

**The Potential for 3D Depth Cameras to Automatically Evaluate Independent Wheelchair
Transfer Techniques**

by

Lin Wei

M.S. Mechanical Engineering, University of Pittsburgh, 2013

M.S. Biotechnology, National Dong Hwa University, 2009

Submitted to the Graduate Faculty of the
Click to choose your school in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy

University of Pittsburgh

2021

UNIVERSITY OF PITTSBURGH

School of Health and Rehabilitation Sciences

This dissertation was presented

by

Lin Wei

It was defended on

March 15, 2021

and approved by

Dan Ding, PhD, Associate Professor, Department of Rehabilitation Science and Technology

Leming Zhou PhD, DSc, Associate Professor, Department of Health Information Management

Brooke A. Slavens, PhD, Associate Professor, Department of Occupational Science and
Technology, University of Wisconsin-Milwaukee

Cheng-Shiu Chung, PhD, Human Engineering Research Laboratories

Dissertation Director: Alicia M Koontz, PhD, RET, ATP, Professor, Department of
Rehabilitation Science and Technology

Copyright © by Lin Wei

2021

The Potential for 3D Depth Cameras to Automatically Evaluate Independent Wheelchair Transfer Techniques

Lin Wei, MS

University of Pittsburgh, 2021

Wheelchair users rely heavily on their upper extremities to complete common but essential activities of daily living such as getting in and out of bed, and transferring to a toilet, a shower, and a car seat. The use of good transfer mechanics to avoid pain and injury is important for wheelchair users when performing transfers. The Transfer Assessment Instrument (TAI), is a tool developed to evaluate the transfer technique and help clinicians and users to recognize deficits in the technique. However, there are some limitations when therapists use the TAI as an assessment tool. These barriers decrease the usability of the TAI in clinical settings. An artificial intelligence system that can automatically score the TAI may potentially reduce the barriers associated with TAI's usability. We aim to develop a system that can watch a patient transfer and allow for automating the TAI using marker-less motion capture technology and machine learning algorithms that classify the motions into proper and improper techniques.

Machine learning algorithms were developed and trained using data from 91 full-time wheelchair users to predict proper (low risk) and improper (high risk) wheelchair transfer techniques in accordance with eleven TAI item scores. The transfer data was split into training set (80%) and testing set (20%). The training set was used for classifier selection and model tuning. The test set was excluded from all training processes. Three k-nearest neighbors (KNN) and eight random forest classifiers were selected for 11 TAI items based on model performance. The area

under the receiver operating characteristic curves (AUCs) are .83 to .99 for the training set and .79 to .94 for the test set. In order to avoid the false positive case (i.e. participant performed improper technique but the transfer is labelled as a proper transfer by the classifier), we tuned the models to achieve high precision. The precisions of the models are .87 to .96, and the recalls are .61 to .93.

For a system to automate the scoring of the TAI the system must also be able to distinguish the “setup phase” and “lift phase” of the transfer. On the TAI 4.0, items 1 to 6 are in the wheelchair setup skill group and Items 7 to 15 are in the body setup and flight/landing skill groups. In order to extract the features of each item, the motion data during the transfer needs to be separated into a setup phase and lift phase. We applied and compared a biomechanical variable based threshold method and an ML algorithm to automatically distinguish the time frames of the transfer phases. For the threshold method, the peaks observed in the linear displacement and velocity of one joint center marked by the Kinect, SPINE_BASE, were used for phase delineation. For the ML method, we trained a KNN classifier using 35 features from 81 participant’s transfer data recorded by Kinect. Using the KNN model, each time frame of the transfer was labeled as belonging to either the “setup” or “lift” phase. After further applying a filter algorithm, the method was used to identify the start and end timepoints of the transfer phases. We found that the ML method had less error in identifying the phase times but the threshold method spends less computational time in identifying the points. Although the threshold errors were larger this method had higher accuracy for predicting the TAI scores for items 10, 11, 12, 13, 14, 15 (lift phase items). The ML method had higher accuracy for predicting the TAI scores for items 1, 2 and 7 (wheelchair and body setup items). For items 8 and 9, the two methods showed equal performance. The ML method tended to undershoot the end phase times of the lift phase so it’s possible that tuning the algorithms to

include more of the lift phase data could increase the accuracies of the TAI item scores that deal with the lift phase biomechanics. This will continue to be an area of future work.

Due to the discontinuation of the Kinect v2, we aimed to find another 3D depth sensor that could track full body motion for future research. The Intel® RealSense has superior technical properties in relation to the Kinect v2 and has shown excellent performance for tracking facial and hand motions in previous studies. Although Intel did not make a full body joint tracking algorithm for the sensor at the time of our study, a 3rd party one (Nuitrack™) was available that can be used with a variety of 3D sensor models including RealSense. This solution enabled us to create the same biomechanical features with the RealSense as we had created with the Kinect to quantify transfer technique. To further understand the potential for RealSense to be used as a viable substitute sensor for capturing wheelchair technique biomechanics, we compared the measurement properties of the two sensors. We assessed intra-rater reliability for each sensor, and evaluated the inter-rater reliability and agreements between the Kinect and RealSense with 30 wheelchair users who performed multiple independent transfers (150 trials total). The study found that the Kinect had higher intra-rater reliability than the RealSense for measuring four key kinematic variables related to the wheelchair transfer technique. For the agreement analysis, more than 95% of the data points fell within ± 1.96 standard deviation of the mean differences. However, the inter-rater reliability between two sensors was poor. The low reliability of the RealSense may be due to the lack of robustness of the 3rd party algorithm for skeletal tracking of sitting postures and in general in comparison to the more extensively tested and developed Kinect SDK. Intel just recently introduced a full body skeletal tracking model for their sensors. It's possible that this version of the RealSense SDK may help increase the reliability for future applications.

Packaging these outcomes together into a user-friendly system could aid therapists and patients in identifying harmful motions and learning proper evidence-based transfer practices. After using a 3D Depth Cameras to watch a wheelchair transfer, the system would be able evaluate the TAI more reliably than a therapist rater would and generate objective feedback to the users. Therefore, the results of the current study could increase the usability and feasibility of TAI in a clinical setting.

Table of Contents

Preface.....	xvii
1.0 Introduction.....	1
1.1 Prevalence of Upper Extremities Pain and Injuries in Wheelchair Users	2
1.2 Wheelchair Transfer and Secondary Injuries.....	4
1.3 Biomechanics of the Wheelchair Sitting Pivot Transfer	5
1.4 Transfer Assessment Instrument – An Evaluation Tool for Determining Quality of Transfer	6
1.5 Marker-less Motion Capture	7
1.6 3D Depth Sensors in Research.....	9
1.7 3D Depth Sensing Applications in Rehabilitation Science and Human Motion Evaluation.....	11
1.8 Significance	13
1.9 Thesis Objectives.....	14
2.0 Automating the Clinical Assessment of Independent Wheelchair Sitting Pivot Transfer Techniques.....	16
2.1 Abstract.....	16
2.2 Introduction.....	17

2.3 Methods.....	19
2.3.1 Participants.....	19
2.3.2 Study Protocol	20
2.3.2.1 Experimental Setup	20
2.3.2.2 Transfer Protocol and Evaluation.....	21
2.3.2.3 Data Labeling	22
2.3.2.4 Features.....	23
2.3.4.5 Model Performance Evaluation.....	29
2.4 Results	32
2.4.1 Participants Demographics	32
2.4.2 Model Performance	33
2.5 Discussion.....	36
2.6 Limitations.....	43
2.7 Conclusions.....	44
3.0 Comparing Two Automated Methods to Detect Sitting Pivot Transfer Phases	45
3.1 Introduction.....	45
3.2 Methods.....	48
3.2.1 Participants.....	48
3.2.2 Study Protocol	49
3.2.2.1 Experimental Setup	49
3.2.2.2 Transfer Protocol and Evaluation.....	50
3.2.3 Data Analysis.....	50

3.2.3.1 Visual Method of Phase Delineation	50
3.2.3.2 Threshold Method of Phase Delineation.....	51
3.2.3.3 Machine Learning Method of Phase Delineation.....	53
3.2.3.4 TAI Score Predictions.....	55
3.2.4 Statistical Analysis	56
3.2.4.1 Aim 1: To determine the overall accuracy of each method in predicting the start and end points of lifting phases.....	56
3.2.4.2 Aim 2: To evaluate the accuracy of each method in predicting the TAI item scores	57
3.3 Results	58
3.3.1 Participants.....	58
3.3.2 KNN Model performance.....	59
3.3.3 Start/End timepoints of lift phase (Aim 1).....	59
3.3.4 Transfer Quality Evaluation (aim2).....	64
3.4 Discussion.....	65
3.5 Limitation	68
3.6 Conclusion	69
4.0 Comparison of Two Depth Cameras for Capturing Upper Body Motions During Wheelchair Transfers	70
4.1 Introduction.....	70
4.2 Methods.....	72
4.2.1 Participants.....	72

4.2.2 Study Protocol	73
4.2.2.1 Motion Sensors and Experimental Setup	73
4.2.2.1 Transfer Protocol	74
4.2.3 Data Analysis.....	75
4.2.3.1 Key Variables	75
4.2.3.2 Statistical Analysis	77
4.3 Results	77
4.3.1 Participant	78
4.3.2 Intra-Rater (Sensor) Reliability.....	78
4.3.3 Inter-Rater (Sensor) Reliability and Agreement	79
4.4 Discussion.....	82
4.5 Limitation	86
4.6 Conclusion	86
5.0 Conclusion	88
5.1 Future Works	91
Appendix A. General Questionnaire for Wheelchair Users.....	95
Appendix B. Transfer Assessment Instrument 4.0	96
Appendix C. Model training process and results of each TAI 4.0 item	103
Item 1 Wheelchair Distance	103
Item 2 Wheelchair Angle.....	107

Item7 Feet Placement	111
Item 8 Hip Scooting	115
Item 9 Leading Arm Position before Transfer	119
Item 10 Push-off Handgrip	123
Item 11 Leading Handgrip	127
Item 12 Leading Hand Position after Transfer	131
Item 13 Trunk Leaning	135
Item 14 Flight	139
Item 15 Landing	143
Appendix D. Model Perforce of Each Classifier for the Machine Learning Method	147
Bibliography	148

List of Tables

Table 1. Commercially available 3D depth sensors	11
Table 2. a) Skeleton position detected by Kinect relative to the human body [54], and b) Body Segment Vectors. The joint centers defined by the Kinect sensor *Trunk Anterior is calculated as the cross product of the trunk and shoulder across vectors. Thus, it is a vector that starts at their chest and points out to the front.....	24
Table 3. Universal biomechanics features for all TAI items. Maximum, minimum, range of motion, and average values of each feature are calculated over the lift or the setup phase durations and applied to the models.....	26
Table 4. Specific features for each TAI item. ^a: from the joint center HIP_LEFT to the HIP_RIGHT, ^b: from the joint center SHOULDER_LEFT to the SHOULDER_RIGHT.....	27
Table 5. Hyper-parameters of each item classifier	31
Table 6. Cases of proper and improper transfer technique based on TAI item scores for the 91 wheelchair users in the database	33
Table 7. Summary of the TAI item model performance in the training set and test set.....	35
Table 8. Participant demographics	58
Table 9. The confusion matrix of the test set of the machine learning methods KNN model	59

Table 10. Means and STDs of each participant’s start and end timepoints marked by visual, TH, and ML method. The values indicate when (in seconds) the participant started the lift phase of transfer after the trial began (timepoint = 0)..... 60

Table 11. Accuracy of the TAI predicted outcomes after applying the threshold and the machine learning timepoint marking methods 64

Table 12. Intra rater reliability (ICC_{3,1}) of the Kinect and the RealSense (n=30)..... 79

Table 13. Inter rater reliability (ICC_{2,1}) between the Kinect and the RealSense (n=30)..... 80

List of Figure

Figure 1. Experimental setup of wheelchair, bench and Kinect. The coordinate system follows the right-hand rule.....	21
Figure 2. Phases of a subject transferring from a wheelchair to a bench. a: start of the setup phase, b: transition point between the setup phase and the lift phase, c: end of the lift phase.....	23
Figure 3. Experimental setup of wheelchair, bench and Kinect. The coordinate system follows the right-hand rule.....	49
Figure 4. Phases of a subject transfer from a wheelchair to a level-height bench. a: start of the setup phase, the participant is sitting on the wheelchair; b: transition point between the setup phase and the lift phase, the participant is about to lift his body from the wheelchair seat; c: end of the lift phase, the participant landed on the bench.	51
Figure 5. Skeleton position detected by Kinect relative to the human body [94].....	52
Figure 6. (a) Determine the beginning of the lift phase by evaluating linear velocity of SPINE_BASE, (b) Determine the end of the lift phase by evaluating displacement in the X direction of SPINE_BASE	53
Figure 7. Example showing how to mark the start (a) and end (b) frame of lift phase by applying the machine learning (ML) algorithm. Each time frame from a transfer is labeled by the trained KNN classifier as “1” (lift phase) and “0” (setup phase). a) A frame (#110) is marked as the beginning of lift phase if the following six frames are all	

predicted as “1”, b) A frame (#161) is marked as the end of lift phase if the following six frames are all predicted as “0”..... 55

Figure 8. Histograms of the start timepoint differences between (a) TH and visual methods and (b) ML and visual methods between for each transfer (n=50) 61

Figure 9. Histograms of the end timepoint differences between (a) TH and visual methods and (b) ML and visual methods between for each transfer (n=50). 61

Figure 10. Histograms of the start timepoint differences (absolute values) between (a) TH and visual methods and (b) ML and visual methods between for each transfer (n=50) 62

Figure 11. Histograms of the end timepoint differences (absolute values) between (a) TH and visual methods and (b) ML and visual methods between for each transfer (n=50) 63

Figure 12. Bland-Altman plots for the start time points marked by using the (a) TH and visual methods and (b) ML and visual methods (n=50) 63

Figure 13. Bland-Altman plots for the end time points marked by using the (a) TH and visual methods and (b) ML and visual methods (n=50) 64

Figure 14. Experimental setup of the Kinect and the RealSense. The coordinate system follows the right-hand rule for both sensors..... 74

Figure 15. a) Skeleton position detected by Kinect relative to the human body [94], b) Joint center map in the NuiTrack SDK [109] 76

Figure 16. Bland-Altman plots for the agreement between Kinect and RealSense (n=150 transfers). LPOE, average plane of elevation angle on the leading side shoulder; LE,

average elevation angle on the leading side shoulder; TF, average trunk flexion angle;
STD, standard deviation; deg, degree 81

Preface

For my dearest mother and father,
Mrs. 薛如月 (Syue, Ru-Yue) and Mr. 魏權 (Wei, Cyuan).

3/15/2021

1.0 Introduction

In the United States, there were approximately 294,000 persons with spinal cord injury (SCI) in 2020, and 17,810 new cases occur each year [1]. Over 3.6 million Americans aged 15 and over used a wheelchair in 2010 [2]. Wheelchair users rely heavily on their upper extremities to complete common but essential activities of daily living such as getting in and out of bed, transferring to a toilet or a shower, and transferring in and out of a car. Manual wheelchair users will perform on average 14 to 18 transfers a day, which are extremely physically demanding and can lead to upper extremity pain and injury [3, 4]. Research shows that the prevalence of upper extremity pain, specifically shoulder pain, in wheelchair users ranges between 31 and 73 percent [5]. Unfortunately, shoulder pain leads to decreased quality of life and participation in physical activity [6].

The use of good transfer mechanics to avoid pain and injury is important for wheelchair users when performing transfers. The Transfer Assessment Instrument (TAI), is a tool developed to evaluate independent transfer technique and help clinicians and users to recognize deficits in technique. Research has shown that wheelchair users who received high scores on the TAI had significantly decreased forces and moments on the upper arm joints while performing transfers to and from different surfaces [7]. However, there are some limitations when therapists use the TAI as an assessment tool. The therapists need time for training to learn and practice the tool. Moreover, they have to score all the items while they are watching a patient's transfer. Even with training and practice, results of some items may be influenced by the rater's subjective interpretation. These barriers decrease the usability of the TAI in clinical settings.

An artificial intelligence system that can automatically score the TAI may potentially reduce the barriers associated with TAI's usability. We aim to develop a system that can watch a patient transfer and allow for automating the TAI using marker-less motion capture technology and machine learning algorithms that classify the motions into proper and improper techniques. Ideally the system would be able evaluate the TAI more reliably than a therapist rater would and generate objective feedback to the users. Therefore, the results of the current study could increase the usability and feasibility of TAI in a clinical setting.

1.1 Prevalence of Upper Extremities Pain and Injuries in Wheelchair Users

Shoulder pain is the most common secondary impairment in the upper limbs reported in the wheelchair user population [8]. A review summarized eighteen studies from 1985 to 2001 documenting the prevalence of upper limb injuries in people with spinal cord injury and reported that around 30–60% of wheelchair users experience upper extremity pain at the shoulder, 22–45% at the elbow, and 40–66% at the wrist [9]. In 2004, Girona and colleagues conducted a study with the largest sample size (around 770 subjects) to date and found that 531 (69%) patients with paraplegia reported current upper limb pain, and that the shoulder pain was most severe during wheelchair-related mobility and transportation activities [10].

The high prevalence of shoulder pain and injury in wheelchair users is believed to be due to overuse of the glenohumeral joint, especially during propulsion and transfers [11-14]. The

extrinsic and the intrinsic factors leading to shoulder injury can be exacerbated by overuse [14]. Intrinsic factors are factors associated with degeneration of the tendon. The mechanisms that originate “within” the tendons include age-related degeneration, poor vascularity, and changing of the mechanical properties of the tensile tissue. Extrinsic factors of the rotator cuff tendinopathy are defined as factors that cause compression of the rotator cuff tendons that originate “outside” of the tendon. These mechanisms include the compression (or shear) of the tendon within the subacromial space from anatomical or biomechanical abnormalities, as well as the compression of tendon posteriorly between the humerus and glenoid rim [14, 15].

The main cause of shoulder pain among manual wheelchair users is tendinopathy of the shoulder, especially of rotator cuff tendons [16]. The incidence of shoulder tendinopathy in wheelchair users is four times higher than that of ambulatory individuals (63% vs.15%) [17]. Akbar and colleagues reported that rotator cuff tears were present in 49% of wheelchair users, 70% of which were full thickness and all involved the supraspinatus [18]. Other pathologies, such as glenohumeral instability, biceps tendinitis, capsulitis, acromioclavicular joint degeneration, and distal clavicle osteolysis were also reported [8, 19-21].

At the elbow prevalence of ulnar mononeuropathy in spinal cord injury varies between 22 percent and 45 percent. At the wrist, the most common problem is carpal tunnel syndrome in wheelchair users with paraplegia. The prevalence is between 50% to 73% [22-25].

1.2 Wheelchair Transfer and Secondary Injuries

Wheelchair users rely heavily on their upper extremities to complete common but essential activities of daily living such as getting in and out of bed, transferring to a toilet or a shower, and transferring in and out of a car. For persons who are unable to bear weight through their legs, the majority of the force involved in lifting the body is placed upon the joints in the upper extremities [26]. During the performance of transfers, the loading on the upper extremity joints is greater than any other wheelchair related activity [27]. Excessive forces acting at the shoulder during transfers can lead to the development of shoulder impingement, posterior instability, capsulitis, and tendinitis [8, 19]. Also, repetitive shoulder motion and improper movements may cause muscle imbalance around the shoulder. Additionally, high superior forces generated during transfers are believed to contribute to pain and secondary impairments at the elbow [28]. The extreme wrist extension angles and forces generated during transfers may increase the pressure within the carpal tunnel and lead to median nerve compression which exacerbates carpal tunnel syndrome [23, 29]. Chronic shoulder pain is most evident during transfers, pressure relief, and wheelchair propulsion [30]. During wheelchair transfer, the shoulder is often placed in a position with a combination of flexion, abduction, and internal rotation. Large anterior force and flexion moment are generated on both leading and trailing arms [27]. This position causes the glenohumeral head to shift closer to the acromion's undersurface, and has been identified as a critical factor for impingement of subacromial soft tissue [21]. Moreover, overstress causes impingement of the tendons and diminished blood supply. Inflammatory changes, mechanical impingement, or soft-tissue or bony injury will be more progressive when the causative motion occurs at high frequency without sufficient time for healing. Properties of disorganized tissue in tendon would be exacerbated in

combination with repetitive tensile loading induced with daily activities such as lifting, pulling, or the strain incurred with the follow through of activities of daily living [14].

1.3 Biomechanics of the Wheelchair Sitting Pivot Transfer

Shoulder pain and injuries are believed to be associated with large joint reaction forces and moments during transfers [4, 8, 28]. However, it is difficult to directly measure the joint reaction force and the moments during the transfer without significant invasive procedures. Joint forces and moments are therefore indirectly estimated using inverse dynamics analysis, which use the anatomical movements (kinematics) to back calculate joint reaction forces and moments (kinetics). Using these analyses, it is possible to explore the kinematics that may lead to injury-inducing kinetics. Extreme combinations of shoulder flexion, internal rotation, and abduction are known to create high internal joint forces and are difficult to avoid during transfers [27]. Systematically varying leading hand placement and trunk position during the transfer has been shown to generate different joint forces and moments at the upper limb joints [31]. Using certain transfer skills (e.g. placing both feet on the floor, using head-hips maneuver to pivot the body, using proper handgrip techniques) has been shown to reduce loading across the wrists, elbows and shoulders [32].

1.4 Transfer Assessment Instrument – An Evaluation Tool for Determining Quality of Transfer

Using proper transfer technique can reduce the loading on the upper arm joints and help protect wheelchair users from developing injury and pain [4, 33]. The TAI is a tool used by clinicians and therapists to assess transfer quality and identify problems in wheelchair transfers which can cause increased forces on upper extremity joints [34, 35]. The TAI is based on clinical practice guidelines, current knowledge in the literature, and best clinical practices related to transfers. The TAI measures multiple components of a transfer including proper setup of the wheelchair and body positioning during transfers. The tool is a series of yes or no questions that evaluate both the wheelchair user's overall technique and any weak component skills within the transfer [34, 35]. Higher TAI scores represent better wheelchair transfer technique [35]. In the previous studies, we demonstrated that individuals who score highly on the TAI have lower mechanical loading at the shoulder, elbow and wrist in different transfer configurations [4, 7]. Therefore, wheelchair users who learn to perform transfers that are consistent with a high TAI score may reduce their risk of upper extremity injury and pain by decreasing the joint loading during the transfer.

The TAI has been used successfully in research to evaluate proper transfer technique. It also has been approved that there are limitations in its use for knowledge translation and application in a clinical setting. Firstly, clinicians need to become familiar with the TAI items to know if their patient's transfer was performed correctly or not for each item. The training materials initially developed for clinicians to learn how to use the TAI required significant time to review.

The newest version, TAI 4.0, refined the item statements and aimed to decrease time to administer and need for review of training materials. However, even with these modifications, the intra-rater reliability of the wheelchair setup items is only on the moderate level (ICC = .44 to .55). The inter-rater reliability of the body setup items (ICC = .65 to .72) is also lower than other items [36]. The body mechanics and some setup items are very difficult to evaluate because the therapist needs to watch multiple movements for different items within 0.5 to 1 second during the lifting phase of the transfer. Secondly, according to our interviews in multiple rehabilitation institutions and hospitals, many therapists agreed that the TAI is great tool to identify specific deficits of patient's techniques but takes too long to perform in the clinic. Thirdly, although the wheelchair may improve the transfer technique after a web-based training addressed by the TAI, it still needs the investigators familiar with TAI to manually evaluate the quality of the transfer [37]. Thus, the tool is currently used in research but has not yet been widely adopted into clinical practice. An automated system that can accurately and objectively observe transfer motions and report the TAI outcomes without requiring users to undergo extensive training could be of great benefit to therapists and patients with SCI in the future.

1.5 Marker-less Motion Capture

Marker-less motion capture systems have been widely used in full-body skeletal tracking, facial features recognition, and hand/finger gestural tracking and has found broad application in

the video gaming community and sports, wellness, and healthcare fields. The types of systems range from high-end indoor/outdoor image-based motion capture systems such as the SIMI, to low cost, portable 3D depth sensors such as the Kinect and RealSense. Wearable sensor technologies (e.g. Inertial Measurement Units (IMU)) can also provide monitoring and feedback to wheelchair users on general activity and motion techniques [38-41], and techniques known to cause injuries [42, 43]. A 3D depth sensor has total freedom of movement without the need to hold or wear any sensors or markers on the body during the transfer task thus the setup time. The sensor is also much cheaper than the other options and can provide more detailed motion tracking data (e.g. x, y, z coordinate positions of joint centers) than IMUs which provide relative segment orientations. Barbareschi and colleagues used a chest mounted IMU and built a Support vector machine (SVM) model to predict three item scores of the TAI 3.0: head-hip relationship, flight phase, and landing phase. The performance of the models ranged between 76% to 80% accuracy which is lower than the accuracies we have found so far in our preliminary analyses with the Kinect2 sensor [44]. One possible reason for the lower accuracies with IMUs might be the limitation of the number of features that can be extracted from them. Moreover, it is very difficult to evaluate all the items on TAI by a single IMU sensor since transfer motions are complex and involve multiple body segments. The body skeletal functions of a 3D depth camera can simultaneously track multiple joints on all the body segments.

1.6 3D Depth Sensors in Research

A variety of 3D depth sensors are available on the market and are supported with different platforms, such as Windows, Android and Linux (Table 1). The sensor can create 3D maps of anatomical landmarks for patient size assessments without using multiple cameras [45]. The two most common ones found in the literature include the Microsoft Kinect v1 and v2 and Intel RealSense. The Kinect is the first mass produced, three-dimensional (3D) depth sensing camera that possessed a price point making it available to almost any consumer. It consists of an infrared (IR) light projector, an IR camera, a RGB video camera, and microphones. One of the most powerful aspects of the Kinect is the ability to automatically identify anatomical landmarks in close to real-time using artificial intelligence. The Kinect sensor v1.8 detects and records the body surface with depth and RGB data recorded from the cameras then applies an algorithm of triangulation to automatically identify the location of joint centers of the body in the 3-D space. In 2014, Microsoft released a new version of Kinect for Xbox One with 60% wide-angle time-of-flight camera. The Kinect v2 measures the time it takes a light source to bounce back from a reflected target, and given that the speed of light is constant, the distance to the objects surface can be calculated. Using this different time-of-flight method, the Kinect v2 creates a depth map which promised increased resolution and accuracy. The Microsoft SDK for the Kinect provides the algorithms for full body skeleton tracking. The v2 sensor can be used to detect and record 25 joint centers (compared to 20 joint centers from the v1.8) of the body in X, Y, and Z dimensions at 30 Hz sampling frequency, and has been shown to be a reliable and valid tool for tracking body motions and human posture [46-50].

The official full-body tracking software development kit (SDK) provided by Microsoft allows investigators to develop a program for their needs. As a result, the Kinect v2 has become widely used in academic research to track upper and lower body motions in various biomechanics and rehabilitation applications [46-49, 51]. The Kinect has also become an integral part of some industrial commercial software applications such as Jintronix [52] and VERA released by Reflexion Health [53], which utilize motions captured by the Kinect for therapeutic gaming and functional assessments.

Microsoft announced the discontinuation of Kinect v2 in 2017 and released a new version Azure Kinect DK (Kinect DK) in 2019. The new SDK is supposed to retain the full body tracking function. However, the skeletal tracking algorithm of the new Kinect DK and the outputted raw data differ from the Kinect v2 [54]. Moreover, research related to classification of body skeletal motion with the Kinect DK is not well developed. An alternative solution that has been used is the Open Natural Interaction (OpenNI) framework which provides a library of body tracking algorithms that are compatible with various sensors. The Intel RealSense series has been used in several clinical research applications, however, work related to human motion classification is not as extensive as the Kinect in the published literatures due to only the hand gestures and movement tracking function being supported by the Intel official SDK [55-58]. One study by Mistry et al. developed an approach to translate sign language by using the hand motion data recorded by the RealSense [59]. They recorded 26,000 signs of English alphabet from 10 professional signers and labeled the data into 26 letters. Learning 90 features such as degrees rotation, flexion of each finger and hand openness, the performances of the model were 92.1% and 95.0% by applying CNN and SVM, respectively.

Table 1. Commercially available 3D depth sensors

	Price	Range	RGB image	Depth image	Physical Dimensions
Kinect V2	\$100	0.5 – 4.5m	1920x1080, 30 FPS	512x424, 30 FPS	250x66x67 mm
Azure Kinect DK	\$299	0.5 – 5.46m	3840x2160, 30 FPS	640x576, 30 FPS 1024x1024, 15 FPS	203x39x126 mm
RealSense D435	\$179	0.11 – 10m	1920x1080, 30 FPS	1280x720, up to 90 FPS	90x25x25 mm
RealSense D415	\$149	0.2 – 10m	1920x1080, 30 FPS 2592x1944, 15 FPS	1280x720, up to 90 FPS	99x20x23 mm
Xtion 2	\$598	0.8 – 3.5m	1920x1080, 30 FPS 1280x720, 60 FPS	640x480, 30 FPS	110x35x35 mm
Vico VR	\$249	0.5 – 4.5m	N/A	640x480, 30 FPS	220x120x50 mm
TVico	\$280	0.6 – 8m	1280x720, 30 FPS	640x480, 30 FPS	172x63x56 mm
Occipital Structure	\$499	0.4 – 3.5m	iOS camera resolution	640x480, 30 FPS 320x240, 60 FPS	119.2x28x29 mm
Orbbec Persee	\$240	0.4 – 8m	1280x720, 30 FPS	640x480, 16 bit, 30 FPS	172x63x56 mm
Leap Motion	\$100	0.025 – 0.6m	N/A	20 to 200+ FPS	76x30x17 mm

1.7 3D Depth Sensing Applications in Rehabilitation Science and Human Motion

Evaluation

Many studies have developed the Kinect for assessment of balance and postural control [60-64], dynamic balance tests [65], fall prevention [66, 67], and stroke rehabilitation [68-71]. Tan et al. reported that the Kinect can measure habitual gait speed and timed-up-and-go variables of the motor and Postural Instability and Gait Difficulty (PIGD) subscales under the Unified Parkinson's Disease Rating Scale (UPDRS). Clark et al. also showed that the Kinect can be a reliable tool for the current clinical gait assessments to evaluate step length, foot swing velocity, gait speed, and other variables associated with trunk balance during walking trials of patients with stroke [72].

Some studies showed that the Kinect has the ability to discriminate between populations with different clinical conditions. For example, Dolatabadi et al. used the Kinect and k-nearest neighbors (k-NN) and dynamic time warping (DTW) algorithms to correctly classify a gait pattern into healthy subjects and patients who previously had a stroke with 96% accuracy [73]. Leightley and colleagues showed that the Kinect can be used to evaluate a clinically validated assessment of sit-to-stand techniques and differentiate groups of master athletes, healthy old people, and young adults [74].

The Kinect has also been used in human movement pattern classification and motion evaluation. Protopapadakis and colleagues used the joint center data and multiple machine learning algorithms to identify six types of dance poses [75]. Plantard et al. computed the Rapid Upper Limb Assessment (RULA) scores based on the Kinect skeleton data [76, 77]. The RULA was developed to investigate the exposure of individual workers to risk factors associated with work-related upper limb disorders [78]. Similar to the TAI, the RULA breaks down a motion task into different components and is evaluated by watching the subject's joint angles and position of his/her upper extremities and trunk. The RULA entails a trained rater who watches the task and provides a posture rating based on the observed shoulder angles of motion, wrist and forearm positioning, and trunk twisting and bending. However, some studies also reported that the main weakness of RULA is related to the fair level of inter-rater reliability ($ICC < 0.5$) [79, 80]. Following the outcome measures of the RULA as the gold standard, Manghisi and colleagues developed a real-time software tool based on the Kinect v2 to autoscore the RULA and evaluate the upper extremity musculoskeletal disorder risk factors in workplaces [81].

1.8 Significance

Transfer-related pain and injury can adversely affect the quality of life of wheelchair users. Many studies reported that the pain is a major factor of functional decline and low quality of life in individuals with SCI [82, 83]. This lesser quality of life manifests itself as a loss of autonomy and a decreased ability to participate in society [84-87]. Dalyan et al. documented a significant association between employment status and upper limb pain, with unemployment higher and full-time employment lower in individuals with upper limb pain than those without pain (21.4% vs. 7.1% and 20% vs. 45.2%, respectively) [88].

Proper transfer technique helps wheelchair users to prevent and reduce the risk of upper arm pain and injuries. However, education on proper transfer technique is not well disseminated and therapists have limited time to work with wheelchair users to develop their skills. A marker-less motion capture sensor combined with an artificial intelligence algorithm that can automate the assessment process may be helpful for clinicians and wheelchair users in identifying their deficits and in learning proper transfer technique.

This study fills a gap that exists between research and the practical clinical application of the TAI. The results of this work will support a background algorithm for an automatic transfer evaluation system that provides real-time feedback to both therapists and patients. Also, we aim to show that our models will be robust to the type of 3D depth sensors so that they will still be useful should a specific brand of sensor become obsolete. The system will ultimately be used to

identify problems with transfers and provide guidance on corrective strategies aimed at preserving upper limb function. Most wheelchair users with upper limb pain or injury do not seek medical attention or tend to seek conservative treatment for their problems [89, 90]. Unfortunately, common treatments such as anti-inflammatories are generally unsuccessful. Moreover, most wheelchair users are not willing to undergo surgical treatment because they rely heavily on their upper extremities to perform activities of daily living. Recovery from surgery would require that they rest their arms to heal like able-bodied patients which would require them to rely on others to support daily mobility and functions. Therefore, it is critical for a wheelchair user to preserve their upper extremity function through increased awareness and preventative strategies such as by using evidence-based proper transfer techniques to reduce the risk of injury. As a result, patients with SCI and the healthcare system will benefit from reduced healthcare costs associated with treating transfer related pain, overuse injuries, and personal care assistance.

1.9 Thesis Objectives

The first objective of this dissertation work (specific aim 1 and aim 2) was to develop machine learning models for classifying proper and improper transfer technique using input from the Kinect v2. However, as Microsoft recently announced discontinuation of the Kinect v2 a second objective of this study was to explore using an alternative sensor on the market that can provide motion tracking. Specific aim 3 was to test the feasibility of a surrogate motion capture

sensor available on the market (Intel's RealSense) for measuring key variables of transfer evaluation.

2.0 Automating the Clinical Assessment of Independent Wheelchair Sitting Pivot Transfer Techniques

2.1 Abstract

This chapter has been published in Topics in Spinal Cord Injury Rehabilitation. Full citation: Wei L, Chung CS, Koontz AM. Automating the Clinical Assessment of Independent Wheelchair Sitting Pivot Transfer Techniques. Topics in Spinal Cord Injury Rehabilitation, 2021, Permission was obtained for use.

Introduction: Using proper transfer technique can help to reduce forces and prevent secondary injuries. However, current assessment tools rely on the ability to subjectively identify harmful movement patterns.

Objective: The purpose of the study was to determine the accuracy of using a low-cost markerless motion capture camera and machine learning methods to evaluate the quality of independent wheelchair sitting pivot transfers. We hypothesized that the algorithms would be able to discern proper (low risk) and improper (high risk) wheelchair transfer techniques in accordance with component items on the Transfer Assessment Instrument (TAI).

Methods: Transfer motions of 91 full-time wheelchair users were recorded and used to develop machine learning classifiers that could be used to discern proper from improper technique. The data were labeled using the TAI item scores. Motion variables from the Kinect were inputted as

the features. Random forests and K-nearest neighbors algorithms were chosen as the classifiers. Eighty percent of the data were used for model training and hyperparameter tuning. The validation process was performed using 20% of the data as the test set.

Results: The area under the receiver operating characteristic curve of the test set for each item was over 0.79. After adjusting the decision threshold, the precisions of the models were over 0.87, and the model accuracies over 71%.

Conclusions: The results show promise for the objective assessment of the transfer technique using a low cost camera and machine learning classifiers.

Key Words: Wheelchair Biomechanics, Skeletal Tracking, Machine Learning, Activities of Daily Living, Feature Engineering and Feature Selection, Motion capture

2.2 Introduction

Using proper transfer mechanics can help protect wheelchair users from developing upper limb pain and injury. The Transfer Assessment Instrument (TAI) [36] is a tool that was developed to evaluate transfer technique and help clinicians and users recognize deficits in technique. Research has shown that wheelchair users who received high scores on the TAI had significantly decreased forces and moments on the upper arm joints while performing transfers to and from different surfaces [7]. However, there are some shortcomings with using the tool in clinical practice. A therapist needs to be familiar with the 15 items and be aware of what to look for so when he or

she watches the transfer, which can occur very quickly, the TAI can be scored. Even with training and practice, results are influenced by the rater's subjective interpretation of the wheelchair setup and body movements. The TAI has low to moderate intra-rater reliability (ICC=.44 to .64) for wheelchair setup items and moderate inter-rater reliability of the body setup items (ICC=.65 to .72) [36]. These issues decrease the reliability and usability of the TAI in clinical settings.

Transfer techniques have been objectively quantified using high-tech methods such as 3D motion capture systems (e.g. Vicon) [91-93]. These systems however are expensive, require technical backgrounds or expertise to use and are not readily available in many SCI clinics where transfers would typically be assessed. Recently the use of marker-less 3D depth sensors to detect and quantify movement patterns have emerged as potential surrogates to these higher-tech methods and for applications that may not need the highest level of accuracy that higher-tech methods can provide. Depth sensors allow for creating 3D maps of a human body and to identify anatomical landmarks (e.g. joint center locations) using a small portable low-cost camera and artificial intelligence [94]. Many studies have used the Microsoft Kinect for the clinical assessment of balance and postural control [60-65], fall prevention [66, 67], and stroke rehabilitation [68-71]. Some studies have also shown that the Kinect can discriminate between populations with different clinical conditions. For example, Dolatabadi et al. used the Kinect and k-nearest neighbors (k-NN) and dynamic time warping (DTW) algorithms to correctly differentiate the gait patterns of unimpaired participants and persons with a stroke with 96% accuracy [73]. Leightley and colleagues showed that the Kinect can be used to evaluate sit-to-stand techniques and differentiate between groups of master athletes, healthy older people, and young adults [74].

The potential to use the Kinect to accurately and reliably quantify transfer motions has been recently studied [36]. A study comparing transfer motions recorded using the Vicon and

first-generation Kinect v1 sensor found good to excellent test-retest reliability ($ICC \geq 0.71$) and upper limb and trunk range of motion trajectories that were similar between the two systems (cross-correlation coefficients ranging from 0.71 to 0.97) [95]. In a follow-up study the Kinect 2 sensor was able to discern differences in movement variables among unimpaired individuals who were trained to use proper technique and three variations of improper transfer technique [96]. This study aims to determine Kinect v2's ability to discern differences between proper and improper movement patterns among individuals who routinely perform transfers. We hypothesize that the Kinect 2 and supporting machine learning (ML) algorithms will be able to achieve an area under the receiver operating characteristic curve (AUC) of at least 0.80 and precision of at least 0.90 for each prediction model, values that were felt would yield high clinical acceptability. Our long-term goal is to develop a system that can watch a transfer and automate the TAI scores in real-time. Ideally the system would produce TAI scores with a higher level of reliability and objectivity than the current method of assessment. Such a system may help to reduce therapist burden and facilitate easier and more reliable transfer assessments in clinical settings.

2.3 Methods

2.3.1 Participants

Participants were recruited through research registries, local SCI clinics and rehabilitation hospitals, and at organized recreational sport events. The inclusion criteria were 1) have discernable neurological impairment affecting both lower extremities or persons with transfemoral or transtibial amputation of both lower extremities who do not use prostheses during transfers, 2) at least one-year post-injury or diagnosis, 3) able to independently transfer to/from a wheelchair without human assistance or assistive devices, 4) use a wheelchair for the majority of mobility (over 40 hours/week), and over the age of 18 years. Participants were excluded if they had 1) current or recent history of pressure sores in the last year, 2) history of seizures or angina, or 3) were able to stand unsupported.

2.3.2 Study Protocol

2.3.2.1 Experimental Setup

A Kinect v2 sensor was positioned two meters in front of the participants, 70 centimeters above the floor, and centered between the wheelchair and the bench (Figure 1). A custom graphical user interface (GUI) was programmed in C# using Visual Studio 2012, .NET Framework 4.0, and the Kinect for Windows SDK to collect the 3D joint center position data in a Cartesian coordinate system from the Kinect sensor. The sampling frequency was 30Hz.

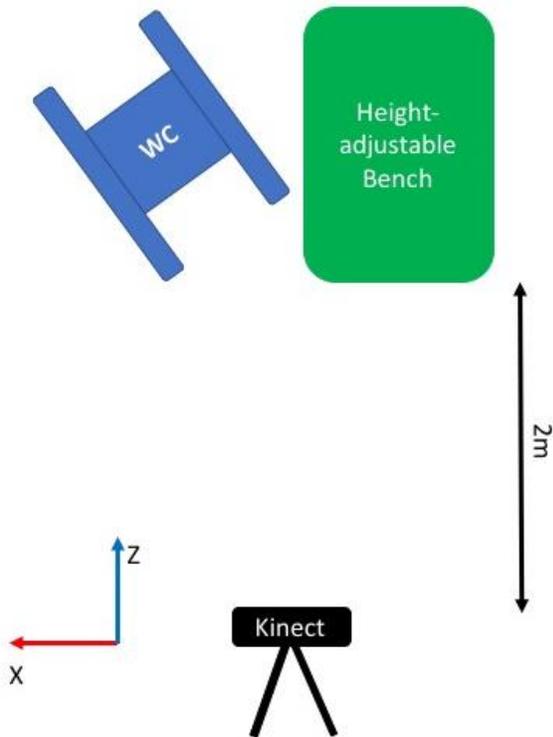


Figure 1. Experimental setup of wheelchair, bench and Kinect. The coordinate system follows the right-hand rule

2.3.2.2 Transfer Protocol and Evaluation

All wheelchair users used their own chairs for the testing and were instructed to perform a transfer to a tub bench (70cm x 55cm) in a habitual way. Before data collection, the participants were provided an opportunity to adjust the position between their wheelchairs and the bench and also practiced the transfers to familiarize themselves with the setup. Participants were asked to transfer up to five times to and from the wheelchair and a level-height bench. A subset of the participants (31 out of 91) also performed five non-level transfers 3 inches higher or lower (randomly assigned) relative to the wheelchair seat. The order of level and non-level was decided by a random number sequence. Up to three clinicians who were experienced in scoring the TAI

evaluated and scored each trial. The participants were asked to sit on the bench for 5 to 10 seconds before they transferred back. A transfer board was provided for the participants who requested it for the transfer. The participant was informed that they could request a break or discontinue the study at any time. Digital video was recorded during all transfers to allow for review and resolution of potential discrepancies in expert scoring if necessary. Opportunities for the participants to rest were built into the protocol. Approximately 3-5 minutes of rest time between trials and the participants were allowed to take more time to rest until they were ready for the next trials.

2.3.2.3 Data Labeling

The TAI 4.0 scores after consensual agreement among raters and video reviews were used as the gold standard (true outcome) for labeling the data [36]. The TAI breaks down a transfer into multiple components and evaluates each one independently. The 15 items are classified into three subdomains: wheelchair setup (4 items), body setup (9 items), and flight/landing (2 items). Items are scored “yes” (1 point) when the participant performs the specified skill correctly and “no” (0 points) when the participant performs the skill incorrectly. A (.5) partial credit is allowed for items 6, 7, 10 and 11. However, because partial credit means there is room for improvement, partial credit techniques were assigned ‘0’ points so that the ML algorithm would flag them as improper techniques. Combining the items scores from all items gives a total score on a continuous scale (0 being worst to 10 being best) that provides an overall measure of transfer performance. Each TAI item has specific movement characteristics (e.g. features) related to the outcomes (e.g. proper or improper technique). Items 1, 2, 7 and 8 pertain to the body position during the setup phase (just before transfer), and items 9, 10, 11, 12, 13, 14 and 15 pertain to the biomechanics in the lift phase. The time frames that define the setup and lift phase were manually (visually) identified for each

of these transfers using the videos captured by the Kinect RGB camera (Figure 2). Separate ML models were developed to predict the outcome of each TAI item. Item 3, 4, 5, and 6 are not included in this study because they cannot be identified by the Kinect.



Figure 2. Phases of a subject transferring from a wheelchair to a bench. a: start of the setup phase, b: transition point between the setup phase and the lift phase, c: end of the lift phase

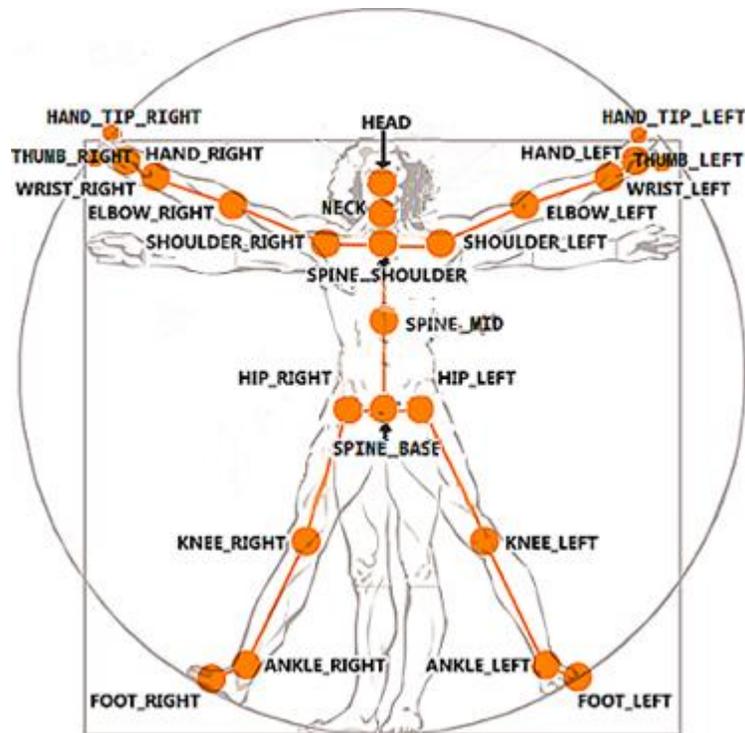
2.3.2.4 Features

Universal features were applied to all items and were comprised of joint motion angles (Table 2, 3). Maximum, minimum, range of motion, and average values of each feature were calculated over the time windows of the setup phase and the lift phase. In addition to the universal features, we designed specific features for each item based on the item statements. These specific features consisted of selected displacements, velocities, accelerations, jerks, triangle areas between three joint centers at the hand, and anthropometric variables (Table 4).

Table 2. a) Skeleton position detected by Kinect relative to the human body [54], and b) Body Segment Vectors. The joint centers defined by the Kinect sensor

*Trunk Anterior is calculated as the cross product of the trunk and shoulder across vectors. Thus, it is a vector that starts at their chest and points out to the front

a)



b)

Body Segment Vectors	Kinect Joint Center
Trunk	SPINE_SHOULDER SPINE_MID
Shoulder Across	SHOULDER_RIGHT SHOULDER_LEFT
Trunk Anterior*	Cross product of: TRUNK and Shoulder Across
Upper Arm	SHOULDER ELBOW
Forearm	ELBOW WRIST
Shoulder to Wrist	SHOULDER WRIST
Hand	WRIST HAND
Thigh	HIP KNEE
Hip Across	HIP_RIGHT HIP_LEFT
Head Hip	Head Spine Base

Table 3. Universal biomechanics features for all TAI items. Maximum, minimum, range of motion, and average values of each feature are calculated over the lift or the setup phase durations and applied to the models

Universal Feature	Definition
Shoulder Plane of Elevation Angle	Angle between Trunk Anterior vector and Upper Arm vector, projected on transverse plane
Shoulder to Wrist Angle	Angle between Trunk Anterior vector and Shoulder to Wrist vector, projected on transverse plane
Shoulder Elevation	Angle between Trunk vector and Upper Arm vector
Elbow Flexion	Angle between Upper Arm vector and Forearm vector
Wrist Flexion	Angle between Forearm vector and Hand vector
Trunk Flexion	Angle between Trunk vector and Thigh vector
Trunk Twist	Angle between Shoulder Across vector and Hip Across vector
Trunk Bending in Transverse	Angle between Shoulder Across vector and Hip Across vector, projected on transverse plane
Trunk Bending in Frontal	Angle between Shoulder Across vector and Hip Across vector, projected on frontal plane

Table 4. Specific features for each TAI item. ^a: from the joint center HIP_LEFT to the HIP_RIGHT, ^b: from the joint center SHOULDER_LEFT to the SHOULDER_RIGHT

TAI Item	Related feature	Definition
1. Distance between wheelchair and the targeted surface is less than 3 inches	Spine Base	Displacement of the SPINE_BASE in X direction during the lift phase
	Displacement in X	
	Spine Mid	Displacement of the SPINE_MID in X direction during the lift phase
	Displacement in X	
	Spine Shoulder	Displacement of the SPINE_SHOULDER in X direction during the lift phase
	Displacement in X	
	Neck Displacement in X	Displacement of the SPINE_NECK in X direction during the lift phase
	Head Displacement in X	Displacement of the SPINE_HEAD in X direction during the lift phase
2. The angle between wheelchair and the targeted surface is 0-19 degrees for power wheelchair, or 20-45 degrees for manual wheelchair	Starting Hip Angle	Angle between the global X and a Hip vector ^a at the start of the setup phase
	End Hip Angle	Angle between the global X and a Hip vector at the end of the setup phase
	Starting Shoulder Angle	Angle between the global X and a shoulder vector ^b at the start of the setup phase
	End Shoulder Angle	Angle between the global X and a shoulder vector at the end of the setup phase
7. Feet are on the floor or the targeted surface	Left / Right Knee	Displacement of the RIGHT / LEFT_KNEE during the setup phase
	Displacement	
	Left / Right Ankle	Displacement of the RIGHT / LEFT_ANKLE during the setup phase
	Displacement	
	Left / Right Foot	Displacement of the RIGHT / LEFT_FOOT during the setup phase
	Displacement	

8. Scooting hips to the front of the seat, at least 1/3 of thigh is off the surface	Spine Base	Displacement of the SPINE_BASE in Y direction during the lift phase
	Displacement in Y	
	Right / Left Hip	Displacement of the HIP_RIGHT/LEFT in Y direction during the lift phase
	Displacement in Y	
	Spine Base	Displacement of the SPINE_BASE during the setup phase
	Displacement	
	Right / Left Hip	Displacement of the HIP_RIGHT/LEFT during the setup phase
	Displacement	
9. Leading Hand position before transfer	Leading Forearm	The average length between ELBOW to WRIST during transfer
	Length	
	Leading Upper Length	The average length between SHOUDLER to ELBOW during transfer
10. Push off hand grips armrest, wheel, frame, cushion or surface edge	Push off Handgrip	Triangle area between HAND_TIP, THUMB, and HAND on the push off side
	Area	
11. Leading hand grips armrest, wheel, frame, cushion or surface edge	Leading Handgrip Area	Triangle area between HAND_TIP, THUMB, and HAND on the leading side
12. Leading Hand position after transfer	No specific features for this item	
13. During transfer, trunk is leaning forward, away from the targeted surface	Trunk Angle at Y Axis	Angle between global Y (vertical-axis) and Head Hip vector

14. Flight movement is in one smooth and fluid motion &	Velocity, Acceleration, and Jerk at SPINE_BASE, SPINE_MID,	Linear velocity, acceleration, jerk at SPINE_BASE, SPINE_MID, SPINE_SHOULDER, NECK, and HEAD during the lift phase. Maximum, minimum, range of motion, and average values of each feature are calculated. Totally 60
15. No excessive movement occurred (no loss of balance) after landing	SPINE_SHOULDER, NECK, and HEAD	features are computed.

2.3.4.5 Model Performance Evaluation

The data were divided into three sets: a training set (80% of transfer trials) used to learn model parameters and to build the cost functions, a validation set (20% of training data set) used to search the best hyper-parameters of the classifier, and a test set (20% of transfer trials) used to assess model performance. A 5-fold cross-validation was applied between the training set and the validation set to tune the models. The AUC and confusion matrixes with true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN) individually for each TAI item were generated. During the feature selection process, the training set's AUC and the average accuracy and standard deviation (STD) of the cross-validation were calculated. For the test set, each model's AUC, model accuracy, false discovery rate (FDR, $(\frac{False\ Positive}{(True\ Positive)+(False\ Positive)})$), precision and recall were computed. The random forest classifier (RF) and K nearest neighbors classifier (KNN) were selected for item models as they showed the best performance after initially testing 16 supervised ML classifiers (see **Appendix C.**). The details of the feature engineering processes are shown below:

- 1) Input the universal features and specific features for each item as potential predictors of each classification model
- 2) Find and remove correlated features:
 - a. Compute absolute values of the coefficient matrix using all the features
 - b. Determine a group of features that have a Spearman's correlation coefficient greater than .9
 - c. Select one feature from each group of correlated features and discard the rest
- 3) Train a preliminary model with the selected features from the previous step
 - a. Scale features via Z-score normalization
 - b. Use Python "GridSearchCV" function from the Scikit-learn library to find the best hyper-parameters for the classifier. 20% of data from the training set is split as the validation set.
 - c. Compute the AUC using both training set and test set data, and the confusion matrix using the test set data
- 4) Determine the number of features by permutation feature importance:
 - a. Permute (shuffle) the values of each feature from the observations
 - b. Re-evaluate model performance by AUC for training set
 - c. Determine the feature importance by the dropped performance
 - d. Keep features that have high permutation importance and re-build a training model using the selected features. 20% of data from the training set is split as the validation set to search the best hyper-parameters (Table 5). The model aims to achieve approximate or better AUC to the preliminary model at step 3)
- 5) Evaluate the model performance

- a. Compute the AUC for both training set and test set, and mean and standard deviation of accuracy from the training set
- b. Tune the decision threshold of the classifier to achieve a better precision score, aim to achieve that the false discovery rate is less than .10 and the precision is greater than .90
- c. Compute the confusion matrix, accuracy, FDR, precision and recall using the test set data

Table 5. Hyper-parameters of each item classifier

Item	Classifier	Hyper-parameters
1. WC distance	KNN	leaf_size=5, n_jobs=-1, n_neighbors=2, p=1
2. WC angle	RF	class_weight=None, criterion='gini', max_depth=15, max_features='log2', min_samples_split=2, n_estimators=50
7. Feet down	RF	class_weight=None, criterion='entropy', max_depth=15, max_features='auto', min_samples_split=3, n_estimators=50
8. Scooting forward	RF	class_weight='balanced', criterion='gini', max_depth=3, max_features='auto', min_samples_split=3, n_estimators=300
9. L-hand position (before)	RF	class_weight='balanced', criterion='entropy', max_depth=3, max_features='auto', min_samples_split=2, n_estimators=200
10. Push-off hand handgrip	KNN	leaf_size=5, n_jobs=-1, n_neighbors=4, p=1
11. L-hand handgrip	KNN	leaf_size=5, n_jobs=-1, n_neighbors=4, p=1
12. L-hand position (after)	RF	class_weight='balanced', criterion='entropy', max_depth=3, max_features='sqrt', min_samples_split=3, n_estimators=50
13. Body leaning	RF	class_weight='balanced', criterion='entropy', max_depth=3, max_features='sqrt', min_samples_split=3, n_estimators=50

14. Flight	RF	class_weight='balanced', criterion='entropy', max_depth=3, max_features='sqrt', min_samples_split=10, n_estimators=200
15. Landing	RF	class_weight='balanced', criterion='entropy', max_depth=3, max_features='sqrt', min_samples_split=10, n_estimators=50

Key: KNN: k-nearest neighbors, L-hand: leading hand, RF: random forest, WC: wheelchair.

2.4 Results

2.4.1 Participants Demographics

Seventy-nine men and 12 women with an average age of 54.9 years (standard deviation (STD)=10.0) contributed a total of 591 transfer trials for the analysis. The group performed on average of 15.4 transfers (STD=6.9, self-reported) per day. Thirty-three (36%) were African Americans, 37 (41%) were Caucasian, 11 (14%) were Hispanic, three were Asian, two denoted mixed race, and five did not answer the question. Seventy-four participants (82%) used a manual wheelchair. Sixty-six (73%) had a spinal cord injury, 15 (16%) had an amputation, 7 (8%) had multiple sclerosis, and others included Guillain barre (n=2) and traumatic brain injury (n=1). The deficit rates of the TAI item scores ranged from 14% (not scooting forward) to 55% (not leaning the trunk forward) (see Table 6).

Table 6. Cases of proper and improper transfer technique based on TAI item scores for the 91 wheelchair users in the database

TAI item	Total	Proper		improper	
	trials	cases	%	cases	%
1. Wheelchair distance	591	413	(70)	178	(30)
2. Wheelchair angle	586	375	(64)	211	(36)
7. Place feet down	581	354	(61)	227	(39)
8. Scooting forward	591	443	(75)	148	(25)
9. Leading hand position (before)	581	482	(83)	99	(17)
10. Push off handgrip	580	342	(59)	238	(41)
11. Leading handgrip	580	336	(58)	244	(42)
12. Leading hand position (after)	591	526	(89)	65	(11)
13. Body leaning	591	420	(71)	171	(29)
14. Flight phase	591	402	(68)	189	(32)
15. Landing phase	591	419	(79)	172	(21)

2.4.2 Model Performance

The AUC of the training set models ranged from .83 to .99 (Table 7). The average accuracy from cross-validation ranged from 72% to 88%. After tuning the decision thresholds of the classification models, the AUC of the test sets were between .79 to .94, the accuracies 71% to 92%, FDR .04 to .13, precisions .90 to .96, and recalls .61 to .95. The average AUCs of the test sets were .86, .89 and .85 for the subdomains of wheelchair setup, body setup, and flight items, respectively. All item models met our hypothesis (the AUC is greater than .80 and the precision is greater than .90) except for item 7, placing feet on the ground during the setup phase

(precision=.87), and item 15, keeping balance while landing on the surface (AUC=.79). More details of the ML training processes and the test set confusion matrix can be found in the “Model training process and results of each TAI 4.0 item” section of the Supplemental Materials.

Table 7. Summary of the TAI item model performance in the training set and test set

TAI Phase	Item	Classifier	Feature:		Training Set			Test Set (after decision thresholds adjustment)					
			# of Feature selected	Top 3 Important Features	AUC	Mean of CV Accuracy	STD	AUC	Accuracy	False Discovery Rate	Precision	Recall	
Phase I Wheelchair Setup	1	WC distance	KNN	26	Average of leading-side shoulder POE angle Maximum of leading-side shoulder elevation angle Minimum of trailing-side shoulder-to-Wrist vector elevation angle	0.96	81%	2%	0.81	71%	0.04	0.96	0.61
	2	WC angle	RF	14	Shoulder-across angle at initial position Minimum of trunk flexion angle Minimum of trailing-side shoulder elevation angle	0.99	87%	4%	0.90	90%	0.09	0.91	0.93
Phase II Body Setup	7	Feet down	RF	29	Average of leading-side elbow flexion angle Maximum of trunk leaning angle Maximum of trailing-side shoulder POE angle	0.99	78%	3%	0.85	71%	0.13	0.87	0.63
	8	Scooting forward	RF	27	Minimum of trailing-side shoulder elevation angle Minimum of leading-side shoulder elevation angle Vertical displacement of spine-base during lift phase	0.96	80%	3%	0.87	82%	0.10	0.90	0.84
	9	L-hand position (before)	RF	27	Shoulder elevation angle at initial position Average of shoulder POE angle Maximum of elbow flexion angle	0.96	84%	5%	0.94	92%	0.04	0.96	0.95
	10	Push-off hand handgrip	KNN	38	Average of hand grip area Shoulder POE angle at initial position Shoulder-to-Wrist vector POE angle at initial position	0.97	72%	3%	0.82	77%	0.08	0.92	0.67
	11	L-hand handgrip	KNN	35	Average of wrist flexion angle Maximum of hand grip area Minimum of trunk flexion angle	0.95	76%	3%	0.87	75%	0.09	0.91	0.63
Phase III Flight	12	L-hand position (after)	RF	23	Minimum of shoulder POE angle Average of shoulder-to-Wrist vector elevation angle Average of shoulder POE angle	0.98	88%	2%	0.85	86%	0.05	0.95	0.89
	13	Body leaning	RF	9	Maximum of trunk leaning angle Average of trunk leaning angle Trunk flexion angle at initial position	0.94	80%	3%	0.89	81%	0.06	0.94	0.78
	14	Flight	RF	40	Average of leading-side elbow flexion angle Leading-side Shoulder-to-Wrist vector POE angle at initial position Average of SPINE_SHOULDER velocity	0.95	73%	3%	0.87	82%	0.09	0.91	0.83
	15	Landing	RF	33	Maximum of SPINE_SHOULDER velocity Minimum of leading-side wrist flexion angle Minimum of trailing-side wrist flexion angle	0.83	80%	2%	0.79	81%	0.09	0.91	0.83

Key: AUC: area under receiver operating characteristic curve, CV: cross-validation, KNN: k-nearest neighbors, L-hand: leading hand, POE: plane of elevation, RF: random forest, STD: standard deviation, WC: wheelchair.

2.5 Discussion

The goal of this study was to determine if Kinect measured transfer motions and ML algorithms could be used to predict TAI item scores. Model performance was tested on 20% of the data that was not used to train and tune the model. Thus, the performance outcomes provide an idea of how good the models would be in predicting the TAI scores for new ‘unseen’ transfers. This study also included one of the largest sample sizes used in an ML modeling application in human movement biomechanics (median = 40 subjects [97]). The wheelchair users tested were also diverse in terms of gender, the types and levels of disabilities, races represented, and wheelchairs used (e.g. manual and power). These factors all help to strengthen the modeling approaches and allow for greater generalizability and prediction success.

The results show that the model predictions across all TAI items have accuracies that range from 71% to 92%. The results for body leaning (item 13: 81%) and smooth landing (item 15: 81%) are better than a recent study that attempted to predict TAI outcomes from accelerometer data [98]. In this previous study a single chest worn accelerometer was used to predict two items that scored the head hips relationship and landing based on the previous TAI 3.0 version with accuracies of 75.9% and 79.9%. Items pertaining to wheelchair setup and arm positioning could not be modeled in their study and there were not enough improper cases in the dataset to model the item related to controlled flight. Unlike accelerometry a depth sensor can quantify static positioning and motion to a finer level of detail (e.g. linear and angular displacements and ranges of motions of body segments) which may explain in part the differences between studies. However, a depth sensor requires that person perform the movements within its field of view. This works well when

performing assessments in a set area or space like a rehabilitation clinic, but the sensor isn't a wearable and therefore can't go with the person and assess their transfer techniques during free-living. It may be possible with a wearable inertial measurement unit sensor (IMU) to measure more items on the TAI and with greater accuracy than accelerometry but based on our results and the types of features that were found to be important in predicting TAI scores it may require that multiple sensors be attached to the arms and trunk to obtain the necessary angles, displacements, etc. This may be impractical if the goal of the system is to monitor real world transfers as compliance with wearing multiple sensors would be an issue for many wheelchair users [99].

Most studies report overall accuracy as a main outcome for judging model performance however imbalanced data between the proper and improper cases can bias this parameter [97]. In this study several items had more proper than improper cases. The AUC unlike the accuracy outcome is not affected by imbalanced data. All of our models achieved AUCs over .80 except for item 15 (AUC=.79). Although the item 15 model fell slightly short of our AUC goal, other outcomes of the test set may be acceptable (accuracy=81%, precision=0.91, recall=0.83). In clinical practice, it is more detrimental to diagnose a false-positive TAI score (e.g. saying that the person is doing it right when they are doing it wrong) than to diagnose a false-negative TAI score (e.g. saying that the person is wrong when they are doing it right). Thus, the decision threshold of the model was tuned to achieve high precision and to minimize the false-positive TAI scores and FDR. To aid the predictions in catching persons who are using improper technique, we tuned the model to achieve precisions that are over .90. The FDR for item 15 after adjusting the decision threshold is 0.09 in the test set. This means that the current model could misjudge 9% of transfer trials as proper if they were improper but would be much more likely to classify an improper landing correctly. However, we felt that erring on the side of classifying a proper one as an

improper one would be more clinically acceptable than scoring an improper landing as a proper one as the ultimate goal of our model development is to develop a system to identify deficits and train individuals on how to perform, or reinforce using proper transfer skills.

In this study, we not only aimed to achieve high model performance, but also were interested in feature interpretation. Thus, we did not select the “black box” models, such as artificial neural network (ANN) or convolutional neural network (CNN), and automatic feature extraction techniques, such as principal component analysis (PCA). These classifiers and training methods might generate high model performance but produce features (principal components that can represent most information of imputed features, or “node” in the hidden layer of the neural network) that are difficult to interpret or not informative with respect to the item score outcomes. For example, if a model scores with high confidence that a patient is using improper transfer technique based on the motion data but offers no insight into the specific features related to patient’s body movements and position, it is unclear how this knowledge could be used to improve the patient’s wheelchair skills.

The item 1 asks the rater to observe the distance between the wheelchair front edge and the transfer surface. The item is scored “1” if the distance of the gap is less than 3 inches. It would be scored “.5” when the gap is three to five inches and scored “0” when the gap is over 5 inches. In this study we sorted the trials that had a gap over three inches as improper technique and labeled these trials “0” in our database. We assumed that the subjects would have different joint angles movements of their upper extremities between the proper and improper groups. We also assumed that subjects who score “0” should create more horizontal displacements in the trunk joint centers than the subjects who scored “1”, thus we added five more features related to the trunk displacements along the horizontal axis. After the model was trained (AUC=.81, Accuracy=71%,

FDR=.04, precision=.96, recall=.61), the most important features selected by the model are the features related to both leading and trailing side shoulder. The trunk displacement features are also selected in the model, but they are less important than joint angle features.

To score the item 2, the rater needs to judge the angle between the subject's wheelchair and the side edge of the transfer surface when the participant places the wheelchair and prepares to transfer (see appendix). If the subject is using manual wheelchair, he would be scored a "1" only if the angle between 20 to 45 degrees. If the subject is using a power wheelchair, he would be scored "1" only if the angle is between 0 to 19 degrees. None of our participants placed their wheelchair over 45 degrees. However, in this study we aimed to build a model that can classify two positions ("0-19 degrees" and "20-45 degrees") of the wheelchair setup. Thus, we labeled the group 0-19 degrees as "0" and 20-45 degrees as "1". We created a Hip Across and Shoulder Across vector parallel to the subject's coronal plane and computed the angle between the global X-axis (Figure 1) and these two vectors at the time frames at the beginning and end of the setup phase and used these angles as the specific features. After the feature selection, we found the most important feature was using the Shoulder Across vector at the beginning of the setup phase. The other specific features of the item 2 might be eliminated due to the high correlation between them. The model has high performance (AUC=.90, Accuracy=90%, FDR=.09, precision=.91, recall=.93), and show that these features have the ability to distinguish two different sets of wheelchair angles through the random forest classifier.

To meet the requirement of the item 7, a manual wheelchair user would need to place both feet on the ground before he transfers to the bench. Power wheelchair users are allowed to place their feet on the footplate. In the TAI, the subject would be scored ".5" if he places only one foot on the ground. Here we labelled the transfer as "0" for the subjects who did not place both

feet on the ground, whether or not they used a manual or power wheelchair. In the future work that the algorithm could be expanded to power wheelchair users by applying a decision tree method. The displacements of the joint centers at the lower body segments during the setup phase, such as the KNEE, the ANKLE, and the FOOT, were calculated and inputted as the specific features for the item 7. However, the feature importance of these features is very low. It might be due to the “occlusion” effect that objects such as the wheelchair frame, wheels, and side guards have on infrared reflections and Kinect’s ability to accurately detect the positions of lower body joint centers when the subject is seated in the wheelchair. Thus, we only achieved moderate model performance on item 7 (AUC=.85, Accuracy=71%, FDR=.13, precision=.87, recall=.63). However, for our long-term goal developing a system that can watch a transfer and automate the TAI scores in real time, this is an item that is more obvious for therapists to score and may not need to be autoscored or supplemented by Kinect evaluation.

Item 8 focuses on hips scooting during the setup phase. The subject needs to scoot his hips to the front of the seat, with at least 1/3 of thigh off the surface. During our first attempt at feature engineering, we attempted to use the displacements of the HIP_RIGHT, the HIP_LEFT, and the SPINE_BASE during the setup phase as the specific features and the universal features to train the model. However, these specific features were not selected, and the model performance of the test set was very low. We noticed that if the wheelchair user did not scoot forward, he would need to lift his buttocks over the wheel during the transfer. Therefore, we added the features that represent the displacements of these three joints during the lift phase but only in the vertical direction (Y-axis). We assumed that the participants who did not scoot forward and transfer above the wheel should generate higher vertical displacements of the HIP_RIGHT, the HIP_LEFT, and the SPINE_BASE. After adjusting the selected features, the model performance increased (AUC=.87,

Accuracy=82%, FDR=.10, precision=.90, recall=.84), and the vertical displacement of the SPINE_BASE is the third most important feature of the trained model.

Both items 9 and 12 deal with the leading hand position during the transfer. For item 9, the proper technique suggests that the leading hand should not be placed behind the hip before the transfer. After the transfer (while the wheelchair user is already on the bench), the item 12 suggests that the distance between the leading hand and hip should not be over than 6 inches. We achieved the desired model performance (AUC=.85, Accuracy=86%, FDR=.05, precision=.90, recall=.84) for item 12 when using only the universal features. However, during our first attempt to train the item 9 model with the same features as item 12, the performance is fair. Thus, we added two features (length of the leading-side forearm and upper arm) for item 9, because we noticed that for persons with shorter arms, they needed to lean their trunk forward and use larger shoulder POE and elevation angles to reach a proper hand position than a person with longer arms. The most important features of items 9 and item 12 are related to the shoulder joint movement. After adding the specific features to item 9, the model achieved very good performance (AUC=.94, Accuracy=92%, FDR=.04, precision=.96, recall=.85).

To score items 10 and 11, the clinician needs to observe the push-off hand and leading hand position and handgrip during the lift phase, respectively. The TAI suggests that the subject should grip the surface edge, wheelchair armrest, frame, or wheel during the transfer rather than make a fist or place a flat hand on the surface. The subject can score “0.5” if he bends his fingers and places his hand on the surface. In the current database, we had the highest deficit rate of item 10 and 11 than any other items. Only 59% and 58% of the total trials were using proper techniques (or scored “1”). The Kinect SDK provides the joint center position of the HAND_TIP, the HAND, and the THUMB. We used their coordinates to calculate the triangle area between the three joints.

We assumed that the subject who performed the proper handgrip should have a larger triangle area than the subject who made a fist or bent their fingers. The hand area feature was the most important feature for item 10 and the second most important feature for item 11. However, the performances of the item 10 and 11 models only achieved moderate level (AUC=.82, Accuracy=77%, FDR=.08, precision=.92, recall=.67; AUC=.87, Accuracy=75%, FDR=.09, precision=.91, recall=.63, respectively). These features might also be affected by Kinect's ability to accurately detect the joint centers when the hand is in contact with a surface.

Item 13 suggests that the subject should lean their trunk forward, and away from the surface he was transferring to while transferring between surfaces. This body motion technique is often referred to as the "head-hip relationship", a motion strategy used for generating momentum for the lift phase. In the universal features we defined the trunk flexion angle using the angle between the trunk vector and the thigh vector. However, this feature might be affected by the height of wheelchair cushion, and the Kinect occlusion. To overcome these limitations, we generated another set of features by computing the angle between the trunk vector and the global Y vector and applied them as the specific features for item 13. The model reached the desired performance of the test set (AUC=.89, Accuracy=81%, FDR=.06, precision=.94, recall=.78) by using only nine features.

2.6 Limitations

The data labeling was based on the TAI which is typically scored after visual observation of the movement patterns and is subject to rater subjectivity and interpretation. To increase the reliability of the TAI score we video recorded each transfer and used these videos in accordance with the in- person scores to resolve any conflicts. For the data splitting process, we randomly separated all the trials into either a training, validation or testing set. This may impact the generalizability of our model results and the ability to make accurate predictions on new data since data generated from the same participant could have been used in more than one of the datasets [97]. It will be important to continue to test and adapt the models as new data becomes available.

Microsoft announced the discontinuation of Kinect v2 in 2017 and released a new version Azure Kinect DK (Kinect DK) in 2019. The new SDK is supposed to retain the full body tracking function however, the skeletal tracking algorithm of the new Kinect DK and the outputted raw data are very different than the Kinect v2 [54]. Moreover, research related to classification of body skeletal motion with the Kinect DK is not well developed. The Intel RealSense series is another type of depth sensor that has been used in several clinical research applications, however, work related to human motion classification is not as extensive as the Kinect in the published literature due to only the hand gestures and movement tracking function being supported by the Intel official SDK [55-58].

2.7 Conclusions

Our study demonstrates the potential of using a low-cost, portable camera and the developed ML models to facilitate autoscoring of the TAI. Our future plan is to embed the developed algorithms into a user-friendly graphic user interface that allows therapists to perform transfer assessments more easily and produce results that support training and education on proper transfer technique. Identifying transfer deficits early and more effectively may help reduce the prevalence of secondary injuries among wheelchair users.

3.0 Comparing Two Automated Methods to Detect Sitting Pivot Transfer Phases

3.1 Introduction

Using proper transfer technique can reduce the loading on the upper arm joints and help protect wheelchair users from developing injury and pain [4, 33]. The TAI is a 15 item scale used by clinicians and therapists to assess transfer quality and identify problems in wheelchair transfers which can cause increased forces on upper extremity joints [34, 35]. The TAI is based on clinical practice guidelines, current knowledge in the literature, and best clinical practices related to transfers. The TAI measures multiple different components of a transfer including proper setup of the wheelchair and body positioning during transfers. The tool is a series of yes or no questions that evaluate both the wheelchair user's overall technique and any weak component skills within the transfer [34, 35]. Higher TAI scores represent better wheelchair transfer technique [35]. Individuals who score highly on the TAI have lower mechanical loading at the shoulder, elbow and wrist in different transfer configurations [4, 7]. Therefore, wheelchair users who learn to perform transfers that are consistent with a high TAI score may reduce their risk of upper extremity injury and pain by decreasing the joint loading during the transfer.

While the TAI has been used successfully in research to evaluate proper transfer technique, there are limitations in its use for knowledge translation and application in a clinical setting. Firstly, clinicians need to become familiar with the TAI items to know if their patient's transfer was

performed correctly or not for each item. The intra-rater reliability of the wheelchair setup items is only on the moderate level (ICC = .44 to .55). The inter-rater reliability of the body setup items (ICC = .65 to .72) is also lower than other items [36]. The body mechanics and some setup items are very difficult to evaluate because the therapist needs to watch multiple movements for different items within 0.5 to 1 second during the lifting phase of the transfer. Secondly, according to our interviews in multiple rehabilitation institutions and hospitals, many therapists agreed that the TAI is great tool to identify specific deficits of patient's techniques but takes too long to perform in the clinic. Thus, the tool is currently used in research but has not yet been widely adopted into clinical practice. An automated system that can accurately and objectively observe transfer motions and report the TAI outcomes without requiring users to undergo extensive training could be of great benefit to therapists and patients with SCI in the future.

The potential to use the Kinect to accurately and reliably quantify transfer motions has been recently studied [36, 95]. The Kinect v2 sensor was able to discern differences in movement variables among unimpaired individuals who were trained to use proper technique and three variations of improper transfer technique [96]. In our previous studies, we demonstrated that the Kinect v2 and supporting machine learning (ML) models achieved an area under the receiver operating characteristic curve (AUC) of at least 0.79 and precision of at least 0.87 for the prediction classifiers of 11 TAI items [100]. Our long-term goal is to develop a system (Transkinect) that can watch a transfer and automate the TAI scores in real-time. Such a system may help to reduce therapist burden and facilitate easier and more reliable transfer assessments in clinical settings.

The TAI item scores are evaluated by the relative components of the wheelchair transfer. On the TAI 4.0, Items 1 to 9 are in the wheelchair setup skill group and items 10 to 15 are in the flight and landing motion skill groups. For a system to automate the scoring of the TAI, the system

must also be able to differentiate the “setup phase” and the flight and landing or “lift phase” of the transfer. In other words, in order to generate the features of the ML prediction classifiers of each TAI item (TAI-classifiers) [100], a process is needed to separate the lift phase from a time series motion capture data of a wheelchair transfer. Desroches and colleagues reported a threshold-based method to define the transfer phases by using a marker-based motion capture system and force sensor data [101]. However, we aim to automate scoring of the TAI using only the Kinect (motion data) to minimize the amount of instrumentation needed to recognize proper from improper technique.

To predict TAI scores using the TAI ML classifiers developed in prior work, the features based on the setup and the lift phases are computed. Any errors caused by mis-labelling the transfer phase could vary the features and then effect the accuracy of the predicted outcomes. The objective of this study was to compare two methods to automate detection of the setup and lift phases, a slightly modified version of the threshold method [101] versus a ML algorithm both using only the motion data from the Kinect v2. Both methods will be used to mark the start and end timepoints of the lift phase. After the start point of the lift phase is marked, the setup phase is defined as a duration between the beginning of the transfer trial and the start point of the lift phase (see Method section for more details). The aims of our study were to:

Aim 1: To determine the overall accuracy of each method in predicting the start and end points of lifting phases.

Hypothesis 1: Using timepoints marked by visual delineation as the gold standard, we hypothesized that the ML method would find start and end points of lift phase that were closer to the gold standard points than the threshold method.

Aim 2: To evaluate the accuracy of each method in predicting the TAI item scores.

Hypothesis 2: Using the true outcomes of the TAI-classifiers as the baseline, we hypothesized that the ML method would have higher accuracy in predicting TAI item scores than the threshold method for each item.

3.2 Methods

3.2.1 Participants

Participants were recruited through research registries, local SCI clinics and rehabilitation hospitals, and at organized recreational sport events. They signed consent forms approved by the Department of Veterans Affairs Institutional Review Board. The inclusion criteria were 1) have discernable neurological impairment affecting both lower extremities or persons with transfemoral or transtibial amputation of both lower extremities who do not use prostheses during transfers, 2) at least one-year post-injury or diagnosis, 3) able to independently transfer to/from a wheelchair without human assistance or assistive devices, 4) use a wheelchair for the majority of mobility (over 40 hours/week), and over the age of 18 years. Participants were excluded if they had 1) current or recent history of pressure sores in the last year, 2) history of seizures or angina, or 3) were able to stand unsupported.

3.2.2 Study Protocol

3.2.2.1 Experimental Setup

A Kinect v2 sensor was positioned two meters in front of the participants, 70 centimeters above the floor, and centered between the wheelchair and the bench (Figure 3). A custom graphical user interface was programmed in C# using Visual Studio 2012, .NET Framework 4.0, and the Kinect for Windows SDK to collect the 3D joint center position data in a Cartesian coordinate system from the Kinect sensor. The sampling frequency was 30Hz.

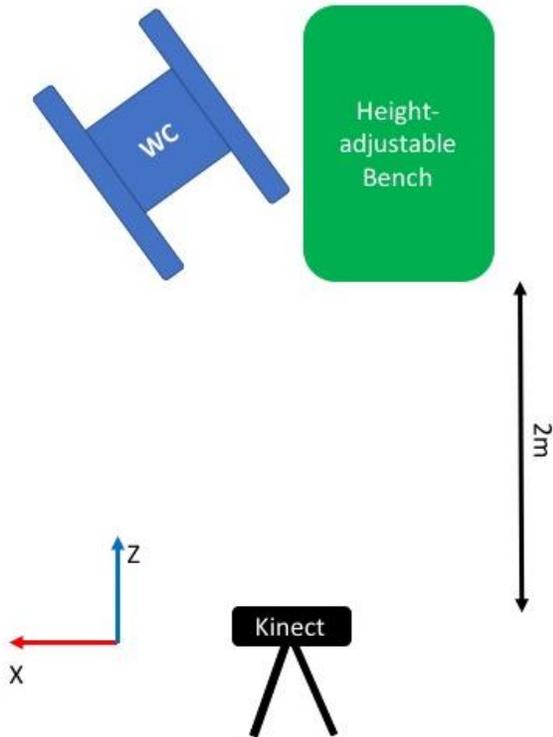


Figure 3. Experimental setup of wheelchair, bench and Kinect. The coordinate system follows the right-hand rule.

3.2.2.2 Transfer Protocol and Evaluation

All wheelchair users used their own chairs for the testing and were instructed to perform a transfer to a tub bench (70cm x 55cm) in a habitual way. Before data collection, the participants were provided an opportunity to adjust the position between their wheelchairs and the bench and also practiced the transfers to familiarize themselves with the setup. Participants were asked to transfer up to five times from the wheelchair to a level-height bench. Up to three raters who were experienced in scoring the TAI evaluated and scored each trial. The raters each had more than 2 years' experience in performing clinical research of wheelchair transfers. The participants were asked to sit on the bench for 5 to 10 seconds before they transferred back. After the participants transferred back to the wheelchair, the investigator would instruct the participants to start the next transfer. A transfer board was provided for the participants who requested it for the transfer. Opportunities for the participant to rest were built into the protocol. Participants could request a break or discontinue the study at any time. Digital video was recorded during all transfers to allow for review and resolution of potential discrepancies in expert scoring if necessary.

3.2.3 Data Analysis

Only the trials transferring from the wheelchair to the bench were analyzed in this study. The phases (start/end points) of the setup and the lift phases were determined using three methods: visual (gold standard), threshold, and ML.

3.2.3.1 Visual Method of Phase Delineation

The videos recorded during each subject's transfer by the Kinect RGB camera were

reviewed and used as the gold standard data of the transfer phases timeline. The “setup phase” was defined as the period when the subject was ready for the transfer trial (beginning of the video file), to a time point right before the subject lifts their buttocks from the sitting surface. The “lift phase” was defined as starting from the end of setup phase to the first time frame when the subject lands on the target surface (Figure 4).



Figure 4. Phases of a subject transfer from a wheelchair to a level-height bench. a: start of the setup phase, the participant is sitting on the wheelchair; b: transition point between the setup phase and the lift phase, the participant is about to lift his body from the wheelchair seat; c: end of the lift phase, the participant landed on the bench.

3.2.3.2 Threshold Method of Phase Delineation

Linear velocity and displacement in the X-axis of the SPINE_BASE (Figure 5), a joint center marker of the Kinect SDK, was chosen to define the beginning and the end of the lift phase, because all the subjects have a similar pattern of the time-series data in these two parameters through a transfer trial.

- Beginning of lift phase:

One negative peak value is observed in time-series data of the SPINE_BASE linear resultant velocity in all subjects. The beginning of the lift phase is determined as the last frame that the velocity is less than 0.01 m/s before the peak occurs (Figure 6a).

- End of lift phase:

The end of the lift phase is determined when the SPINE_BASE reaches the maximum displacement after the beginning of the lift-phase (Figure 6b).

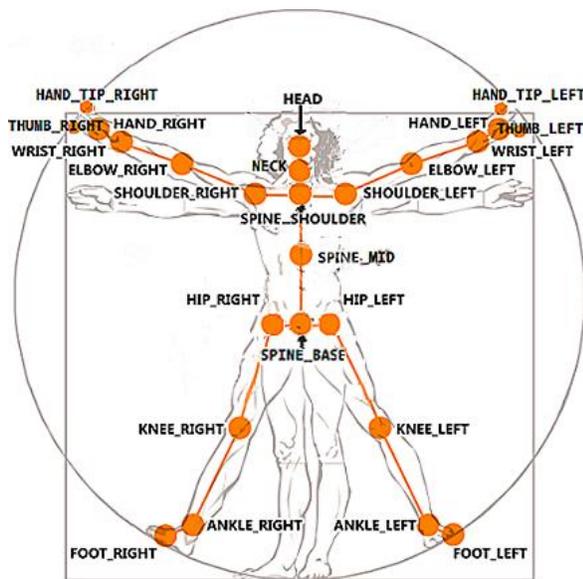
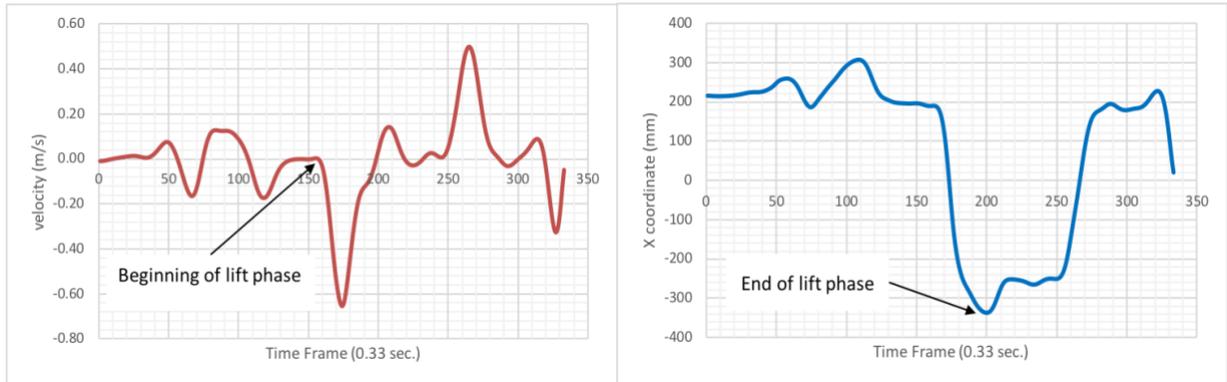


Figure 5. Skeleton position detected by Kinect relative to the human body [94]



(a)

(b)

Figure 6. (a) Determine the beginning of the lift phase by evaluating linear velocity of SPINE_BASE, (b) Determine the end of the lift phase by evaluating displacement in the X direction of SPINE_BASE

3.2.3.3 Machine Learning Method of Phase Delineation

- Database:

Data collected from 95 transfers performed by 95 wheelchair users (52,777 time frames, about 30 minutes) was analyzed.

- Labeling:

The videos recorded during each subject’s transfer by the Kinect RGB camera were reviewed as the “ground truth” of the transfer phases. The timelines of the Kinect motion data were synchronized with the timeline of the video files. Time frames in a transfer trial were manually labelled as either setup or lift phase.

- Features:

The Kinect v2 can track 25 skeleton joint centers in 30 Hz (Figure 5). Each frame’s linear displacements, velocity, acceleration, jerk of the X, Y, Z coordinates of the SPINE_BASE, SPINE_MID, SPINE_SHOULDER, and HEAD, were calculated as the features of the ML model.

Other features, such as shoulder plane of elevation angles, shoulder abduction/adduction angles, elbow, wrist and trunk flexion angle on both arms were computed from the joint center data (see **2.3.2.4 Feature**, universal features, page 22). There was a total of 35 features imputed into the feature engineering analysis.

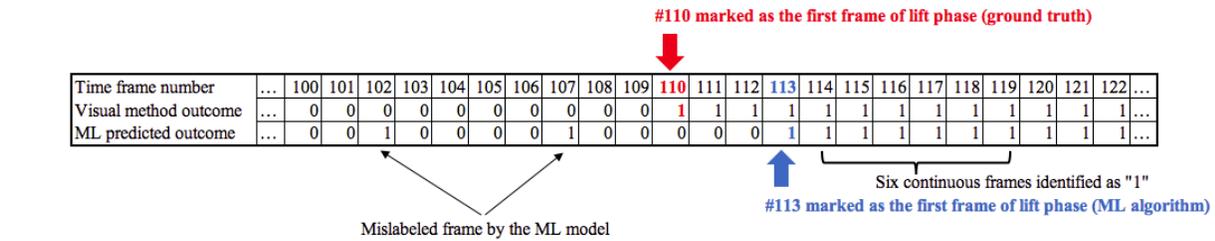
- Training Approaches:

The data were divided into three sets: a training set (80% of transfer trials) used to learn model parameters and to build the cost functions, a validation set (20% of training data set) used to search the best hyper-parameters of the classifier, and a test set (20% of transfer trials) used to assess model performance. A 5-fold cross-validation was applied between the training set and the validation set to tune the models. The K nearest neighbors classifier (KNN) was selected for item models as it showed the best performance after initially testing 16 supervised ML classifiers (see **Appendix D**). The number of the features was determined by the recursive feature elimination with cross validation analysis. Feature engineering, model training, and cross validation were practiced by applying the Scikit-learn python library. For the test set, the model's area under the receiver operating characteristic curve (AUC), and model accuracy were computed to report the model performance.

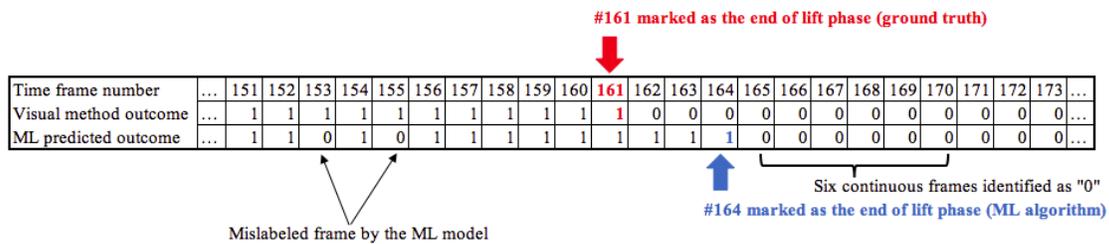
- Timepoints:

After applying the trained KNN model to a time series data of a transfer trial, each time frame would be labeled as either “1” (lift phase) or “0” (setup phase). To enhance the accuracy of the ML prediction method another program was applied as a filter to remove mislabeled frames. The program started out by searching the first frame that had six consecutive frames that were predicted as “1”. This frame would be marked as the beginning of the lift phase. Then the program

searched for the next six consecutive frames labeled as “0”. This frame would be marked as the end of the lift phase (Figure 7).



(a)



(b)

Figure 7. Example showing how to mark the start (a) and end (b) frame of lift phase by applying the machine learning (ML) algorithm. Each time frame from a transfer is labeled by the trained KNN classifier as “1” (lift phase) and “0” (setup phase). a) A frame (#110) is marked as the beginning of lift phase if the following six frames are all predicted as “1”, b) A frame (#161) is marked as the end of lift phase if the following six frames are all predicted as “0”.

3.2.3.4 TAI Score Predictions

Using the timepoints marked by the visual, threshold, and ML methods, the 11 item TAI-classifiers developed in previous work were applied to predict the TAI item scores (see **Chapter**

2.0, page 15). The time points and the joint center data served as the inputs to a program coded in MATLAB R2020a to calculate the features for the TAI-classifiers [100]. The TAI predicted outcomes are generated after applying these features to the TAI-classifiers.

3.2.4 Statistical Analysis

3.2.4.1 Aim 1: To determine the overall accuracy of each method in predicting the start and end points of lifting phases

Fifty trials performed by a randomly selected ten wheelchair users from our database were analyzed for the main outcomes of this study. The average and standard deviation (STD) of the start and end timepoints of the lift phase from five transfers performed by each participant was computed. Time = 0 indicated the initiation of the transfer (i.e. the start point of the setup phase). In order to examine the errors of timepoint marking for the two sensors, a mean of differences of the start/end lift phase timepoints between each automated method and visual method was computed:

$$MD = \sum_{i=1}^n \frac{x_i - x_g}{n} \quad (1)$$

, where MD is the mean of differences; n is sample size; i indicates the trial number, x_i is each subject's start or end timepoint (in seconds) marking by the threshold or ML methods, x_g is the timepoints marked by the visual method. An absolute value of the MD (AMD) was also calculated

to examine the scalar of differences (ignores if the timepoint is marked before or after the gold standard):

$$AMD = \sum_{i=1}^n \frac{|x_i - x_g|}{n} \quad (2)$$

Histograms of the MDs and AMDs for all transfers were created. The Bland-Altman plots were used to investigate the agreement between the visual and the TH method, and between the visual and the ML method. The “agreement” in this study is the characteristic that describes how close the two measurements are. The infimum (inf) and supremum (sub) of the agreement will be set as:

$$inf = MD - 1.96 \times STD, \quad (3a)$$

$$sub = MD + 1.96 \times STD, \quad (3b)$$

3.2.4.2 Aim 2: To evaluate the accuracy of each method in predicting the TAI item scores

The 50 trials performed by the randomly selected ten wheelchair users were excluded from the ML training and model tuning process. To generate the predicted TAI outcomes, we applied the TAI-classifiers for 11 TAI items. Each model generates a dichotomized (Yes/No) predicted item score. The accuracies of each TAI-classifier item ranges from 71% to 92% [100]. Using the TAI scores rated by the clinicians as the true outcome and comparing with the predicted scores from the visual method, all the true positive and true negative trials were identified and used as the “baseline” trials for testing the accuracy of the two methods. The false positive and false negative trials scored by the TAI-classifiers were eliminated from the analysis. The accuracy of the threshold or ML method in predicting the TAI scores was defined as:

$$Accuracy = \frac{T}{BL} \times 100\% \quad (4)$$

where BL is the number of baseline trials; T is the number of trials correctly scored (i.e. the predicted TAI scores matched the BL) by applying the timepoints from the threshold method or

the ML method.

3.3 Results

3.3.1 Participants

The demographics of the ten wheelchair users who were randomly selected from the database of 81 participants and used to examine the accuracy of the threshold and ML methods are shown in Table 8.

Table 8. Participant demographics

P	Disability	Gander	Ethnic	WC	Age	Year	Transfer	TAI	
1	SCI, T4 complete	Male	Caucasian	Manual	44	2	10	5.5	
2	Amputation, left leg	Male	African American	Power	57	2	10	1.8	
3	SCI, T10 complete	Female	Hispanic	Manual	48	22	24	6.4	
4	Amputation, below knee	Male	African American	Power	61	13	12	8.2	
5	SCI, C5-6, incomplete	Male	African American	Power	70	11	14	5.5	
6	Amputation, below knee	Male	Caucasian	Power	47	24	8	4.0	
7	SCI, T12 incomplete	Male	Caucasian	Manual	57	14	4	7.3	
8	SCI, L incomplete	Male	(no answer)	Manual	62	27	6	9.1	
9	Amputation, below knee	Male	Caucasian	Manual	69	13	4	8.2	
10	Poliomyelitis	Female	Hispanic	Power	55	48	9.5	4.5	
					Mean	57.0	17.6	10.2	6.0
					STD	8.9	13.6	5.9	2.2

Key: P, participant number; WC, wheelchair type; Year, years using wheelchairs; Transfer, times of transfer per day (self-reported); TAI, total TAI scores (average of item scores multiply 10, “yes” scores 1 and “no” scores 0); STD, standard deviation.

3.3.2 KNN Model performance

Seventy-six participants’ transfer data were split into the training set and 19 participants’ data were split into of the test set. The model was tuned to the achieve the best performance of AUC (n_neighbors=5, p=1, leaf_size=25) by only using the training set data. The average accuracy of the validation set was .97 (STD = .00146), and the AUC of the tuned training set was .99. Each data (~13,987 frames) in the test set were predicted by KNN model as “setup” or “lift” phase. The AUC of the test set was .99. The confusion matrix of test set is shown in the Table 9.

Table 9. The confusion matrix of the test set of the machine learning methods KNN model

		Predicted phase	
		Setup	Lift
Ture phase	Setup	10107	98
	Lift	163	3619

3.3.3 Start/End timepoints of lift phase (Aim 1)

Table 10 shows each participant’s mean start/end timepoints of the lift phase over the five trials. The average duration of the lift phase in our database is 2.45 seconds.

Table 10. Means and STDs of each participant’s start and end timepoints marked by visual, TH, and ML method. The values indicate when (in seconds) the participant started the lift phase of transfer after the trial began (timepoint = 0)

P	Start time point of lift phase (unit: sec.)						End time point of lift phase (unit: sec.)					
	Visual		TH		ML		Visual		TH		ML	
	Mean	STD	Mean	STD	Mean	STD	Mean	STD	Mean	STD	Mean	STD
1	10.13	3.39	9.87	3.59	8.34	3.50	11.09	3.81	11.04	3.52	10.27	2.89
2	2.16	0.76	2.43	0.55	2.75	0.83	3.93	0.72	5.55	2.20	3.85	0.46
3	4.38	1.70	6.31	1.51	8.44	3.76	7.24	1.71	8.73	2.59	9.15	3.77
4	4.07	1.16	3.67	0.84	4.06	1.11	4.83	1.27	5.39	1.43	4.61	0.83
5	7.99	4.02	11.29	2.98	7.57	2.94	11.17	3.50	16.33	4.39	9.41	3.06
6	3.70	1.07	4.85	1.81	2.70	1.00	4.81	1.02	6.75	1.65	4.43	2.30
7	4.09	1.10	3.40	1.96	2.72	1.43	7.13	1.36	8.85	1.52	5.69	2.82
8	2.01	0.81	2.59	1.23	2.44	0.97	4.01	0.91	5.23	1.53	3.95	1.14
9	1.12	0.86	1.34	1.17	1.29	0.73	3.55	0.69	3.50	0.74	3.37	0.81
10	2.33	1.96	6.37	1.89	3.34	2.95	8.51	2.98	9.84	4.46	6.72	2.82

Key: P, participant number; TH, threshold method; ML, machine learning method; STD, standard deviation.

For the start timepoint of the lift phase, the MD with the threshold method was 1.01 (STD=1.99, median=0.42) second and the ML method was 0.17 (STD=1.96, median=0.07) second, in comparison with the visual method. For the end timepoints, the MD between the threshold method was 1.49 (STD=1.92, median=0.58) second and the ML method was -0.48 (STD=1.96,

median=0.13) second, in comparison with the visual method. The histograms for all transfers (n=50) are shown in Figure 8 and 9.

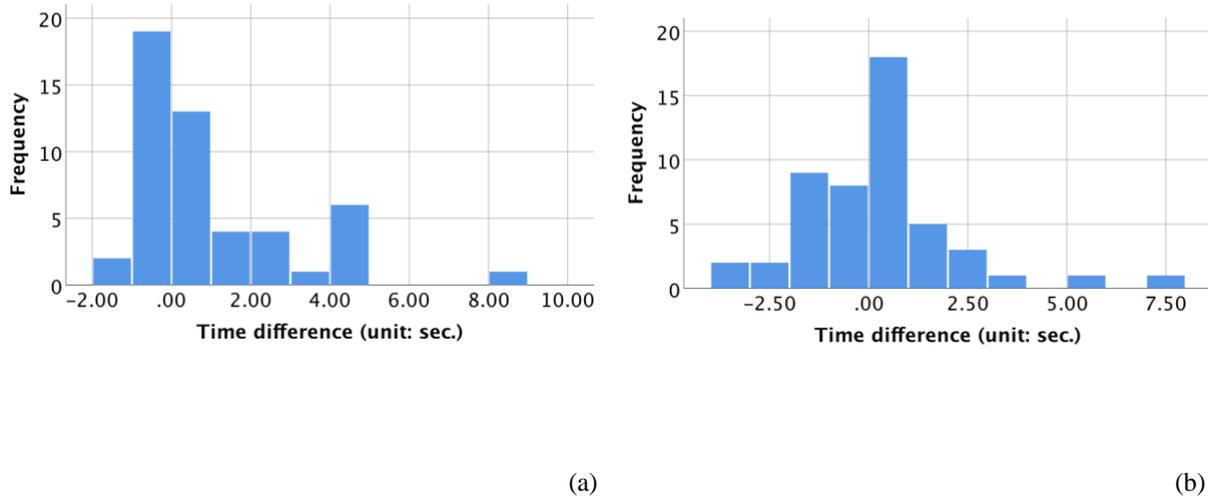


Figure 8. Histograms of the start timepoint differences between (a) TH and visual methods and (b) ML and visual methods between for each transfer (n=50)

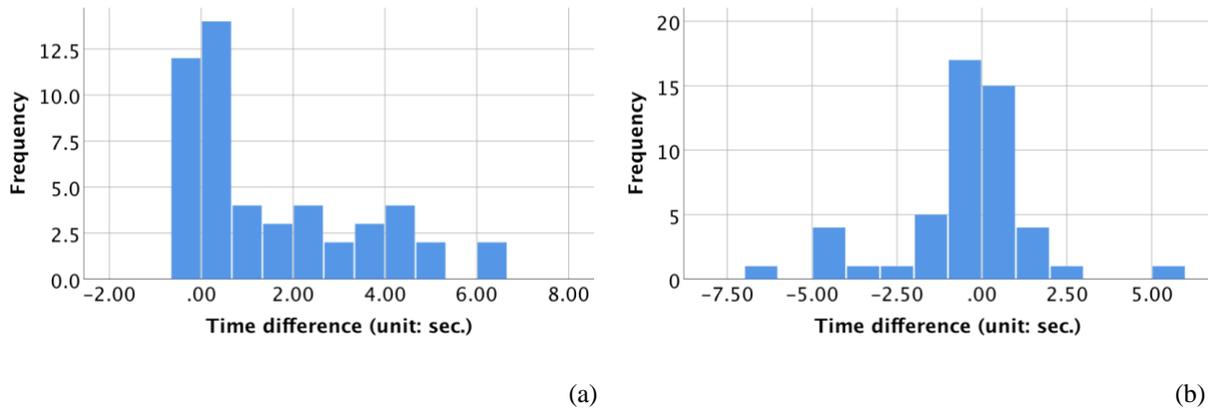


Figure 9. Histograms of the end timepoint differences between (a) TH and visual methods and (b) ML and visual methods between for each transfer (n=50).

For the start timepoint of the lift phase, the AMD with the threshold method was 1.42 (STD=1.72, median=0.68) seconds and the ML method was 1.30 (STD=1.46, median=0.77) seconds, in comparison with the visual method. For the end timepoints, the AMD between the threshold method was 1.58 (STD=1.85, median=0.58) second and the ML method was 1.21 (STD=1.61, median=0.50) seconds, in comparison with the visual method. The histograms (n=50) are shown in Figure 10 and 11. The Bland-Altman plots for the agreement of the start and end time points between the visual and TH methods, and between visual and ML methods are shown in Figure 12 and 13.

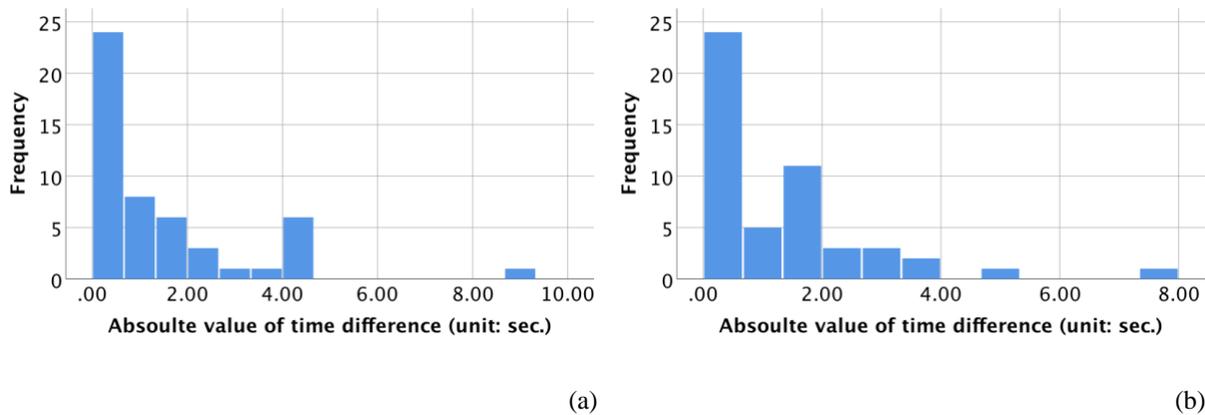
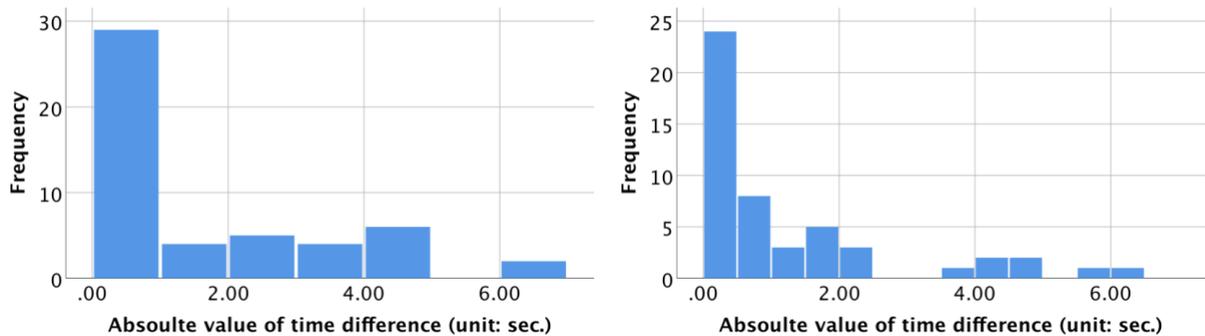


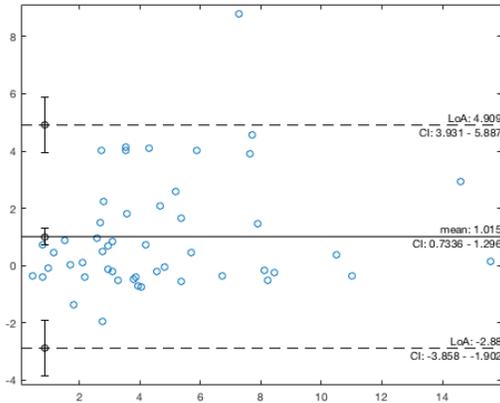
Figure 10. Histograms of the start timepoint differences (absolute values) between (a) TH and visual methods and (b) ML and visual methods between for each transfer (n=50)



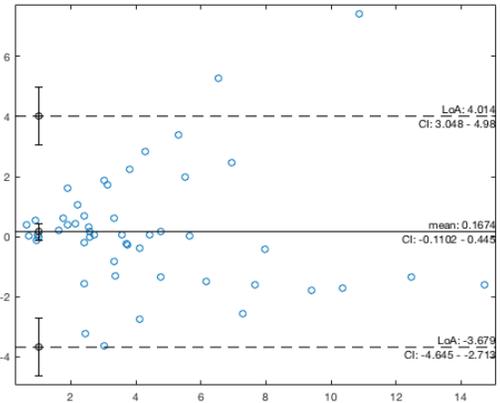
(a)

(b)

Figure 11. Histograms of the end timepoint differences (absolute values) between (a) TH and visual methods and (b) ML and visual methods between for each transfer (n=50)

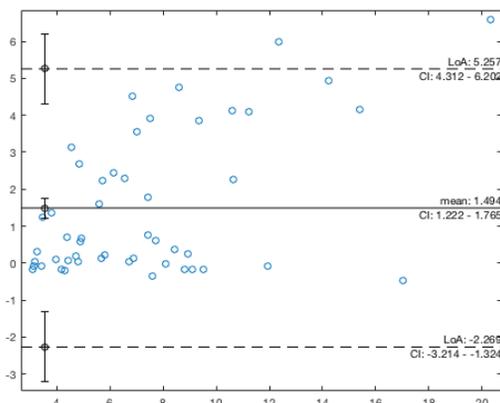


(a)

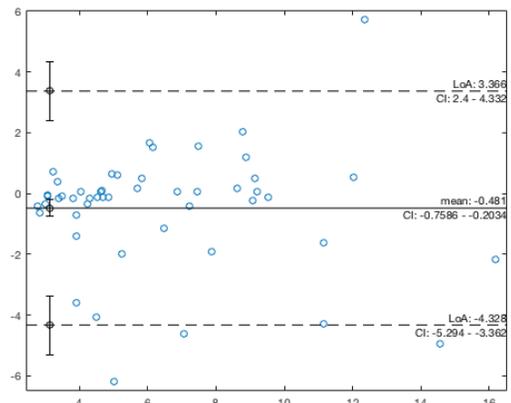


(b)

Figure 12. Bland-Altman plots for the start time points marked by using the (a) TH and visual methods and (b) ML and visual methods (n=50)



(a)



(b)

Figure 13. Bland-Altman plots for the end time points marked by using the (a) TH and visual methods and (b) ML and visual methods (n=50)

3.3.4 Transfer Quality Evaluation (aim2)

The average accuracy of the threshold method is slightly higher than the ML method for predicting TAI scores (91% vs 88%) (Table 11). For some items related to the setup phase (item 1, 2, 7) and trunk flexion (item 13), the ML method generated a more accurate outcome. However, the threshold method had higher accuracy of the items related to the lift phase (item 10, 11, 12, 14, 15). The two methods have the same accuracy to item 8 (scouting forward) and item 9 (leading hand position before transfer).

Table 11. Accuracy of the TAI predicted outcomes after applying the threshold and the machine learning timepoint marking methods

TAI item		N	Method	Accuracy
Wheelchair			TH	95%
Setup	item 1	44	ML	98%
	WC distance			
	item 2	43	TH	91%
			WC angle	ML
Body			TH	95%
Setup	item 7	39	ML	97%
	Feet down			
	item 8	33	TH	91%
			Scouting forward	ML
	item 9	38	TH	92%
			L-hand position (before)	ML

Flight	item 10	Push-off hand handgrip	35	TH	92%
				ML	84%
	item 11	L-hand handgrip	34	TH	88%
				ML	75%
	item 12	L-hand position (after)	49	TH	100%
				ML	92%
	item 13	Body leaning	39	TH	85%
				ML	87%
	item 14	Flight	48	TH	85%
				ML	81%
	item 15	Landing	47	TH	81%
				ML	72%
Average				TH	91%
				ML	88%

Key: N, number of transfer trials correctly scored by the TAI-classifiers out of 50 trials performed by the 10 participants; TH, threshold method; ML, machine learning method.

3.4 Discussion

Applying an automated method that can extract the setup and lifting phase timing is an essential step to developing an automated TAI scoring system. In this study we compared two automated methods, one based on salient peaks in the motion data recorded during the transfer and one based on ML methods. Provided that the features used in our TAI-classifiers are not solely

based on these time points but rather the data that lies in between them it's possible that even if the timepoints are mis-labelled, the approaches could still yield accurate TAI scores, thus we aimed to explore not just the mean differences in setup and lift phase timings but the resultant predictions in the TAI scores.

The MD of the start/end points are 1.01 and 1.49 seconds for the TH method, 0.17 and -0.48 seconds for ML method. The positive and negative values mean that the method either overestimated (positive) or underestimated (negative) the timepoints on average. The histograms (Figure 8 and 9) show that most timepoints marked by the TH and ML methods are within ± 1 second of the gold standard points. However, the errors of underestimating and overestimating can cancel each other out. For example, if the time differences compared to the gold standard of two trials are -2.00 and +2.00 seconds with method A, and -0.5 and +1 second with method B, the MD of method A is 0, and MD of method B is 0.25. In this case, method A would appear to outperform method B yet method B is actually better than method A for these data.

Thus, alternatively we calculated the AMDs to further examine the accuracy of the TH and ML methods. The AMD is the absolute value of the average variance between the gold standard and the TH/ML methods, thus it accounts for the variance from both from underestimating and overshooting the targets. The AMD of the start/end points were 1.42 and 1.58 second for the TH method, 1.30 and 1.21 seconds for ML method. The data distribution and outliers are similar for the two methods (Figure 10 and 11).

The ML method showed less error in phase identification (lower MDs and AMDs than the TH method) but achieved slightly lower overall accuracy (88%) in the TAI score predictions. The flight item (item 10 to 15) also has lower accuracy (72% - 92%) while applying the ML method. The decreased accuracy might be due to the ML method underestimating the end of the lift phase

by an average of 0.48 seconds. The TH method has less underestimating error (Figure 13a; notice that almost all dots are above horizontal line equal to zero) compared to the ML method (Figure 13b). These results illustrate that the ML method might end the lift phase earlier than the gold standard for several trials. As noted offsets in the start and end points of the lift phase could change the values of features and then vary the predicted outcomes of the TAI-classifier. Underestimated end timepoints of the end of the lift phase might acutely alter the features of the TAI-classifier related to the lift phase (e.g. the maximum, minimum, range of motion, and average joint angles) since the data only includes a portion of the lift phase duration. Tuning the model of the ML method to achieve higher recall (increasing the possibility that the model labels the setup phase frames as the lift phase) could reduce the underestimating the end of the lift phase. Further studies are needed to understand how much the specific features change with regard to phase timing. Moreover, since both methods are not 100% perfect for all items, the decreased accuracy is potentially added error onto the errors associated with the TAI classifiers. Future studies that investigate the tolerance of timepoints mis-labelling for each TAI-classifier are needed.

One specific defect of the threshold method that we noticed is mis-identifying the phases when a wheelchair user performs multiple ‘scoots’ to complete the transfer. According to our database, this transfer technique occurs more often with power wheelchair users than manual wheelchair users. In this study, participant 1 and 10 were scooting multiple times to transfer from their power wheelchair to the bench. The AMD of the threshold method in the two participants are approximately 3.5 seconds, and the ML method error is 1.2 seconds. Thus an ML method could be a better approach to use for evaluating power wheelchair users.

Another factor to consider in addition to the accuracy of the TAI score predictions is the computation time. The threshold method uses only two kinematic variables, linear velocity and

displacement of one joint center to identify the start and end frames of the lift phase. In comparison, 35 features/variables need to be computed for the ML method. In addition to computing the features, the ML method needs to perform multiple post-processing steps before identifying the time frame. Using the ML method requires 60-120 seconds (Intel i9 3.1GHz CPU with 32GB memory) more than the threshold method to generate the TAI predicted scores. As a result, the threshold method is less computationally expensive. This factor may need to be considered in the development of an automated TAI scoring system.

3.5 Limitation

One limitation of this study is that a small sample size (n=10 wheelchair users) was used to evaluate the accuracy of the TAI score prediction. The results may not generalize to the broader wheelchair user population or other variants of transfer technique. Another possible limitation was that a visual (video) method was used as the gold standard to identify start and end points of the lift phase. This was the best method available to us given that we lacked a fully instrumented setup (e.g. force plates or sensors) which would have provided more accurate information on the “ground truth” of the time points [101]. To support recruiting a large enough sample for training the ML models, participants were recruited and the study protocol was conducted outside of the lab at several organized recreational events and it was not possible to transfer our fully instrumented lab

setup to these environments for the data collection. However recent work does support the use of video methods for enhancing the reliability and validity of TAI scoring [102].

3.6 Conclusion

To automate a real-time TAI score using computer methods requires identifying when the setup and lift phases of the transfer occur. Our research tested two possible ways to do this using only motion data recorded by the Kinect sensor during transfers. The ML method had less error in phase identification but has lower accuracy in TAI score prediction and took longer to process. Tuning the model of the ML methods to have higher recall (avoiding cases of ending the lift phase too early) may increase the accuracies of the predicted TAI outcomes. Further research is needed to study the overall accuracies associated with combining the automated phase identification methods with the TAI classifiers that have been developed to predict TAI item scores.

4.0 Comparison of Two Depth Cameras for Capturing Upper Body Motions During Wheelchair Transfers

4.1 Introduction

In the United States, there were approximately 282,000 persons with spinal cord injury (SCI) in 2016 and 12,500 new cases occur each year [103]. Over 3.6 million Americans aged 15 and over used a wheelchair in 2010 [2]. Wheelchair users (WUs) rely heavily on their upper extremities to complete common but essential activities of daily living such as getting in and out of bed, transferring to a toilet or a shower, and transferring in and out of a car. Manual WUs will perform on average 14 to 18 transfers a day, which are extremely physically demanding and can lead to upper extremity pain and injury [3, 4]. Research shows that the prevalence of upper extremity pain, specifically shoulder pain, in WUs ranges between 31 and 73 percent [5]. Unfortunately, shoulder pain leads to decreased quality of life and participation in physical activity [6].

The Transfer Assessment Instrument (TAI) was developed to evaluate the quality of sitting-pivot wheelchair transfer techniques and identify any deficits in component skills [36]. Higher scores on the TAI (e.g. using better hand/arm and trunk positions to perform transfers) translate to less mechanical loading on the upper extremities joints [4]. The TAI measures many different components of a transfer including proper setup of the wheelchair and body positioning

during transfers. Using proper transfer technique can reduce the loading on the upper arm joints and help protect WUs from developing injury and pain [4, 33]. Previous work has shown that the Microsoft Kinect, a portable inexpensive markerless 3D depth motion sensor, can be used to distinguish proper and improper wheelchair transfer techniques [31, 96, 104-106]. A 3D depth sensor allows total freedom of movement without the need to hold or wear any sensors or markers on the body during the transfer task thus reducing setup time and effort. A 3D depth camera is also cheaper than other forms of motion capture and can provide more detailed motion tracking data (e.g. x, y, z coordinate positions of joint centers) in comparison to other portable sensors (e.g. inertial measurement units (IMUs) which provide relative segment orientations--raw position data are possible but post-processing methods are required). For these reasons, many researchers have developed clinical applications with the Kinect for assessment of balance and postural control [60-64], dynamic balance tests [65], fall prevention [66, 67], and upper extremity motor and functional recovery [68-71] .

Microsoft discontinued the Kinect v2 sensor in 2017. After some time had passed the company eventually introduced Microsoft Azure a more compact and technologically advanced 3D depth sensor. In addition, there are many other alternative 3D depth sensors on the market. The Intel® RealSense™ D435 for example, is a higher resolution depth camera than the Kinect v2. The Intel RealSense series has been used in several clinical research applications, however, work related to human motion classification is not as extensive as the Kinect in the published literature due to only the hand gestures and movement tracking function being supported by the Intel official SDK [55-58]. Due to the discontinuation of the Kinect v2, which was originally used to develop our algorithms for detecting proper from improper technique, we aimed to investigate if the Intel RealSense could be used as potential surrogate sensor as the Azure was still in production at the

time of this study. The Intel RealSense series has been used in several clinical research applications, however, work related to human motion classification is not as extensive as the Kinect in the published literatures [55-58]. One study by Mistry et al. developed an approach to translate sign language by using the hand motion data recorded by the RealSense [59].

In this study we aimed to compare the performance of the RealSense D435 sensor to the Kinect v2 for tracking wheelchair transfer motions. We hypothesized that the RealSense-measured motion variables would be statistically repeatable within subjects over five transfer trials (intra-rater reliability > 0.8) and that the repeatability would be similar between the Kinect and the RealSense for each variable (intra-class correlation coefficient (ICC) differences < 0.1) (Hypothesis A). In addition, we expected that the motions measured by the RealSense would relate to those measured by the Kinect (inter-rater reliability > 0.8) and be within the range of 95% limits of agreement as the mean difference plus and minus 1.96 times the standard deviation of the differences (Hypothesis B).

4.2 Methods

4.2.1 Participants

The study was approved by the Department of Veterans Affairs Institutional Review Board. The participant recruitment and the testing were conducted at the National Veterans Wheelchair

Games (NVWG) in 2019. All participants provided informed consent prior to participating in any research activities. The inclusion criteria were 1) have discernable neurological impairment affecting both lower extremities or persons with transfemoral or transtibial amputation of both lower extremities and who do not use prostheses during transfers, 2) at least one-year post-injury or diagnosis, 3) able to independently transfer to/from a wheelchair without human assistance or assistive devices, 4) use a wheelchair for the majority of mobility (over 40 hours/week), and over the age of 18 years. Participants were excluded if they had 1) current or recent history of pressure ulcers in the last year, 2) history of seizures or angina, or 3) were able to stand unsupported.

4.2.2 Study Protocol

4.2.2.1 Motion Sensors and Experimental Setup

The Kinect v2 sensor was positioned two meters in front of the participants, 70 centimeters above the floor, and centered between the wheelchair and the bench. A custom software program was created using Windows Kinect SDK to collect the 3D joint center position data. A graphical user interface (GUI) was programmed in C# using Visual Studio 2012, .NET Framework 4.0, and the Kinect for Windows SDK to collect the 3D joint center position data in a Cartesian coordinate system from the Kinect system. The RealSense D435 camera was positioned in a similar proximity in front of the participant (Figure 14) in accordance with the manufacturer specifications for camera placement [107] and was operated by a third-party NuiTrack™ SDK released by the 3DiVi Inc. Each sensor was connected to its own computer. The raw data of the two sensors were saved as a .csv (comma separated value) file, with columns of joint-centers position data, in rows indexed by time in milliseconds, also referred to as frames. The sampling frequency was 30Hz for both

sensors. The two sensor's data were synchronized to start and end simultaneously by using the timestamps data.

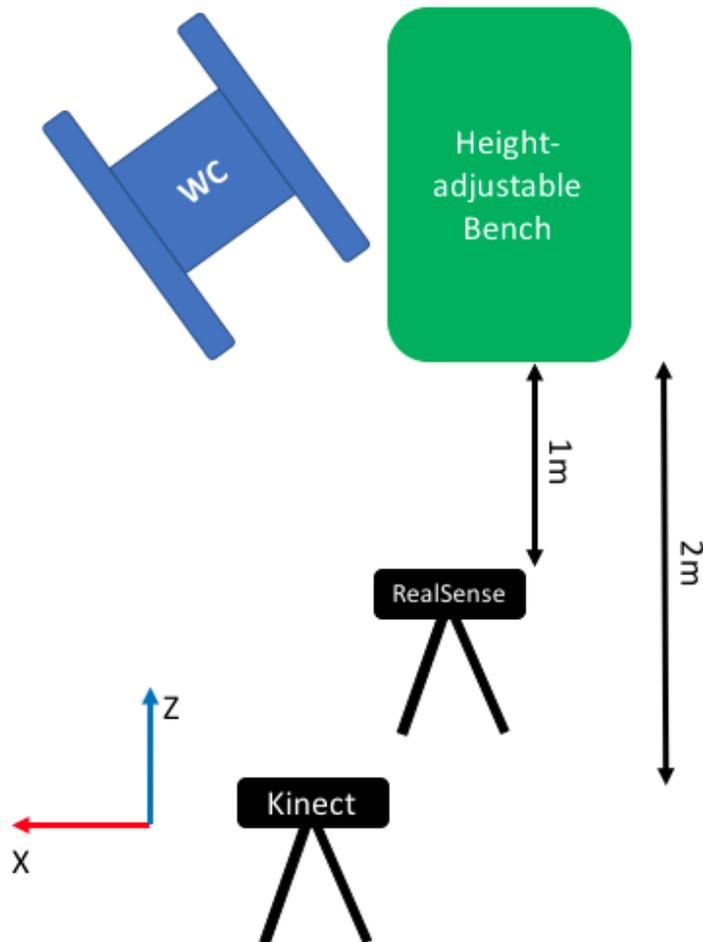


Figure 14. Experimental setup of the Kinect and the RealSense. The coordinate system follows the right-hand rule for both sensors.

4.2.2.1 Transfer Protocol

After informed consent, participants completed a general questionnaire. Then the participant's wheelchair was positioned next to a level-height bench (70cm x 55cm) based on their transfer preferences. The bench was placed on the participant's left-hand side. The participants

were asked to transfer from their wheelchair to a level-height bench. The participant repeated the transfer up to five times for a maximum total of 10 transfers. Approximately 3-5 minutes of rest time was provided between trials and the participants were allowed to take more time to rest until they were ready for the next trials.

4.2.3 Data Analysis

4.2.3.1 Key Variables

The Kinect v2 tracks the 3D positions of 25 skeleton joint centers at 30 Hz [94, 108]. Similarly, using the NuiTrack SDK, the RealSense tracks the 3D positions of 19 joints in a global coordinate system [109] (Figure 15). Four key kinematic variables related wheelchair transfer biomechanics [4] and the TAI [100] were analyzed. All the variables are computed from the “lift phase” of the transfer motion time series data which was determined using the video files synchronized with the time frames to the Kinect and RealSense (see **3.2.3 Data Analysis**). The key variables included:

- SPINE_BASE/WAIST displacement in the horizontal direction (DSBW): X-axis component displacement of SPINE_BASE (Kinect) and WAIST (RealSense)
- Average plane of elevation angle on the leading side shoulder (LPOE): The average value of left shoulder horizontal flexion/extension during the transfer [110]. We used the joint centers at the left shoulder and left elbow to create a vector that represents the upper arm, and used the SPINE_SHOULDER/COLLAR and the joint center of the left shoulder to create another vector. The LPOE is determined as the angle between the two vectors projected onto the transverse plane

- Average elevation angle on the leading side shoulder (LE): The average value of the left shoulder elevation angle during the transfer. The LE is determined by the angle between a trunk vector (from SPINE_SHOULDER/COLLAR to SPINE_BASE/WAIST) and the upper arm vector (from LEFT SHOUDLER to LEFT ELBOW)
- Average trunk flexion angle (TF): The average value of the trunk flexion angle during the transfer. The TF is determined as the angle between a trunk vector (from SPINE_BASE/WAIST to SPINE_SHOULDER/COLLAR) and a normal vector perpendicular to the transverse plane

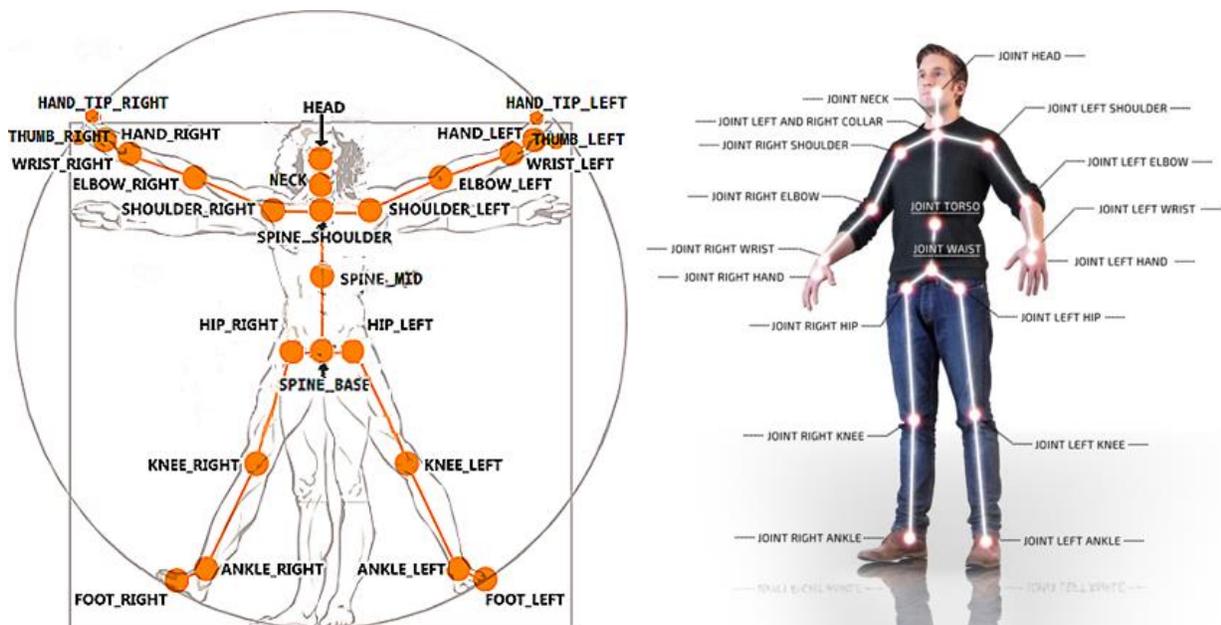


Figure 15. a) Skeleton position detected by Kinect relative to the human body [94], b) Joint center map in the NuiTrack SDK [109]

4.2.3.2 Statistical Analysis

All the statistical analysis was performed using SPSS 26 (Chicago, IL). For the *Hypothesis a*, the intra-rater reliability for the Kinect and the RealSense was calculated using ICC_{3,1} (two-way mixed effects model with absolute agreement for a single measure) for the five repeated trials.

For the *Hypothesis b*, the inter-rater reliability between the Kinect and the RealSense were calculated using ICC_{2,1} (two-way random effects model with absolute agreement for a single measure) [111]. In this study, repeatability was characterized as excellent (ICC>0.8), good (ICC 0.6–0.79), moderate (ICC 0.4–0.59), fair (ICC 0.2– 0.39) or poor (ICC<0.2). To test agreement between the key variables (i.e. DSBW, LPOW, LE, and TF) measured by the Kinect and the RealSense, the Bland-Altman plots were used. The “agreement” in this study is the characteristic that describes how close the two measurements are. The infimum (inf) and supremum (sub) of the agreement will be set as:

$$\begin{aligned} \text{inf} &= m - 1.96 \times STD, \\ \text{sub} &= m + 1.96 \times STD, \end{aligned} \quad (1)$$

where m is the mean of differences of both measures (Kinect and RealSense) and STD is standard deviation. Based on the Gaussian hypothesis, if the 95% of the data are within the range between the inf and sub, it is valid to affirm that the two methods are interchangeable [112].

4.3 Results

4.3.1 Participant

Twenty-six men and 4 women with an average age of 56.6 years (standard deviation (STD)=11.8) contributed a total of 150 transfer trials for the analysis. The group performed on average of 13.0 transfers (STD=10.9, self-reported) per day. Participants had 16.8 years (STD=5.76) experience in using wheelchairs and used their wheelchair for 13.2 hours (STD=5.76) per day. Nine (30%) were African Americans, 12 (40%) were Caucasian, three (10%) were Hispanic, two were Asian, and one denoted mixed race, and three did not answer the question. Twenty-three participants (77%) used a manual wheelchair. Nineteen (63%) had a spinal cord injury, six (20%) had an amputation, two (7%) had multiple sclerosis, and others included Guillain barre (n=1), traumatic brain injury (n=1), and poliomyelitis (n=1).

4.3.2 Intra-Rater (Sensor) Reliability

The ICCs of the intra rater reliability for the Kinect and RealSense are shown in the Table 12. The Kinect has higher reliability (ICC = .60 - .82) than the RealSense (ICC = .25 - .70) for the four key variables. Both Kinect and RealSense have high reliability for the TF (ICC = .75, 95%CI = .63 - .85; ICC.70, 95%CI = .56 -.82). The DSBW has the largest difference in ICCs between the two sensors (difference = .57).

Table 12. Intra rater reliability (ICC_{3,1}) of the Kinect and the RealSense (n=30)

Variables	Kinect		RealSense		ICC Difference
	ICC	95%CI	ICC	95%CI	
DSBW	.82	.72 - .90	.25	.11 - .44	.57
LPOE	.81	.71 - .89	.60	.44 - .75	.21
LE	.60	.44 - .75	.38	.22 - .57	.22
TF	.75	.63 - .85	.70	.56 - .82	.05

Key: ICC, intra-class correlation coefficient; 95%CI, 95% confidence interval, DSBW, SPINE_BASE/WAIST displacement in horizontal direction; LPOE, average plane of elevation angle on the leading side shoulder; LE, average elevation angle on the leading side shoulder; TF, average trunk flexion angle.

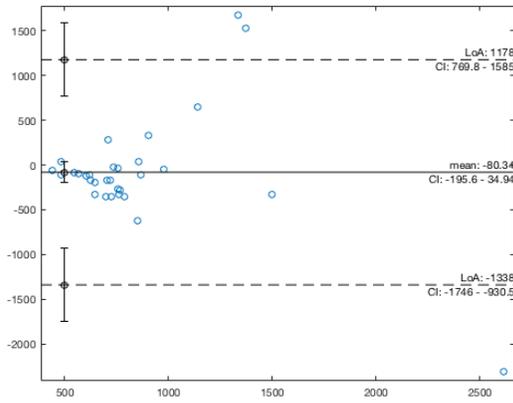
4.3.3 Inter-Rater (Sensor) Reliability and Agreement

The inter-rater reliability between the Kinect and RealSense is shown in the Table 13. The LPOE and TF have moderate reliability (ICC = .52, 95%CI = .01 - .73; ICC = 0.51, 95%CI = .18 - .70). DSBW and LE have low reliability (ICC = .13, 95%CI = .03 - .29; ICC = 0.17, 95%CI = .03 - .35). The Bland-Altman plots for the agreement between sensors are shown in Figure 16.

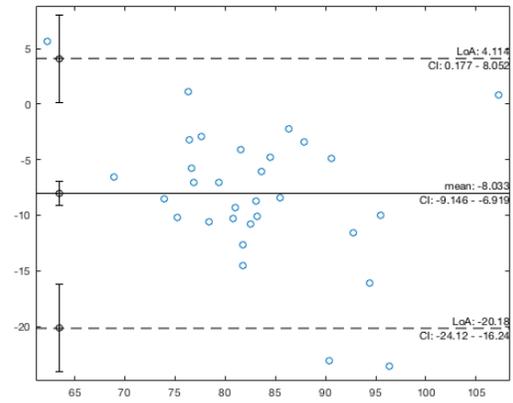
Table 13. Inter rater reliability (ICC_{2,1}) between the Kinect and the RealSense (n=30)

Variables (unit)	ICC	95%CI		Mean	STD
DSBW (mm)	.25	.11 - .55	Kinect	863	542
			RealSense	751	596
LPOE (deg)	.57	.07 - .84	Kinect	86.99	11.51
			RealSense	79.28	9.96
LE (deg)	.13	.10 - .40	Kinect	45.77	8.67
			RealSense	36.15	9.27
TF (deg)	.63	.06 - .85	Kinect	27.79	10.51
			RealSense	21.86	8.22

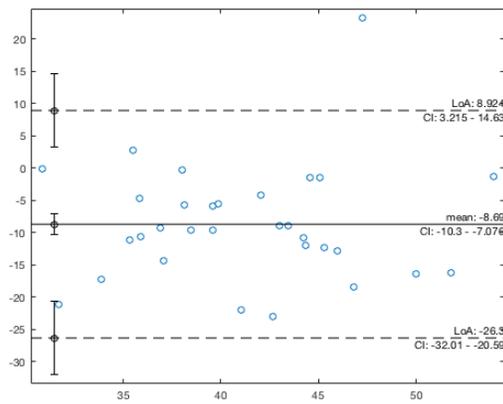
Key: ICC, intra-class correlation coefficient; 95%CI, 95% confidence interval, DSBW, SPINE_BASE/WAIST displacement in horizontal direction; LPOE, average plane of elevation angle on the leading side shoulder; LE, average elevation angle on the leading side shoulder; TF, average trunk flexion angle; STD, standard deviation; deg, degree.



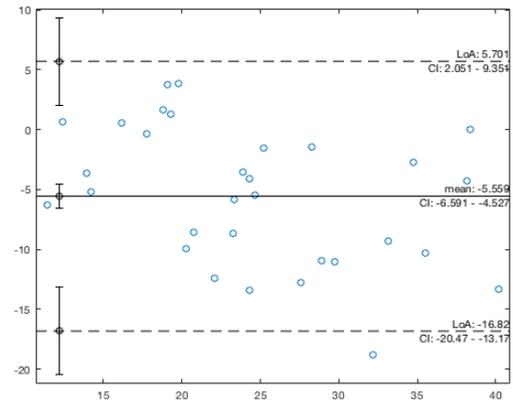
(a) DSBW



(b) LPOE



(c) LE



(d) TF

Figure 16. Bland-Altman plots for the agreement between Kinect and RealSense (n=150 transfers). LPOE, average plane of elevation angle on the leading side shoulder; LE, average elevation angle on the leading side shoulder; TF, average trunk flexion angle; STD, standard deviation; deg, degree

4.4 Discussion

Using proper wheelchair technique to transfer between two sitting surfaces is critical to avoiding secondary injury. Obtaining key biomechanical factors extracted from a marker-less motion capture sensor can provide a tool for rehabilitation clinicians to gauge the injury-risk of body motions. Our research team has demonstrated that the Kinect v2 has the ability to recognize at risk body motions during independent sitting pivot transfers [96, 100, 104]. The Intel RealSense depth sensor could be a solution for Kinect v2's discontinuation because the full-body tracking SDK is available providing similar 3D joint center locations as the Kinect. In this study, we computed and compared some key kinematic variables related to the quality of the wheelchair transfer and reported the intra-rater, inter-rater reliability, and limits of agreement from the Kinect and the RealSense. These methods could similarly be used to evaluate the performance of other 3D depth sensor models as well.

The quality of wheelchair transfer techniques can be evaluated by multiple components related to wheelchair setup, body setup, and flight movements using the latest version of the TAI 4.0 [36]. The kinematic variables related to the upper extremities and trunk motion have high correlation with joint force and moment at shoulder, elbow and wrist during the transfer [4, 113]. Learning to align the upper limb and trunk motions with TAI principles can reduce upper limb joint forces and moments [33]. In our previous studies we applied kinematic variables as the features of machine learning classifiers to predict TAI scores [100]. The DSBW is a feature that is important for identifying if the person uses the correct wheelchair distance between transfer surfaces (TAI 4.0 item 1), scoots forward to the edge of the sitting area before transfer (item 8),

and has good body balance during flight (item 14, 15). The DSBW is also useful for detecting when the lift phase of the transfer process occurs which is required for the application of the TAI prediction classifiers (see **3.2.3.2 Threshold Method of Phase Delineation**, page 51). The LPOE and LE are important for identifying correct technique related to positioning of the arm (items 9, 11, 12). The TF is a key feature for measuring correct trunk leaning motion (item 13). Using this movement pattern can reduce the upper extremities joint loading during transfer [4, 36, 100].

We examined the intra-rater reliability of the Kinect and RealSense by analyzing the five repeated trials performed by the same participant. Some intra trial variability is expected because the individuals could have varied their technique for each transfer. However, because both sensors are watching the same transfers, the variation from the participants would have been controlled for in the analysis. Thus, the intra-rater reliability from the repeated transfers provides an indication of how reliable the sensors are relative to each other in detecting the motions. The intra-rater reliability was assessed by ICC_{3,1} and the 95% CI are also reported as suggested from previous studies [111, 114]. Although there was variance of body movement between transfers, the Kinect's intra-rater reliability was good to excellent for the four key variables (ICC > .6). All the lower boundaries of 95% confidence interval (95% CI) are greater than .63 except for the LE (95% CI = .44 - .75). However, the RealSense's intra-rater reliability was only fair (DSBW and LE, ICC >.2) to good (LPOE and TF, ICC >.6).

In addition to repeatability of the sensors, we were also interested in the agreement and correlations of the variables measured by the Kinect and RealSense. We conducted the inter-rater reliability and the agreement analysis using ICC_{2,1} and Bland-Altman plot. For the agreement analysis, more than 95% of the data points fell within ± 1.96 standard deviation of the mean differences (Figure 16). These results suggest that the values of the variables measured by the two

sensors are close. However, the mean of differences between the two sensors for the four key variables are non-zero and all negative (DSBW, -80.34 mm; LPOE, -8.03 degree; LE, -8.69 degree; TF, -5.56 degree, Figure 16), indicating that RealSense overestimates the displacements and angles relative to the Kinect. The plots also show that three variables related to the joint angles (i.e. LPOE, LE, TF) have similar variability in agreements. The similar LOA's indicated that the STD of the means differences are close.

Harkel et al. using the RealSense SDK demonstrated high reliability for static facial tracking when compared to gold standard measures using the 3dMD imaging system [115]. The SDK was used to label 14 facial landmarks. The average intra- (ICC= .83) and inter-rater reliability (ICC=.80) of the facial landmarks was high for the RealSense. Because the hardware specifications of RealSense (e.g. resolution) are better than that of the Kinect 2 suggests that the errors may be due to either the data acquisition process and/or the body joint tracking algorithms that were used in our current study. However, as noted until recently there was no Intel specific skeletal tracking algorithm and we are not aware of any studies that have reported the reliability of body tracking motion data for the RealSense. Instead we used a 3rd party product's algorithms to obtain the joint centers. In comparison, the Kinect SDK algorithms are based on several years of development and as a result are likely to be more robust for joint center detection than the 3rd party option.

Using multiple RealSense sensors to simultaneously record a motion trial may also increase the reliability. Labuguen and colleagues used three synchronized RealSense sensors to record the freestyle popping motion capture data of a professional dancer [116]. In comparison with a marker-based motion capture system, the results showed that the errors of the body joint movement were less than 150 mm except for the wrists and elbows. Using multiple sensors theoretically increases

the sampling frequency, thus it may help to reduce errors and increase the RealSense' reliabilities for detecting full body motion.

4.5 Limitation

Intel® released the skeleton tracking SDK in 2020 and it supports tracking of 18 joints and multiple subjects simultaneously [117]. The RealSense SDK was not available when we conducted the study. Using the new skeleton tracking SDK may increase the reliability of the RealSense. Due to the NuiTrack™ having a less robust joint center tracking algorithm than the Kinect SDK likely introduced some errors with regards to reliability and agreement between the two sensors. Further study is needed using the new SDK and different 3D depth sensors.

4.6 Conclusion

The Microsoft Kinect can be a useful tool for evaluating the quality of independent wheelchair sitting pivot transfers. However, finding an alternative 3D depth sensor is critical due to the discontinuation of the Kinect v2. The Intel RealSense series has been used in several clinical research applications, however, the body of work has been limited to tracking facial and upper extremity motion. Using a 3rd party SDK designed to work with a variety of 3D depth sensors to support the skeleton tracking function, we found the intra-rater reliability and inter-rater reliability of the RealSense to be fair for tracking the body movements during wheelchair transfers. The

newly released SDK by Intel may help increase the reliability of using RealSense in future applications.

5.0 Conclusion

Transfers are the gateway to independence among individuals with spinal cord injury who rely on wheelchairs for mobility. Lifting and moving the body between surfaces using only the arms however can result in excessive upper limb joint loading, pain and injuries. Using proper transfer technique can help reduce forces and prevent secondary injuries however current assessment tools rely on the ability to subjectively identify harmful movement patterns. Using a low-cost markerless motion capture camera and machine learning (ML) methods can evaluate the quality of independent wheelchair sitting pivot transfers and may serve as a tool that could be used to assist clinicians and patients with identifying deficits in technique.

Our study demonstrates the potential of using the Microsoft Kinect and the developed ML models to facilitate autoscoring of the Transfer Assessment Instrument (TAI). The ML algorithms trained from 91 full-time wheelchair users are able to evaluate proper (low risk) and improper (high risk) wheelchair transfer techniques in accordance with the eleven TAI item scores independently. The transfer data was split into training set (80%) and testing set (20%). The training set was used for classifier selection and model tuning. The test set was excluded from all training processes. Three k-nearest neighbors (KNN) and 8 random forest classifiers were selected for each TAI item. The area under the receiver operating characteristic curves (AUCs) are .83 to .99 for the training set and .79 to .94. for the test set. In order to avoid the false positive case (i.e. participant performed improper technique but the transfer is labelled as a proper transfer by the classifier), we tuned the models to achieve high precision. The precisions of the models are .87 to .96, and the recalls are .61 to .93. The future plan aims to embed the developed algorithms into

a user-friendly graphic user interface that allows therapists to perform transfer assessments more easily and produce results that support training and education on proper transfer technique.

The TAI item scores are evaluated by the relative components of the wheelchair transfer. For a system to automate the scoring of the TAI the system it must also be able to distinguish the “setup phase” and “lift phase” of the transfer. On the TAI 4.0, items 1 to 6 are in the wheelchair setup skill group and Items 7 to 15 are in the body setup and flight/landing skill groups. In order to extract the features of each item, the motion data during the transfer needs to be separated into a setup phase and lift phase. We defined the “setup phase” as the period when the subject starts moving their body for transfer preparation, to a time point right before the subject lifts their buttocks from the sitting surface. The “lift phase” was defined as starting from the end of setup phase to the first time frame when the subject lands on the target surface. We applied and compared a biomechanical variable based threshold method and an ML algorithm to automatically distinguish the time frames of the transfer phases. For the threshold (TH) method, the peaks observed in the linear displacement and velocity of one joint center marked by the Kinect, SPINE_BASE, were used for phase delineation. For the ML method, we trained a KNN classifier using 35 features from the Kinect data. Using the KNN model, each time frame of the transfer was labeled as belonging to either the “setup” or “lift” phase. After further applying a filter algorithm, the method was used to identify the start and end timepoints of the transfer phases. We found that the ML method had less error in identifying the phase times but the threshold method spends less computational time in identifying the points. Although the threshold errors were larger this method had higher accuracy for predicting the TAI scores for items 10, 11, 12, 13, 14, 15. The ML method had higher accuracy for predicting the TAI scores for items 1, 2 and 7. For items 8 and 9, the two methods showed equal performance. Because the ML method tended to undershoot the end phase

times it's possible that tuning the algorithms to include more of the lift phase data could increase the accuracies of the TAI item scores that rely more heavily on the lift versus setup phase biomechanics.

Due to the discontinuation of the Kinect v2 in 2017, we aimed to find another 3D depth sensor that could track full body motion for future research. The Intel[®] RealSense could be an ideal surrogate sensor because the RealSense has superior technical properties relative to the Kinect v2 and has shown excellent performance for tracking facial and hand motions in previous studies. Although Intel did not make a full body joint tracking algorithm for the sensor at the time of our study, a 3rd party one was available that can be used with a variety of 3D sensor models (e.g. NuiTrack[™]). This solution enabled us to create the same biomechanical features with the RealSense as we had created with the Kinect to quantify transfer technique. To further understand the potential for RealSense to be used as a substitute sensor for capturing wheelchair technique biomechanics, we compared the measurement properties of the two sensors. We assessed intra-rater reliability for each sensor, and evaluated the inter-rater reliability and agreements between the Kinect and RealSense. The study found that the Kinect had higher intra-rater reliability than the RealSense for measuring four key kinematic variables related to the wheelchair transfer technique. For the agreement analysis, more than 95% of the data points fell within ± 1.96 standard deviation of the mean differences. However, the inter-rater reliability between two sensors was poor. The low reliability of the RealSense may be due to the lack of robustness of the 3rd party algorithm for skeletal tracking of sitting postures and in general in comparison to the more extensively tested and developed Kinect SDK. Using the current (2020) version of the RealSense SDK for skeletal tracking may help increase the reliability for future applications.

5.1 Future Works

To improve the generalizability of the TAI classifiers, refining the data splitting method during the training process is necessary. In our current protocol, each individual did five transfer trials with each surface. The transfer data was randomly split into the training and test set by trial. Thus, the five transfers performed by the same wheelchair user might be split into both training and test set (e.g. three to the training set and two to the test set). Due to the similarity of the motion pattern from the same individual, the training set might “see” the data from the test set and decrease the model generalizability [97]. Splitting the data by the individuals is a better strategy to assess the generalizability. All the trials from the same subject should be in the same set. Thus, the data in the test set would be “new data” to examine the model performance.

Future work may also include collecting more transfer trials to refine the ML of the TAI scoring for each item and the phase time distinguishing algorithms. A dataset that includes not only a sufficient sample size but also various types of wheelchair users and transfer motions can increase the performance of predicting TAI outcomes. Currently, all partial credits (0.5) are labeled as improper techniques for the TAI-classifiers. For the items allowed to be scored “0.5”, a new model can be trained by labeling three outcomes (i.e. 0, 0.5, 1). This approach needs sufficient data for each label. For some items the deficit rates among wheelchair users might be very low [36, 118]. As a result, it may be difficult to collect enough improper and partial credits cases from the wheelchair user population. One possible solution is to use “pseudo-patients” (e.g. able-bodied participants who are trained to perform the specific partial credit transfer techniques) to enlarge the sample size. Moreover, the ML practitioner can also apply over-sampling or down-sampling strategies to overcome the limitation with imbalanced labeling datasets.

Not all TAI 4.0 items were evaluated using the TAI-classifiers. Items that are related to the participant's transfer habits (e.g. item 16 which asks if they always lead with the same arm or if they alternate arms when they transfer) and assistive technology use (items 17 and 18) were not included. Other items such as locking the wheelchair or turning it off before transfer (item 3), removing the armrest (item 4) and clothing/sides guards (item 5) before transfer were excluded. Transferring to a level surface if it is possible (item 6) was also excluded due to the protocol design. The items 3, 4, 5 and 6 have the highest inter-rater reliability (ICC=.94) [36] and therefore are more easily observed by raters than some of the items that involve interpreting the body motions. In order to evaluate these items, the motion pattern and related kinematic variables (i.e. specific joint angles, displacements) would need to show differences between the proper and improper techniques. However, some differences may not occur during the setup or lift phases defined in this study. For example, the participant may lock the brakes of the wheelchair before the investigators start observing or recording a transfer trial using the Kinect. Motion data before the wheelchair setup (i.e. pre-setup phase) may be important to collect in the future and investigated to identify potential features that can be used to predict these items.

The TransKinect project aims to apply the current study outcomes to the clinical setting. Packaging these outcomes together into a user-friendly system could aid therapists and patients in identifying harmful motions and learning proper evidence-based transfer practices. The graphical user interface (GUI) of the TransKinect is designed to highlight the items which are easier and those that are more difficult to evaluate by the therapists. For the future plan, firstly we aim to collect feedback from clinicians on usability of the system. We also aim to understand how much error (false positive cases) of the ML model can be accepted by clinicians. Secondly, the system will be refined following the feedback and tested in a lab-controlled environment to further

understand the system's usability and validity. Clinician raters will use both the TransKinect and TAI to evaluate the wheelchair users' transfers. The ML models may be further refined using the new collected data. Thirdly and lastly, we aim to conduct a field trial with the TransKinect in multiple clinical settings to explore its utility, practicality and effectiveness as a clinical tool. Using a simple 3D Depth Camera to watch a wheelchair transfer and produce a score in real-time could provide a valuable evaluation and education tool for evaluating and training proper technique. Patients in the clinic who are assessed with TransKinect will receive their results and education on proper technique. Afterwards the quality of their transfers will be reevaluated by follow-up TransKinect assessments. A feature to compare side by side the pre and post training assessment reports has been built into the TransKinect application per feedback from initial expert therapist review of the application to facilitate comparisons in client performance between assessments.

It may be possible to extend the current work to evaluate assisted or dependent transfers. An evaluation tool similar to the TAI has been developed for dependent transfers which evaluates proper and improper caregiver techniques (e.g. Caregiver Assisted Transfer Technique Instrument). However, the tool is still undergoing validity and reliability which is required before it can be used for data labelling. In lieu of the tool it may be possible to label the data based on a biomechanical analysis of the caregiver techniques (e.g. identifying techniques that cause loading at the lumbar spine to exceed certain levels or the lifting index set by the National Institute for Occupational Safety and Health (NIOSH)) [119]. Assuming a solid gold standard could be identified then pilot studies would be needed to determine whether the 3D depth motion sensor(s) can detect differences in caregiver assisted transfer kinematic variables between the proper and improper techniques. If so then the same methods we employed for item model training and time-series phase delineation could be used to develop and test ML classifiers for predicting the quality of

dependent transfer techniques. Refined algorithms could further be implemented in a similar platform as TransKinect for evaluating caregiver techniques in the future.

Appendix A. General Questionnaire for Wheelchair Users

DEMOGRAPHICS AND TRANSFER ACTIVITIES:

Age: _____

Type of disability _____

If SCI, at what level (please indicate)? _____ Incomplete / Complete (Circle

One) Approximate month/day/year of onset or diagnosis _____

Gender _____ Race _____

Handedness _____

How long have you used a wheelchair? _____ / _____ months/yr

How many hours a day do you spend in your wheelchair? _____ hours

Make of wheelchair (ie, Quickie, TiLite, Invacare): _____

Model of wheelchair (i.e., Q7, Aero, etc): _____

Arm rests ___ YES ___ NO Are they removable? ___ YES ___

_____ NO Foot rest type _____ Fixed_

_____ Removable

How many level transfers do you do each day on average? _____

How many non-level transfers do you do each day on average? _____

Do you use an assistive device (e.g., transfer board or lift) with any of your daily transfers? YES / NO, if YES, please list: _____

Have you ever received formal training on how to transfer? ___ YES ___ NO

If YES, please list: _____

approximately how many hours of training? _____ Hours

Additional Comments:

Appendix B. Transfer Assessment Instrument 4.0

Transfer Assessment Instrument 4.0 – Independent Transfers

This tool is designed to objectively unassisted transfers. The tool breaks down the transfer into three components: wheelchair setup, body set up, and flight/landing. Scoring differs based on whether the user is transferring from a manual or power wheelchair. Each section of the tool should be completed before advancing to the next phase. It is written in user-centered language but can be utilized by end users, their caregivers, and clinicians. Space is allotted to score a transfer to and from a surface (2 total transfers), however only one transfer is necessary to use the tool.

Select Wheelchair Type: Power wheelchair Manual wheelchair

Do you use a sliding board when transferring? Yes No

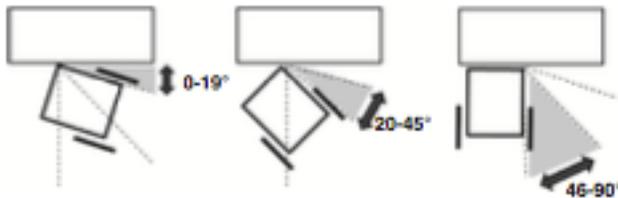
Phase I: Wheelchair Setup

Set up your wheelchair to transfer to another surface (bed, mat table, couch), then answer the following questions.

1. Using a ruler (if available), measure the distance from the front corner of your wheelchair to the object to which you are transferring. What is the distance?
 - a. Less than 3 inches [score 1]
 - b. 3-5 inches [score 0.5]
 - c. Greater than 5 inches [score 0]



2. What is the angle between your wheelchair and the mat? Use an angle measurement tool if possible (see appendix).
 - a. 0-19 degrees [score 0 mwc, score 1 pwc]
 - b. 20-45 degrees [score 1 mwc, score 0 pwc]
 - c. 46-90 degrees [score 0]



3. Did you lock the brakes on your wheelchair?
 - a. Yes, I engaged the brakes (manual wheelchair) [score 1]
 - b. Yes, I turned my wheelchair off (power wheelchair) [score 1]
 - c. No [score 0]
 - d. Not applicable, my wheelchair does not have brakes [score N/A]

Scoring	
Transfer 1 Transfer 2	
Transfer 1 Transfer 2	
Transfer 1 Transfer 2	

This supplementary digital content accompanies
Top Spinal Cord Inj Rehabil 2018;24(3):217-226
www.acijournal.com

doi: 10.1310/sci2403-217

4. Did you remove the armrest from your chair?
- a. Yes [score 1]
 - b. No, but my wheelchair does have armrests [score 0]
 - c. Not applicable, my wheelchair does not have armrests [score N/A]
 - d. Not possible, my wheelchair has armrests but they cannot be removed (bolted in, welded) [score N/A]
5. Did you remove the clothing/sides guards or postural supports (thigh guides, lateral supports) from your chair?
- a. Yes [score 1]
 - b. No [score 0]
 - c. No, however my clothing/sideguards don't go any higher than my wheel [score N/A]
 - d. Not applicable, my wheelchair does not have clothing/side guards [score N/A]
 - e. Not possible, my wheelchair has clothing/side guards but they cannot be removed (bolted in, welded) [score N/A]
6. Was your transfer set up to be level (top of cushion is level with the surface you are transferring to)? Use a ruler (if possible) to measure the difference in height between the top of the front corner your cushion and the surface you are transferring to. What is the distance?
- a. My cushion height is within 1 inch of the surface [score 1]
 - b. My cushion is more than 1 inch higher [score 0]
 - c. My cushion is more than 1 inch lower [score 0.5]
 - d. Not possible to adjust height of wheelchair/transfer destination [score N/A]



Scoring	
Transfer 1	Transfer 2
<input type="text"/>	<input type="text"/>
Transfer 1	Transfer 2
<input type="text"/>	<input type="text"/>
Transfer 1	Transfer 2
<input type="text"/>	<input type="text"/>

Phase II: Body Setup

Position your body for the transfer, adjusting your hips and legs as you normally would, then answer the following questions

7. Where are your feet?
- a. Both on footplate [score 0 mwc, 1 pwc]
 - b. One on footplate, one on the ground [score 0.5]
 - c. Both up on the surface I am transferring to [score 1]
 - d. One up on the surface I am transferring to and one on the ground/footplate [score 0.5]
 - e. Both on the floor [score 1]
 - f. One on floor (I am a single limb amputee) [score 1]
 - g. One on footplate (I am a single limb amputee) [score 0 mwc, 1 pwc]
 - h. Not touching any surface (I am a double limb amputee) [score N/A]
 - i. Not touching any surface (I am a double limb amputee) [score N/A]
8. Did you scoot to your hips to the front of your seat, so at least 1/3 of your thigh was off the surface?
- a. Yes [score 1]
 - b. No [score 0]
 - c. Not possible, I am unable to maintain my balance in this position [score N/A]
9. Where is your leading arm (see images) once you position it to transfer?
- a. On the surface I am transferring to behind my hip [score 0]
 - b. On the surface I am transferring to between my hip and knee [score 1]
 - c. On the surface I am transferring to past my knee [score 1]
 - d. It is in my lap, I don't use this arm during my transfer [score N/A]

Scoring	
Transfer 1	Transfer 2
<input type="text"/>	<input type="text"/>
Transfer 1	Transfer 2
<input type="text"/>	<input type="text"/>
Transfer 1	Transfer 2
<input type="text"/>	<input type="text"/>



Phase III: Flight

Complete your transfer between surfaces, then answer the questions on the following pages

10. Identify the type of surface you transferred **FROM**. Then, check the box of the type of hand position that most closely represents the hand position you used for your **push off hand** when **your hips were moving between surfaces**. **CHECK ALL THAT APPLY.**

Wheelchair	Armrest	Wheel	Cushion edge/frame	Fist on cushion	Flat on cushion	Bent fingers on cushion
	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
score:	1	1	1	0	0	0.5
Firm Surface	Fist on surface	Surface edge	Flat on surface	Fingers bent on surface		
	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>		
score:	0	1	0	0.5		
Soft Surface (Bed/Couch)	Edge of soft surface	Flat on soft surface	Fist on soft surface	Bent fingers soft surface		
	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>		
score:	1	0	0	0.5		
Chair	Chair Arm	Edge of Chair	Flat on Chair	Fist on Chair	Bent fingers on chair	
	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
score:	1	1	0	0	0.5	
Bathroom	Edge of Toilet Seat	Horizontal Grab Bar	Vertical Grab Bar			
	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>			
	1	1 if <6" from BOS	1 if <6" from BOS			
Not Used	Hand on leg					
	<input type="checkbox"/>					
	N/A					

Scoring

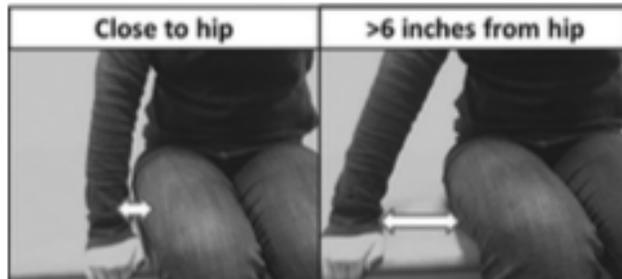
Transfer 1	Transfer 2
A) Total Score: _____	A) Total Score: _____
B) Total # Positions Used: _____	B) Total # Positions Used: _____
Score = A/B: _____	Score = A/B: _____

11. Identify the type of surface you transferred **TO**. Then, check the box of the type of hand position that most closely represents the hand position you used for your **leading hand** when **your hips were moving between surfaces**. **CHECK ALL THAT APPLY.**

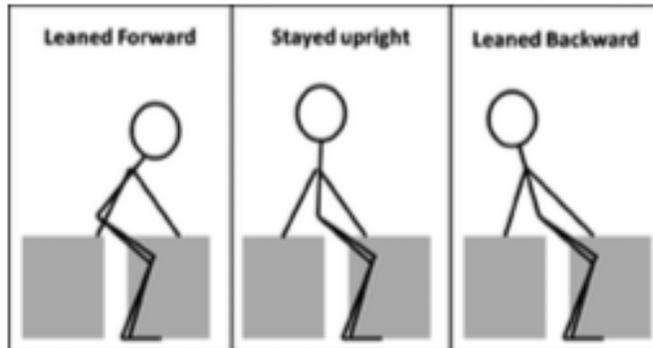
Wheelchair	Armrest	Wheel	Cushion edge/frame	Flat on cushion	Flat on cushion	Rest fingers on cushion		
score:	1	1	1	0	0	0.5		
Sliding Board	Edge sliding board	Fingers curled under slide board	Flat on sliding board	Flat on sliding board	Rest fingers on board			
score:	1	0	0	0	0.5			
Firm Surface	Flat on surface	Surface edge	Flat on surface	Fingers bent on surface	Bathroom	Edge of Toilet Seat	Horizontal Grab Bar	Vertical Grab Bar
score:	0	1	0	0.5	1	1 if <6" from BOS	1 if <6" from BOS	
Soft Surface (Bed/Couch)	Edge of soft surface	Flat on soft surface	Flat on soft surface	Rest fingers soft surface	Not Used	Hand on lap		
score:	1	0	0	0.5	N/A			
Chair	Chair Arm	Edge of Chair	Flat on Chair	Flat on Chair	Rest fingers on chair			
score:	1	1	0	0	0.5			

Scoring	
Transfer 1	Transfer 2
A) Total Score: _____	A) Total Score: _____
B) Total # Positions Used: _____	B) Total # Positions Used: _____
Score = A/B: _____	Score = A/B: _____

12. When you **finished** the transfer, where was your **leading hand** (see image) in relation to your hips?
- Close to my hip [score 1]
 - More than 6 inches from my hip [score 0]



13. When you transferred between surfaces which way were you leaning?
- I was leaning backward, towards the surface I was transferring to [score 0]
 - My body remained upright [score 0]
 - I was leaning forward, away from the surface I was transferring to [score 1]
 - I'm not sure [score 0]



14. When you transferred between surfaces did you (please circle all that apply):
- Perform the movement in one smooth and fluid motion [score 1]
 - Use multiple 'scoots' to complete my transfer [score 1]
 - Land or rest on the tire [score -1 mwc, N/A pwc]
 - Perform the movement in an abrupt manner where you had to change directions or body positions in rapid manner to avoid falling (unintentionally landing on an undesired surface) [score -0.5]
 - Experienced a near fall (unintentionally landing on an undesired surface) [score -1]
 - Experience a fall (unintentionally landing on an undesired surface) [score -1]

Scoring

Transfer 1	Transfer 2

Transfer 1	Transfer 2

Transfer 1	Transfer 2
----- (lowest possible score: 0)	----- (lowest possible score: 0)

Transfer 1	Transfer 2
----- (lowest possible score: 0)	----- (lowest possible score: 0)

15. When you landed on the target surface (please circle all that apply):
- a. No excessive movement occurred (no loss of balance) [score 1]
 - b. Experienced excessive movement (loss of balance) but did not unintentionally land on an undesired surface (experience a fall) [score -0.5]
 - c. Experienced excessive movement (loss of balance) and some part of your body unintentionally landed on an undesired surface (experience a fall) [score -1]
16. When you transfer do you:
- a. Always lead with the same arm [score 0]
 - b. Alternate which arm you lead with [score 1]

Assistive Technology

17. I use assistive technology (such as a transfer board or lift) :
- a. Never, I am strong enough to perform the transfer without a struggle [score 1]
 - b. Never, however I sometimes feel that I struggle to complete the transfer [score 0]
 - c. Sometimes when I feel tired or weak [score 1]
 - d. Sometimes, due to pain [score 1]
 - e. Sometimes for safety or to preserve my arms [score 1]
 - f. All of the time because of fatigue or strength limitations [score 1]
18. If a sliding is used, when using the board do you:
- a. Perform the transfer as multiple 'lifts'/'scoots', picking up your hips and placing them over several steps [score 1]
 - b. Slide your hips along the board [score 0]

Scoring	
Transfer 1	Transfer 2
[]	[]
Transfer 1	Transfer 2
[]	[]
Transfer 1	Transfer 2
[]	[]

To calculate total score for Independent Transfers:

1. Sum all scores for Transfer 1 items (Box A)
2. Count # N/A items for Transfer 1 (Box B)
3. Calculate a score for Transfer 1 (Box C):
 - a. $100\% \times A / (19-B)$

Ignore remaining steps if only 1 transfer completed

4. Sum all scores for Transfer 2 items (Box D)
5. Count # N/A items for Transfer 2 (Box E)
6. Calculate a score for Transfer 2 (Box F):
 - a. $100\% \times D / (19-E)$
7. Calculate the total score:
 - a. $C + F / 2$

Transfer 1 Total Score	Transfer 1 # items N/A	Transfer 1 TOTAL
A	B	C
Transfer 2 Total Score	Transfer 2 # items N/A	Transfer 2 TOTAL
D	E	F
TOTAL		
[] %		

Appendix C. Model training process and results of each TAI 4.0 item

Item 1 Wheelchair Distance

The KNN classifier was chosen for the item 1 model (Figure A1a). Four hundred and seventy-two trials were split into the training set and 119 trials were into the test set. The transfer technique deficit rate was 30%. Eighty-five features were selected into the feature engineering process and 26 of them were applied into the final model (Figure A1b). For the model from the training set, the AUC was .98, and the mean accuracy from CV was 75% (STD = 3%). After tuning the model decision threshold, the test set AUC was .85, and precision was .96. Figure A1c shows the ROC curve, and the relationship between precision, recall when adjusting decision threshold of the outcomes from the test set.

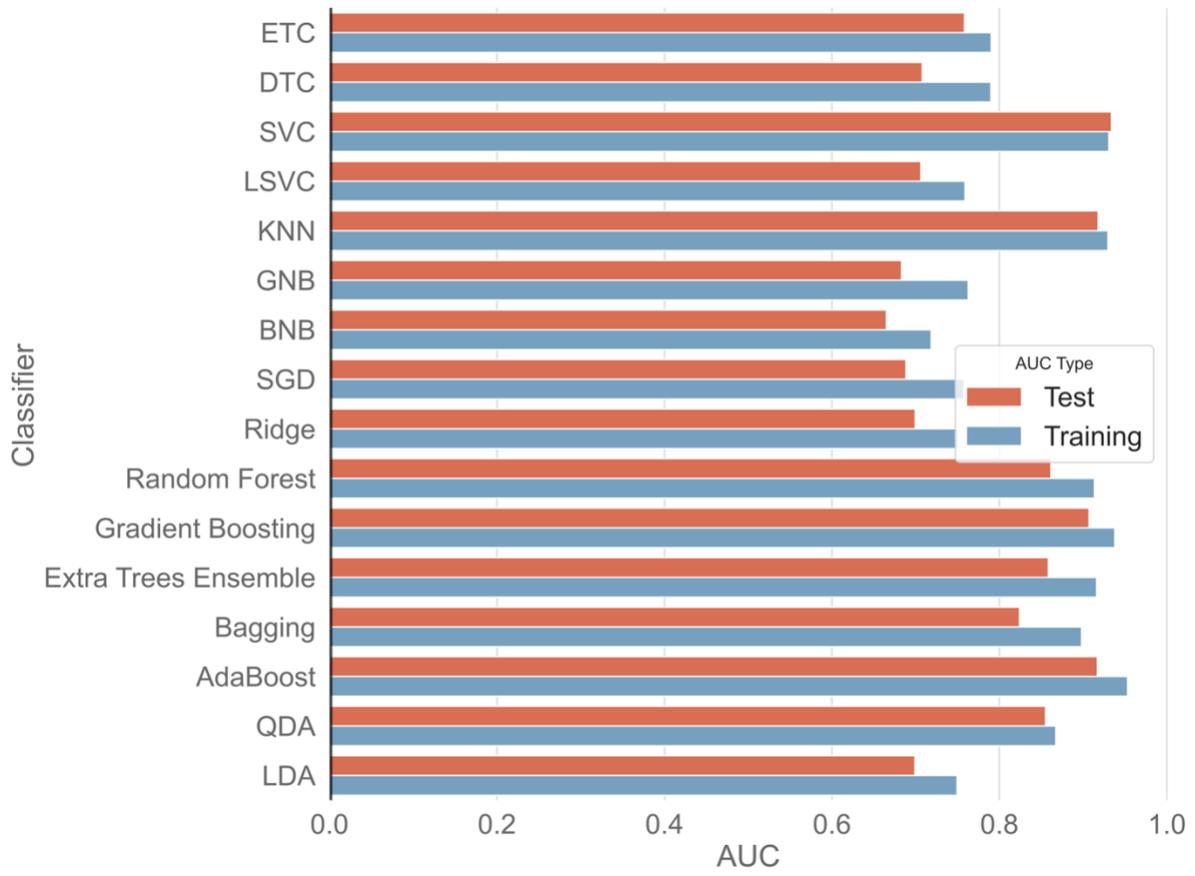


Figure A.1a. Model performance of each classifier for item 1

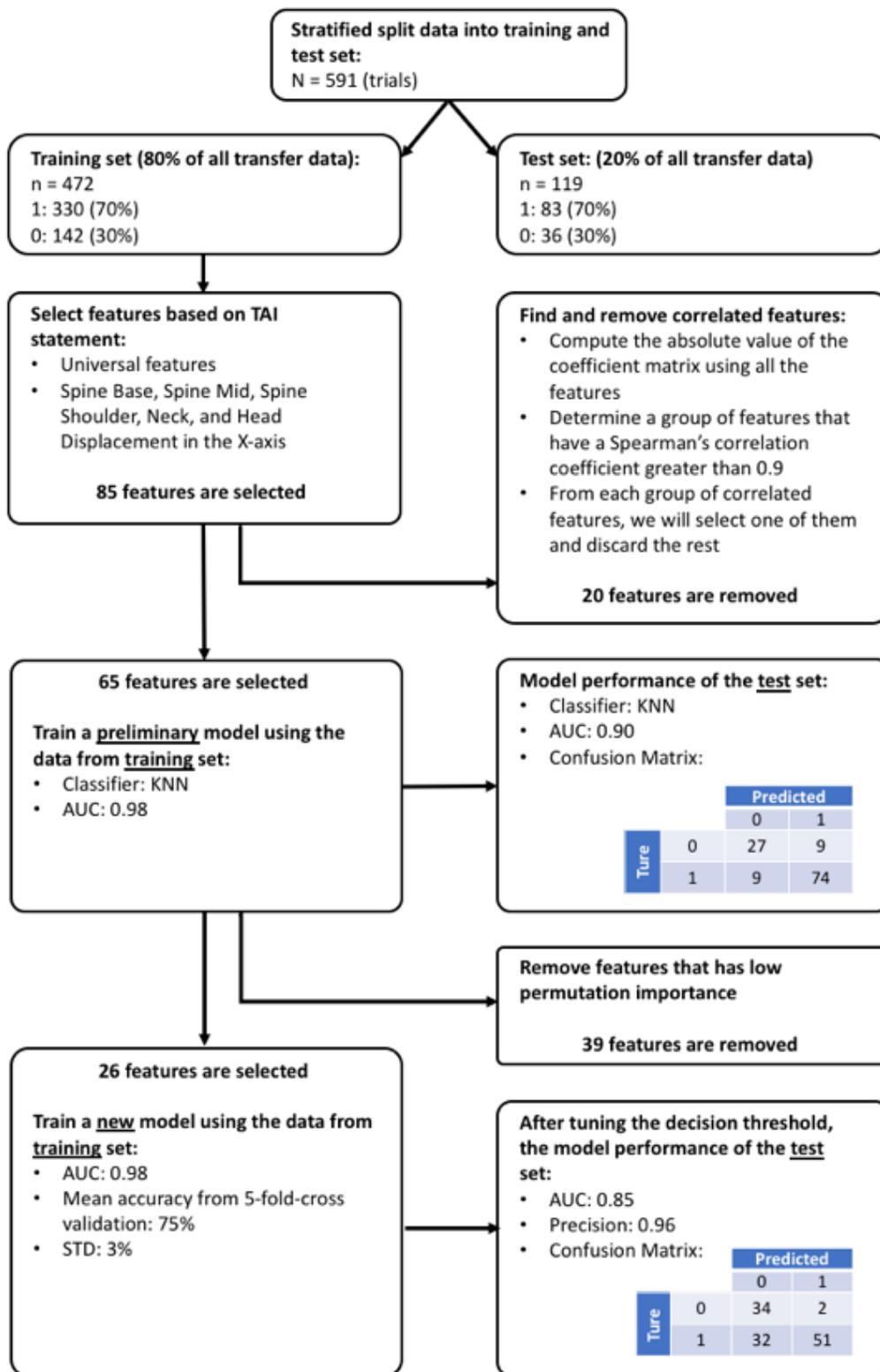


Figure A1b. Feature selection and training process of item 1 model

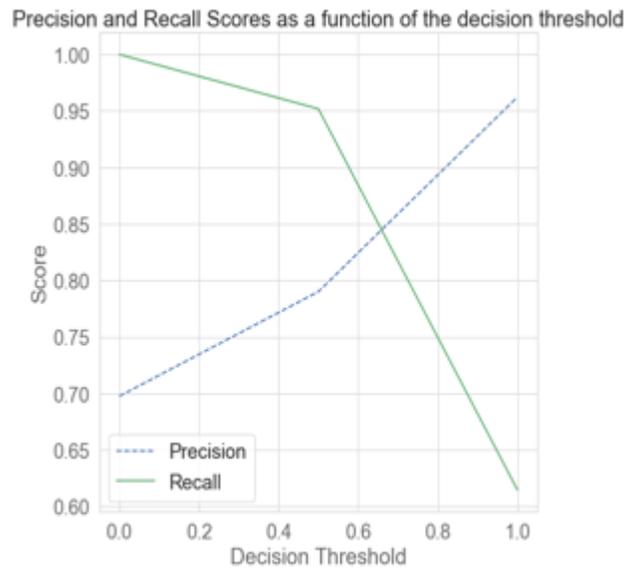
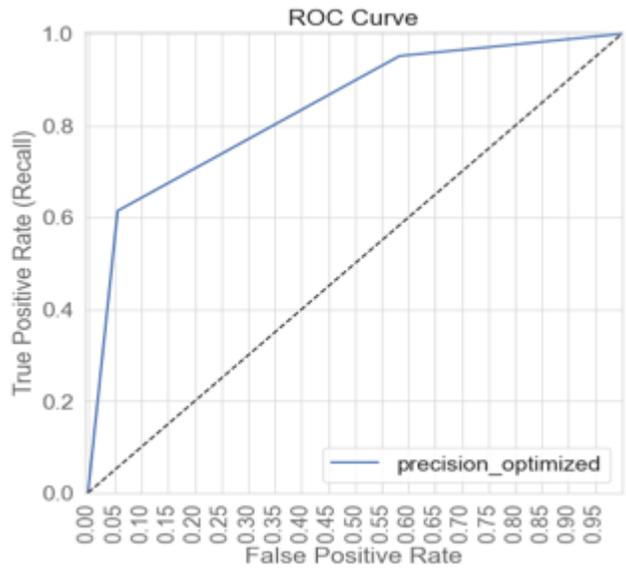


Figure A1c. ROC curve, and relationship between precision, recall and decision threshold values of the item 1 model

Item 2 Wheelchair Angle

The random forest classifier was chosen for the item 2 model (Figure A2a). Four hundred and sixty-eight trials were split into the training set and 118 trials were into the test set. The transfer technique deficit rate was 36%. Eighty-four features were selected into the feature engineering process and 14 of them were applied into the final model (Figure A2b). For the model from the training set, the AUC was .99, and the mean accuracy from CV was 87% (STD = 4%). After tuning the model decision threshold, the test set AUC was .90, and precision was .91. Figure A2c shows the ROC curve, and the relationship between precision, recall when adjusting decision threshold of the outcomes from the test set.

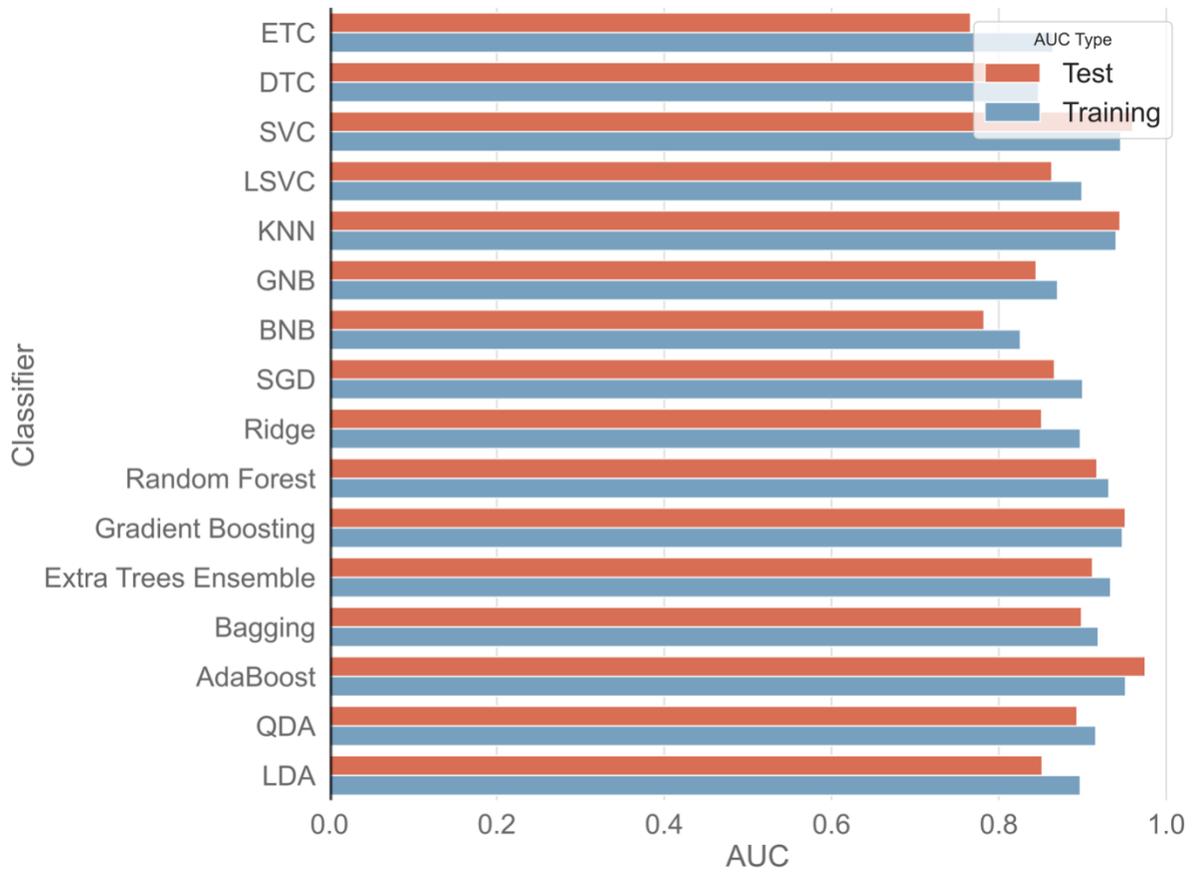


Figure A2a. Model perforce of each classifier for item 2

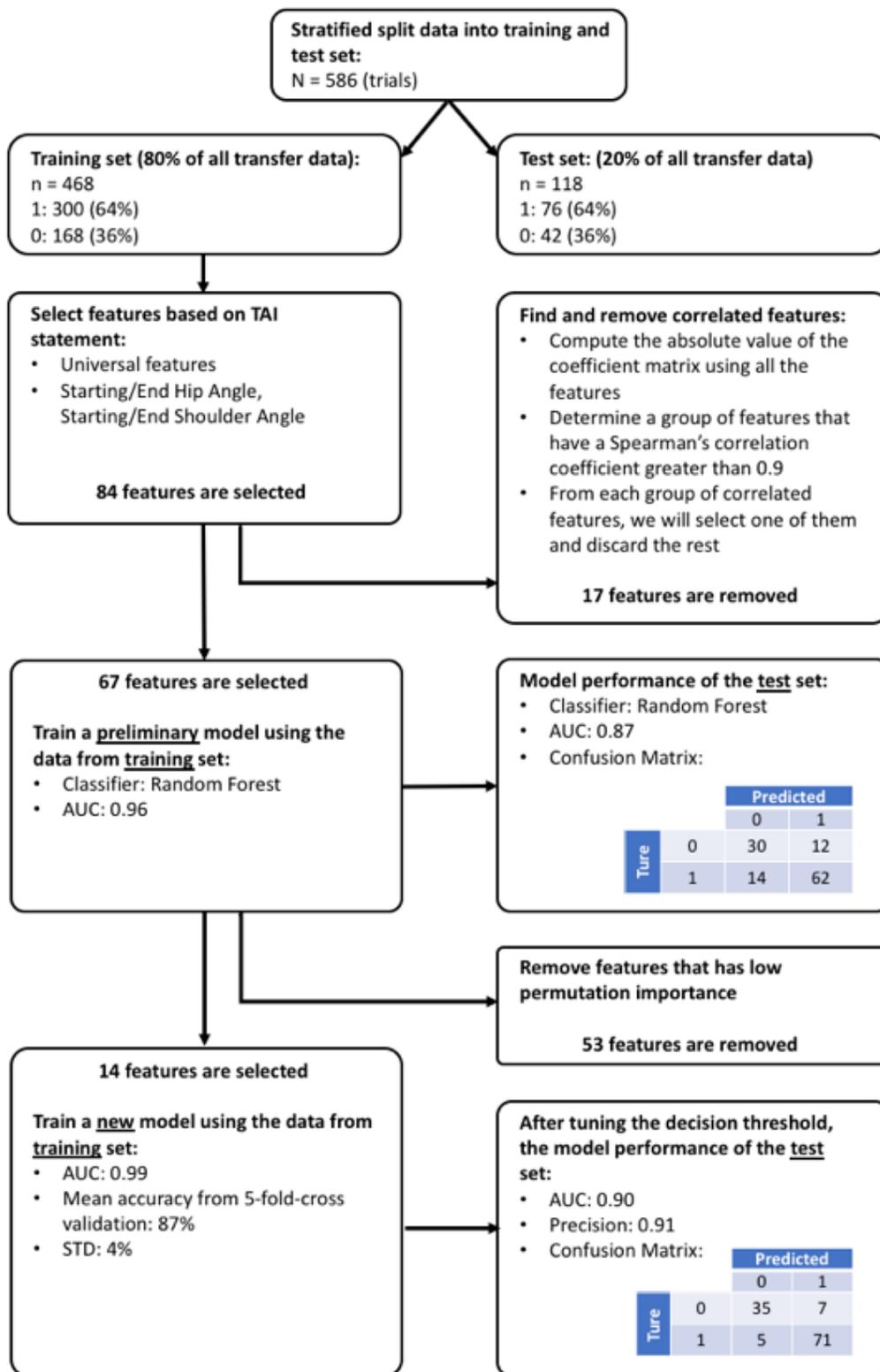


Figure A2b. Feature selection and training process of item 2 model

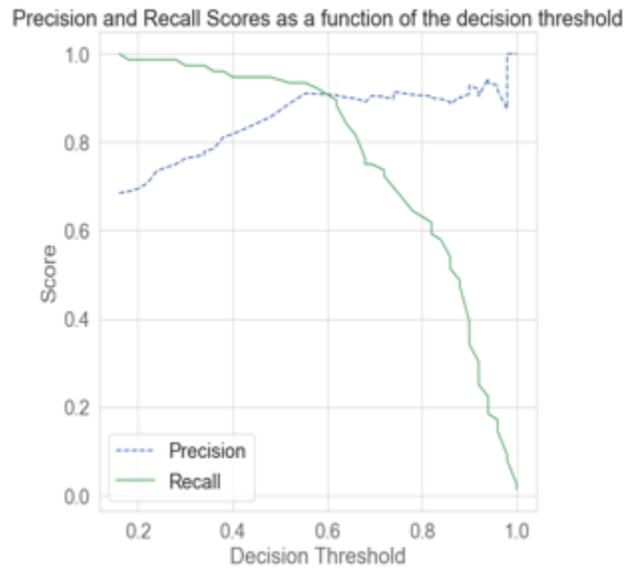
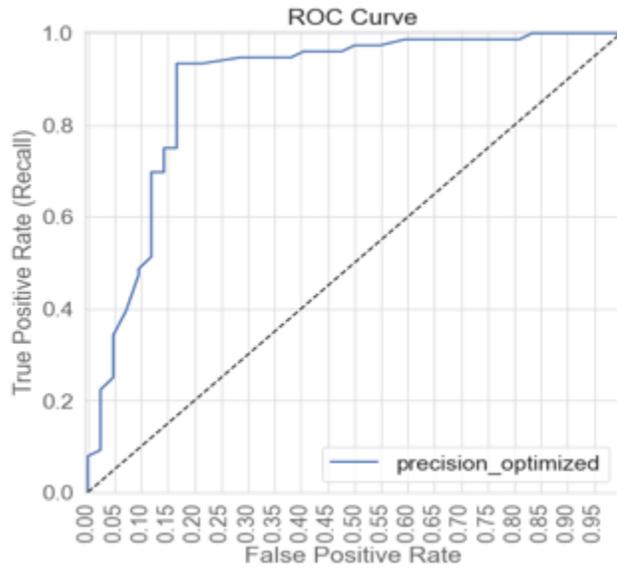


Figure A2c. ROC curve, and relationship between precision, recall and decision threshold values of the item 2 model

Item7 Feet Placement

The random forest classifier was chosen for the item 7 model (Figure A3a). Four hundred and sixty-four trials were split into the training set and 117 trials were into the test set. The transfer technique deficit rate was 39%. Eighty-nine features were selected into the feature engineering process and 29 of them were applied into the final model (Figure A3b). For the model from the training set, the AUC was .99, and the mean accuracy from CV was 78% (STD = 3%). After tuning the model decision threshold, the test set AUC was .85, and precision was .87. Figure A3c shows the ROC curve, and the relationship between precision, recall when adjusting decision threshold of the outcomes from the test set.

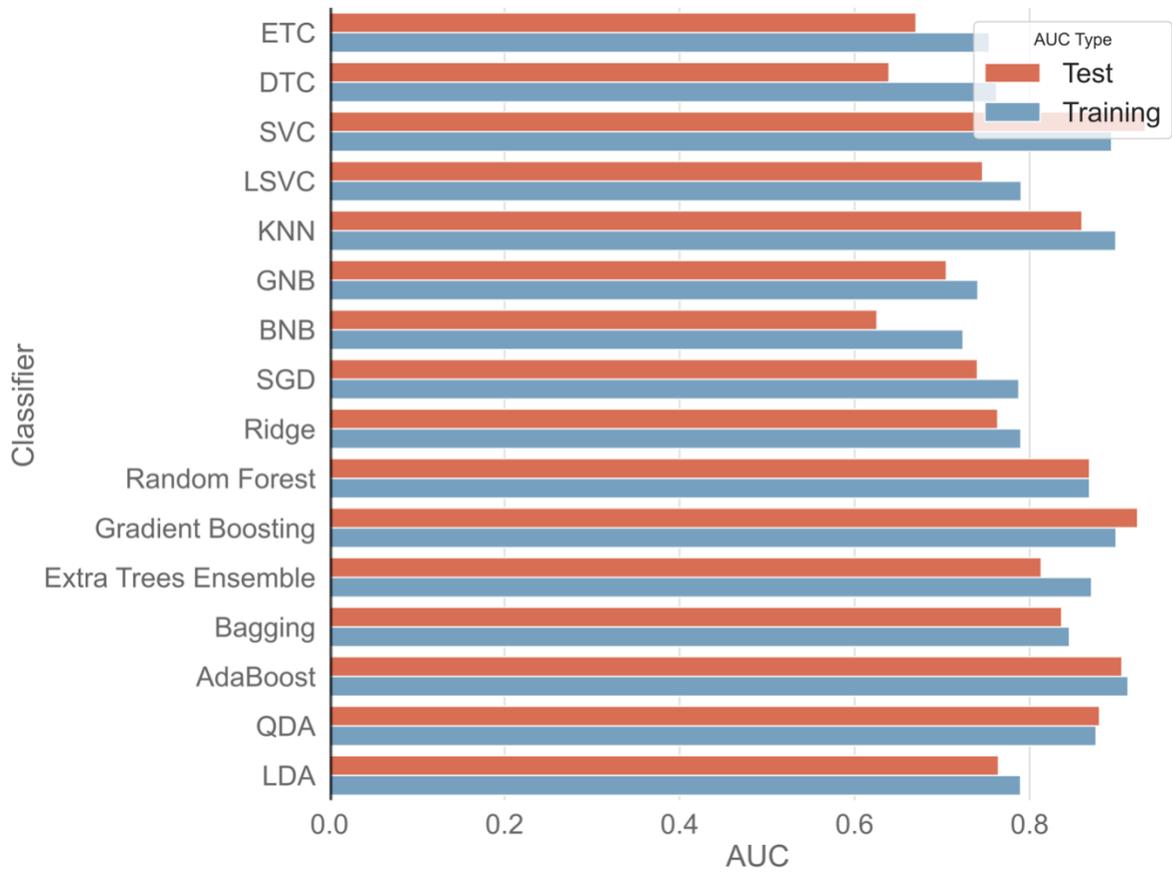


Figure A3a. Model perforce of each classifier for item 7

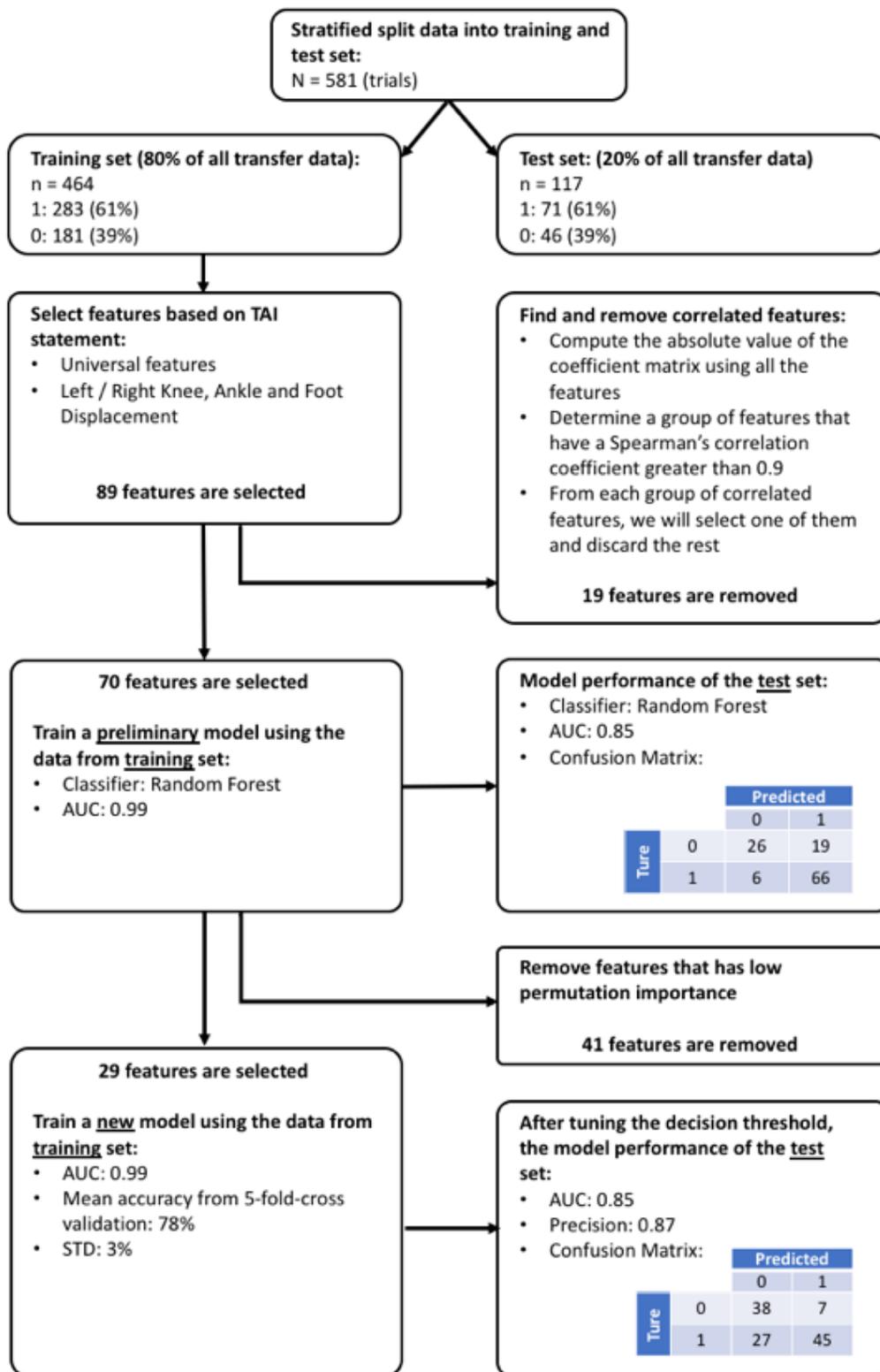


Figure A3b. Feature selection and training process of item 7 model

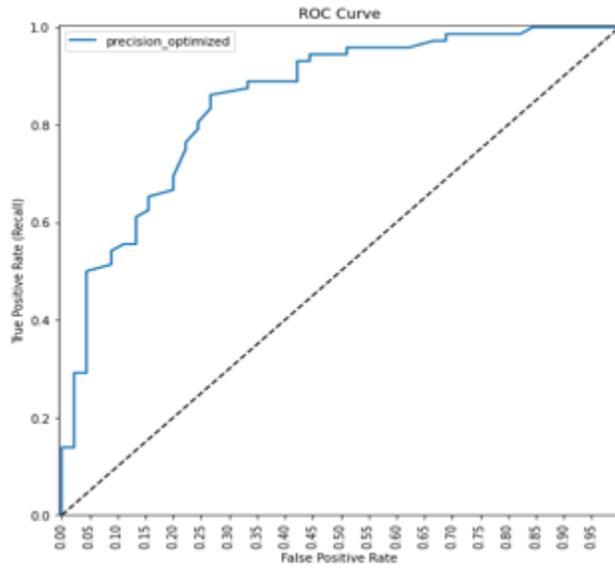


Figure A3c. ROC curve, and relationship between precision, recall and decision threshold values of the item 7 model

Item 8 Hip Scooting

The random forest classifier was chosen for the item 8 model (Figure A4a). Four hundred and sixty-two trials were split into the training set and 119 trials were into the test set. The transfer technique deficit rate was 25%. Eighty-six features were selected into the feature engineering process and 27 of them were applied into the final model (Figure A4b). For the model from the training set, the AUC was .96, and the mean accuracy from CV was 80% (STD = 3%). After tuning the model decision threshold, the test set AUC was .87, and precision was .90. Figure A4c shows the ROC curve, and the relationship between precision, recall when adjusting decision threshold of the outcomes from the test set.

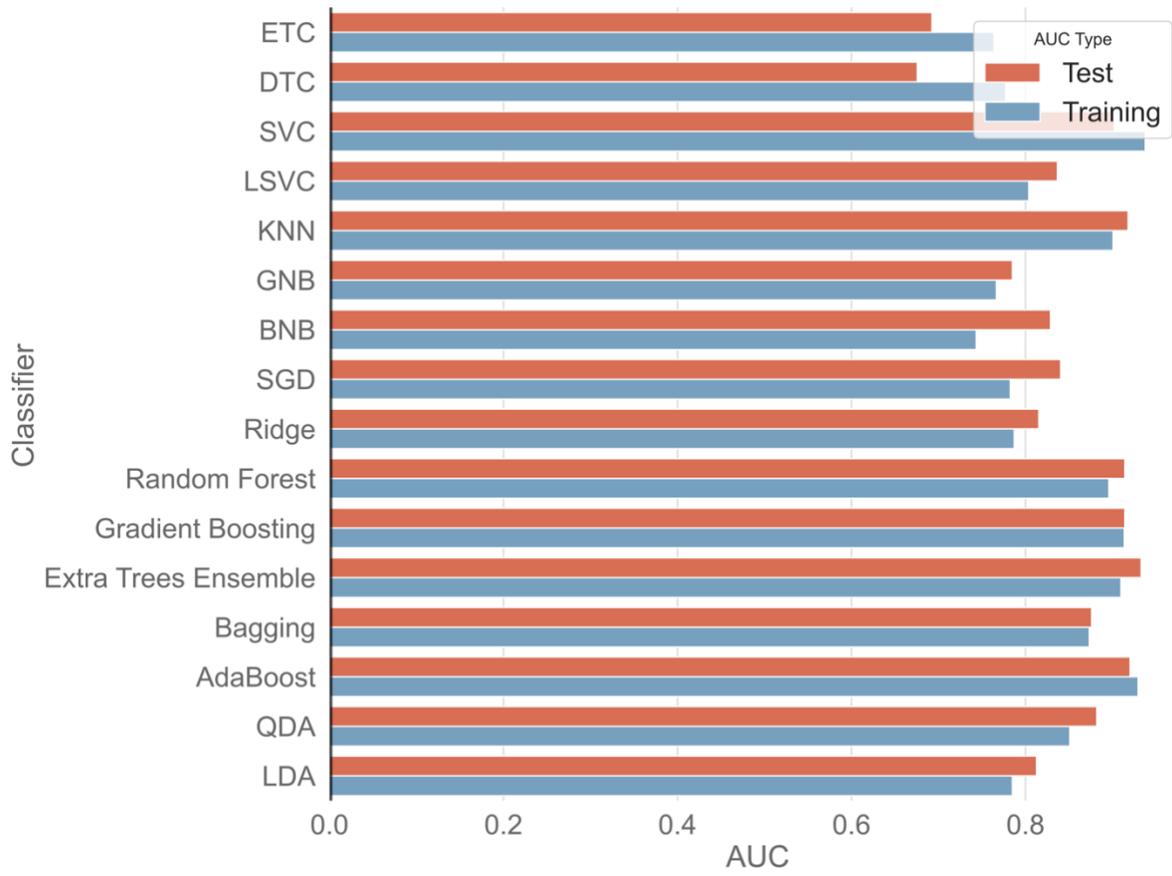


Figure A4a. Model performance of each classifier for item 8

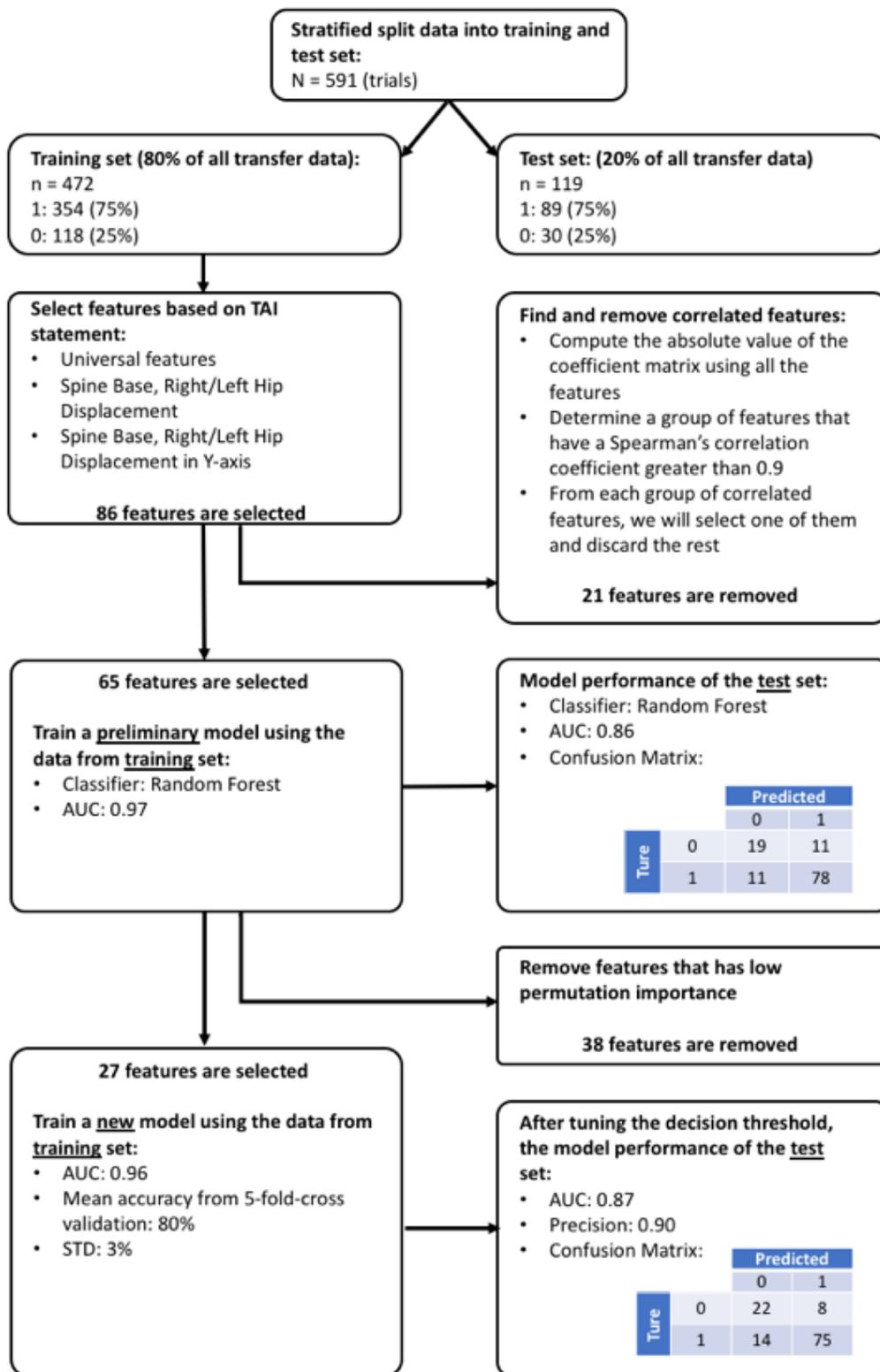


Figure A4b. Feature selection and training process of item 8 model

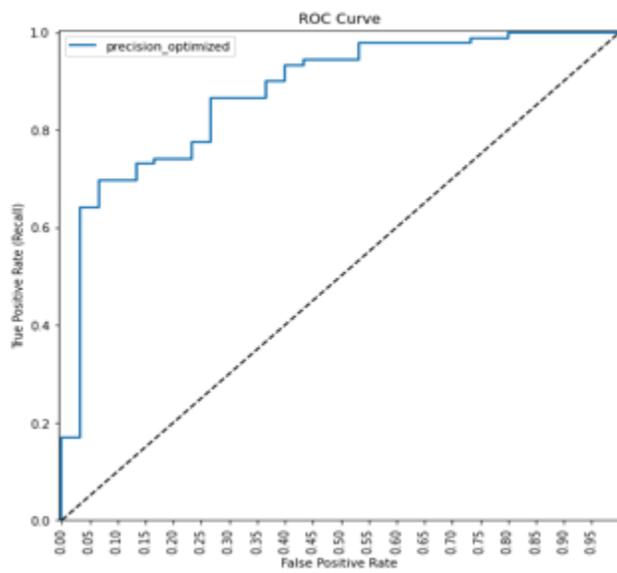
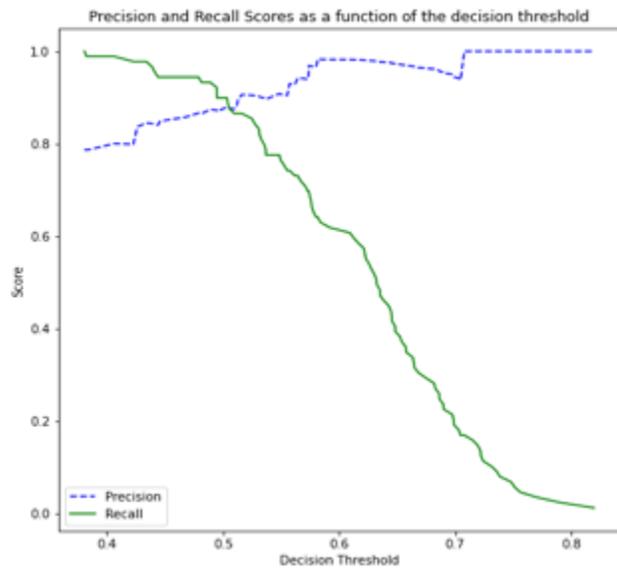


Figure A4c. ROC curve, and relationship between precision, recall and decision threshold values of the item 8 model

Item 9 Leading Arm Position before Transfer

The random forest classifier was chosen for the item 9 model (Figure A5a). Four hundred and sixty-four trials were split into the training set and 117 trials were into the test set. The transfer technique deficit rate was 17%. Sixty-eight features were selected into the feature engineering process and 27 of them were applied into the final model (Figure A5b). For the model from the training set, the AUC was .96, and the mean accuracy from CV was 84% (STD = 5%). After tuning the model decision threshold, the test set AUC was .94, and precision was .96. Figure A5c shows the ROC curve, and the relationship between precision, recall when adjusting decision threshold of the outcomes from the test set.

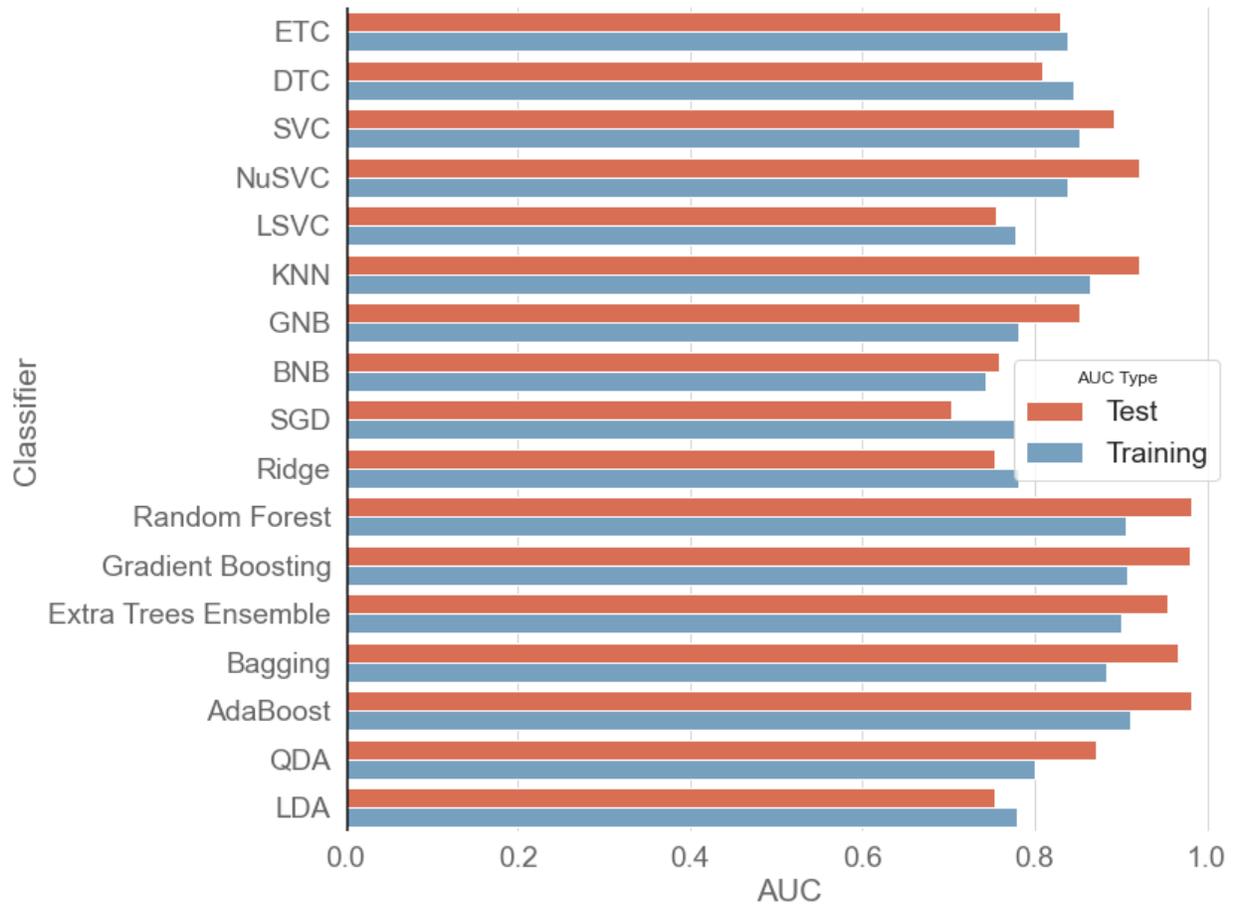


Figure A5a. Model performance of each classifier for item 9

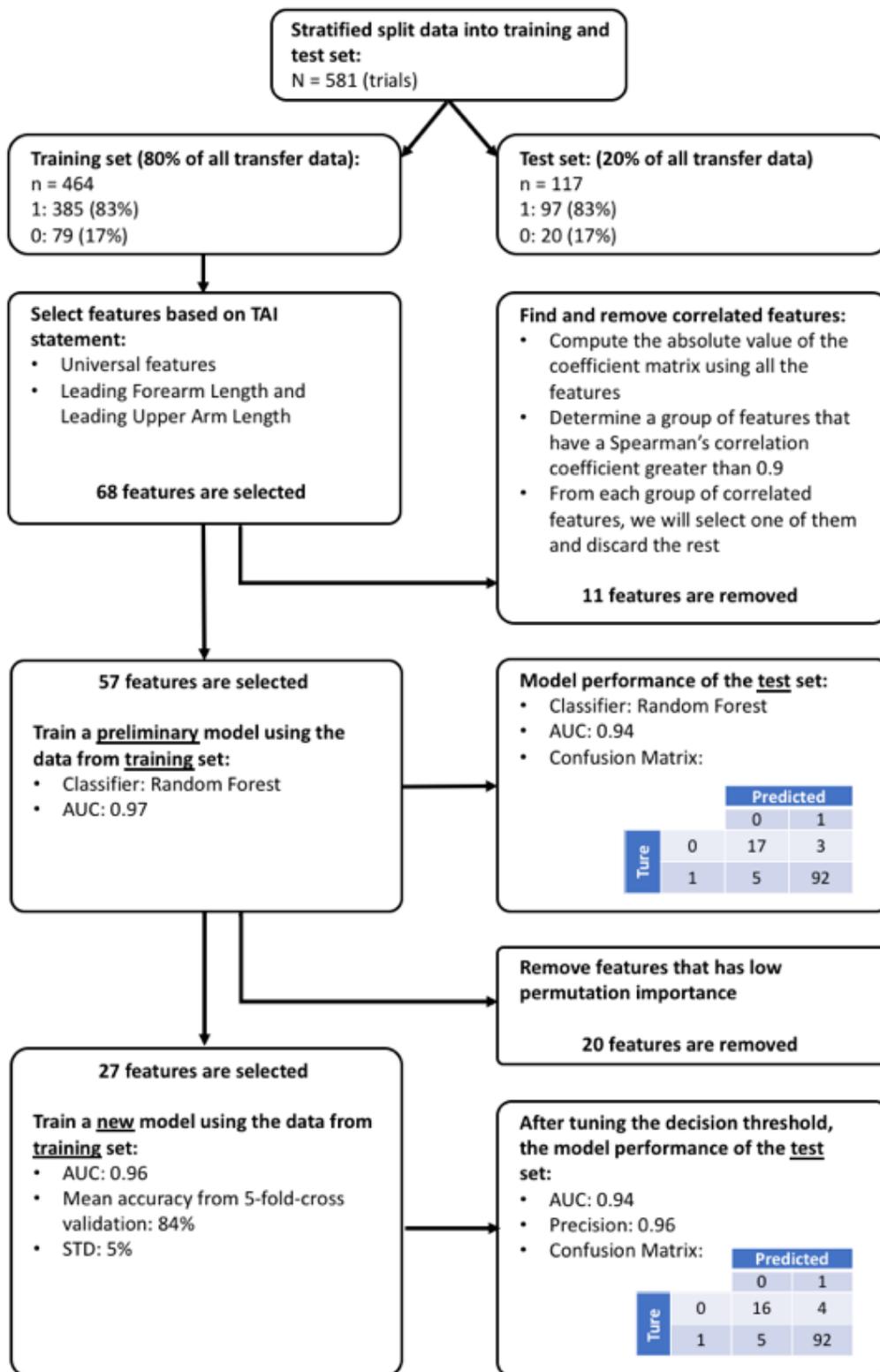


Figure A5b. Feature selection and training process of item 9 model

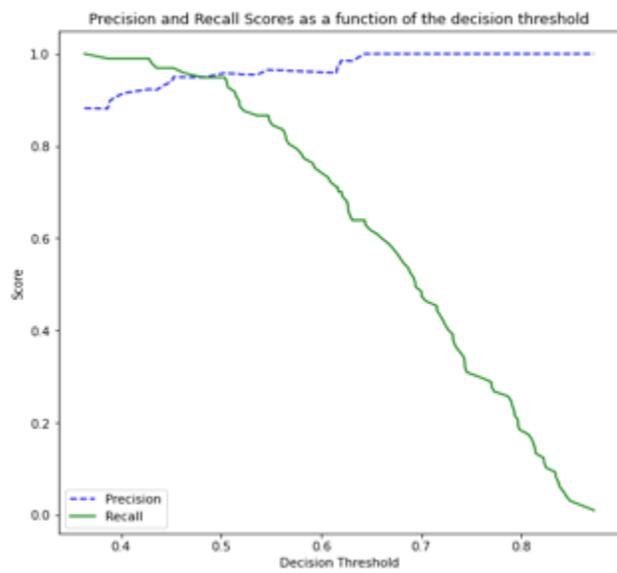
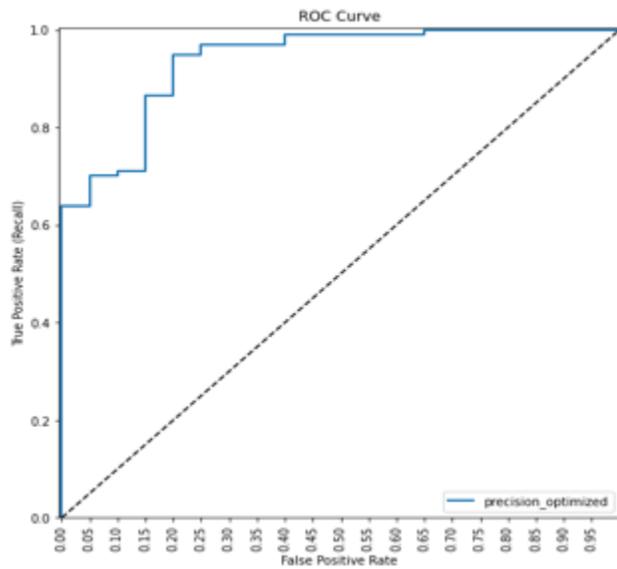


Figure A5c. ROC curve, and relationship between precision, recall and decision threshold values of the item 9 model

Item 10 Push-off Handgrip

The KNN classifier was chosen for the item 10 model (Figure A6a). Four hundred and sixty-four trials were split into the training set and 116 trials were into the test set. The transfer technique deficit rate was 41%. Seventy features were selected into the feature engineering process and 38 of them were applied into the final model (Figure A6b). For the model from the training set, the AUC was .97, and the mean accuracy from CV was 72% (STD = 3%). After tuning the model decision threshold, the test set AUC was .82, and precision was .92. Figure A6c shows the ROC curve, and the relationship between precision, recall when adjusting decision threshold of the outcomes from the test set.

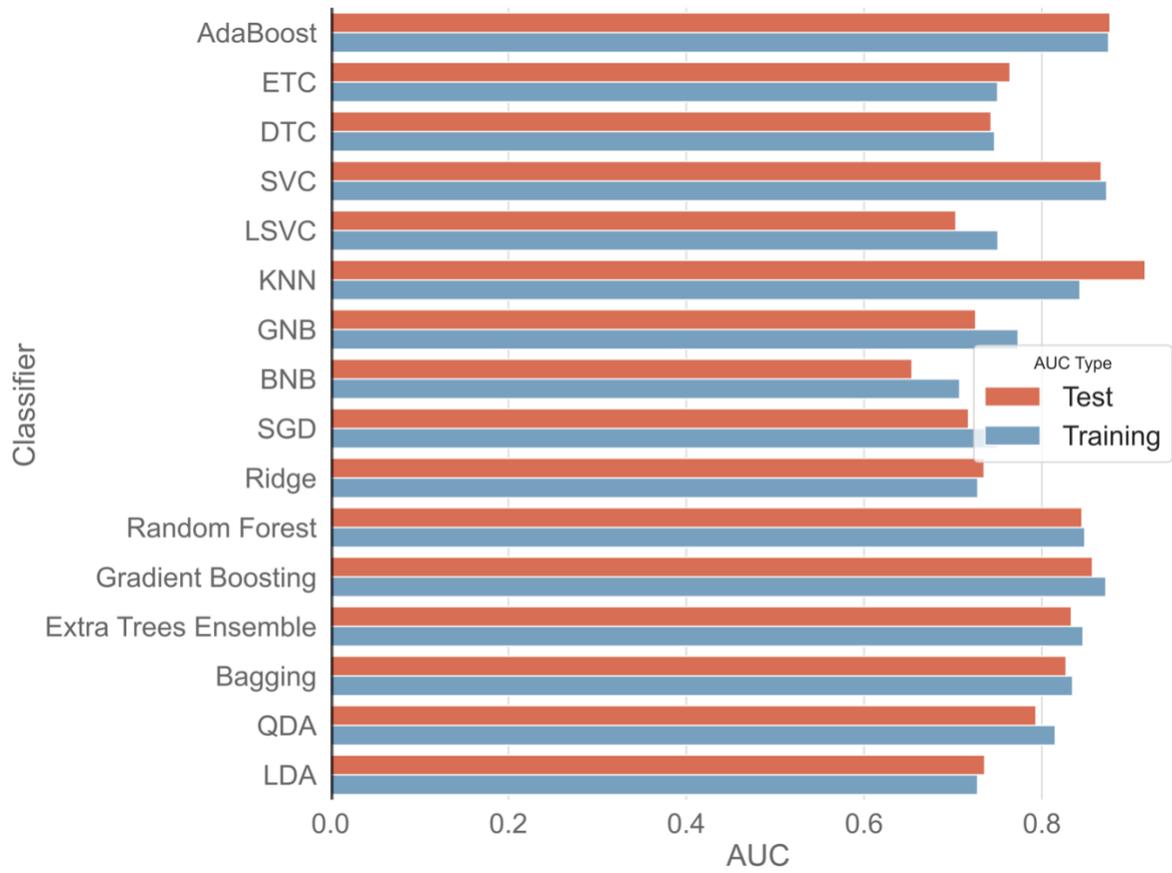


Figure A6a. Model perforce of each classifier for item 10

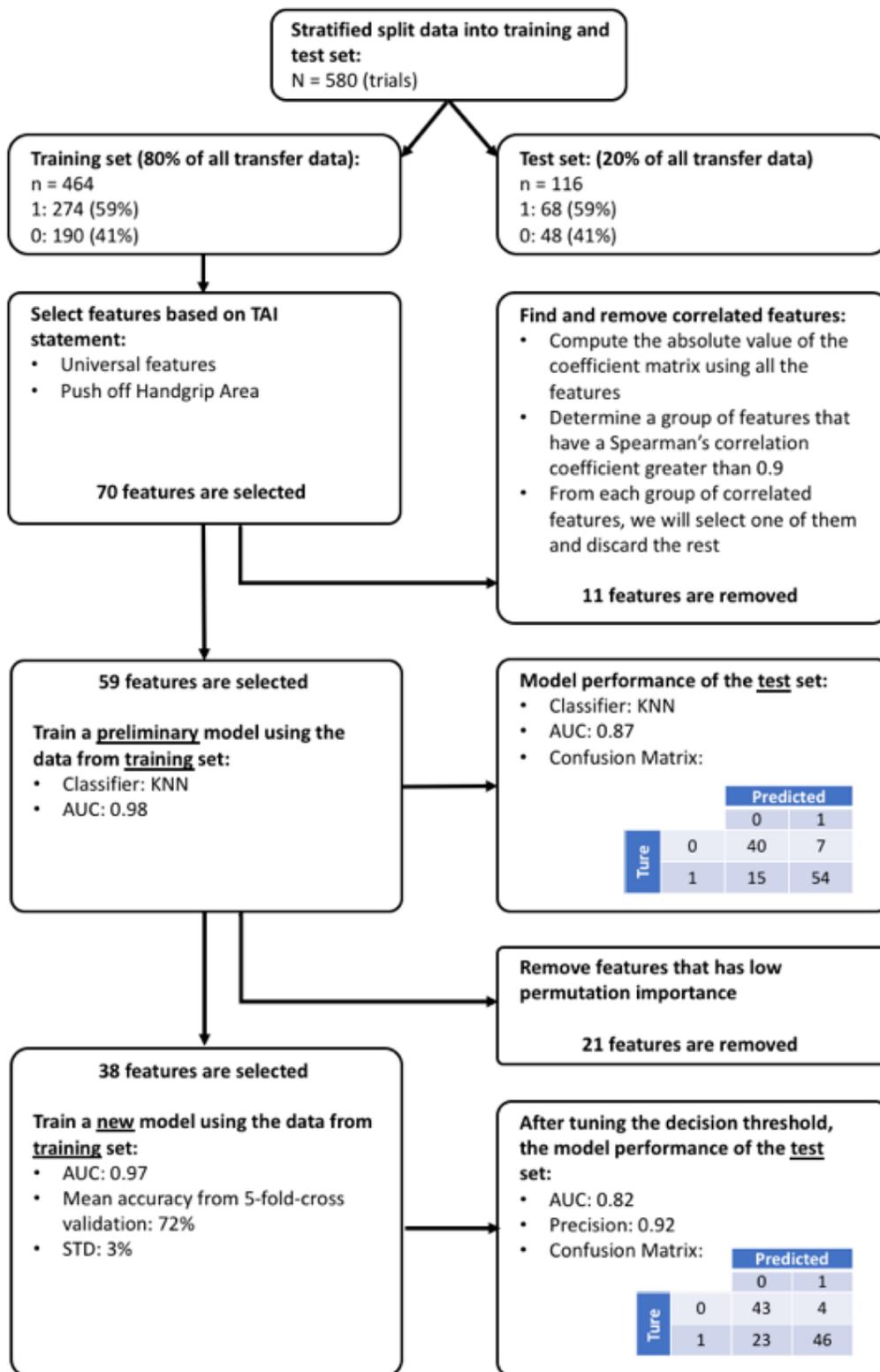


Figure A6b. Feature selection and training process of item 10 model

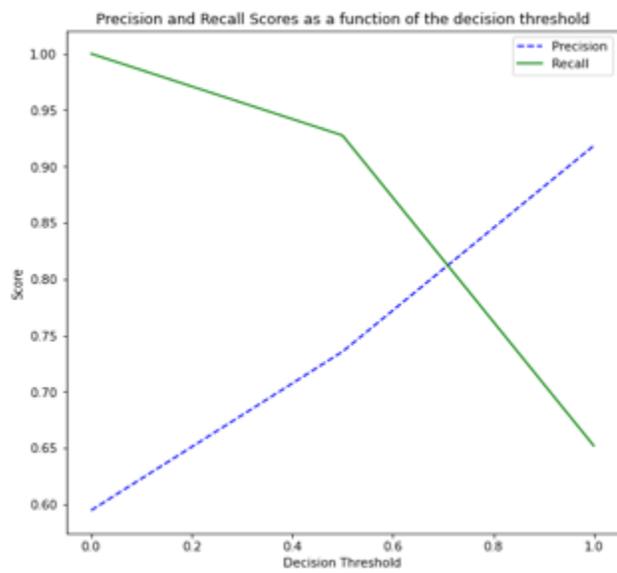
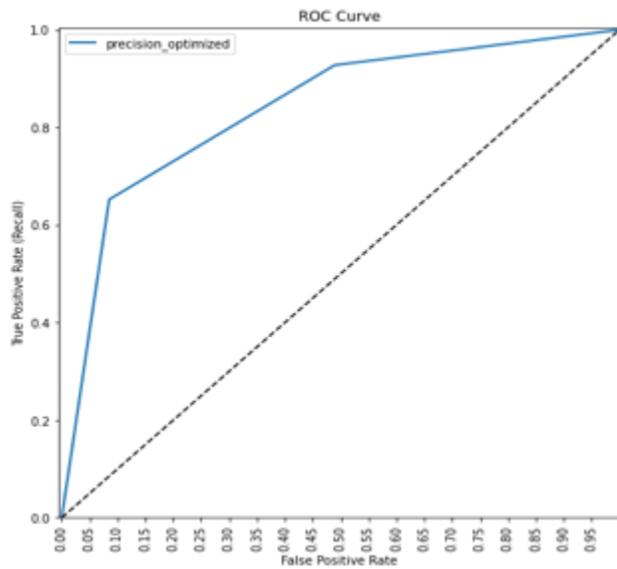


Figure A6c. ROC curve, and relationship between precision, recall and decision threshold values of the item 10 model

Item 11 Leading Handgrip

The KNN classifier was chosen for the item 11 model (Figure A7a). Four hundred and sixty-four trials were split into the training set and 116 trials were into the test set. The transfer technique deficit rate was 42%. Seventy features were selected into the feature engineering process and 35 of them were applied into the final model (Figure A7b). For the model from the training set, the AUC was .95, and the mean accuracy from CV was 76% (STD = 3%). After tuning the model decision threshold, the test set AUC was .87, and precision was .91. Figure A7c shows the ROC curve, and the relationship between precision, recall when adjusting decision threshold of the outcomes from the test set.

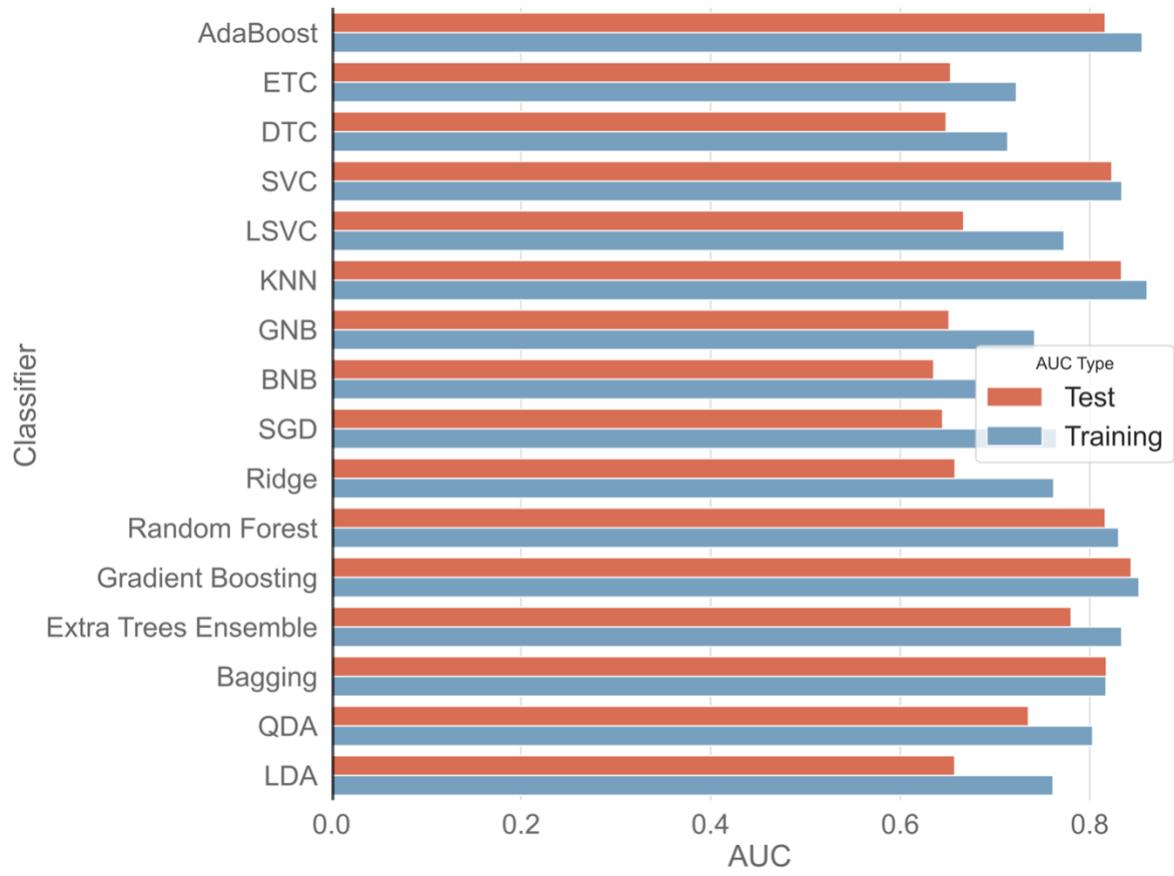


Figure A7a. Model performance of each classifier for item 11

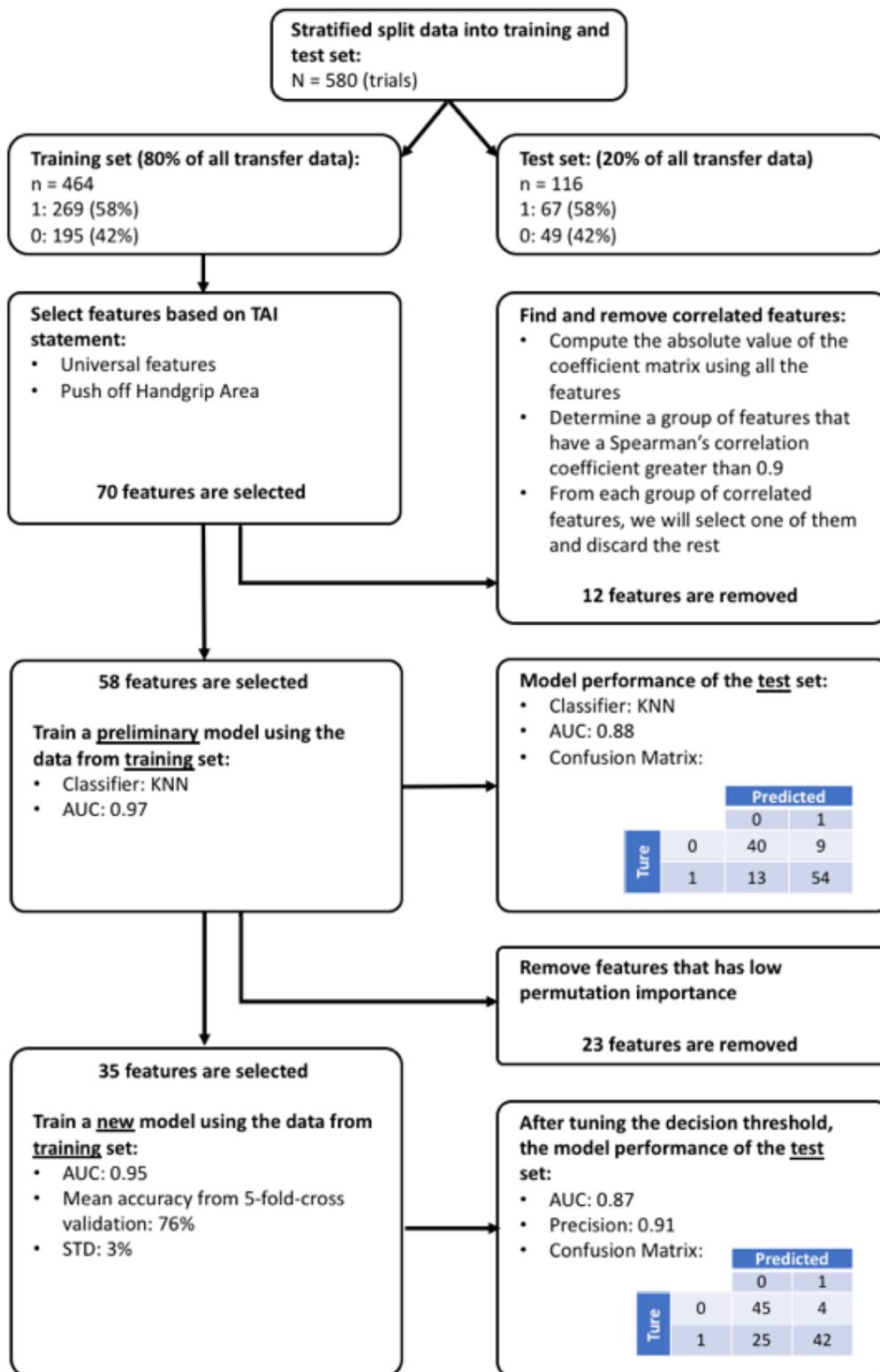


Figure A7b. Feature selection and training process of item 11 model

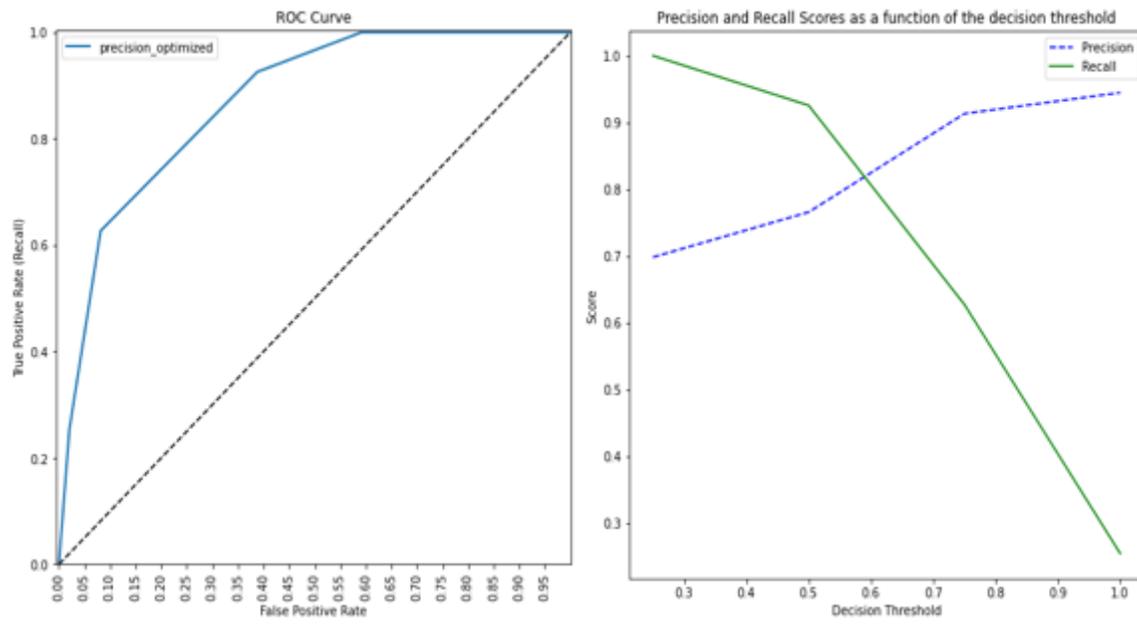


Figure A7c. ROC curve, and relationship between precision, recall and decision threshold values of the item 11 model

Item 12 Leading Hand Position after Transfer

The random forest classifier was chosen for the item 12 model (Figure A8a). Four hundred and seventy-two trials were split into the training set and 119 trials were into the test set. The transfer technique deficit rate was 11%. Eighty features were selected into the feature engineering process and 23 of them were applied into the final model (Figure A8b). For the model from the training set, the AUC was .98, and the mean accuracy from CV was 88% (STD = 2%). After tuning the model decision threshold, the test set AUC was .85, and precision was .95. Figure A8c shows the ROC curve, and the relationship between precision, recall when adjusting decision threshold of the outcomes from the test set.

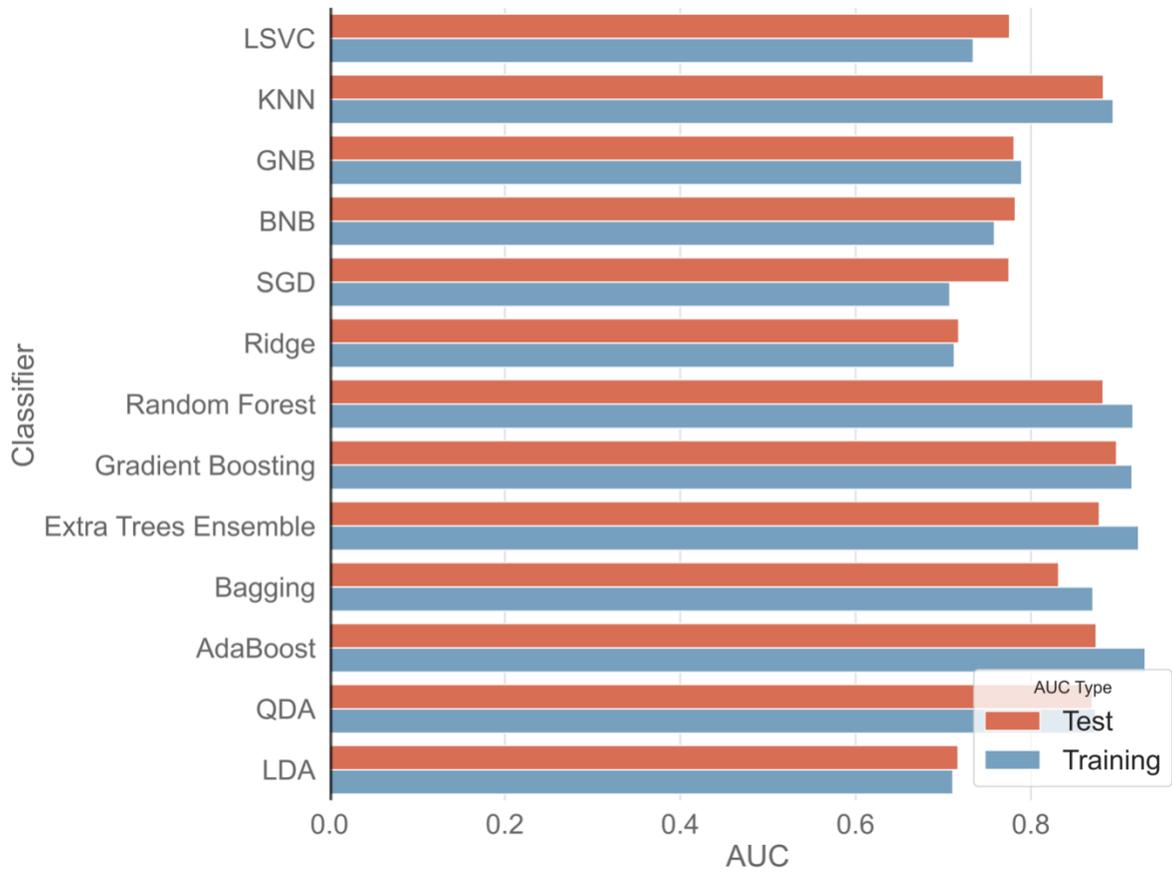


Figure A8a. Model perforce of each classifier for item 12

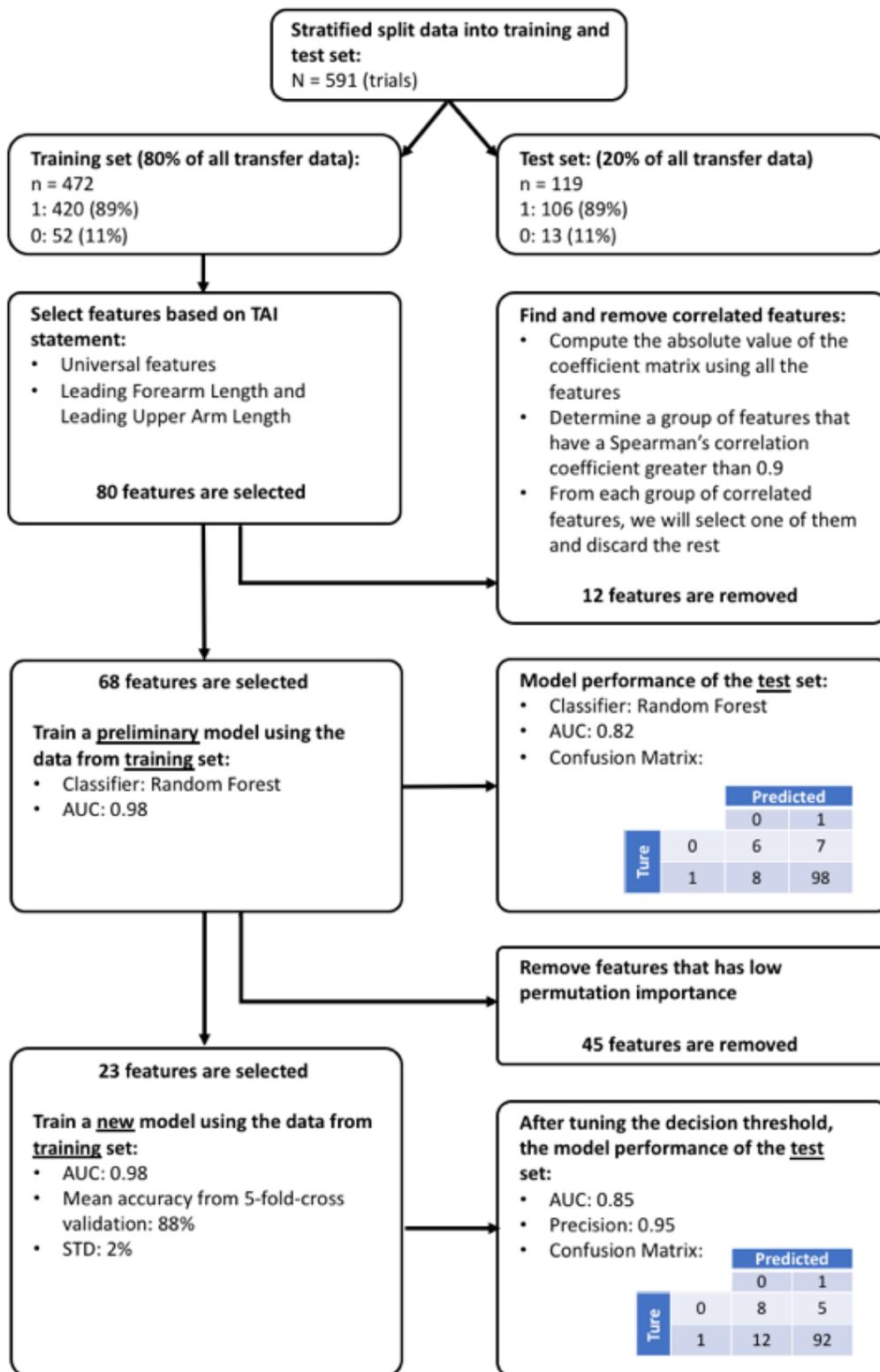


Figure A8b. Feature selection and training process of item 12 model

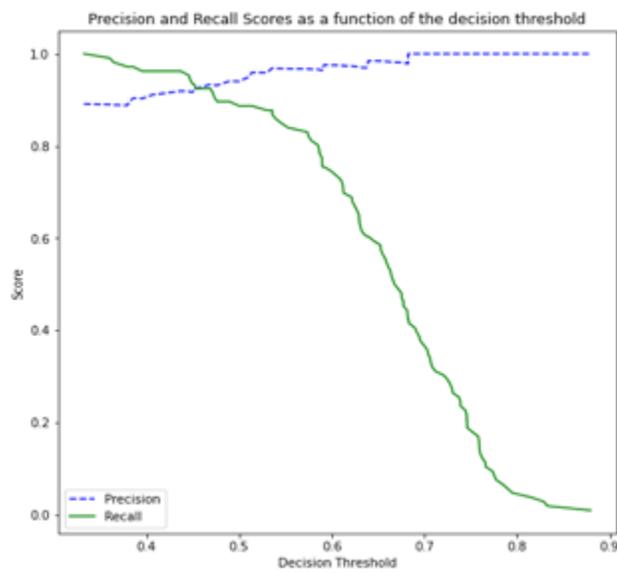
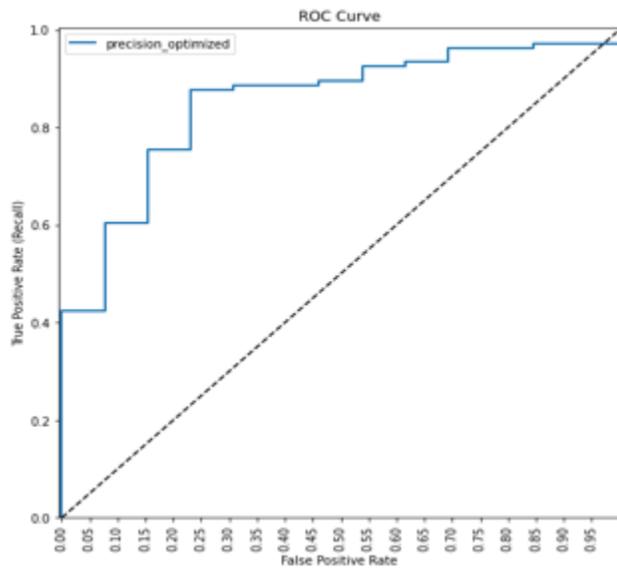


Figure A8c. ROC curve, and relationship between precision, recall and decision threshold values of the item 12 model

Item 13 Trunk Leaning

The random forest classifier was chosen for the item 13 model (Figure A9a). Four hundred and seventy-two trials were split into the training set and 119 trials were into the test set. The transfer technique deficit rate was 29%. Thirty-one features were selected into the feature engineering process and 9 of them were applied into the final model (Figure A9b). For the model from the training set, the AUC was .94, and the mean accuracy from CV was 80% (STD = 3%). After tuning the model decision threshold, the test set AUC was .89, and precision was .94. Figure A9c shows the ROC curve, and the relationship between precision, recall when adjusting decision threshold of the outcomes from the test set.

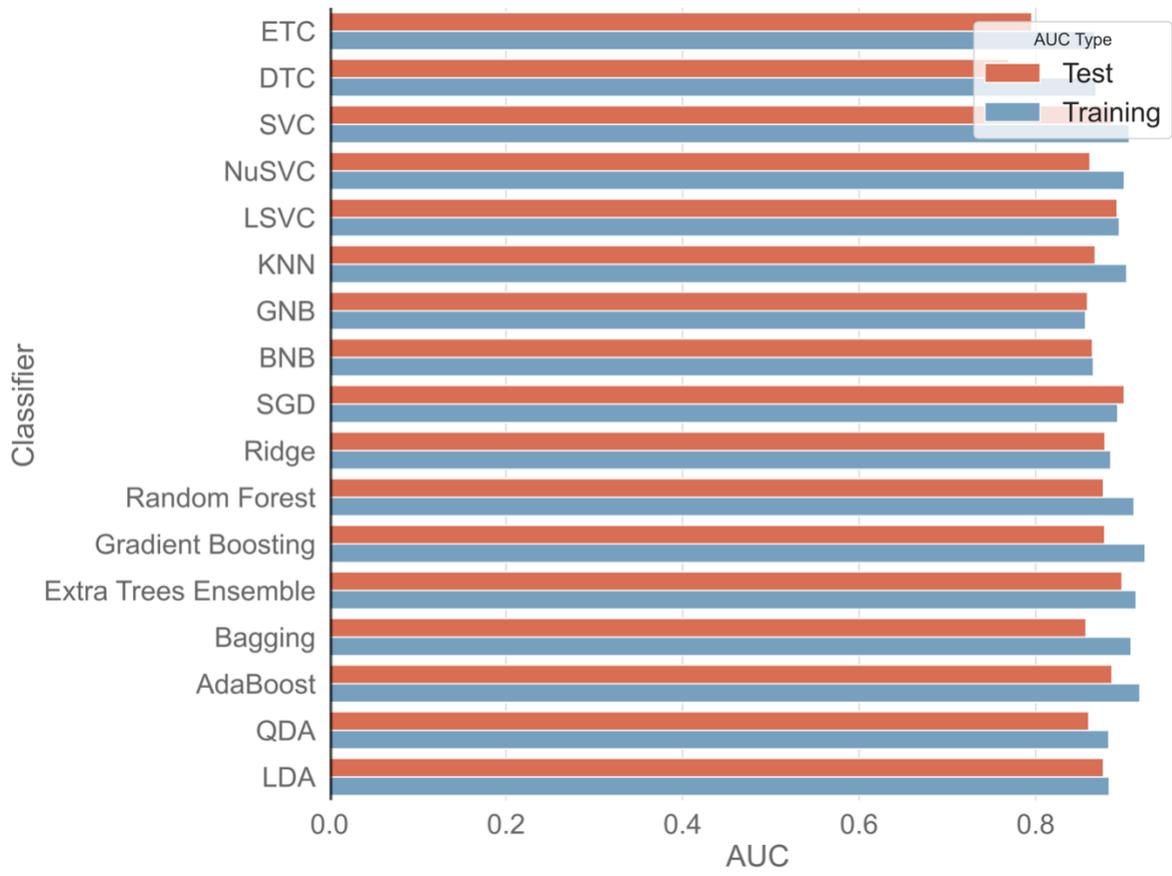


Figure A9a. Model perforce of each classifier for item 13

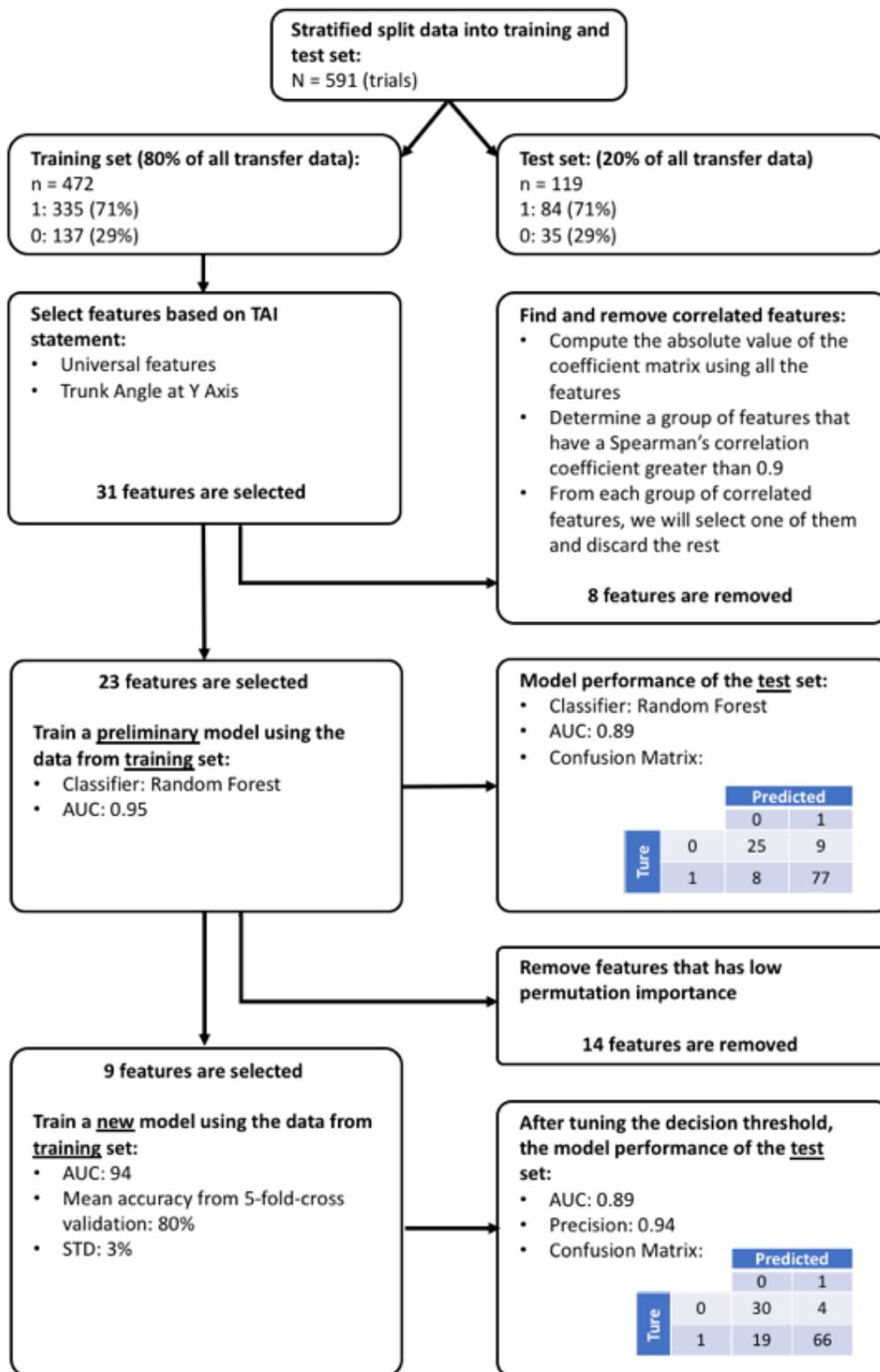


Figure A9b. Feature selection and training process of item 13 model

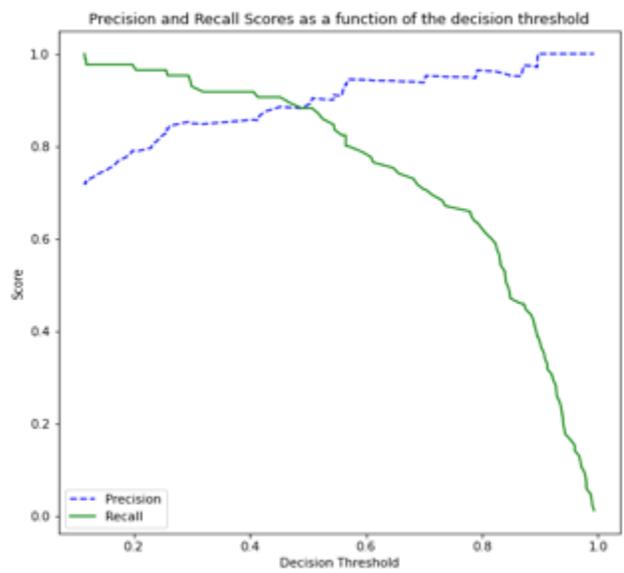
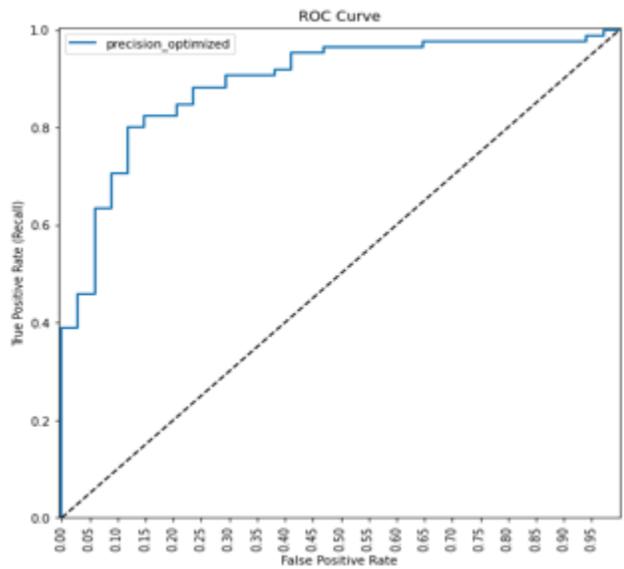


Figure A9c. ROC curve, and relationship between precision, recall and decision threshold values of the item 13 model

Item 14 Flight

The random forest classifier was chosen for the item 14 model (Figure A10a). Four hundred and seventy-two trials were split into the training set and 119 trials were into the test set. The transfer technique deficit rate was 32%. One-hundred and seventy-five features were selected into the feature engineering process and 40 of them were applied into the final model (Figure A10b). For the model from the training set, the AUC was .95, and the mean accuracy from CV was 73% (STD = 3%). After tuning the model decision threshold, the test set AUC was .87, and the precision was .91. Figure A10c shows the ROC curve, and the relationship between precision, recall when adjusting decision threshold of the outcomes from the test set.

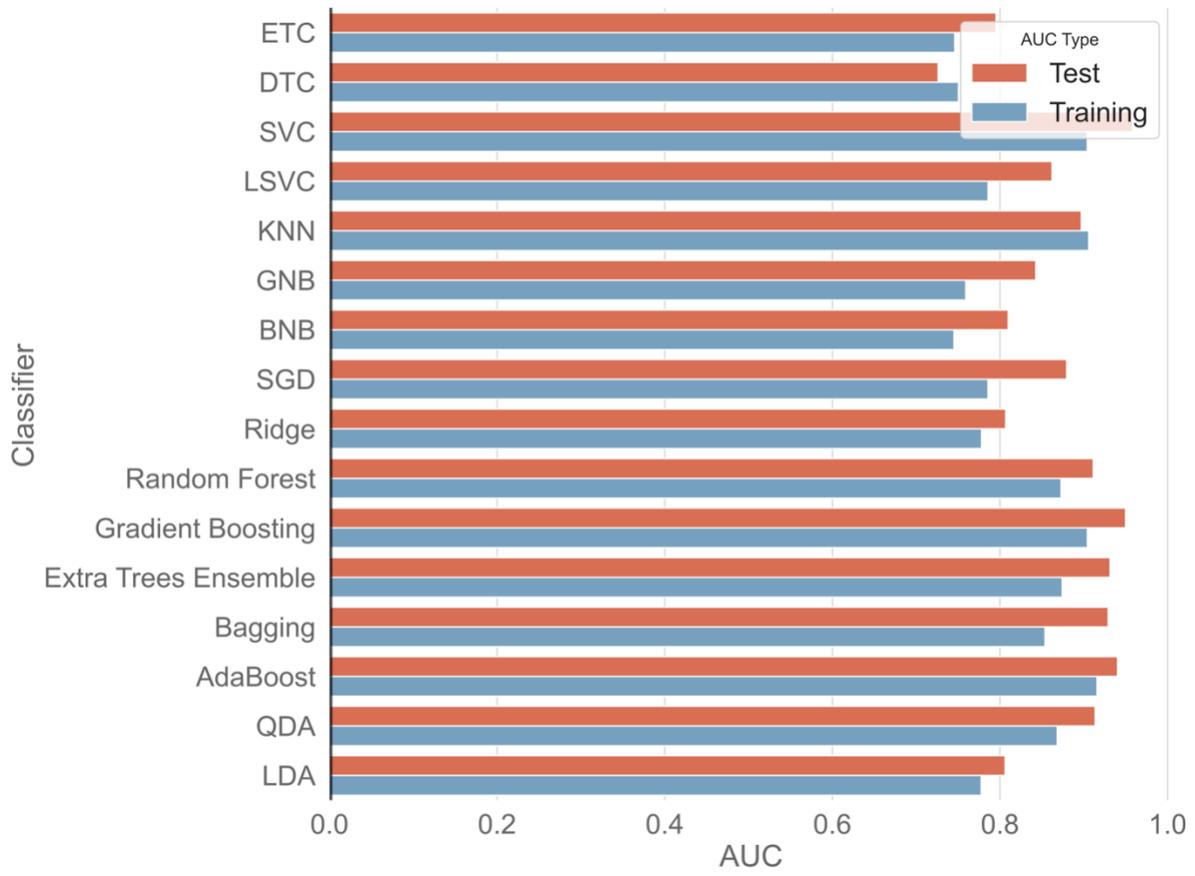


Figure A10a. Model performance of each classifier for item 14

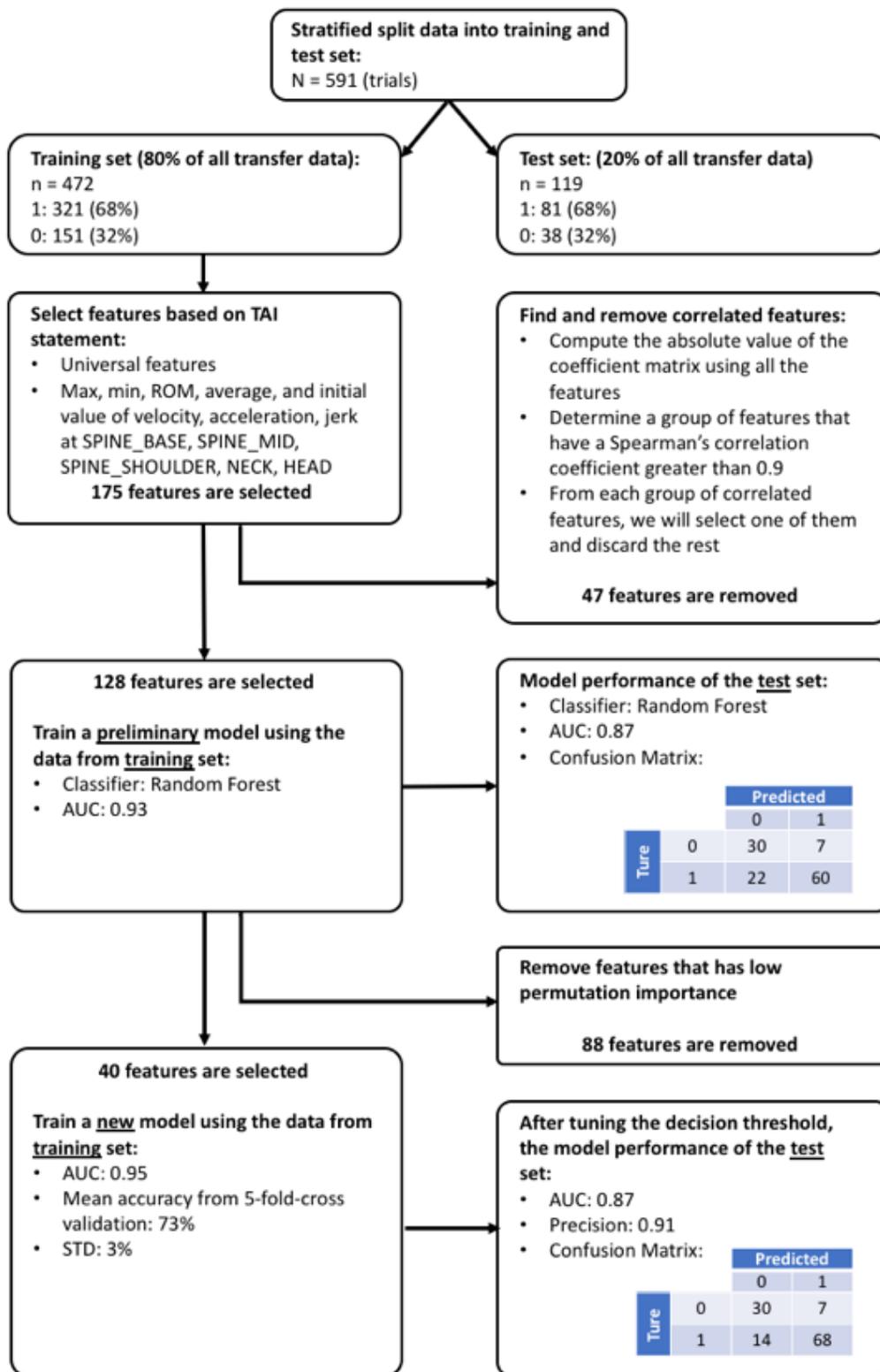


Figure A10b. Feature selection and training process of item 14 model

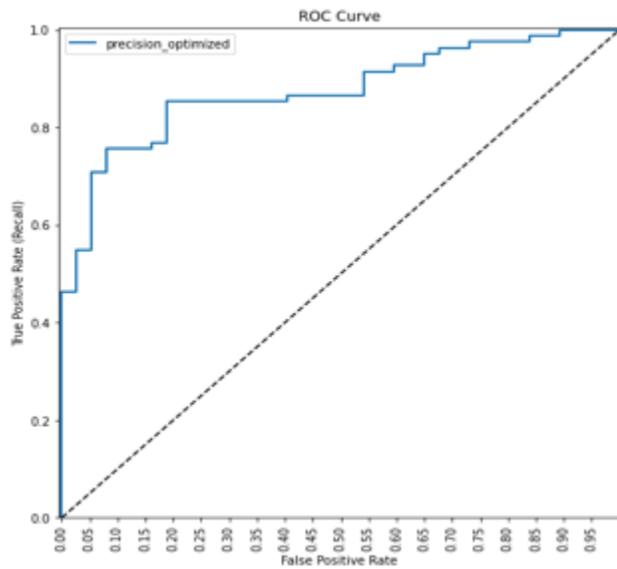


Figure A10c. ROC curve, and relationship between precision, recall and decision threshold values of the item 14 model

Item 15 Landing

The random forest classifier was chosen for the item 15 model (Figure A11a). Four hundred and seventy-two trials were split into the training set and 119 trials were into the test set. The transfer technique deficit rate was 21%. One-hundred and seventy-five features were selected into the feature engineering process and 33 of them were applied into the final model (Figure A11b). For the model from the training set, the AUC was .83, and the mean accuracy from CV was 80% (STD = 2%). After tuning the model decision threshold, the test set AUC was .79, and the precision was .91. Figure A11c shows the ROC curve, and the relationship between precision, recall when adjusting decision threshold of the outcomes from the test set.

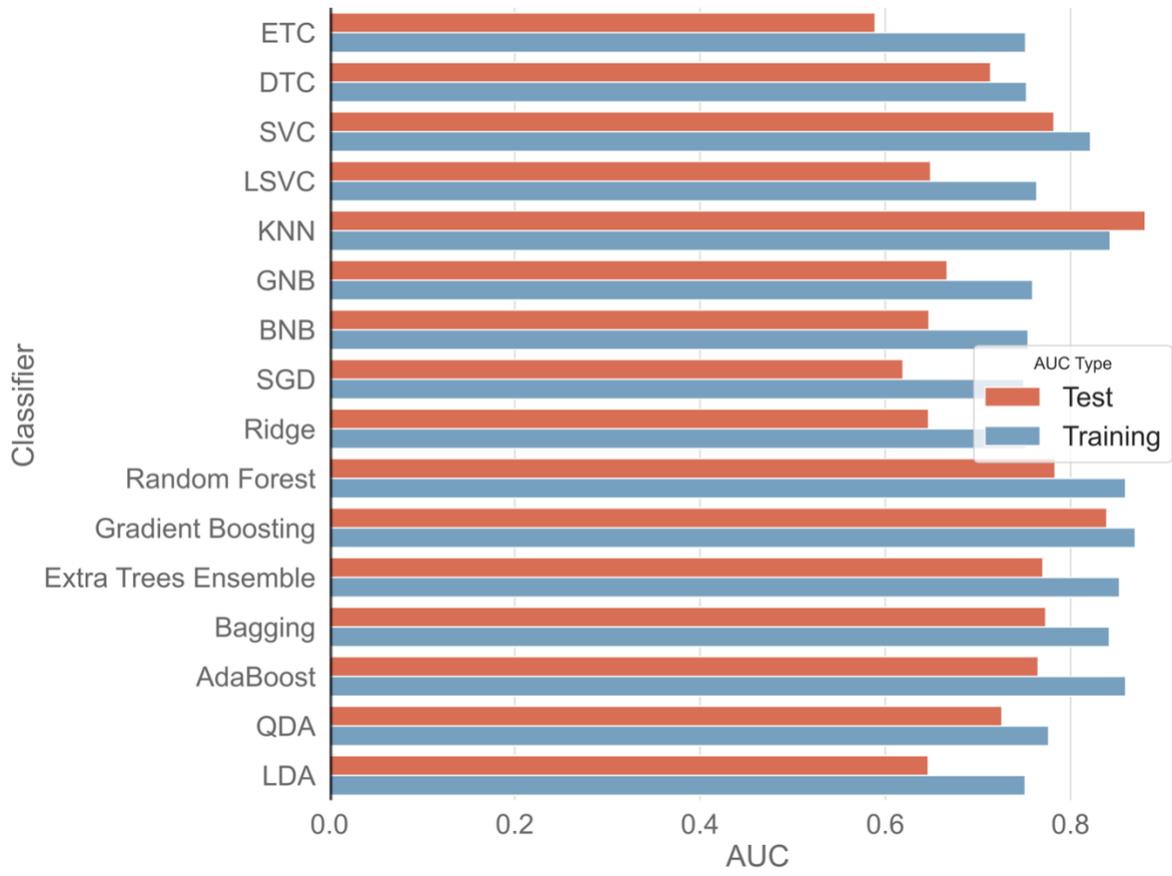


Figure A11a. Model performance of each classifier for item 15

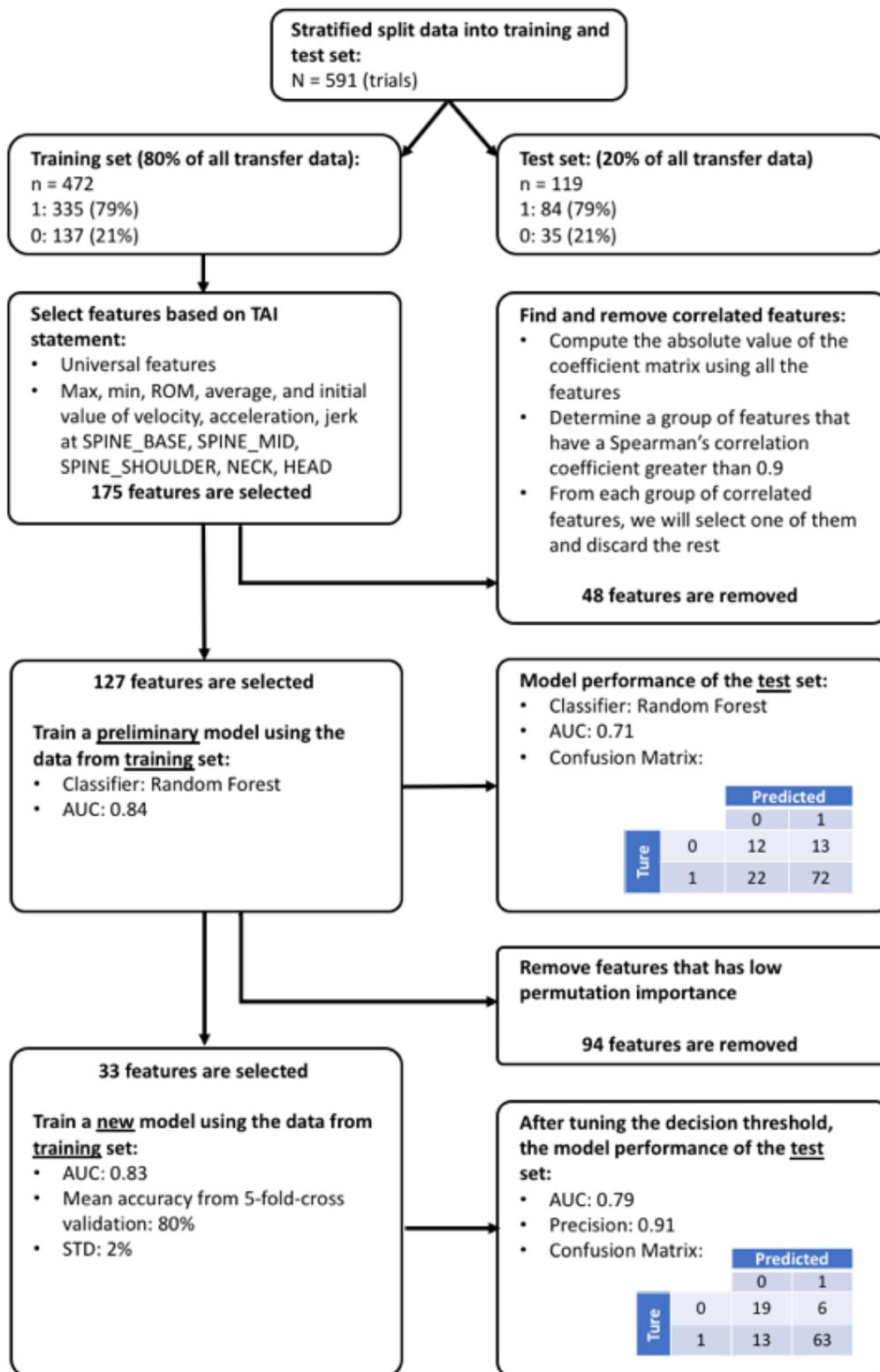


Figure A11b. Feature selection and training process of item 15 model

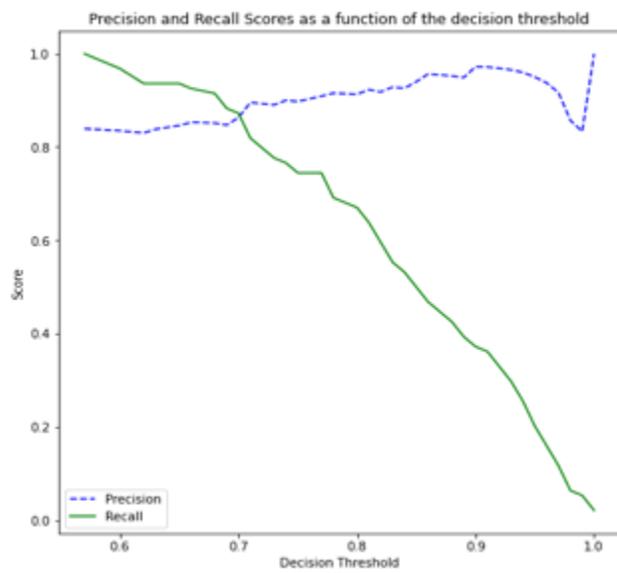
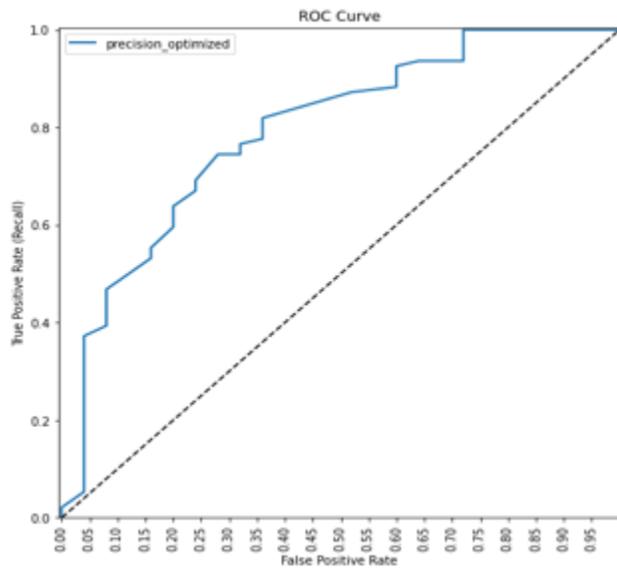


Figure A11c. ROC curve, and relationship between precision, recall and decision threshold values of the item 15 model

Appendix D. Model Performance of Each Classifier for the Machine Learning Method

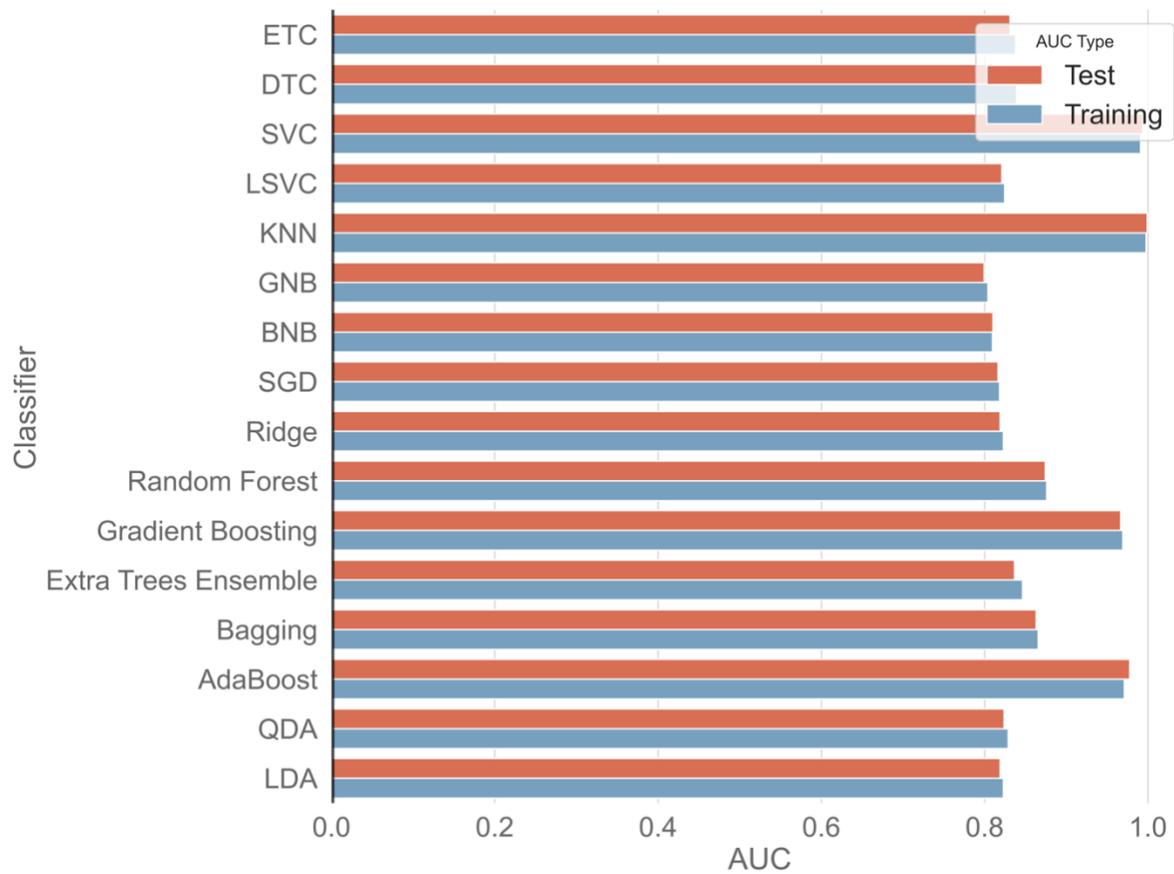


Figure D1. Model performance of each classifier for the machine learning method (**Chapter 3, Section 3.2.3.3 Machine Learning Method of Phase Delineation**)

Bibliography

1. *Facts and Figures at a Glance*, N.S.C.I.S. Center, Editor. 2020: Birmingham, AL: University of Alabama at Birmingham.
2. *Americans With Disabilities*. 2010; Available from: <http://www.census.gov/prod/2012pubs/p70-131.pdf>.
3. Hogaboom, N.S., L.A. Worobey, and M.L. Boninger, *Transfer Technique Is Associated With Shoulder Pain and Pathology in People With Spinal Cord Injury: A Cross-Sectional Investigation*. Arch Phys Med Rehabil, 2016. **97**(10): p. 1770-6.
4. Tsai, C.Y., N.S. Hogaboom, M.L. Boninger, and A.M. Koontz, *The relationship between independent transfer skills and upper limb kinetics in wheelchair users*. Biomed Res Int, 2014. **2014**: p. 984526.
5. Cooper, R.A., M.L. Boninger, S.D. Shimada, and B.M. Lawrence, *Glenohumeral joint kinematics and kinetics for three coordinate system representations during wheelchair propulsion*. Am J Phys Med Rehabil, 1999. **78**(5): p. 435-46.
6. Gutierrez, D.D., L. Thompson, B. Kemp, S.J. Mulroy, N. Physical Therapy Clinical Research, R. Rehabilitation, and D. Training Center on Aging-Related Changes in Impairment for Persons Living with Physical, *The relationship of shoulder pain intensity to quality of life, physical activity, and community participation in persons with paraplegia*. J Spinal Cord Med, 2007. **30**(3): p. 251-5.
7. Tsai, C.Y., M.L. Boninger, S.R. Bass, and A.M. Koontz, *Upper-limb biomechanical analysis of wheelchair transfer techniques in two toilet configurations*. Clin Biomech (Bristol, Avon), 2018. **55**: p. 79-85.
8. Finley, M.A. and M.M. Rodgers, *Prevalence and identification of shoulder pathology in athletic and nonathletic wheelchair users with shoulder pain: A pilot study*. J Rehabil Res Dev, 2004. **41**(3B): p. 395-402.
9. Paralyzed Veterans of America Consortium for Spinal Cord, M., *Preservation of upper limb function following spinal cord injury: a clinical practice guideline for health-care professionals*. J Spinal Cord Med, 2005. **28**(5): p. 434-70.
10. Girona, R.J., M.E. Clark, B. Neugaard, and A. Nelson, *Upper limb pain in a national sample of veterans with paraplegia*. J Spinal Cord Med, 2004. **27**(2): p. 120-7.
11. Pentland, W.E. and L.T. Twomey, *Upper limb function in persons with long term paraplegia and implications for independence: Part I*. Paraplegia, 1994. **32**(4): p. 211-8.
12. Kulig, K., S.S. Rao, S.J. Mulroy, C.J. Newsam, J.K. Gronley, E.L. Bontrager, and J. Perry, *Shoulder joint kinetics during the push phase of wheelchair propulsion*. Clin Orthop Relat Res, 1998(354): p. 132-43.
13. Neer, C.S., *Anterior acromioplasty for the chronic impingement syndrome in the shoulder*. 1972. J Bone Joint Surg Am, 2005. **87**(6): p. 1399.
14. Fu, F.H., C.D. Harner, and A.H. Klein, *Shoulder impingement syndrome. A critical review*. Clin Orthop Relat Res, 1991(269): p. 162-73.
15. Seitz, A.L., P.W. McClure, S. Finucane, N.D. Boardman, 3rd, and L.A. Michener, *Mechanisms of rotator cuff tendinopathy: intrinsic, extrinsic, or both?* Clin Biomech (Bristol, Avon), 2011. **26**(1): p. 1-12.

16. Dyson-Hudson, T.A. and S.C. Kirshblum, *Shoulder pain in chronic spinal cord injury, Part I: Epidemiology, etiology, and pathomechanics*. J Spinal Cord Med, 2004. **27**(1): p. 4-17.
17. Akbar, M., G. Balean, M. Brunner, T.M. Seyler, T. Bruckner, J. Munzinger, T. Grieser, H.J. Gerner, and M. Loew, *Prevalence of rotator cuff tear in paraplegic patients compared with controls*. J Bone Joint Surg Am, 2010. **92**(1): p. 23-30.
18. Akbar, M., M. Brunner, G. Balean, T. Grieser, T. Bruckner, M. Loew, and P. Raiss, *A cross-sectional study of demographic and morphologic features of rotator cuff disease in paraplegic patients*. J Shoulder Elbow Surg, 2011. **20**(7): p. 1108-13.
19. Campbell, C.C. and M.J. Koris, *Etiologies of shoulder pain in cervical spinal cord injury*. Clin Orthop Relat Res, 1996(322): p. 140-5.
20. Yanai, T., F.K. Fuss, and T. Fukunaga, *In vivo measurements of subacromial impingement: substantial compression develops in abduction with large internal rotation*. Clin Biomech (Bristol, Avon), 2006. **21**(7): p. 692-700.
21. Dany Gagnon, A.M.K., Sara J. Mulroy, Debbie A. Nawoczenski, Emelie Butler-Forslund, Anna Granstrom, Sylvie Nadeau, and a.M.L. Boninger, *Biomechanics of Sitting Pivot Transfers Among Individuals with a Spinal Cord Injury: A Review of the Current Knowledge*. Top Spinal Cord Inj Rehabil, 2009. **15**(2): p. 33-58.
22. Davidoff, G., R. Werner, and W. Waring, *Compressive mononeuropathies of the upper extremity in chronic paraplegia*. Paraplegia, 1991. **29**(1): p. 17-24.
23. Sie, I.H., R.L. Waters, R.H. Adkins, and H. Gellman, *Upper extremity pain in the postrehabilitation spinal cord injured patient*. Arch Phys Med Rehabil, 1992. **73**(1): p. 44-8.
24. Burnham, R.S. and R.D. Steadward, *Upper extremity peripheral nerve entrapments among wheelchair athletes: prevalence, location, and risk factors*. Arch Phys Med Rehabil, 1994. **75**(5): p. 519-24.
25. Akbar, M., S. Penzkofer, M.A. Weber, T. Bruckner, M. Winterstein, and M. Jung, *Prevalence of carpal tunnel syndrome and wrist osteoarthritis in long-term paraplegic patients compared with controls*. J Hand Surg Eur Vol, 2014. **39**(2): p. 132-8.
26. Fliess-Douer, O., Y.C. Vanlandewijck, and L.H. Van der Woude, *Most essential wheeled mobility skills for daily life: an international survey among paralympic wheelchair athletes with spinal cord injury*. Arch Phys Med Rehabil, 2012. **93**(4): p. 629-35.
27. Gagnon, D., S. Nadeau, L. Noreau, P. Dehail, and F. Piote, *Comparison of peak shoulder and elbow mechanical loads during weight-relief lifts and sitting pivot transfers among manual wheelchair users with spinal cord injury*. J Rehabil Res Dev, 2008. **45**(6): p. 863-73.
28. Koontz, A.M., P. Kankipati, Y.S. Lin, R.A. Cooper, and M.L. Boninger, *Upper limb kinetic analysis of three sitting pivot wheelchair transfer techniques*. Clin Biomech (Bristol, Avon), 2011. **26**(9): p. 923-9.
29. Keir, P.J., R.P. Wells, D.A. Ranney, and W. Lavery, *The effects of tendon load and posture on carpal tunnel pressure*. J Hand Surg Am, 1997. **22**(4): p. 628-34.
30. Sinnott, K.A., P. Milburn, and H. McNaughton, *Factors associated with thoracic spinal cord injury, lesion level and rotator cuff disorders*. Spinal Cord, 2000. **38**(12): p. 748-53.
31. Kankipati, P., M.L. Boninger, D. Gagnon, R.A. Cooper, and A.M. Koontz, *Upper limb joint kinetics of three sitting pivot wheelchair transfer techniques in individuals with spinal cord injury*. J Spinal Cord Med, 2015. **38**(4): p. 485-97.

32. Tsai, C.-Y., M.L. Boninger, J. Hastings, R.A. Cooper, L. Rice, and A.M. Koontz, *Immediate Biomechanical Implications of Transfer Component Skills Training on Independent Wheelchair Transfers*. Archives of physical medicine and rehabilitation, 2016. **97**(10): p. 1785-92.
33. Tsai, C.Y., M.L. Boninger, J. Hastings, R.A. Cooper, L. Rice, and A.M. Koontz, *Immediate Biomechanical Implications of Transfer Component Skills Training on Independent Wheelchair Transfers*. Arch Phys Med Rehabil, 2016. **97**(10): p. 1785-92.
34. McClure, L.A., M.L. Boninger, H. Ozawa, and A. Koontz, *Reliability and validity analysis of the transfer assessment instrument*. Arch Phys Med Rehabil, 2011. **92**(3): p. 499-508.
35. Tsai, C.Y., L.A. Rice, C. Hoelmer, M.L. Boninger, and A.M. Koontz, *Basic psychometric properties of the transfer assessment instrument (version 3.0)*. Arch Phys Med Rehabil, 2013. **94**(12): p. 2456-64.
36. Worobey, L.A., C.K. Zigler, R. Huzinec, S.K. Rigot, J. Sung, and L.A. Rice, *Reliability and Validity of the Revised Transfer Assessment Instrument*. Top Spinal Cord Inj Rehabil, 2018. **24**(3): p. 217-226.
37. Worobey, L.A., S.K. Rigot, N.S. Hogaboom, C. Venus, and M.L. Boninger, *Investigating the Efficacy of Web-Based Transfer Training on Independent Wheelchair Transfers Through Randomized Controlled Trials*. Arch Phys Med Rehabil, 2018. **99**(1): p. 9-16 e10.
38. French B, S.A., Siewiorek D, Ambur V, Tyamagundlu D, *Classifying Wheelchair Propulsion Patterns with a Wrist Mounted Accelerometer*, in Proc ICST 3rd Int Conf Body Area Netw. 2008, ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering): Brussels, Belgium. p. 1–20.
39. Garcia-Masso, X., P. Serra-Ano, L.M. Garcia-Raffi, E.A. Sanchez-Perez, J. Lopez-Pascual, and L.M. Gonzalez, *Validation of the use of Actigraph GT3X accelerometers to estimate energy expenditure in full time manual wheelchair users with spinal cord injury*. Spinal Cord, 2013. **51**(12): p. 898-903.
40. Garcia-Masso, X., P. Serra-Ano, L.M. Gonzalez, Y. Ye-Lin, G. Prats-Boluda, and J. Garcia-Casado, *Identifying physical activity type in manual wheelchair users with spinal cord injury by means of accelerometers*. Spinal Cord, 2015. **53**(10): p. 772-7.
41. Learmonth, Y.C., D. Kinnett-Hopkins, I.M. Rice, J.L. Dysterheft, and R.W. Motl, *Accelerometer output and its association with energy expenditure during manual wheelchair propulsion*. Spinal Cord, 2016. **54**(2): p. 110-4.
42. Hiremath, S.V., S.S. Intille, A. Kelleher, R.A. Cooper, and D. Ding, *Detection of physical activities using a physical activity monitor system for wheelchair users*. Med Eng Phys, 2015. **37**(1): p. 68-76.
43. Holloway C, H.B., Barbareschi G, Nicholson S, Hailes S, *Street rehab: Linking accessibility and rehabilitation*, in 2016 38th Annu Int Conf IEEE Eng Med Biol Soc. 2016, EMBC 2016. p. 3167–3170.
44. Barbareschi G, H.C., Bianchi-Berthouze N, Sonenblum S, Sprigle S, *Use of a Low-Cost, Chest-Mounted Accelerometer to Evaluate Transfer Skills of Wheelchair Users During Everyday Activities: Observational Study*. JMIR Rehabil Assist Technol, 2018. **5**(2 e11748).

45. Cook, T.S., G. Couch, T.J. Couch, W. Kim, and W.W. Boonn, *Using the Microsoft Kinect for patient size estimation and radiation dose normalization: proof of concept and initial validation*. J Digit Imaging, 2013. **26**(4): p. 657-62.
46. Xu, X., R.W. McGorry, L.S. Chou, J.H. Lin, and C.C. Chang, *Accuracy of the Microsoft Kinect for measuring gait parameters during treadmill walking*. Gait Posture, 2015. **42**(2): p. 145-51.
47. Clark, R.A., Y.H. Pua, A.L. Bryant, and M.A. Hunt, *Validity of the Microsoft Kinect for providing lateral trunk lean feedback during gait retraining*. Gait Posture, 2013. **38**(4): p. 1064-6.
48. Clark, R.A., Y.-H. Pua, C.C. Oliveira, K.J. Bower, S. Thilarajah, R. McGaw, K. Hasanki, and B.F. Mentiplay, *Reliability and concurrent validity of the Microsoft Xbox One Kinect for assessment of standing balance and postural control*. Gait & posture, 2015. **42**(2): p. 210-3.
49. Kiselev, J., M. Haesner, M. Govercin, and E. Steinhagen-Thiessen, *Implementation of a home-based interactive training system for fall prevention: requirements and challenges*. J Gerontol Nurs, 2015. **41**(1): p. 14-9.
50. Reither, L.R., M.H. Foreman, N. Migotsky, C. Haddix, and J.R. Engsborg, *Upper extremity movement reliability and validity of the Kinect version 2*. Disabil Rehabil Assist Technol, 2018. **13**(1): p. 54-59.
51. Dehbandi, B., A. Barachant, D. Harary, J.D. Long, K.Z. Tsagaris, S.J. Bumanlag, V. He, and D. Putrino, *Using Data From the Microsoft Kinect 2 to Quantify Upper Limb Behavior: A Feasibility Study*. IEEE J Biomed Health Inform, 2017. **21**(5): p. 1386-1392.
52. Jintronix. 2019; Available from: <http://www.jintronix.com>.
53. Reflexion Health. 2019; Available from: <https://reflexionhealth.com>.
54. Microsoft. *Azure Kinect body tracking joints*. Available from: <https://docs.microsoft.com/en-us/azure/kinect-dk/body-joints>.
55. Baldominos, A., Saez, Y., and del Pozo, C.G., *An approach to physical rehabilitation using state-of-the-art virtual reality and motion tracking technologies*. Procedia Computer Science, 2015. **64**: p. 10-16.
56. Ferche, O., Moldoveanu, A., and Moldoveanu, F., *Evaluating Lightweight Optical Hand Tracking for Virtual Reality Rehabilitation*. Romanian Journal of Human-Computer Interaction, 2016. **9**(2).
57. Marina A. Cidota, S.G.L., Paul Dezentje, Paulina J. M. Bank, Heide K. Lukosch, Rory M. S. Clifford, *Serious Gaming in Augmented Reality using HMDs for Assessment of Upper Extremity Motor Dysfunctions*. i-com, 2016. **15**(2): p. 155-169.
58. Chhor, J., Gong, Y., and Rau, P.L., *Breakout: Design and Evaluation of a Serious Game for Health Employing Intel RealSense*, in *International Conference on Cross-Cultural Design*, Springer, Editor. 2017. p. 531-545.
59. Jayan Mistry, B.I., *An Approach to Sign Language Translation using the Intel RealSense Camera*, in *10th Computer Science and Electronic Engineering Conference*. 2018, IEEE.
60. Clark, R.A., Y.H. Pua, K. Fortin, C. Ritchie, K.E. Webster, L. Denehy, and A.L. Bryant, *Validity of the Microsoft Kinect for assessment of postural control*. Gait Posture, 2012. **36**(3): p. 372-7.
61. Galen, S.S., V. Pardo, D. Wyatt, A. Diamond, V. Brodith, and A. Pavlov, *Validity of an Interactive Functional Reach Test*. Games Health J, 2015. **4**(4): p. 278-84.

62. Shih, M.C., R.Y. Wang, S.J. Cheng, and Y.R. Yang, *Effects of a balance-based exergaming intervention using the Kinect sensor on posture stability in individuals with Parkinson's disease: a single-blinded randomized controlled trial*. J Neuroeng Rehabil, 2016. **13**(1): p. 78.
63. Dehbandi, B., A. Barachant, A.H. Smeragliuolo, J.D. Long, S.J. Bumanlag, V. He, A. Lampe, and D. Putrino, *Using data from the Microsoft Kinect 2 to determine postural stability in healthy subjects: A feasibility trial*. PLoS One, 2017. **12**(2): p. e0170890.
64. Eltoukhy, M.A., C. Kuenze, J. Oh, and J.F. Signorile, *Validation of Static and Dynamic Balance Assessment Using Microsoft Kinect for Young and Elderly Populations*. IEEE J Biomed Health Inform, 2018. **22**(1): p. 147-153.
65. Galna, B., G. Barry, D. Jackson, D. Mhiripiri, P. Olivier, and L. Rochester, *Accuracy of the Microsoft Kinect sensor for measuring movement in people with Parkinson's disease*. Gait Posture, 2014. **39**(4): p. 1062-8.
66. Gabel, M., R. Gilad-Bachrach, E. Renshaw, and A. Schuster, *Full body gait analysis with Kinect*. Conf Proc IEEE Eng Med Biol Soc, 2012. **2012**: p. 1964-7.
67. Stone, E.E. and M. Skubic, *Capturing habitual, in-home gait parameter trends using an inexpensive depth camera*. Conf Proc IEEE Eng Med Biol Soc, 2012. **2012**: p. 5106-9.
68. Saposnik, G., M. Levin, and G. Outcome Research Canada Working, *Virtual reality in stroke rehabilitation: a meta-analysis and implications for clinicians*. Stroke, 2011. **42**(5): p. 1380-6.
69. Bo, A.P.L., M. Hayashibe, and P. Poignet, *Joint angle estimation in rehabilitation with inertial sensors and its integration with Kinect*. Conference proceedings : Annual International Conference of the IEEE Engineering in Medicine and Biology Society IEEE Engineering in Medicine and Biology Society Annual Conference, 2011. **2011**: p. 3479-83.
70. Da Gama, A., P. Fallavollita, V. Teichrieb, and N. Navab, *Motor Rehabilitation Using Kinect: A Systematic Review*. Games for health journal, 2015. **4**(2): p. 123-35.
71. Pastor, I., H.A. Hayes, and S.J.M. Bamberg, *A feasibility study of an upper limb rehabilitation system using Kinect and computer games*. Conference proceedings : Annual International Conference of the IEEE Engineering in Medicine and Biology Society IEEE Engineering in Medicine and Biology Society Annual Conference, 2012. **2012**: p. 1286-9.
72. Clark, R.A., S. Vernon, B.F. Mentiplay, K.J. Miller, J.L. McGinley, Y.H. Pua, K. Paterson, and K.J. Bower, *Instrumenting gait assessment using the Kinect in people living with stroke: reliability and association with balance tests*. J Neuroeng Rehabil, 2015. **12**: p. 15.
73. Dolatabadi, E., B. Taati, and A. Mihailidis, *Automated classification of pathological gait after stroke using ubiquitous sensing technology*. Conf Proc IEEE Eng Med Biol Soc, 2016. **2016**: p. 6150-6153.
74. Leightley, D. and M.H. Yap, *Digital Analysis of Sit-to-Stand in Masters Athletes, Healthy Old People, and Young Adults Using a Depth Sensor*. Healthcare (Basel), 2018. **6**(1).
75. Eftychios Protopapadakis, A.G., Anastasios Doulamis, Nikos Grammalidis, *FOLK DANCE PATTERN RECOGNITION OVER DEPTH IMAGES ACQUIRED VIA KINECT SENSOR*, in *The International Archives of the Photogrammetry*. 2017: Nafplio, Greece.

76. Plantard, P., E. Auvinet, A.S. Pierres, and F. Multon, *Pose estimation with a Kinect for ergonomic studies: evaluation of the accuracy using a virtual mannequin*. Sensors (Basel), 2015. **15**(1): p. 1785-803.
77. Plantard, P., H.P.H. Shum, A.S. Le Pierres, and F. Multon, *Validation of an ergonomic assessment method using Kinect data in real workplace conditions*. Appl Ergon, 2017. **65**: p. 562-569.
78. McAtamney, L. and E. Nigel Corlett, *RULA: a survey method for the investigation of work-related upper limb disorders*. Appl Ergon, 1993. **24**(2): p. 91-9.
79. Robertson, M., B.C. Amick, 3rd, K. DeRango, T. Rooney, L. Bazzani, R. Harrist, and A. Moore, *The effects of an office ergonomics training and chair intervention on worker knowledge, behavior and musculoskeletal risk*. Appl Ergon, 2009. **40**(1): p. 124-35.
80. Dockrell, S., E. O'Grady, K. Bennett, C. Mullarkey, R. Mc Connell, R. Ruddy, S. Twomey, and C. Flannery, *An investigation of the reliability of Rapid Upper Limb Assessment (RULA) as a method of assessment of children's computing posture*. Appl Ergon, 2012. **43**(3): p. 632-6.
81. Manghisi, V.M., A.E. Uva, M. Fiorentino, V. Bevilacqua, G.F. Trotta, and G. Monno, *Real time RULA assessment using Kinect v2 sensor*. Appl Ergon, 2017. **65**: p. 481-491.
82. Lundqvist, C., A. Siosteen, C. Blomstrand, B. Lind, and M. Sullivan, *Spinal cord injuries. Clinical, functional, and emotional status*. Spine (Phila Pa 1976), 1991. **16**(1): p. 78-83.
83. Gerhart, K.A., E. Bergstrom, S.W. Charlifue, R.R. Menter, and G.G. Whiteneck, *Long-term spinal cord injury: functional changes over time*. Arch Phys Med Rehabil, 1993. **74**(10): p. 1030-4.
84. Lundqvist, C., A. Siosteen, C. Blomstrand, B. Lind, and M. Sullivan, *Spinal cord injuries. Clinical, functional, and emotional status*. Spine, 1991. **16**(1): p. 78-83.
85. Gerhart, K.A., E. Bergstrom, S.W. Charlifue, R.R. Menter, and G.G. Whiteneck, *Long-term spinal cord injury: functional changes over time*. Archives of physical medicine and rehabilitation, 1993. **74**(10): p. 1030-4.
86. Rintala, D.H., P.G. Loubser, J. Castro, K.A. Hart, and M.J. Fuhrer, *Chronic pain in a community-based sample of men with spinal cord injury: prevalence, severity, and relationship with impairment, disability, handicap, and subjective well-being*. Arch Phys Med Rehabil, 1998. **79**(6): p. 604-14.
87. Mortenson, W.B., W.C. Miller, C.L. Backman, and J.L. Oliffe, *Association between mobility, participation, and wheelchair-related factors in long-term care residents who use wheelchairs as their primary means of mobility*. J Am Geriatr Soc, 2012. **60**(7): p. 1310-5.
88. Dalyan, M., D.D. Cardenas, and B. Gerard, *Upper extremity pain after spinal cord injury*. Spinal Cord, 1999. **37**(3): p. 191-5.
89. Alm, M., H. Saraste, and C. Norrbrink, *Shoulder pain in persons with thoracic spinal cord injury: prevalence and characteristics*. J Rehabil Med, 2008. **40**(4): p. 277-83.
90. Brose, S.W., M.L. Boninger, B. Fullerton, T. McCann, J.L. Collinger, B.G. Impink, and T.A. Dyson-Hudson, *Shoulder ultrasound abnormalities, physical examination findings, and pain in manual wheelchair users with spinal cord injury*. Arch Phys Med Rehabil, 2008. **89**(11): p. 2086-93.

91. Koontz, A.M., Y.S. Lin, P. Kankipati, M.L. Boninger, and R.A. Cooper, *Development of custom measurement system for biomechanical evaluation of independent wheelchair transfers*. J Rehabil Res Dev, 2011. **48**(8): p. 1015-28.
92. Gagnon, D., S. Nadeau, P. Desjardins, and L. Noreau, *Biomechanical assessment of sitting pivot transfer tasks using a newly developed instrumented transfer system among long-term wheelchair users*. J Biomech, 2008. **41**(5): p. 1104-10.
93. Gagnon, D., S. Nadeau, L. Noreau, J.J. Eng, and D. Gravel, *Trunk and upper extremity kinematics during sitting pivot transfers performed by individuals with spinal cord injury*. Clin Biomech (Bristol, Avon), 2008. **23**(3): p. 279-90.
94. Microsoft. Available from: <https://msdn.microsoft.com/en-us/library/microsoft.kinect.jointtype.aspx>.
95. Xu, X. and R.W. McGorry, *The validity of the first and second generation Microsoft Kinect for identifying joint center locations during static postures*. Appl Ergon, 2015. **49**: p. 47-54.
96. Lin Wei, T.B., Hyun Ka, Alicia M. Koontz, *Evaluating Wheelchair Transfer Technique by Microsoft Kinect*, in *International Seating Symposium*. 2017: Nashville, TN.
97. Halilaj, E., A. Rajagopal, M. Fiterau, J.L. Hicks, T.J. Hastie, and S.L. Delp, *Machine learning in human movement biomechanics: Best practices, common pitfalls, and new opportunities*. J Biomech, 2018. **81**: p. 1-11.
98. Barbareschi, G., C. Holloway, N. Bianchi-Berthouze, S. Sonenblum, and S. Sprigle, *Use of a Low-Cost, Chest-Mounted Accelerometer to Evaluate Transfer Skills of Wheelchair Users During Everyday Activities: Observational Study*. JMIR Rehabil Assist Technol, 2018. **5**(2): p. e11748.
99. van den Berg-Emons, R.J., A.A. L'Ortye, L.M. Buffart, C. Nieuwenhuijsen, C.F. Nooijen, M.P. Bergen, H.J. Stam, and J.B. Bussmann, *Validation of the Physical Activity Scale for individuals with physical disabilities*. Arch Phys Med Rehabil, 2011. **92**(6): p. 923-8.
100. Wei, L.C., C. S.; Koontz, A. M., *Automating the Clinical Assessment of Independent Wheelchair Sitting Pivot Transfer Techniques*. Topics in Spinal Cord Injury Rehabilitation (under review), 2021.
101. Desroches, G., D. Gagnon, S. Nadeau, and M.R. Popovic, *Effects of sensorimotor trunk impairments on trunk and upper limb joint kinematics and kinetics during sitting pivot transfers in individuals with a spinal cord injury*. Clin Biomech (Bristol, Avon), 2013. **28**(1): p. 1-9.
102. Worobey, L.A., S.K. Rigot, M.L. Boninger, R. Huzinec, J.H. Sung, K. DiGiovine, and L.A. Rice, *Concurrent Validity and Reliability of the Transfer Assessment Instrument Questionnaire as a Self-Assessment Measure*. Arch Rehabil Res Clin Transl, 2020. **2**(4): p. 100088.
103. *Spinal Cord Injury Facts and Figures at a Glance*. J Spinal Cord Med, 2014. **37**(3): p. 355-6.
104. Lin Wei, H.W.K., Chung-Ting Tsai, and Alicia M. Koontz, *Can the Kinect detect differences between proper and improper wheelchair transfer techniques?*, in *Proceedings of the Rehabilitation Engineering and Assistive Technology Society of North America Conference*. 2016: Washington, D.C.

105. Hwang, S., C.Y. Tsai, and A.M. Koontz, *Feasibility study of using a Microsoft Kinect for virtual coaching of wheelchair transfer techniques*. Biomed Tech (Berl), 2017. **62**(3): p. 307-313.
106. Nicholas Sallinger, L.W., Sarah Bass, Hyun Ka, Alicia Koontz, *Real Time Transfer Technique Assessment Using the Kinect2 Sensor*, in *Proceedings of the Rehabilitation Engineering and Assistive Technology Society of North America Conference*. 2018: Arlington, VA.
107. Intel®. *Intel® RealSense™ Depth Camera D435*. 2021; Available from: <https://www.intelrealsense.com/depth-camera-d435/>.
108. Jamhoury, L. *Understanding Kinect V2 Joints and Coordinate System*. 2018; Available from: <https://medium.com/@lisajamhoury/understanding-kinect-v2-joints-and-coordinate-system-4f4b90b9df16>.
109. *Nuitrack: Overview*. 2019; Available from: (https://download.3divi.com/Nuitrack/doc/Overview_page.html).
110. Wu, G., F.C. van der Helm, H.E. Veeger, M. Makhsous, P. Van Roy, C. Anglin, J. Nagels, A.R. Karduna, K. McQuade, X. Wang, F.W. Werner, B. Buchholz, and B. International Society of, *ISB recommendation on definitions of joint coordinate systems of various joints for the reporting of human joint motion--Part II: shoulder, elbow, wrist and hand*. J Biomech, 2005. **38**(5): p. 981-992.
111. Koo, T.K. and M.Y. Li, *A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research*. J Chiropr Med, 2016. **15**(2): p. 155-63.
112. Lopez, N., E. Perez, E. Tello, A. Rodrigo, and M.E. Valentinuzzi, *Statistical Validation for Clinical Measures: Repeatability and Agreement of Kinect-Based Software*. Biomed Res Int, 2018. **2018**: p. 6710595.
113. Lin Wei, C.-Y.T., and Alicia M. Koontz, *The Relationship between Joint Ranges of Motion and Joint Kinetics during Sitting Pivot Wheelchair Transfers*, in *Proceedings of the Rehabilitation Engineering and Assistive Technology Society of North America Conference*. 2018.
114. Portney, L.G., *Foundations of clinical research : applications to evidence-based practice*. 2020, F.A. Davis, Philadelphia. p. 1 online resource.
115. Harkel, T.C.T., S. Vinayahalingam, K. Ingels, S.J. Berge, T.J.J. Maal, and C.M. Speksnijder, *Reliability and Agreement of 3D Anthropometric Measurements in Facial Palsy Patients Using a Low-Cost 4D Imaging System*. IEEE Trans Neural Syst Rehabil Eng, 2020. **28**(8): p. 1817-1824.
116. R. T. Labuguen, S.B.N., T. Kogami, W. E. M. Ingco and T. Shibata, *Performance Evaluation of Markerless 3D Skeleton Pose Estimates with Pop Dance Motion Sequence*, in *Joint 9th International Conference on Informatics, Electronics & Vision (ICIEV) and 2020 4th International Conference on Imaging, Vision & Pattern Recognition (icIVPR)*. 2020: Kitakyushu, Japan. p. 1-7.
117. Intel®. *Skeleton Tracking SDK for Intel® RealSense™ Depth Cameras*. 2021; Available from: <https://www.intelrealsense.com/skeleton-tracking/>.
118. Koontz, A.M., C.Y. Tsai, N.S. Hogaboom, and M.L. Boninger, *Transfer component skill deficit rates among Veterans who use wheelchairs*. J Rehabil Res Dev, 2016. **53**(2): p. 279-94.
119. *The National Institute for Occupational Safety and Health (NIOSH)*. 2018 May 25, 2018; Available from: <https://www.cdc.gov/niosh/>.

