase in BMI, participants were 18% less likely to pass (OR 82). This finding matches previous research from Sheridan (2019) who, after conducting a meta-analysis, suggested that elevated BMI may be linked with decreased occupational performance. This is a particularly relevant finding for the paramedic population as Studnek et al. (2010) noted that over half of working EMS professionals are classified as “overweight” or “obese” based on their BMI score. In addition to the general health and wellness concerns with elevated BMI, there now appears to be a detrimental effect on physical occupational performance. Collectively interpreting our findings related to the effects of body mass and BMI on performance, simply increasing mass may not be a viable solution. Since mass is the numerator in the calculation of BMI, there appears to be a limit to the benefit of increasing ones’ mass for a given height, before the increase in BMI becomes a detriment. As previously mentioned, a potential performance benefit from increased mass is the increase in muscle mass and functional strength. Alterations in body composition can be achieved through resistance training interventions, as shown by Pawlak et al. (2015) who used physical training to reduce body fat percentage and BMI in military professionals. More importantly, physical training has also been linked to improved occupational performance. Both Pawlak et al. (2015) and Armstrong et al. (2019b) showed that introducing a resistance training program improved physical occupational performance in both military professionals and paramedics. Employing resistance training interventions to improve lean body mass and strength may be a viable solution to improve female paramedic performance.

Despite the noted significant predictors, the overall model fit accounted for 11% of the variance in outcomes (McFadden’s pseudo $R^2 = 0.11$). This suggests that other factors may play a more dominant role in predicting OPPAT™ performance among female participants. For comparison, when predicting performance in a law enforcement physical ability test, Dawes and colleagues (2017) found that three fitness tests scores (vertical jump, one-minute sit up test and

20-meter multistage shuttle run test) accounted for 54% of the variance in outcomes. The poor predictive ability of our current model indicates that additional factors need to be considered when predicting physical occupational performance in female paramedics, where those factors might be more closely aligned with underlying physical fitness constructs related to physically demanding paramedic work.

4.4.1 Limitations

This study relied on a secondary analysis with no ability to obtain alternative measures beyond those that were routinely captured during OPPAT™ performance. As a result, it was not feasible to measure height and mass directly. Measuring these demographics, rather than relying on self-reported values, could have influenced the fidelity of these recorded metrics. Additionally, we were not able to directly measure lean body mass or other characteristics of strength and fitness, where such variables could improve the ability to predictive performance.

4.4.2 Future Directions

Future research may also consider movement competency as a factor of interest in performing physical paramedic duties. Movement competency, or the ability to select a movement solution that minimizes exposure (e.g., keep the load close when lifting) or increases the bodies tolerance to withstand an exposure (e.g., minimize spine flexion when lifting), has been noted as an area that should be explored in paramedic research (Armstrong et al., 2019a). Differences in movement competency might explain why experience was a significant performance factor from our study. For example, when performing occupational tasks, novices and experts display distinctly different lifting techniques, with novices performing lifts with greater lumbar and upper trunk flexion angles (Plamondon et al., 2010), and a greater reliance on passive tissues rather than active muscle (Gagnon et al., 2018). Future research should include a kinematic comparison of successful and unsuccessful performances in the OPPAT™ to identify movement strategies that are favourable for meeting the physical demands of paramedic duties.

4.5 Conclusion:

Female paramedics continue to be underrepresented in the workforce, while simultaneously experiencing elevated rates of work-related injuries. This investigation found that body mass, BMI, college type and employment status were the strongest predictors of success for females in the OPPAT™; however, 89% of the variance in female performance was not accounted for by the predictors tested in the current model. To help develop and target interventions to improve the physical readiness for duty of female incumbent and working paramedics, more information is needed to uncover the driving factors that influence readiness for duty. As we continue to research further these data support that hands-on training / practice with lifting and strength training are options that can likely improve physical readiness for duty as measured by the successful completion of the OPPAT™.

5 Comparing lift duration and lower body coordination between females performing Barbell and Scoop lifts in the Ottawa Paramedic Physical Abilities Test™

Authors: Alexander. L. Malone[1], Daniel. P. Armstrong[1], Renée. S. MacPhee[2] and Steven. L. Fischer[1]

[1]Department of Kinesiology, Faculty of Health Sciences, University of Waterloo, Waterloo, Canada
[2]Kinesiology & Physical Education and Health Sciences, Wilfrid Laurier University, Waterloo, Canada

Highlights:

- No significant differences were found between passing and failing participants
- Markerless motion capture compared live physical employment standard performance
- Unsupervised machine learning could remove bias in future comparisons

5.1 Introduction

A sex disparity exists in both workplace representation and prevalence of injury among paramedics. More than 50% of physical injuries to paramedics are experienced by females (Maguire, 2011), despite comprising only one quarter of the paramedic workforce (Crowe et al., 2020; Maguire, 2011). In addition to higher injury rates, females have been found to have reduced success rate compared to males in physically demanding occupational tasks (Gumieniak et al., 2018; Jamnik et al., 2010). What could explain the elevated injury rates among females is a gap in the ability to meet the physical demands of the job. However, not every female employed in paramedic services will suffer an injury, nor does every female candidate fail their physical evaluation. Perhaps we can learn more by studying females who are successful as a mechanism to develop interventions for those who have been unsuccessful. To adequately support female paramedics, more information is needed to identify performance differences between successful and unsuccessful females during physically demanding paramedic tasks.

Physical employment standards (PES) are used to evaluate one's physical ability to perform work-related duties and are particularly common as entry-to-practice test prior to being hired within first responder work. PES are used by employers to ensure that prospective and current employees (i.e., those returning to work following an absence) have the physical capacity to perform job related tasks (Reilly, et al., 2016). In comparison to using general physical fitness scores, work related performance can be a better indicator of injury risk (Harbin and Olson, 2005; Legge, Burgess-Limerick & Peeters, 2013). A PES for paramedics already exists in the form of the Ottawa Paramedic Physical Abilities Test (OPPAT^{TM}). The OPPAT^{TM} is a legally defensible PES through which physical readiness for duty can be evaluated for prospective and current paramedics (Sinden, MacPhee, & Fischer, 2015).

Although the OPPAT™ provides an evaluation of the ability to complete paramedic duties, the outcome is binary (i.e., pass or fail). Using test outcomes alone may limit the ability to identify differences in strength capabilities or movements strategies that underpin to success. Further, by focusing on differences during specific high demands tasks, like lifting a scoop board (Armstrong et al., 2020a), or lifting a simulated stair chair, we might gain the greatest insights regarding movement strategy differences. Further, since muscular stress while lifting and/or carrying was found to be associated with injury among working paramedics, accounting for 37-

44% of workers compensation claims (Maguire et al., 2014), it remains prudent to focus on high demand lifting and carrying tasks.

Confounding variables such as demographic determinants of performance should also be considered, such as body mass, BMI and experience, when comparing differences in movement performance during lifting. In a recent evaluation of predictors for OPPAT™ performance among female candidates, both employment status and paramedic college type were found to be significant predictors of success (Malone, Armstrong, MacPhee & Fischer, 2023). In addition to pass-fail groupings, both employment status and paramedic college type should be included as factors that may inform movement difference during performances among females. Other anthropometric measures, such as body mass and BMI have been linked to occupational performance and were also found to be influential in paramedic physical performance (Malone et al., 2023). In addition to the accounting for previously established determinants of performance, we also wanted to capture authentic, unincumbered movements from current and future paramedics so our findings translate seamlessly into the field.

Real-time motion capture offers an opportunity to collect valuable data while also mitigating influences on ones' natural movement selection. Researchers have expressed concern over the lack of external validity from lab-based kinematic evaluations with paramedic tasks (Armstrong et al., 2020a; Gagnon et al., 2018). Capturing kinematics via traditional 3D motion capture can be challenging as it is often cumbersome, time consuming and expensive (Colyer et al., 2018; Ota et al., 2020). However, it is possible to obtain real-time movement data without modifying the constraints that make the OPPAT™ a legally defensible physical employment standard. Markerless motion capture has been considered a less restrictive method to capture movement patterns compared to a traditional markered approach due to the minimal interference imposed on the individual's movement (Colyer et al., 2018). The use of markerless motion capture to evaluate real-time PES performance is novel. Incorporating validated motion capture tools to obtain unencumbered, real-time kinematics increases the confidence that movements being compared are authentic and externally valid.

When comparing performance in the Barbell and Scoop Lift, previous research should guide the selection of dependent variables to best compare performance. Lentz et al., (2019) found that leg power was lower in emergency responders who become injured, and Makhoul et al. (2017) noted that increased lower body joint power was associated with a decrease in the low-back spine moments when paramedics performed lifting tasks. As an alternative measure that does not require a measure of force, lift duration has also been used when comparing occupational lifting performance between males and females (Lindbeck & Kjellberg, 2001). In addition to lift duration, the relative movement of joints during a lift can provide insight into performance differences. When comparing movement coordination between males and females, Plamondon et al. (2014) found that females adopted asynchronous joint coordination between the knees and hips during lifting tasks. Using a measure of movement coordination could provide additional insight on the technique differences employed between female performers who are and are not successful in performing the OPPAT™.

The objective of this study was to compare movement characteristics among successful and unsuccessful female participants in the OPPAT™ while accounting for factors that contribute to movement competency. We hypothesized that successful performers would lift faster and with less time between peak hip and knee joint angular velocities (i.e., more synchronous) when completing the Barbell Lift and the Scoop Board Lift. We also hypothesized that there would be a significant interaction between public/private college type and Pass-Fail status as categorical factors and body mass and BMI as covariates.

5.2 Methods:

A retrospective analysis of data from females completing the OPPAT™ between January 2021 and February 2022 was used. Participants included in this study needed to be eligible to attempt the OPPAT™, indicate female on their demographics form, complete both the Barbell Lift (Figure 1) and Scoop Lift (Figure 2) during the test, and consented to have their demographics and performance video used for research purposes. Additional OPPAT™ details can be found in Armstrong et al., 2019b. The collection and secondary analysis of OPPAT™ participant data was approved by the ethics review board from the University of Waterloo and Wilfrid Laurier University.


FIGURE 1: Barbell Lift from the OPPAT

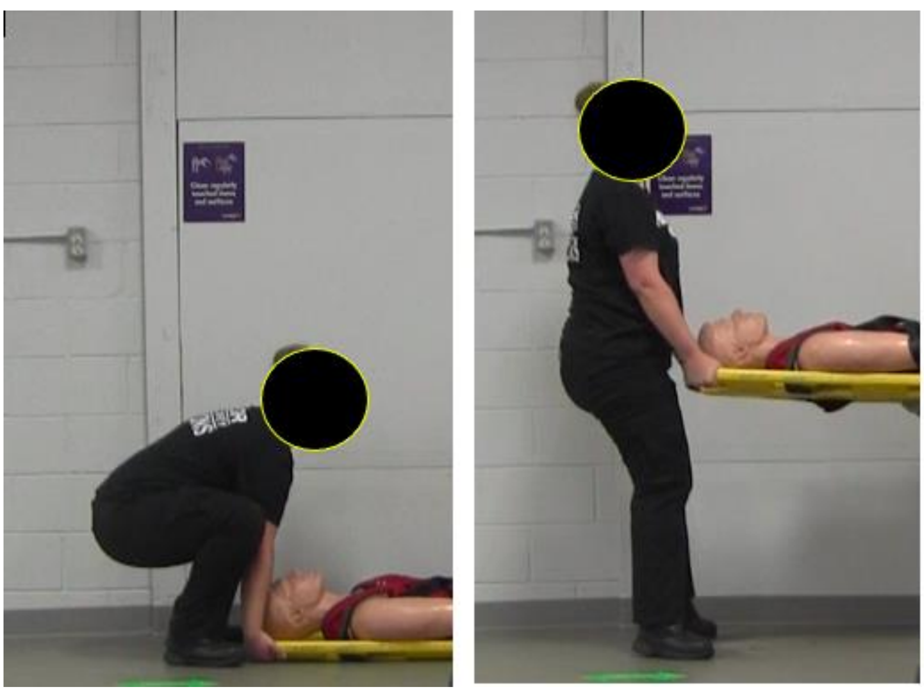


FIGURE 2: Scoop Lift from the OPPAT

5.2.1 Collection

Employment and paramedic college of training were self-reported in the intake questionnaire. Employment status and College type were dichotomized in Employed or Unemployed and Public College or Private College, respectively. Body mass was self-reported and BMI was calculated using the reported body mass and height. Pass-Fail status from the OPPAT™ was judged and provided by the OPPAT™ examiner. Barbell Lift and Scoop Lift performances were captured using a Canon VIXIA HFM40 Camera at frame rate of 30 Hz. In Python (Version 4.0.1), an open-access Google MediaPipe script was adapted as a pose estimation tool to produce kinematic outputs of estimated joint centres from the lifting videos.

5.2.2 Data Processing

From MediaPipe, 2D x-axis and y-axis trajectories for the ankle, knee, hip and shoulder joint centre data were extracted. Available marker locations are displayed in Figure 3. Raw joint position outputs were automatically gap filled and interpolated via the MediaPipe processing script. MediaPipe position outputs were then low-pass filtered through a dual-pass Butterworth filter with an effective cut-off frequency of 4 Hz. This cut-off frequency was selected in alignment with the cut-off frequency used by Armstrong and colleagues (2020a) when evaluating biomechanical exposure in paramedic lifting tasks.

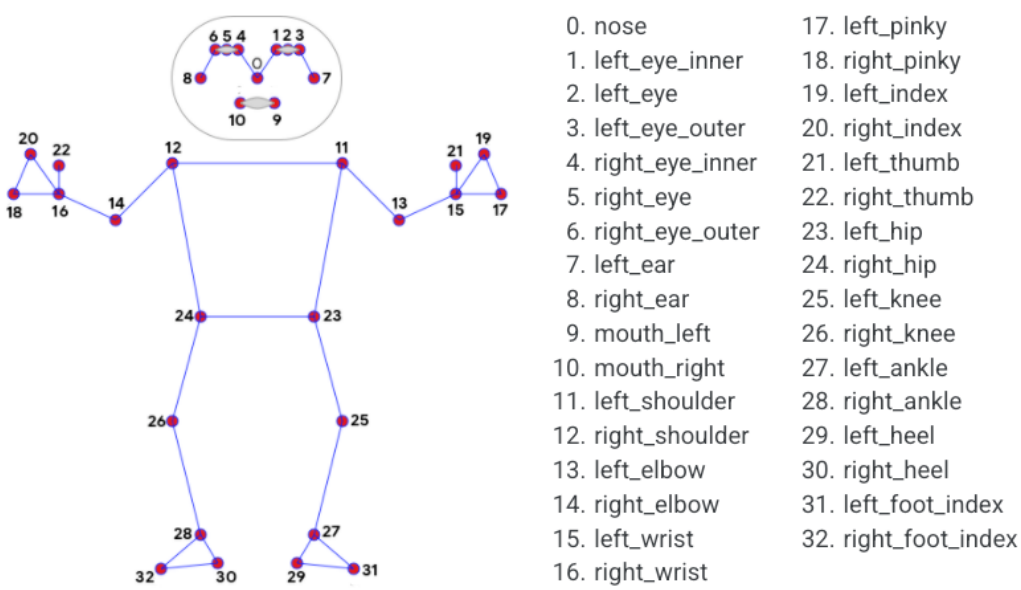


FIGURE 3: Joint centres and landmark outputs from Google MediaPipe Pose

5.2.3 Calculations

The start of each lift was defined as the instant when the shoulder joint experienced in increase in vertical (y-axis) velocity after reaching its minimum. The end of the lift was defined as the instant when shoulder vertical velocity returned to zero after the start of the lift. Lift duration was defined as the time difference between the start of the lift and the end of the lift. Linear vertical velocity of the shoulder joint was calculated using the central differences method.

Unilateral joint angles were created using joint position data from the limb most proximal to the camera (right). Hip joint angles were calculated as the angle between the thigh and torso segments, where the thigh was defined as the vector between the knee and hip joint position and the torso defined as the vector between the hip and shoulder joint position. Knee joint angles were calculated as the angle between the shank and thigh segments, where the shank was defined as the vector between the ankle and knee joint position. Joint angular velocity was calculated by differentiating continuous joint angle data using the central difference method. The time of the peak of the hip and knee joint angular velocity during the lift were identified, then the absolute difference in timings between these peaks was used as the peak hip vs. knee joint angular velocity timing variable (HvK). Greater difference in timing of these peaks represented more asynchronous movement.

Figure 4 shows a sample plot of a Barbell Lift performance where shoulder joint velocity and position are plotted against frame numbers in the x-axis, which frames indicate the start of the lift, the end of the lift, and when peak knee and hip joint angular velocities occurred during the lift. Figure 5 shows a sample plot from the same Barbell Lift performance from the start to the end of the lift where knee and hip joint angular velocity are plotted against frames numbers in the x-axis, as well as which frames indicate when peak knee and hip joint angular velocity occur.

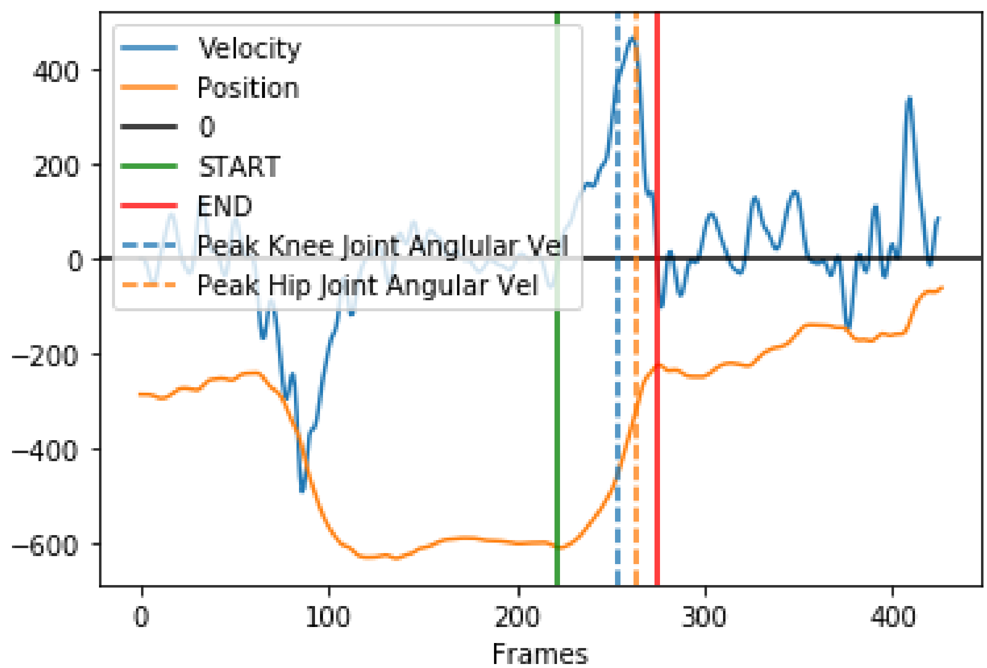


FIGURE 4: Sample plot of a Barbell Lift performance with vertical shoulder joint velocity (blue), position (orange), the start of the lift (green), the end of the lift (red), peak knee joint angular velocity during the lift (dashed blue), peak hip joint angular velocity during lift (dashed orange) with frame numbers on the x-axis

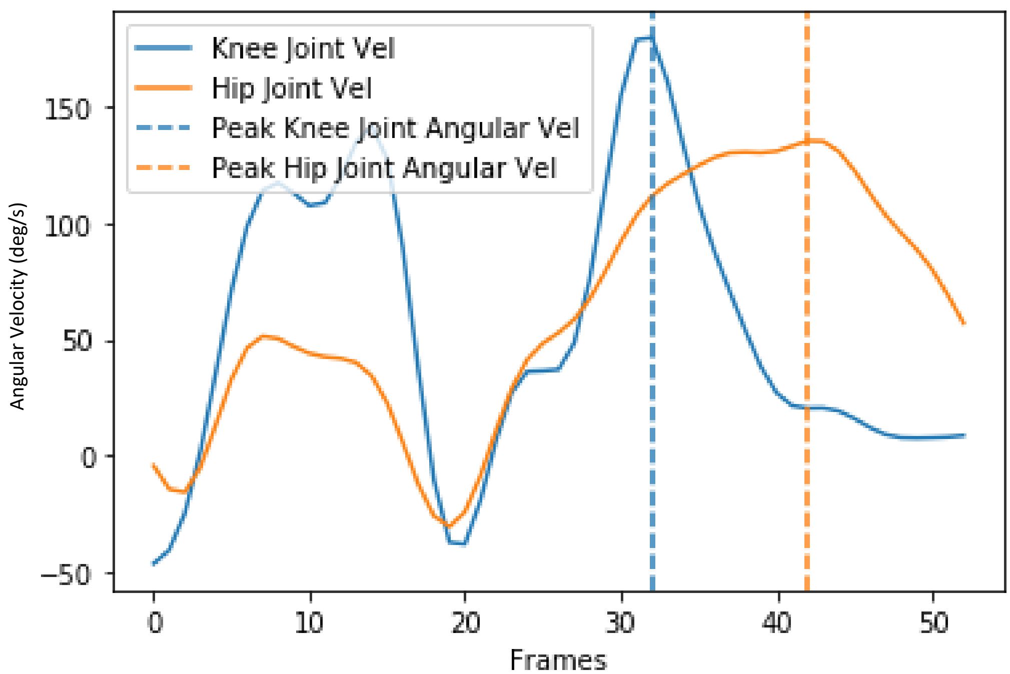


FIGURE 5: Sample plot of a Barbell Lift performance from start to end with joint angular velocities for the knee (blue) and hip (orange), the frame where the peak joint angular velocity for the knee (dashed blue) and hip (dashed orange) occur, and frame numbers on the x-axis

5.2.4 Statistical Tests

For the statistical analysis, four 3-way between groups factorial ANCOVAs were conducted to compare lift duration and HvK between Pass-Fail status, College type and Employment type groups, with body mass and BMI as covariates for the Barbell Lift and Scoop Lift ($p < 0.05$). Any significant differences identified by the ANCOVAs were compared using a pair-wise Mann-Whitney U test ($p < 0.05$). All statistical analyses were completed using R (The R Foundation for Statistical Computing, Vienna Austria).

5.3 Results:

Descriptive statistics are reported in Table 3 for the 137 participants that met the inclusion criteria for this study. The ANCOVA tests account for the relationship of continuous covariates (i.e. body mass and BMI) when comparing dependent variables (i.e. lift duration and Hip vs Knee Peak Joint Angular Velocity timing) between factor groups (i.e. Pass-Fail Status, College Type, Employment Status). Tables 4-7 show the ANCOVA comparisons from this study. The Main Effects portion of each ANCOVA table show results for differences in all determinants (factor groups and covariates) in relation to the dependent variable. The 2-way and 3-way interaction sections show difference results from comparing factor groups while accounting for the influence of the covariates. Tables 4 and 5 show ANCOVA results for the Barbell lift performance comparisons for lift duration and HvK, respectively. Tables 6-7 ANCOVA results for the Scoop lift performance comparisons for lift duration and HvK, respectively.

The tests performed in this study showed no significant differences between passing and failing groups in lift duration or lower extremity inter-joint coordination. The ANCOVA comparison of Lift Duration (s) for the Scoop Lift (Table 6 and Figure 6) identified a significant interaction between Pass-Fail status, college type and employment [$F = 5.879$, $p=0.017$]. After using Mann-Whitney U pairwise comparisons to decompose the interactions, one comparison remained statistically significant indicating that passing, private college graduate who were unemployed completed the Scoop lift faster than the failing, private college graduates who were unemployed (Table 8). However, the comparison group sizes and p-value adjustments from the interaction decomposition resulted in no significant differences in lift duration.

TABLE 3: Descriptive statistics (± standard deviation) and frequency counts for duration and hip vs knee peak angular velocity timing (HvK) stratified by Pass-Fail status, Public-Private College, and Employed-Unemployed status groups

		Barbell		Scoop	
	Count	Duration Mean ± SD (s)	HvK Mean ± SD (s)	Duration Mean ± SD (s)	HvK Mean ± SD (s)
Pass	94	1.63 ± 0.39	0.26 ± 0.29	1.68 ± 0.35	0.56 ± 0.37
Fail	43	1.73 ± 0.43	0.25 ± 0.23	1.68 ± 0.40	0.63 ± 0.42
Public College	94	1.70 ± 0.41	0.27 ± 0.28	1.71 ± 0.36	0.62 ± 0.38
Private College	43	1.59 ± 0.40	0.25 ± 0.26	1.63 ± 0.37	0.50 ± 0.39
Employed	57	1.66 ± 0.39	0.24 ± 0.27	1.63 ± 0.34	0.54 ± 0.36
Unemployed	80	1.66 ± 0.42	0.28 ± 0.28	1.72 ± 0.38	0.62 ± 0.40

TABLE 4: 3-Way Between Groups Factorial ANCOVA comparing Lift Duration (s) from the Barbell Lift between Pass-Fail status, College type, Employment type groups with Body Mass and BMI as covariates

	df	F Statistic	p value
MAIN EFFECTS			
Body Mass	1	0.292	0.590
BMI	1	0.849	0.358
Pass-Fail Status	1	2.517	0.115
College	1	2.753	0.100
Employment	1	0.206	0.651
2-WAY INTERACTIONS			
Pass-Fail Status x College	1	0.009	0.925
Pass-Fail Status x Employment	1	0.054	0.817
College x Employment	1	0.515	0.474
3-WAY INTERACTIONS			
Pass-Fail Status x College x Employment	1	0.130	0.719
Residuals	127		

*indicates significance (p<0.05)

TABLE 5: 3-Way Between Groups Factorial ANCOVA comparing HvK (s) from the Barbell Lift between Pass-Fail status, College type, Employment type groups with Body Mass and BMI as covariates

	df	F Statistic	p value
MAIN EFFECTS			
Body Mass	1	0.012	0.913
BMI	1	0.100	0.752
Pass-Fail Status	1	0.038	0.846
College	1	0.101	0.752
Employment	1	0.990	0.322
2-WAY INTERACTIONS			
Pass-Fail Status x College	1	3.465	0.065
Pass-Fail Status x Employment	1	0.091	0.763
College x Employment	1	0.034	0.855
3-WAY INTERACTIONS			
Pass-Fail Status x College x Employment	1	0.040	0.841
Residuals	127		

*indicates significance (p<0.05)

TABLE 6: 3-Way Between Groups Factorial ANCOVA comparing Lift Duration (s) from the Scoop Lift between Pass-Fail status, College type, Employment type groups with Body Mass and BMI as covariates

	df	F Statistic	p value
MAIN EFFECTS			
Body Mass	1	0.338	0.562
BMI	1	0.991	0.321
Pass-Fail Status	1	0.229	0.633
College	1	0.505	0.478
Employment	1	1.640	0.203
2-WAY INTERACTIONS			
Pass-Fail Status x College	1	0.057	0.812
Pass-Fail Status x Employment	1	3.004	0.085
College x Employment	1	0.337	0.562
3-WAY INTERACTIONS			
Pass-Fail Status x College x Employment	1	5.879	0.017*
Residuals	127		

*indicates significance (p<0.05)

TABLE 7: 3-Way Between Groups Factorial ANCOVA comparing HvK (s) from the Scoop Lift between Pass-Fail status, College type, Employment type groups with Body Mass and BMI as covariates			
	df	F Statistic	p value
MAIN EFFECTS			
Body Mass	1	1.047	0.308
BMI	1	0.358	0.551
Pass-Fail Status	1	0.142	0.707
College	1	2.562	0.112
Employment	1	0.776	0.380
2-WAY INTERACTIONS			
Pass-Fail Status x College	1	2.835	0.095
Pass-Fail Status x Employment	1	1.705	0.194
College x Employment	1	0.578	0.448
3-WAY INTERACTIONS			
Pass-Fail Status x College x Employment	1	2.094	0.150
Residuals	127		

*indicates significance (p<0.05)

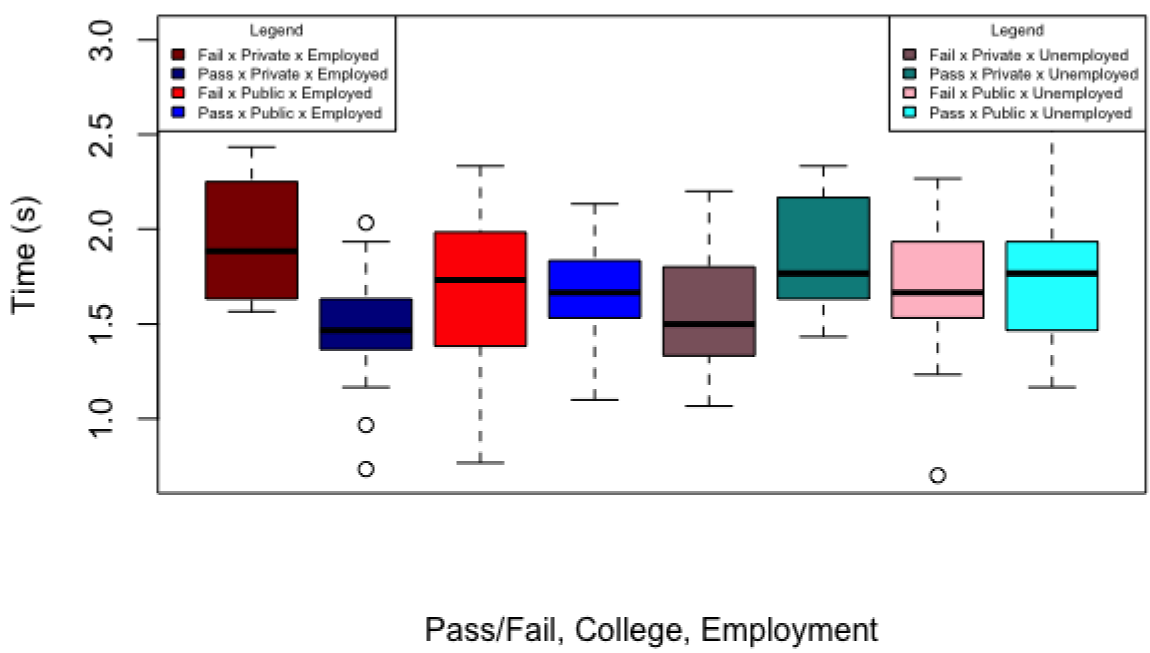


FIGURE 6: Scoop Lift Duration (s) for Pass/Fail, College Type and Employment Status Groups

TABLE 8: Pairwise Comparison using Mann-Whitney U Test for Lift Duration (s) in the Scoop Lift between Passing and Failing groups, who were from a Private College and Unemployed						
Groups	Count	Median	Median Difference	Stat	p-value	Benjamini-Hochberg Adjusted p-value
Fail x Private x Employed	4	1.88	0.41	59	0.028*	0.448
Pass x Private x Employed	17	1.47				

*indicates significance (p<0.05)

5.4 Discussion

The objective of this study was to assess whether differences existed when performing the Barbell and Scoop lift between females who passed or failed the OPPAT^{TM} while accounting for interacting factors such as paramedic college type and employment status as categorical variables, and body mass and BMI as covariates. Initially, there was significant difference in Lift Duration (s) during the Scoop Lift between pass and fail groups who were from a private college and employed, which aligned with previous occupation research. Studies from Hydren and colleagues (2017) and Lentz et al. (2019) showed greater lower body power was predictive of increased physical occupational performance and reduced injury risk, respectively. However, in the Mann-Whitney U comparison of duration between groups in the Scoop Lift the failing group contained only 4 participants (Table 8). This provoked concerns regarding the external validity of the sample outputs in this comparison. Additionally, p-value adjustments from the 3-way interaction resulted in a minimum of 16 planned comparisons (passing groups compared to failing groups), which would result a non-significant difference after a liberal adjustment from p = 0.028 to p = 0.448 using the Benjamini-Hochberg Adjustment (Table 8). We expected significant differences in lift duration and lower body joint coordination with the Barbell and Scoop Lift considering the final test results for passing and failing participants were different. However, the null findings in kinematic performance differences could be attributed to our selection of movements, dependent variables and/or motion capture technologies to compare between passing and failing female participants.

Focusing on Barbell and Scoop lift performance, while informed by previous research, may have limited our ability to identify significant differences in performance between successful and unsuccessful female participants. In alignment with other established PESs, outcomes from the

OPPAT™ are a binary pass or fail, therefore limiting our knowledge of the demand that each task from OPPAT™ imparts on the individual. Tasks requiring higher effort can dimensionally compress the number of degrees of freedom that a performer has access to when selecting a successful strategy to complete a task (Kay, 1988). For example, a passing individual whose capacity greatly exceeds the demands of the Barbell Lift could perform the lift in a variety of ways and still be successful. The movements selected across the passing individuals could have so much variety and overlap with failing individuals that they are indiscernible from failing individuals via statistical difference tests. Referring back to Table 3, the greatest mean difference in lift duration between initial categorical demographic groups is 0.11s, but the standard deviation of these values are 3 to 4 times greater than this difference. This is similar to largest HvK mean difference (0.12s) and the standard deviation range (0.23s-0.42s).

Another movement, the Stretcher Lift, has been noted as difficult task from the OPPAT™, which may further limit the available movement options for success. Previous studies have found that the loads in the average stretcher lift exceeds NIOSH lifting limits (Fischer et al., 2017) and, compared to other paramedic tasks, involves the greatest cumulative damage to the low back as estimated with the Lifting Fatigue Failure Tool (Armstrong, 2020a). This movement has also been used to evaluate performance differences paramedics in the past (Barnekow-Bergkvist et al., 2004; Makhoul et al., 2017). Since the null findings from this study the use of a stretcher lift may be more informative considering the heavy loads from this task, which may cluster the degrees of freedom available during performance, indicate that future research may seek to compare movement outcomes with this task rather than the Barbell and Scoop Lift.

The kinematic variables that were selected for comparison between pass and fail groups were informed, not only by previous research, but also by the availability and reliability of trajectories from markerless motion capture. Since paramedic performance is multifaceted and requires more exploration, we extrapolated previously used kinematic outputs and adapted them to fit the study parameters. While this provided confidence that the kinematic outcomes have been previously established as different, we did not explore other differences that may arise while attempting the selected OPPAT™ lifting tasks. The incorporation of markerless motion capture to obtain and compare kinematics from live PES performance was a novel approach to performance research but may have contributed to the null findings in this study. Wade and colleagues (2020) note a variety of opportunities and limitations regarding the use of current markerless motion capture for real-world evaluations. The use of a single camera (monocular), and reliability the training data sets have been mentioned as areas concern (Wade et al., 2020), and are relevant to our investigation. Among the concerns with monocular cameras, parallax error and z-axis data assumptions could have influenced the position data in this study. Based on visual observations, when performing the Barbell Lift, the participants appeared to move primarily in the x-y plane during their performance, but occasionally participants changed their position in the z-axis (moving closer or further from the camera) while lifting.

Since this study used only x-axis and y-axis outputs from MediaPipe, this movement in the z-axis to would change y-axis position values without adjustment based on z-axis data. As a participant moves closer to the camera all the estimated trajectories above the midline of the camera become have an increase y-axis value, while all trajectories below the midline are interpreted as a decreased y-axis position value, despite the reality of the participants y-axis positing relative to the ground remaining unchanged. As a result, these errors to delay the identification of lift termination based on the parameters described for this study. Figure 7 shows an example of this z-axis change resulting in alterations to y-axis position data. Additionally, monocular camera issues with z-axis readings have been noted to potentially contribute to joint centre location inaccuracies in 2D outputs (Wade et al., 2020). While the use of MediaPipe as a pose estimation tool has been proven to have agreement with markered motion capture and outperform other commonly used markerless motion capture technologies (Lafayette et al., 2022), the training data set that MediaPipe uses has been questioned before. Little is known about the processes employed in MediaPipe to label training data, leading researchers to question the influence of human error in annotating anatomical key points (Bittner et al., 2022). Bittner and colleagues (2022) also note that while the internal gap filling of position data from MediaPipe yields better results that OpenPose, there is still a lack of information regarding the interpolation techniques. However, there is hope with updates to existing markerless motion capture tools. Incorporating multicamera hardware, with appropriate software updates to account for multiple camera inputs, could mitigate issues with z-axis errors (Bittner et al., 2022; Wade et al., 2022). However, while multi-camera inputs may improve the accuracy of joint centre estimation and still leave the participants unencumbered, this may increase the awareness of evaluation and limit the external application. Additionally, further updates to training data could improve the accuracy and precision of markers, as well as expand the available markers.

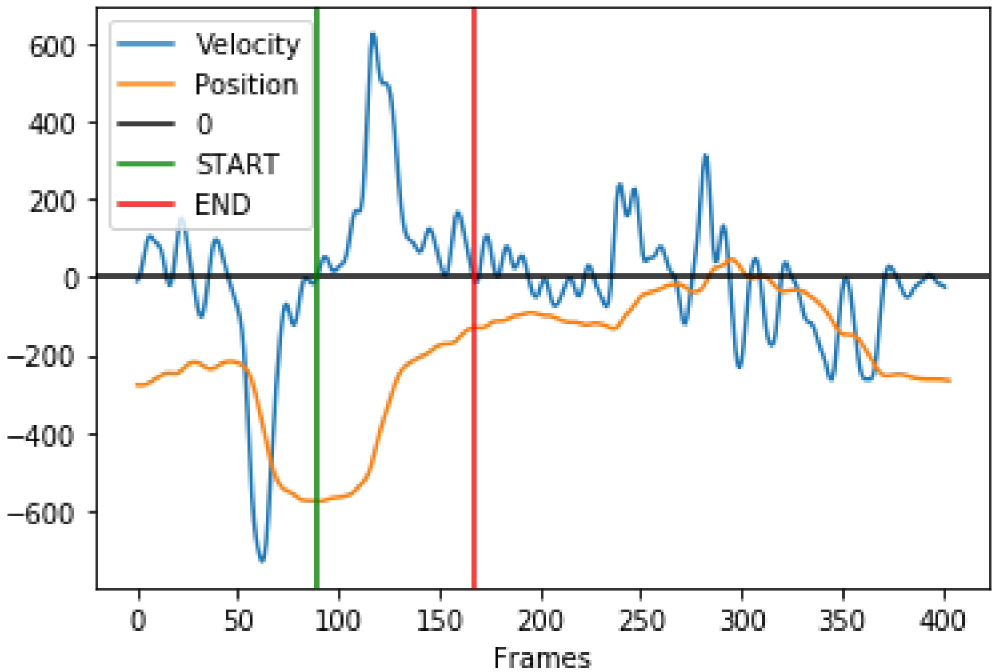


FIGURE 7: Sample plot of a Barbell Lift performance with shoulder joint velocity (blue) and position (orange), as well as lift start (green) and delayed lift end (red) times with frame numbers on the x-axis

5.4.1 Future Research

As an alternative, leveraging emerging machine learning techniques could remove the *a priori* dependent variable selection bias. Principal Component Analysis to reduce the feature space, followed by linear discriminant analysis is an example of an unsupervised machine learning approach suggested to aid in differentiating groups via kinematics and, in turn, inform targeted training interventions (Remedios et al., 2020). Rather than relying on the biases of the researcher and prior knowledge, unsupervised machine learning can be used to cluster data and identify patterns that best separate selected groups. Unsupervised machine learning has been used to successfully identify and cluster participants based on kinematic outcomes among individuals performing screening movements (Remedios et al., 2020) and lifting tasks (Beaudette et al., 2019; Hawley, Hamilton-Wright & Fischer, 2023). For this investigation, we know that there is a difference in performance that is resulting in some groups passing, while others fail, but we have limited information on what these differences are. Leveraging unsupervised machine learning to identify these differences may be a more appropriate method to identify dependent kinematic variables between passing and failing groups while this current gap exists in physical paramedic performance. However, along with the opportunities with emerging technologies, limitations should also be considered.

5.5 Conclusion

This study aimed to identify movement differences when lifting between females who passed and failed the OPPAT™ while accounting for additional factors known to influence physical performance. This study is the first to incorporate the use of markerless kinematics to capture movements during live PES performance. However, despite the novel approach, no significant differences were found between individuals who pass and fail when comparing lift duration or hip vs. knee peak joint angular velocity timing while accounting for the interaction of paramedic college type and employment status, as well as body mass and BMI as covariates. The null findings from this study may be a result of the tasks, dependent measures and/or the technology used to capture and compare performance differences between participant groups. Future research should explore the use a machine learning to identify dependent movement variables and seek to add validate the use of novel pose estimation technology to evaluate physical occupational performance.

6 General Discussion

The purpose of this thesis was to identify determinants of success for females completing the OPPAT^{TM}, and by extension, completing paramedic duties in the field. This thesis was divided into two research questions, where Research Question 1 sought to identify demographic determinants of OPPAT^{TM} success for female participants using heart rate response, height, body mass, BMI, paramedic college type, and employment status. Research Question 2 aimed to identify differences in lift duration and lower body joint coordination during the Barbell and Scoop lift between successful and unsuccessful female participants while accounting for significant predictors from Research Question 1. The findings from Research Question 1 indicated that there were demographics that significantly predict pass/fail outcomes for females performing the OPPAT^{TM}. While body mass, BMI employment status and paramedic college type were able to significantly predict pass/fail outcomes in the OPPAT^{TM}, these determinants accounted for less than a quarter of the variance in outcomes from the PES, suggesting that other determinants should be explored when predicting when females can complete the OPPAT^{TM}. Significant predictors from Research Question 1 were incorporated into Research Question 2, with paramedic college type and employment status as interacting categorical factors, and body mass and BMI as continuous covariates. Research Question 2 revealed no significant difference in lift duration or hip vs knee peak joint angular velocity timing between pass and fail females when performing the Barbell or Scoop lift from the OPPAT^{TM}.

6.1 Public Paramedic College

Findings from Research Question 1 align with previous literature on demographic determinants of physical performance differences, but the inclusion of paramedic college type is a novel contribution. There is preliminary evidence indicating that various anthropometric measures influence ones' ability to complete physical occupational tasks, such as elevated body mass and reduced BMI (Malone et al., 2023). Paramedic college type was a new inclusion which was found to be a significant predictor of female success in the OPPAT^{TM}, where those from a public paramedic program were more likely to successfully complete the OPPAT^{TM} compared to those with private college training. We suspect that this association with paramedic college type and OPPAT^{TM} success is attributed to greater practical physical experience with paramedic tasks from public paramedic college graduates. Practical experience with physical occupational task has been linked to performance outcomes in other sectors, which substantiates my claim. Sedliak and colleagues (2021) conducted a study that compared physical performance among military personnel between one group who was deployed for 6 months and a control group who were not deployed but had a guided training protocol with the support of a strength and conditioning professional. Upon re-testing anthropometrics and physical performance, the deployed group showed significant reductions in body fat percentage, shuttle run time and 5 kilometer run time, as well as an increase in their pull-up test scores, while the control group showed no significant changes (Sedliak et al., 2021). In alignment with our findings, additional practical experience in the deployed group appears to be beneficial to physical performance. However, this is not to say that more work is the solution for gaps in performance. In fact, the systematic review from Kearney et al. (2022) linked elevated paramedic call volumes with elevated rates of injury in the paramedic workforce, indicating that more work can have a detrimental effect. Perhaps exposure to specific work duties should be considered when evaluating this effect. In another evaluation of military personnel, risk of injury was evaluated after exposure to various military tasks and found that certain activities (road marching and obstacle courses) have a higher risk of injury than others (Lovalekar et al., 2021). With this in mind, Sedliak and colleagues (2021) provided additional context to their findings, noting that the physical demand of the deployed group was noted as lower compared to other studies (Sedliak et al., 2021), potentially explaining the increase physical performance rather than an increase in injury or decrease in performance in the deployed group. It has been suggested that there is a need to quantify work related exposure (Lovalekar et al., 2021) which could be beneficial in future research on determinants of performance. If exposures can be quantified, perhaps future research can find a balance between sufficient exposure for familiarity and adaptation before overexposing the individuals and increasing injury risk.

Previous research with PES explores a balance between exposure and performance, where additional exposure to PES testing has resulted in improved physical performance. Armstrong and colleagues (2019c) exposed working and future paramedics to the OPPAT™ 6 times over 7 days, resulting in female pass rate increasing from 71% to 90%. Similarly, Gumieniak, Gledhill and Jamnik (2018) compared completion time in a firefighting fitness circuit between a group exposed to the firefighting circuit 7 times, another group who had a physical training plan as well as the 7 exposures to the circuit, and a final control group with only 2 exposures to the circuit. While the control group did not show a significant change in circuit completion time, the circuit exposure group improved completion time by 12.2% (females) and 9.8% (males), and the group exposure to physical training and repeated bouts of the circuit improved by 19.8% (female) and 16.9% (male) (Gumieniak, Gledhill & Jamnik, 2018). The inclusion of paramedic college type as well as employment status in the likelihood of success for females in the OPPAT™, in addition to the effects of repeated testing in the occupational testing (Armstrong et al., 2019c; Gumieniak et al., 2018) and occupational exposure (Sedliak et al., 2021), indicate that a certain amount of exposure to occupation specific tasks will be beneficial in occupational performance. Accounting for the suggestions from Lovalekar and colleagues (2021) and our own findings, future research should investigate quantification of task exposures and injury risk in addition to the performance benefits of occupational practice.

6.2 Variance in OPPAT Outcomes

Additional exposure to occupational tasks could be beneficial, however, this may be a result of improved physical capacity rather than just task familiarization alone. Reflecting on the findings from Gumieniak et al. (2018), the group that was exposed to physical training in addition to repeated attempts with the firefighting circuits showed a greater improvement in performance compared to the group with repeated circuit exposure alone. Stimuli from both the physical training and circuit attempts could summate into an increase in overall capacity and improve performance. Linking to another component of Research Question 1 that should be expanded on in future research, the limited Pseudo R^2 could be supplemented with the addition of physical capacity testing scores in the predictive demographics. While the logistic regression model including body mass, BMI, college type and employment status was able to significantly predict the likelihood of success among female participants, these determinants only accounted for about 10% of the variance in outcomes (Malone et al., 2023). This could be a limitation from the specific determinants we selected, and we believe that physical paramedic performance is dictated by more than anthropometrics and demographics. Previous research has found physical capacity scores to be some of best predictors of performance. For example, Reilly et al. (2016) conducted a similar study with military personnel and their physical employment standard. While lean body mass was a significant predictor of performance, the strongest predictor incorporated a measure of lifting capacity (Reilly et al., 2016). A measure of lifting capacity could be helpful in determining determinants of successful performance and connect to some of the most challenging paramedic duties i.e., the lifting tasks (Armstrong et al., 2020a; Fischer, Sinden, MacPhee, 2017). Other physical capacity measures could also provide insight into what else determines successful physical occupational performance. When predicting time to complete a physical employment standard for police officers, Eduardo and colleagues (2019) found that agility test scores yielded an R^2 of 0.45. Additionally, combining relative aerobic power, upper limb strength and agility yielded an R^2 of 0.81 (Eduardo et al., 2019). Physical capacity tests could be used to improve the ability of future models to predict success in complete physical occupational tasks and inform interventions.

The use of a training program to improve participants physical capacity could be beneficial. A 4-week training program was successfully implemented between attempts at the OPPAT™, where the intervention group increased grip strength and peak lower body power, as well as reduced time to complete the OPPAT™ (Armstrong et al., 2019b). Improved physical capacity has been linked to injury risk as well as performance. Among police officers, grip strength scores had a linear, positive relationship to passing the TACOPS (tactical operations) test completion time and were inversely related to injury risk during the study (Orr et al., 2017). These findings, in tandem with the limited pseudo R^2 from Research Question 1, indicate that physical capacity may narrow the gap in our understanding of determinants for physical performance in paramedic related tasks. However, this is not to say that physical capacity is the only component to be focused on when seeking to improve occupational performance and reduce injury risk. In a systematic review, Lentz and colleagues (2019) found limited evidence for a relationship between physical fitness test scores and injury risk. Physical performance, like injury risk, is multifactorial, so while physical training interventions and task specific exposure can aid in performance, researchers should be open to exploring various domains in the pursuit of performance determinants for paramedic work.

6.3 Incorporate Emerging Technology

To aid the exploration of performance differences in the paramedic workforce, future investigations should also consider the benefits and limitations of emerging technology such as markerless motion capture and machine learning. Research Question 2 leveraged markerless motion capture through Google MediaPipe. Previous investigations have indicated strong agreement between kinematics from gold-standard markered motion capture and outputs from MediaPipe. For example, MediaPipe outputs for the lower body and upper body were classified as “Excellent” and “Good” agreement, respectively, with regards to Qualysis markered system (Lafayette et al., 2022). Although researchers have suggested that MediaPipe can be confidently applied to biomechanical evaluation of joint angles (Lafayette et al., 2022), our own validation of the MediaPipe outputs raised some concerns in using the software outputs for kinematic comparisons (Appendix A). In a preliminary comparison for our investigation, the root mean squared differences between MediaPipe and Qualysis showed statistically significant correlation between the mocap systems but the joint angle root mean square difference ranged from 6.92 to 8.95 degrees, with the standard deviation of the error ranging from 7.17 to 11.19 degrees (see Appendix A Table 9). In contrast, Wade and colleagues (2022) indicated that difference between markerless and markered motion capture larger than 3-11 degrees can be considered too large for real world applications for kinematic comparisons for gait. Additionally, our Bland-Altman plots (see Appendix A Figures 8-16) indicated an inconsistent error in the average and maximum joint angle selection. The RMSD and Bland-Altman data from our preliminary validation created hesitation when interpreting the comparison results in Research Question 2 due to both the magnitude in inconsistent direction of the error. As markerless motion capture technologies continue to develop, their incorporation into movement comparisons should be met with some caution. While MediaPipe is considered to be in alignment with gold standard motion capture technology and outperform other markerless indicated gaps in alignment that could have influenced kinematic comparisons in Research Question 2.

The use of machine learning to cluster performance data is another opportunity where emerging technology could change how motion capture informs comparisons between successful and unsuccessful groups in physical paramedic performance. The dependent variables used for comparison in Research Question 2 were adapted from other measures of power and coordination based on the limitations of our minimally intrusive design and the available data from our selected motion capture system. Using previous research added to our confidence in selecting dependent variables, but both the need to adapt these measures to fit our parameters and the limited pool of previous research in physical paramedic performance indicate a need for alternative methods to conduct performance comparisons at this stage. Remedios and colleagues

(2020) suggest that machine learning may be beneficial to identify and screen movement patterns that differentiate groups and inform targeted training interventions. These researchers were able to successfully identify distinct movement clusters among individuals performing typical screening movements such as a deep squat and hurdle step using unsupervised machine learning (Remedios et al., 2020). Considering the existing gaps in knowledge around the kinematics in successful paramedic performance, reliance on previous movement comparison studies may not be the appropriate method. Alternatively, removing the bias of the researcher and allowing unsupervised machine learning algorithms to identify and cluster real-time performance data may uncover strategies that have yet to be identified.

6.4 Conclusion

This thesis included an investigation of the demographic determinants of performance for females in the OPPAT™ as well as a comparison of lift duration and lower body joint coordination between females that passed and failed the evaluation. This research is the first to identify demographics linked to physical occupational performance for paramedics, as well as the first to compare live performances in physical employment standard. When testing the likelihood of passing the OPPAT™ among females remaining significant predictors of success included body mass, BMI, employment, and paramedic college type. Of the significant predictors from Research Question 1, paramedic college type was a novel discovery where public college graduates had a higher likelihood of success compared to private college graduates. This could be linked to a higher exposure to paramedic occupational tasks in public colleges, but future research should explore this further. Our hypothesis for Research Question 1 expected that height and employment status would be remaining predictors of success likelihood, so we failed to reject the null hypothesis. When comparing differences in lift duration and lower body coordination no significant differences were found between females who passed and failed the OPPAT™ when performing the Barbell or Scoop Lift. Our hypothesis for Research Question 2 expected that significant interactions between selected factors and the pass/fail groups, as well as significantly shorter lift duration and time between hip and knee peak joint angular velocities during the Barbell and Scoop Lift for the pass group. Therefore, we failed to reject the null hypothesis. Future research should continue to advance markerless motion capture technologies and leverage machine learning systems to identify performance differences between successful and unsuccessful performers.

References

Armstrong, D. P., Makhoul, P. J., Sinden, K. E., & Fischer, S. L. (2020a). Ranking Stretcher and Backboard Related Paramedic Lifting Tasks Based on Their Biomechanical Demand on the Low Back. *IISE Transactions on Occupational Ergonomics and Human Factors*, *8*(1), 9-19.

Armstrong, D. P., Pegg, C. E., & Fischer, S. L. (2020b). Is deep squat movement strategy related to floor-to-waist height lifting strategy: implications for physical employment testing. *Ergonomics*, *63*(2), 152-162.

Armstrong, D. P., Ross, G. B., Graham, R. B., & Fischer, S. L. (2019a). Considering movement competency within physical employment standards. *Work*, *63*(4), 603-613.

Armstrong, D. P., Sinden, K. E., Sendsen, J., MacPhee, R. S., & Fischer, S. L. (2019b). Evaluating the effect of a strength and conditioning program to improve paramedic candidates' physical readiness for duty. *Work, 63*(4), 623-633.

Armstrong, D. P., Sinden, K. E., Sendsen, J., MacPhee, R. S., & Fischer, S. L. (2019c). The Ottawa Paramedic Physical Ability Test: test-retest reliability and analysis of sex-based performance differences. *Ergonomics, 62*(8), 1033-1042.

Avesta Systems Inc & the American Ambulance Association. (2019). *AAA/Avesta 2019 Ambulance Industry Employee Turnover Study*. https://ambulance.org/wp-content/uploads/2019/07/AAA-Avesta-2019-EMS-Employee-Turnover-Study-Final.pdf

Barnekow-Bergkvist, M., Aasa, U., Ängquist, K. A., & Johansson, H. (2004). Prediction of development of fatigue during a simulated ambulance work task from physical performance tests. *Ergonomics*, *47*(11), 1238-1250.

Bittner, M., Yang, W. T., Zhang, X., Seth, A., van Gemert, J., & van der Helm, F. C. (2022). Towards Single Camera Human 3D-Kinematics. *Sensors*, *23*(1), 341. https://doi.org/10.3390/s23010341

Charkazi, A., Fazli, L., Alizadeh, F., Fazelnia, A., Koochaki, G. M., & Bakhsha, F. (2014). Regular physical activity based on transtheoretical model among health and paramedic schools of Golestan University of medical sciences. *Journal of health education and health promotion, 1*(4), 57-68.

Colyer, S. L., Evans, M., Cosker, D. P., & Salo, A. I. (2018). A review of the evolution of vision-based motion analysis and the integration of advanced computer vision methods towards developing a markerless system. *Sports medicine-open*, *4*(1), 1-15.

Courtney, J. A., Francis, A. J., & Paxton, S. J. (2010). Caring for the carers: Fatigue, sleep, and mental health in Australian paramedic shiftworkers. *The Australasian Journal of Organisational Psychology*, *3*, 32-41.

Crowe, R. P., Krebs, W., Cash, R. E., Rivard, M. K., Lincoln, E. W., & Panchal, A. R. (2020). Females and minority racial/ethnic groups remain underrepresented in emergency medical services: a ten-year assessment, 2008–2017. *Prehospital Emergency Care*, *24*(2), 180-187.

Dawes, J. J., Lindsay, K., Bero, J., Elder, C., Kornhauser, C., & Holmes, R. (2017). Physical fitness characteristics of high vs. low performers on an occupationally specific physical agility test for patrol officers. *The Journal of Strength & Conditioning Research*, *31*(10), 2808-2815.

Eduardo, F. M., Leo, C., Charles, B., J. Jay, D., & Del Vecchio, F. B. (2019). Aerobic fitness, upper-body strength and agility predict performance on an occupational physical ability test among police officers while wearing personal protective equipment. *The Journal of sports medicine and physical fitness*, *59*(11), 1835-44.

Fischer, S. L., Sinden, K. E., & MacPhee, R. S. (2017). Identifying the critical physical demanding tasks of paramedic work: Towards the development of a physical employment standard. *Applied ergonomics*, *65*, 233-239.

Gagnon, D., Plamondon, A., & Larivière, C. (2018). A comparison of lumbar spine and muscle loading between male and female workers during box transfers. *Journal of biomechanics*, *81*, 76-85.

Gumieniak, R. J., Gledhill, N., & Jamnik, V. K. (2018). Physical employment standard for Canadian wildland firefighters: examining test–retest reliability and the impact of familiarisation and physical fitness training. *Ergonomics*, *61*(10), 1324-1333.

Harbin, G., & Olson, J. (2005). Post-offer, pre-placement testing in industry. *American journal of industrial medicine*, *47*(4), 296-307.

Hawley, S. J., Hamilton-Wright, A., & Fischer, S. L. (2023). Detecting subject-specific fatigue-related changes in lifting kinematics using a machine learning approach. *Ergonomics, 66*(1), 113-124.

Hydren, J. R., Borges, A. S., & Sharp, M. A. (2017). Systematic review and meta-analysis of predictors of military task performance: maximal lift capacity. *Journal of strength and conditioning research*, *31*(4), 1142-1164.

Jamnik, V. K., Gumienak, R., & Gledhill, N. (2013). Developing legally defensible physiological employment standards for prominent physically demanding public safety occupations: a Canadian perspective. *European journal of applied physiology*, *113*, 2447-2457.

Jamnik, V. K., Thomas, S. G., & Gledhill, N. (2010). Applying the Meiorin Decision requirements to the fitness test for correctional officer applicants; examining adverse impact and accommodation. *Applied Physiology, Nutrition, and Metabolism*, *35*(1), 71-81.

Kay, B. A. (1988). The dimensionality of movement trajectories and the degrees of freedom problem: A tutorial. *Human Movement Science*, *7*(2-4), 343-364.

Kearney, J., Muir, C., & Smith, K. (2022). Occupational injury among paramedics: a systematic review. *Injury prevention*, *28*(2), 175-184.

Kranz, C., Lee, K., Jadhav, P., Vestlin, L., Barker, M., Jacques, A., & Netto, K. (2021). Kinematic and perceptual responses in heavy lifting and pulling: Are there differences between males and females?. *Applied Ergonomics*, *90*, 103274.

Lafayette, T. B. D. G., Kunst, V. H. D. L., Melo, P. V. D. S., Guedes, P. D. O., Teixeira, J. M. X. N., Vasconcelos, C. R. D., Teichrieb, V. & da Gama, A. E. F. (2022). Validation of Angle Estimation Based on Body Tracking Data from RGB-D and RGB Cameras for Biomechanical Assessment. *Sensors*, *23*(1), 3.

Legge, J., Burgess-Limerick, R., & Peeters, G. (2013). A new pre-employment functional capacity evaluation predicts longer-term risk of musculoskeletal injury in healthy workers: a prospective cohort study. *Spine*, *38*(25), 2208-2215.

Lentz, L., Randall, J. R., Gross, D. P., Senthilselvan, A., & Voaklander, D. (2019). The relationship between physical fitness and occupational injury in emergency responders: A systematic review. *American journal of industrial medicine*, *62*(1), 3-13.

Lovalekar, M., Hauret, K., Roy, T., Taylor, K., Blacker, S. D., Newman, P., Yanovich, R., Fleischmann, C., Nindl, B.C., Jones, B., & Canham-Chervak, M. (2021). Musculoskeletal injuries in military personnel—Descriptive epidemiology, risk factor identification, and prevention. *Journal of Science and Medicine in Sport*, *24*(10), 963-969.

MacQuarrie, S. (2019). *Fit for duty: context and correlates of paramedic health status and job* *performance.* [Unpublished doctoral dissertation]. Charles Sturt University.

Maguire, B. J. (2011). Transportation-related injuries and fatalities among Emergency Medical Technicans and Paramedics. *Prehospital and disaster medicine*, *26*(5), 346.

Maguire, B. J., Hunting, K. L., Guidotti, T. L., & Smith, G. S. (2005). Occupational injuries among emergency medical services personnel. *Prehospital Emergency Care, 9*(4), 405-411.

Maguire, B. J., O'Meara, P. F., Brightwell, R. F., O'Neill, B. J., & Fitzgerald, G. J. (2014). Occupational injury risk among Australian paramedics: an analysis of national data. *Medical journal of Australia*, *200*(8), 477-480.

Maguire, B. J., & Smith, S. (2013). Injuries and fatalities among emergency medical technicians and paramedics in the United States. *Prehospital and disaster medicine*, *28*(4), 376-382.

Makhoul, P. J., Sinden, K. E., MacPhee, R. S., & Fischer, S. L. (2017). Relative contribution of lower body work as a biomechanical determinant of spine sparing technique during common paramedic lifting tasks. *Journal of applied biomechanics*, *33*(2), 137-143.

Malone, A. L., Armstrong D. P., MacPhee, R. S., & Fischer, S. L. (2023) *Predictors of Test Outcomes for Females on the Ottawa Paramedic Physical Ability Test^{TM}*. [Unpublished manuscript].

McKinnon, C. D., Sonne, M. W., & Keir, P. J. (2022). Assessment of joint angle and reach envelope demands using a video-based physical demands description tool. *Human Factors*, *64*(3), 568-578.

Morgan, M. V., Allison, S., & Duhon, D. (2012). Heart rate changes in functional capacity evaluations in a workers' compensation population. *Work*, *42*(2), 253-257.

Newell, K. M. (1986). Constraints on the development of coordination. In M. G. Wade & H. T. A. Whiting (Eds.), Motor development in children: Aspects of coordination and control (pp. 341–360). Dordrecht: Martinus Nijhoff.

Orr, R., Pope, R., Stierli, M., & Hinton, B. (2017). Grip strength and its relationship to police recruit task performance and injury risk: A retrospective cohort study. *International journal of environmental research and public health*, *14*(8), 941.

Ota, M., Tateuchi, H., Hashiguchi, T., Kato, T., Ogino, Y., Yamagata, M., & Ichihashi, N. (2020). Verification of reliability and validity of motion analysis systems during bilateral squat using human pose tracking algorithm. *Gait & posture*, *80*, 62-67.

Pasma, C. (2020). *Under pressure: a statistical report on paramedic services in Ontario.* Ontario: Canadian Union of Public Employees. https://cupe.ca/under-pressure-statistical-%20report-paramedic-services-ontario

Plamondon, A., Denis, D., Delisle, A., Larivière, C., Salazar, E., & IRSST MMH research group. (2010). Biomechanical differences between expert and novice workers in a manual material handling task. *Ergonomics*, *53*(10), 1239-1253.

Plamondon, A., Lariviere, C., Denis, D., St-Vincent, M., Delisle, A., & IRSST MMH Research Group. (2014). Sex differences in lifting strategies during a repetitive palletizing task. *Applied ergonomics*, *45*(6), 1558-1569.

Poplin, G. S., Roe, D. J., Peate, W., Harris, R. B., & Burgess, J. L. (2014). The association of aerobic fitness with injuries in the fire service. *American journal of epidemiology*, *179*(2), 149-155.

Reilly, T., Spivock, M., Prayal-Brown, A., Stockbrugger, B., & Blacklock, R. (2016). The influence of anthropometrics on physical employment standard performance. *Occupational Medicine*, *66*(7), 576-579.

Remedios, S. M., Armstrong, D. P., Graham, R. B., & Fischer, S. L. (2020). Exploring the application of pattern recognition and machine learning for identifying movement phenotypes during deep squat and hurdle step movements. *Frontiers in Bioengineering and Biotechnology*, *8*, 364.

Remedios, S. M., & Fischer, S. L. (2021). Towards the use of 2D video-based markerless motion capture to measure and parameterize movement during functional capacity evaluation. *Journal of Occupational Rehabilitation*, *31*, 754-767.

Roberts, D., Gebhardt, D. L., Gaskill, S. E., Roy, T. C., & Sharp, M. A. (2016). Current considerations related to physiological differences between the sexes and physical employment standards. *Applied physiology, nutrition, and metabolism*, *41*(6), S108-S120.

Schenk, P., Klipstein, A., Spillmann, S., Strøyer, J., & Laubli, T. (2006). The role of back muscle endurance, maximum force, balance and trunk rotation control regarding lifting capacity. *European journal of applied physiology*, *96*(2), 146-156.

Sedliak, M., Sedliak, P., & Vaara, J. P. (2021). Effects of 6-month military deployment on physical fitness, body composition, and selected health-related biomarkers. *The Journal of Strength & Conditioning Research*, *35*(4), 1074-1081.

Sheridan, S. (2019). Paramedic health status, fitness and physical tasks: A review of the literature. *Australasian Journal of Paramedicine*, *16*, 1-7.

Sinden, K. E., MacPhee, R., & Fischer, S. (2015). The Ottawa Paramedic Physical Ability Test (OPPAT): A Review of Process and Development. In *Proceedings of the 2nd International Conference on Physical Employment Standards*. Canmore, Alberta, August.

Studnek, J. R., Bentley, M., Mac Crawford, J., & Fernandez, A. R. (2010). An assessment of key health indicators among emergency medical services professionals. *Prehospital Emergency Care*, *14*(1), 14-20.

U.S. Bureau of Labor Statistics. (2022, November 9). EMPLOYER-REPORTED WORKPLACE I NJURIES AND I LLNESSES – 2021 [Media release]. https://www.bls.gov/news.release/pdf/osh.pdf

Wade, L., Needham, L., McGuigan, P., & Bilzon, J. (2022). Applications and limitations of current markerless motion capture methods for clinical gait biomechanics. *PeerJ*, *10*, e12995.

Appendix A – MediaPipe Validation Test with Paramedic Tasks

Prior to Research Question 2, a preliminary validation of markerless motion capture trajectories in comparison to gold-standard markered motion capture trajectories was conducted. This analysis used 3D passive, optoelectrical motion capture system (Qualisys: Gothenburg, Sweden) to collect landmark position data during Scoop Board and Stretcher lifting tasks for paramedics. 2D video was collected during the performance occupational lifts performed by two paramedics. Video data was available for 9 Scoop Board lifts and 4 Stretcher lifts, for total of 13 lifts. MediaPipe Pose was used as a pose-estimation software to extract joint position data from the video recordings.

Joint angles were defined as the angle between adjacent segments. The knee joint was defined as the angle between the shank, a rigid segment created between the ankle joint and the knee joint, and the thigh, a rigid segment created from the knee joint to the hip joint. The hip joint was defined as the angle between the thigh and the torso, a rigid segment created from the hip joint to the shoulder (MediaPipe Pose) or C7 (Qualisys). Angular velocities were calculated using the central difference method. Only 2D (x, y) outputs were used to define segments and calculate joint angles and angular velocities. The start of each lift was defined as the instant where the vertical velocity over 0 frames/second of the shoulder (MediaPipe Pose) or C7 (Qualisys) exceeded 0 frames/s after achieving a minimum velocity value. The end of each lift was defined as the first instance where the vertical lift velocity returned to 0 frames/s after the start of the lift. Lift duration was defined as the time (s) between the start and end of each lift.

Average and maximum joint angles were extracted and compared for the knee (KJA) and hip (HJA). Additionally, average and maximum joint angular velocity were extracted and compared for the knee (KJAV) and hip (HJAV). Lift duration was also compared. Comparisons were made using the Root Mean Squared Difference (RMSD) of outputs from each measurement system to quantify magnitude of differences between each motion capture tool. A Spearman Rank Correlation test (r) was used to measure the relationship between outputs of the two motion capture methods (Table 9). Additionally, Bland-Altman plots were used to evaluate the range and consistency of agreement between the motion capture tools (Figures 8-16).

TABLE 9: RMSD and Spearmen Rank Correlation of outputs between 3D Markered Motion Capture from Qualisys and 2D Markerless Motion Capture from MediaPipe

Kinematic Outputs	RMSD ± SD	Spearman Rank Correlation		
		r		p-value
Average KJA (deg)	7.47 ± 9.88	**0.95**		**< 0.01***
Max KJA (deg)	6.92 ± 7.17	-0.20		0.52
Average HJA (deg)	8.95 ± 9.89	**0.96**		**< 0.01***
Max HJA (deg)	8.13 ± 11.19	0.39		0.19
Average KJAV (deg/s)	16.83 ± 29.89	**0.84**		**< 0.01***
Max KJAV (deg/s)	30.61 ± 36.39	0.57		0.04
Average HJAV (deg/s)	7.02 ± 11.43	**0.89**		**< 0.01***
Max HJAV (deg/s)	18.15 ± 23.75	**0.82**		**< 0.01***
Average TA (deg)	8.97 ± 6.78	0.61		0.03
Lift Duration (s)	0.21 ± 0.23	**0.86**		**< 0.01***

* Asterisked p-values indicate significance with an alpha = 0.01

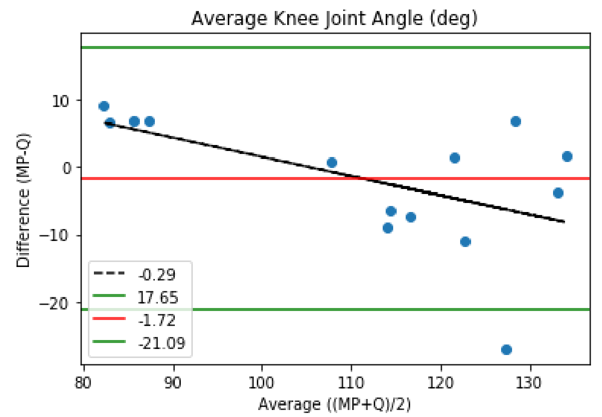


FIGURE 8: Average Knee Joint Angle Bland-Altman plot comparing MediaPipe (MP) and Qualisys (Q) outputs

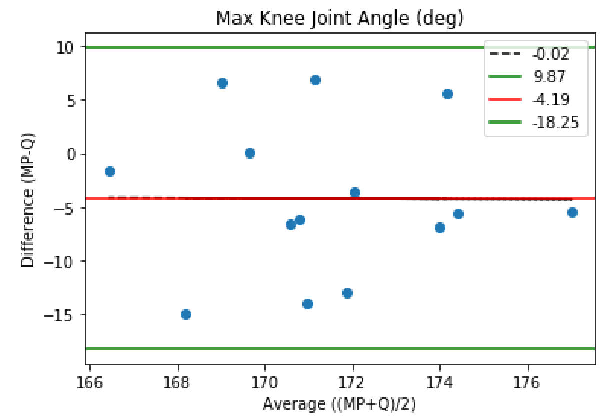


FIGURE 9: Max Knee Joint Angle Bland-Altman plot comparing MediaPipe (MP) and Qualisys (Q) outputs

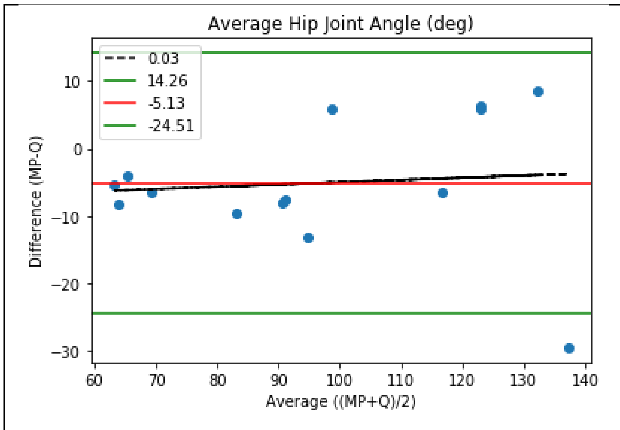


FIGURE 10: Average Hip Joint Angle Bland-Altman plot comparing MediaPipe (MP) and Qualisys (Q) outputs

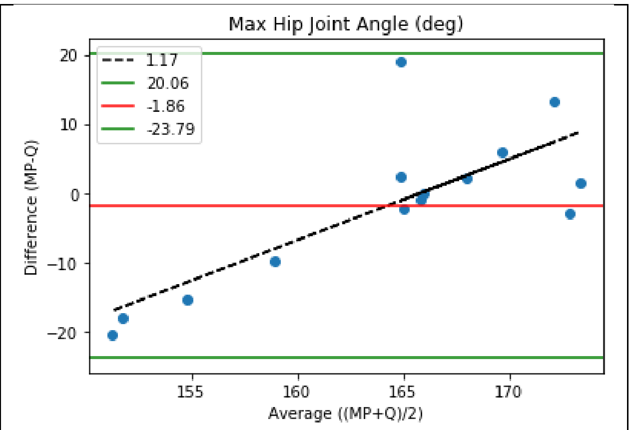


FIGURE 11: Max Hip Joint Angle Bland-Altman plot comparing MediaPipe (MP) and Qualisys (Q) outputs

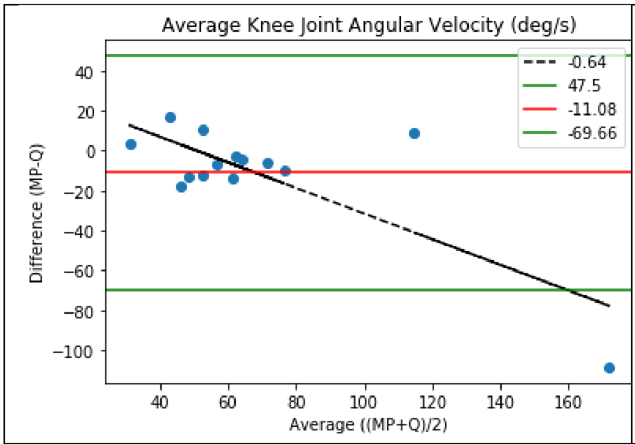


FIGURE 12: Average Knee Joint Angular Velocity Bland-Altman plot comparing MediaPipe (MP) and Qualisys (Q) outputs

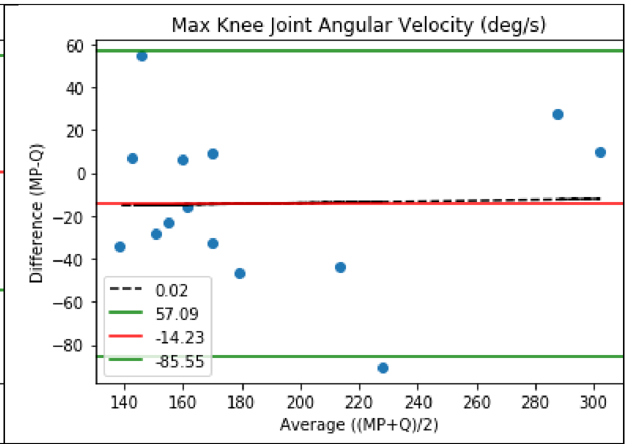


FIGURE 13: Max Knee Joint Angular Velocity Bland-Altman plot comparing MediaPipe (MP) and Qualisys (Q) outputs

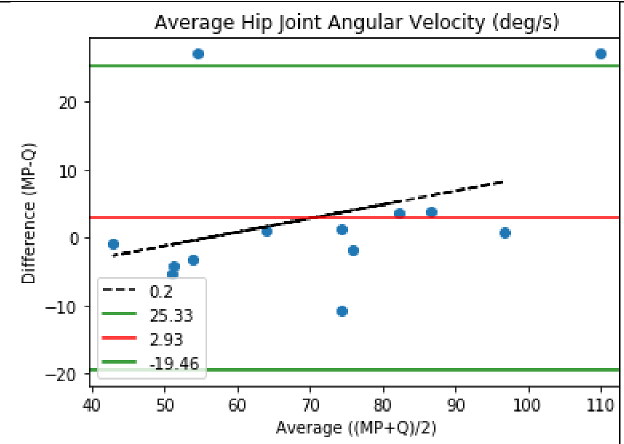


FIGURE 14: Average Hip Joint Angular Velocity Bland-Altman plot comparing MediaPipe (MP) and Qualisys (Q) outputs

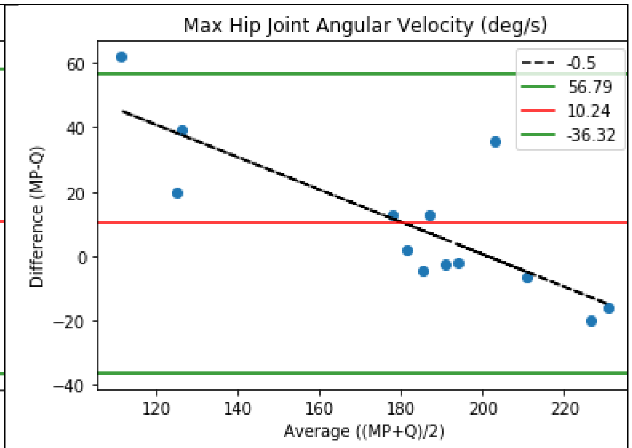


FIGURE 15: Max Hip Joint Angular Velocity Bland-Altman plot comparing MediaPipe (MP) and Qualisys (Q) outputs

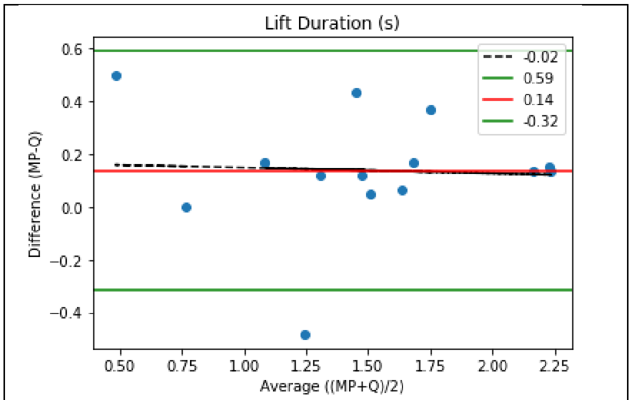
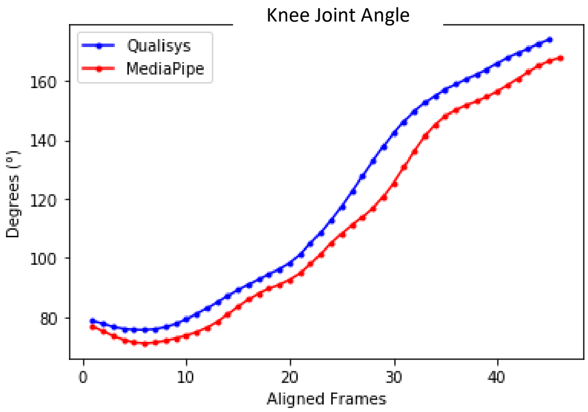


FIGURE 16: Lift Duration Bland-Altman plot comparing MediaPipe (MP) and Qualisys (Q) outputs

In addition to the RSMD and Bland-Altman comparisons, Figures 17-20 display continuous joint angle and joint velocity comparisons between markered Qualisys motion capture and markerless MediaPipe pose estimation outputs for one individual performing a Scoop lift.



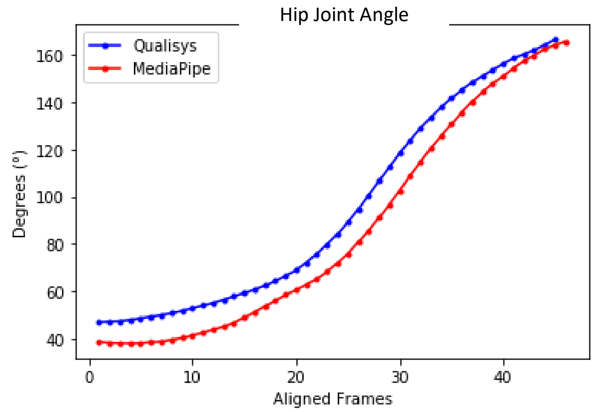
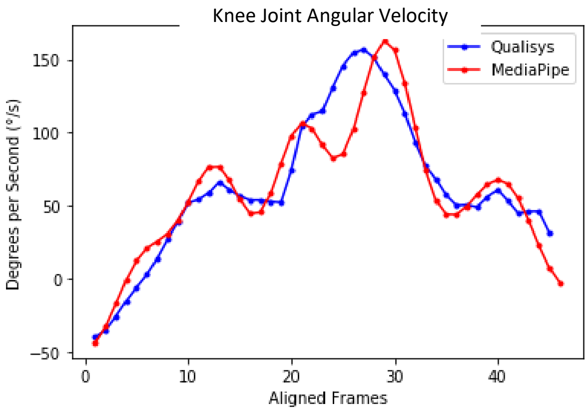


FIGURE 17: Sample comparison of continuous Knee Joint Angle outputs between MediaPipe and Qualisys	FIGURE 18: Sample comparison of continuous Hip Joint Angle outputs between MediaPipe and Qualisys



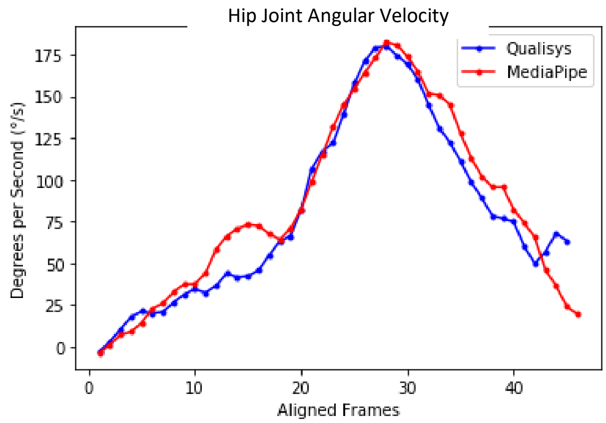


FIGURE 19: Sample comparison of continuous Knee Joint Angular Velocity outputs between MediaPipe and Qualisys	FIGURE 20: Sample comparison of continuous Knee Joint Angular Velocity outputs between MediaPipe and Qualisys