Enhancing the Measurement of Clinical Outcomes Using Microsoft Kinect

Philip Breedon and Francesco Luke Siena
Design for Health and Wellbeing Research Grp
Nottingham Trent University
Nottingham, UK
philip.breedon@ntu.ac.uk; luke.siena@ntu.ac.uk

Bill Byrom and Willie Muehlhausen
Product Innovation
ICON Clinical Research
Marlow, UK
bill.byrom@iconplc.com; willie.muehlhausen@iconplc.com

Abstract— **There is a growing body of applications leveraging Microsoft Kinect and the associated Windows Software Development Kit in health and wellness. In particular, this platform has been valuable in developing interactive solutions for rehabilitation including creating more engaging exercise regimens and ensuring that exercises are performed correctly for optimal outcomes.**

Clinical trials rely upon robust and validated methodologies to measure health status and to detect treatment-related changes over time to enable the efficacy and safety of new drug treatments to be assessed and measured. In many therapeutic areas, traditional outcome measures rely on subjective investigator and patient ratings. Subjective ratings are not always sensitive to detecting small improvements, are subject to inter- and intra-rater variability and limited in their ability to record detailed or subtle aspects of movement and mobility. For these reasons, objective measurements may provide greater sensitivity to detect treatment-related changes where they exist.

In this review paper, we explore the use of the Kinect platform to develop low-cost approaches to objectively measure aspects of movement. We consider published applications that measure aspects of gait and balance, upper extremity movement, chest wall motion and facial analysis. In each case, we explore the utility of the approach for clinical trials, and the precision and accuracy of estimates derived from the Kinect output.

We conclude that the use of games platforms such as Microsoft Kinect to measure clinical outcomes offer a versatile, easy to use and low-cost approach that may add significant value and utility to clinical drug development, in particular in replacing conventional subjective measures and providing richer information about movement than previously possible in large scale clinical trials, especially in the measurement of gross spatial movements. Regulatory acceptance of clinical outcomes collected in this way will be subject to comprehensive assessment of validity and clinical relevance, and this will require good quality peer-reviewed publications of scientific evidence.

Keywords—Microsoft Kinect; Motion Capture; Clinician Reported Outcomes (ClinRO); Performance Outcomes (PerfO); Clinical Trials.

I. INTRODUCTION

Over the past two decades, clinical drug development programs have become increasingly complex, involved more clinical trials with higher numbers of patients, each including a greater volume and variety of assessments and clinical procedures and operating with a larger global footprint. This has been associated with spiraling costs. Developing a new drug can take around 12 years from patent application at a cost of $1.4 billion in direct costs ($2.6 billion capitalized costs) [1].

Clinical trials rely upon robust and validated methodologies to measure health status and to detect treatment-related changes over time to enable the efficacy and safety of new drug treatments to be assessed and measured. In many therapeutic areas, the outcomes measures used rely on subjective ratings by the investigators at each study research site. Subjective ratings may include measures of symptoms through direct observation of the patient, for example depression rating using the Hamilton Depression Rating Scale [2], or aspects of Parkinson's Disease symptoms using the Unified Parkinson's Disease Rating Scale (UPDRS) [3]; or subjective assessment of performance, balance, movement or mobility based on observation of the patient conducting a specified movement or activity. Examples include the Fugl-Meyer Assessment (FMA) which measures aspects of arm, hand and shoulder movement on 3-point verbal response scales [4], the Tinetti Balance Assessment measuring aspects of balance and walking movement on 2- and 3-point verbal response scales [5], the Berg Functional Balance Scale measuring aspects of balance and movement on 5-point verbal response scales [6], the Dynamic Gait Index measuring aspects of gait during a simple walking test on 4-point verbal response scales [7], and many others. Subjective ratings, such as the assessment of a patient using a verbal response or numeric rating scales, are not always sensitive to detecting small improvements. Different raters may vary in their interpretation of the scale requirements which may add variability into the outcomes measured (inter-rater variability). In addition, subjective scales may be prone to conscious or unconscious bias, especially when changes over time are expected or treatment is unblinded through an open design or revealed though the pharmacologic profile of the treatment (see [8] for example). It is sometimes difficult to implement measurements that are sensitive enough to detect treatment-related changes and able to conclusively show treatment effects when they exist. In addition, using investigator observation it is less likely that detailed or subtle aspects of movement and mobility can be recorded. For these reasons, accurate objective measurements, where possible, are preferred over subjective assessments.

CPS
Conference Publishing Services

Some solutions already exist that measure aspects of movement objectively and enable greater detail to be captured and explored. GAITRite™ (CIR systems Inc., Franklin, NJ, USA), for example, uses a walking circuit of pressure pads to learn more about gait and balance. While effective, this solution is relatively expensive which may limit its use in large scale clinical studies. In addition, a further solution would be required to look at other aspects of movement such as upper body and limb movements.

In this paper, we review the utility of different motion-based gaming platforms and depth cameras to enable low-cost approaches to objectively measure aspects of movement for patients in clinical trials. We consider published applications developed using the Microsoft Kinect® (Microsoft Corp., Redmond, Washington, USA) platform to measure: (i) aspects of gait and balance, with specific reference to stroke, multiple sclerosis and Parkinson's disease; (ii) upper extremity movement, with reference to stroke, multiple sclerosis and adhesive capsulitis; (iii) chest wall motion analysis for respiratory conditions; and (iv) facial analysis for facial paralysis in stoke and Bell's Palsy. In each case, we explore the utility of the approach for clinical trials, in particular considering the practical aspects of test conduct, the clinical relevance of outcomes measured, and the precision and accuracy of estimates derived from the Kinect output.

II. Motion-Capture Platforms For Use in Healthcare Applications

There is a growing body of applications utilizing motion capture technology to study or encourage movement in wellness, healthcare and clinical research. Much of this work has been accomplished in the area of rehabilitation to provide an engaging environment through which to conduct a regular exercise regimen to enable patient feedback and correction to ensure that exercises are being performed correctly for optimal benefit, and to enable remote assessment and adjustment of exercise regimens between clinic visits. While wearable technologies such as accelerometers and gyroscopes, and other "non-seeing" sensor systems such as pressure pads (e.g. GAITRite) provide useful instrumentation and analysis of specific performance tests, motion capture systems offer the potential to study a comprehensive digital representation of movement, potentially providing a richness of information beyond the capability of other systems.

We focus our review on the use of these optoelectronics systems, systems that are able to spatially track movement. These systems typically use a variety of methods including using detectable markers attached to the body at key joints that the optical system can track, and using cameras detecting shapes, contrast or depth to interpret position and motion [9].

In this section, we review the technical capabilities of the Microsoft Kinect gaming platform, the Intel® RealSense™ (Intel Corp., Santa Clara, CA, USA) camera range, and Leap Motion (Leap Motion Inc., San Francisco, CA, USA), CREATIVE® SENZ3D (Creative Technology Ltd., Jurong East, Singapore) and Xtion Pro Live (ASUSTek Computer Inc., Taipei, Taiwan) solutions with regard to development of healthcare applications. We focus on these as inexpensive

hardware solutions supported by versatile Software Development Kits (SDKs). These potentially enable development of health measurement solutions that could be practical for implementation in large scale multicenter clinical trials across a variety of clinical settings and not confined to movement labs.

A. Microsoft Kinect

The Kinect sensor was originally developed for the Xbox video game console to provide immersive motion-based gaming experiences, and comprises an RGB camera, a depth sensor and a multi-array microphone. The depth sensor enables additional information such as the distance of surfaces of a 3D object to be determined. The associated SDK contains skeletal tracking and facial analysis components, and these enable the position and orientation of body joints and facial expression to be determined. The Microsoft Kinect is available in two versions, 1.0 and 2.0. A comparison of the technical specifications of both Kinects can be found in Table I.

TABLE I.   Comparison of Kinect 1.0 and 2.0 Hardware Capabilities

Capability	Kinect model	
	Kinect 1.0	Kinect 2.0
RGB Camera (Pixel)	1280×1024 or 640×480	1920×1080
Depth Camera (Pixel)	640×480	512×424
Sampling Rate (FPS / Hz)	30	30
Skeletal Joint Definition Points	20	26
Full Skeleton Tracking	2	6
Depth Range (m)	30 FPS	30 FPS
Horizontal Field Of Vision (Degrees)	57	70
Vertical Field Of Vision (Degrees)	43	60
Tilt Motor	Yes	No
Audio Stream	4-Mic-Array	4-Mic-Array
SDK 1.7 Compatibility	Yes	No
SDK 1.8 Compatibility	Yes	No
SDK 2.0 Compatibility	No	Yes
Connectivity (USB)	2.0 or 3.0	3.0
Approx. price (USD) [a]	45	190

[a.] Prices as of May 2016

The Microsoft Kinects are low-cost camera systems which capture real time depth measurements through triangulation alongside RGB and IR imaging. Pagliari and Pinto [10] discuss that one of the key drawbacks of the Kinect 1.0 sensor is the low geometric quality of the delivered data, noise and low repeatability. They also identify that the RGB camera is of poor quality in comparison to a standard webcam. Additionally the depth data registered by the Kinect 1.0 is of poor quality due to the structured light approach not always being robust enough to provide a high level of completeness of a framed scene [10]. When comparing the two Kinect platforms there is a significant

increase in performance quality, yet the same number of sensors are used in the construction of the device. The improved resolution and field of vision of the Kinect 2.0 may be important in improving accuracy and increasing the utilizable area that can be used for performance test measurements in patients.

The Microsoft Kinects do have their limitations with regards to the capture speeds which may be realized in the measurement of quick and sharp movements. This is due primarily to the modest sampling rate of 30 frames per second (30 Hz). Some authors have suggested that this can be overcome to a certain degree using spline interpolation algorithms which can resample collected data to increase the accuracy. For example, Clark et al [11] use this approach to reprocess Kinect data to estimate a 100 Hz sampling rate. There are a number of alternative hardware solutions which have higher capture rates to enable the measurement of quicker movements, but these are typically found in machine vision systems and motion capture devices which are associated with a significantly greater price which may limit their large-scale application in clinical trials. The novelty of the Microsoft Kinect sensors is that they provide a low-cost solution for the real time measurement of image data which allows the tracking body and facial movements for analysis and manipulation.

TABLE II.   COMPARISON OF KINECT 1.0 AND KINECT 2.0 FUNCTIONS

Function	Kinect model	
	Kinect 1.0	Kinect 2.0
Face Tracking	Yes	Yes
Expression Recognition	No (Possible with Additional Algorithms)	Yes
Bone Orientations	No	Yes
Body Joint Forces	No	Yes
Hand Tracking	No (Possible with Additional Algorithms)	Yes
Muscle Simulation	No	Yes
Heart Rate Measurement	No	Yes

As shown in Table II, the Kinect 2.0 can perform a number of tasks that the Kinect 1.0 cannot and this is due to the SDK's utilized and how the sensor measures and captures the information scanned. Pagliari and Pinto [10] highlight that the depth measurement system utilized in the Kinect 2.0 utilizes a new measurement principle in order to perform more precise skeleton tracking and gesture recognition.

B. Intel® RealSense™ SR300

The Microsoft Kinect platforms are the most commonly used low-cost gaming camera systems currently utilized to develop a wide variety of applications, especially within the health care arena. However, there are a number of other 3D camera alternatives which offer similar or improved functionality dependent on the uses. One alternative is the Intel® RealSense™ SR300 which provides users the opportunity to work with dynamic background segmentation, 3D scanning, facial recognition and hand gesture recognition.

A comparison of the RealSense™ SR300 and Kinect 2.0 can be found in Table III.

TABLE III.   COMPARISON OF INTEL REALSENSE SR300 AND KINECT 2.0

Capability/Function	Model	
	Intel RealSense SR300	Kinect 2.0
RGB Camera (Pixel)	1080p at 30 FPS, 720p at 60 FPS	1920×1080 at 30 FPS
Depth Camera (Pixel)	Up to 640 x 480 at 60 FPS (Fast VGA, VGA), HVGA at 110 FPS	512×424 at 30 FPS
Skeletal Joint Definition Points	22	26
Face Tracking and Recognition	Yes	Yes
Expression Recognition	Yes	Yes
Gesture Recognition	Yes	Yes
Hand Tracking	Yes	Yes
Audio Stream	Dual Array Microphones	4-Mic-Array
Connectivity (USB)	3.0	3.0
Approx. price (USD) [b]	130	190

b. Prices as of May 2016

The Intel camera offers greater resolution and sampling rate in comparison to Kinect 2.0, which may offer advantages when tracking fine or fast movements. One of the novelties of the RealSense 3D camera range is its versatility for integration into a variety of platforms, yet at the same time it remains affordable. Intel have developed a number of camera systems which can be integrated into a variety of platforms including PCs, external camera systems, smartphone and tablet kits and a robotics development kit which gives makers an interface board to control cross-platform and open source systems.

C. Leap Motion

Leap Motion is a user controlled hand gesture control device which is constructed of two monochromatic IR cameras and three infrared LEDs. Leap Motion tracks infrared light with a wavelength of 850 nanometers and with sub-millimeter levels of control, providing a user with an accurate gesture and position tracking system (Table IV). Leap Motion is a novel touch-free sensor which is used in realistic stereo 3D interaction systems where direct selection of stereoscopically-displayed objects is possible [12]. Hondori and Khademi [9] indicate that Leap Motion can detect and track both fingers and tools (such as pen) providing developers with hand and digit information such as fingertip position, hand velocity, and hand/finger direction.

D. CREATIVE® SENZ3D

CREATIVE® SENZ3D is an interactive depth and gesture control camera which detects hand gestures and head movement through a 720p HD webcam and dual-array microphone. The SENZ3D camera is used alongside FastAccess 3D facial recognition software which allows the user to capture face tracking data to such a degree of detail that facial contours can be captured [13]. Combining movement

and gesture, facial recognition and voice interaction the SENZ3D camera provides an interactive gaming platform which has utility in medicine including rehabilitation systems. However, the limitation of operating only a single camera or voice enabled application at a time reduces its scope of use. A full technical specification of the SENZ3D camera can be found in Table IV. Although now discontinued, the SENZ3D provided developers a low-cost solution to create motion-sensing applications and games. The replacement CREATIVE® LIVE! CAM CONNECT HD retains similar smart face tracking capabilities.

TABLE IV. COMPARISON OF LEAP MOTION, CREATIVE SENZ3D AND XTION PRO LIVE

Capability/ Function	Model		
	Leap Motion	SENZ3D	Pro Live
RGB Camera (Pixel)	N/A	1280×720	1280×1024 (SGXA)
Depth Camera (Pixel)	N/A	320×240 QVGA	640 × 480 VGA at 30 FPS, 320 × 240 QVGA at 60 FPS
Sampling Rate/Frame Rate (FPS)	30	30	30/60
Depth Sensing Range	0.025 to 0.6m	0.15 to 1m	0.8 to 3.5m
Face Tracking and Recognition	No	Yes	No
Expression Recognition	No	No	No
Body Tracking	No	No	Yes
Hand Tracking	Yes	Yes	Yes
Gesture Control	Yes	Yes	Yes
Connectivity (USB)	2.0 or 3.0	2.0 or 3.0	2.0 or 3.0
Approx. price (USD) c	62	Discontinued	Discontinued

c. Prices as of May 2016

E. Xtion Pro Live

Xtion uses an infrared sensor and adaptive depth detection utilizing the OpenNI NITE development middleware SDK to track precise body movements therefore providing a platform which can detect and track the whole body including detailed tracking of hands and gestures [9]. This middleware SDK allows developers to create motion-sensing applications using RGB PC motion sensing software (Table IV). This solution has since been discontinued.

In conclusion, the use of 3D camera systems and sensors has great potential to enable inventive application across a wide range of industries, especially within health care and clinical research. Hardware specification improvements may still be required when considering accurate tracking of fine or rapid movements, and therefore the sampling rates associated with the capture of this data may need to improve in future versions. However, this has not limited the application of this technology in clinical and home health care applications, especially within the rehabilitation sector where platforms such as Neuroforma

(Titanis Sp. z o.o., Warsaw, Poland) [14], JINTRONIX (Jintronix, Montreal, Canada) [15], and Face To Face [16] have recently been developed; therefore suggesting that 3D cameras and sensors will continue to be used in clinical platforms for the foreseeable future.

III. MEASURING OBJECTIVE HEALTH OUTCOMES USING MICROSOFT KINECT

The Kinect platform has been valuable in developing interactive solutions for rehabilitation including novel ways of engaging patients in regular exercise regimens and solutions to ensure that exercises are performed correctly for optimal outcomes. Rehabilitation solutions using Kinect have been reported as providing therapeutic benefits in the areas of Stroke [17], Parkinson's Disease [18], Multiple Sclerosis [19] and Cerebral Palsy [20], for example. Through measurement within an immersive gaming environment, Kinect has also been shown to enable assessments of elements of cognition and range of motion [21].

The use of games platforms such as Microsoft Kinect to measure clinical outcomes offers a versatile, easy to use and low cost approach that may add significant value and utility to clinical drug development, in particular in replacing conventional subjective measures. Regulatory acceptance of clinical trial data will be subject to comprehensive assessment of the validity of the data collected in this way, and this will require good quality peer-reviewed publications of validation evidence. As seen in the area of rehabilitation, serious gaming and motion-based games platforms in particular have the potential to enable researchers to understand more, and characterize more precisely, the effects of treatment in a wide range of disease indications.

In this section, we review a number of published examples of solutions for objective measurement of movement using Microsoft Kinect focusing on the areas of gait and balance [11, 22-23], upper extremity movement [24-26], chest wall motion analysis [27] and facial analysis [16] (Table V). Our aim is not to provide a comprehensive review, but to illustrate the growing application and potential of the Kinect platform to enable objective clinical endpoint assessment.

A. Gait and Balance

The ability to track the position of the major body joints is useful in tracking gross spatial movements and this information can be used to derive descriptive analytics useful in summarizing health outcomes. We review three studies using the Kinect depth camera and windows SDK to track 20 body joints (Figure 1) to derive gait and balance outcomes measures during short performance tests performed by the patient.

Clinical assessments of gait often rely upon functional fitness tests such as timed corridor walking tests or treadmill tests for example the six-minute walking test (6MWT); pressure pad and balance plate technology; and/or subjective tests of aspects of gait such as the Dynamic Gait Index as described earlier. However, all these methods lack the richness of information that can be derived by full body motion analysis.
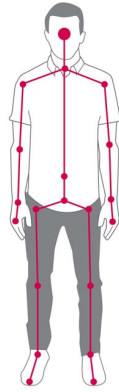
Fig. 1. Body joint skeletal tracking using Microsoft Kinect.

Clark et al. [11] reported the use of the Kinect platform to derive gait parameters in comparison to a standard 10 meter walking test. Their study in 30 stroke patients used a 6 meter straight walkway with the Kinect camera placed at the end of the route. Patients were instructed to walk towards the camera stopping 0.5 m in front of it, and this enabled the recording of a minimum of a full gait cycle (stride) per limb. The authors used spline interpolation to estimate a 100 Hz sampling frequency from the 30 Hz data provided by the Microsoft Kinect system, which they reported useful in overcoming sampling irregularities. Gait event times for ground contact and toe-off were based on the velocity of the ankle joint center. These enabled estimation of outcome measures including step length and foot swing velocity for the affected and unaffected limbs, and anterior displacement of the shoulder center enabled mean and peak gait velocity estimation. Operationally, using the Kinect-based assessment system offered advantages to other instrumented approaches such as GAITRite, as it required minimal setup and minimal interference to the patient – being conducted in normal clothes (not requiring shorts or bare feet). The authors also concluded that the richness of information enabled a broad range of outcomes measures to be derived, some of which are not possible with other instrumented approaches. For example, while pressure mat systems can provide some spatiotemporal measures of gait such as stride length, they cannot provide other potentially important measures such as foot swing velocity or variability in trunk motion. Limitations of their approach arose from the field of vision of the Kinect camera.

Accurate assessment of asymmetry and gait variability requires the recording of multiple gait cycles, which was not possible in a corridor-based test where accurate joint tracking requires subjects to remain in the field of vision of the depth camera. In addition, there was some suggestion that, due to the setup of their test, patients may decelerate as they approach the camera before completing a full gait cycle. Despite this their study showed good concordance of Kinect-derived outcome measures compared to gold standard approaches, and robust test-retest reliability.

A second study developed a short maximum speed walking test to assess gait parameters in patients suffering from multiple sclerosis [22]. Patients were requested to walk as fast as they could towards a Kinect camera positioned on a pole, starting from a point 2 m outside the camera detection region. It was assumed that this distance was sufficient to achieve maximum walking speed. Using only the hip-center joint, the authors derived speed, speed deviation, and 3D left/right and up/down deviation measures that were able to distinguish between MS patients and healthy volunteers. Importantly, this study identified limitations in the Kinect SDK's in-built error detection system (SDK Recognition Quality) which works to identify data artefacts affecting the ability to track the position of each joint. The authors developed a custom error correction technique that was able to significantly improve accuracy in their results compared to the Kinect SDK which tended to fail to accurately identify data artefacts and in these cases falsely report joint detections.

A study using a battery of functional performance tests in Parkinson's disease patients showed promise in the use of the Kinect depth camera in measuring some but not all clinical outcomes measures [23]. The tests included quiet standing, multidirectional reaching, stepping and walking on the spot, and a number of functional test from the UPDRS: hand clasping, finger tapping, foot/leg agility, chair rising and hand pronation. In comparison to estimates obtained using a Vicon 3D motion analysis system (Vicon Motion Systems Ltd., Oxford, UK), the authors reported that the Kinect system was able to accurately measure the timing (ICCs: 0.940-0.999) and gross spatial characteristics of clinically relevant movements but not provide the same spatial accuracy for smaller movements, such as hand clasping (ICC = 0.009) or toe tapping (ICC = 0.038).

B. Upper Extremity Movement

Two studies examined the use of Microsoft Kinect to make objective measurements of shoulder range of motion [24, 25]. The first, in stroke patients with chronic hemiparesis [24], compared Kinect estimates to those obtained using a gold standard motion capture tool and the Fugl-Meyer clinician reported outcome scale. The second [25], explored range of motion angles measured from shoulder, trunk and arm vectors using Kinect to those obtained by manual goniometer readings by qualified clinical readers. Both studies reported strong concordance with the comparator methodologies.

Olesh et al. [24] reported that reliable estimates were obtained using Kinect when each movement was repeated three times. While traditional subjective clinician ratings typically request only a single repetition of each movement, this may be primarily due to practical time constraints. The use of a low-cost, portable solution to make objective measures, such as Kinect, may enable more accurate measures and time saving compared to goniometer measurement. Lee et al. [25] identified that this approach, in addition to providing accurate estimates of shoulder abduction, flexion and rotation angles, was able to provide clinical differentiation between patients suffering from adhesive capsulitis and healthy volunteers.

In their study of arm and finger movement exercises in MS patients, Kontschieder et al. [26] used a random forest machine learning approach to develop algorithms for analyzing four arm and finger movement tests: finger to nose, finger to finger, drawing squares and Truncal Ataxia assessment in which both arms were held outstretched for 5 seconds. These were developed using a large (training and analysis) dataset of 1,041 depth videos from MS patients and healthy volunteers. From this work, they concluded the potential of this approach for depth-camera supported clinical assessments for a range of conditions.
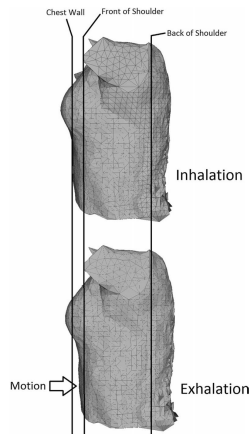
C. Chest Wall Motion Analysis

Fig. 2. Chest wall visualisation and change in volume analysis using Microsoft Kinect[d]

One study [27] investigated the use of a prototype system using four Kinect sensors positioned perpendicularly, each 1 m from the patient, to create a 3D temporal representation of a patient's torso. Results from the prototype were compared to spirometry in healthy volunteers and cystic fibrosis patients. Using multiple cameras enabled the estimation of volumes (Figure 2), as opposed to solely surface changes as provided by single camera approaches, important in estimation of respiratory outcomes measures. Spirometry and Kinect data were collected during a short performance test requiring the subject to perform quiet breathing for 20 s, followed by a relaxed vital capacity manoeuver (maximum inspiration and expiration) and followed by 20 s of quiet breathing, and was repeated three times. The authors concluded good validity of their approach compared to spirometry based on correlations (r > 0.8656) and that their system could accurately assess chest wall motion even in moving subjects. This may make this approach particularly valuable in assessing the effects of pharmacological and physical therapy treatment.

D. Facial Analysis

While the facial analysis capabilities of Kinect have been used successfully in delivering rehabilitation exercise regimens and feedback (see [16] for example), its use to translate digital tracking data into health outcome measures that can track treatment effects has been less well studied. However, the potential to achieve this is well demonstrated by these approaches. For example, the Face to Face solution [16] (Figure 3), funded by the UK NIHR Invention for Innovation Programme, is a rehabilitation system for facial paralysis in stroke patients. It recognizes facial expressions by tracking movement across the face and applying the recognized motion onto an onscreen representation of the user. How the user performs a series of facial exercises is assessed by the system and scored according to how well the user can undertake each of the defined set of expressions. This will enable shorter consultations through the presentation of quantified improvement, in addition to optimizing the performance of the exercises. Numeric measures of facial expression and symmetry enable the tracking of temporal improvements arising from pharmacological or physical therapy treatment.
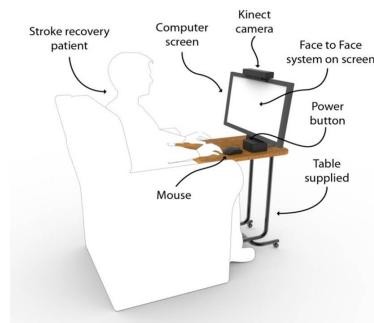

Fig. 3. Face to Face system for at-home rehabilitation

The ability of Kinect to identify eye opening and closure events may also facilitate measurement of blink rate. Blink rates are believed to reflect the activity of the central dopaminergic systems [28]. Patients with Huntingdon's Disease or Schizophrenia show higher blink rates than normal subjects; and patients with Parkinson's disease and progressive supranuclear palsy lower blink rates [29]. As an average blink takes 300-400 ms to complete, theoretically the ability to capture this event should be possible within the sampling frequency of the Kinect camera (30 frames per second), although more frequent sampling capabilities, such as that of the Intel RealSense SR300, may provide more robust measures.

Facioscapulohumeral muscular dystrophy may present as a weakness of the facial muscles. For example, a patient's eyes may remain slightly open when asleep, or they may be unable to tightly close their eyes. Patients may also experience difficulty pursing their lips. Using the tracking points around the mouth and lips it may be possible to instrument a lip pursing test, along with other tests of facial muscle strength, which could provide valuable outcome measures in these and other patient groups.

Patients with Huntingdon's Disease are often assessed using the UHDRS [30], a component of which considers measurement of aspects of motor impersistence, specifically difficulty keeping the tongue fully protruded upon command. The scale's tongue protrusion task is performed by asking the

TABLE V. SUMMARY OF SAMPLE OF STUDIES USING MICROSOFT KINECT TO MEASURE HEALTH OUTCOMES

Reference	Disease Indication	Measure	No. Patients	Comparator	Validation / Research Findings
Gait and balance					
[11]	Stroke	Six meter walking test: Step length, step length asymmetry, foot swing velocity, foot swing velocity asymmetry, peak and mean gait speed, percentage difference between the peak and mean gait speed.	30	10 meter walk test (10MWT), Timed Up and Go (TUG) test, functional reach, step tests, Force platform balance tests.	• Good test re-test reliability (ICCs > 0.8). • Kinect parameters explained at least an equal and in some cases greater amount of variance compared to clinically-derived measures. • Instrumenting gait using the Kinect is reliable and provides insight into the dynamic balance capacity of people living with stroke. • This system provides a minimally intrusive method of examining potentially important gait characteristics in people living with stroke.
[22]	MS	Short maximum speed walking test: Average walking speed, speed deviation, 3D direction deviation, Left/right deviation, Up/down deviation.	22 + 22 [e]	Timed 25-foot walking test (T25WT)	• Good test re-test reliability for speed, speed deviation and up/down deviation (ICCs > 0.74), but less robust for 3D direction deviation and left/right deviation (ICCs: 0.072-0.533). • Modest significant correlation between SMSW average walking speed and T25FW (r = -0.447, p = 0.042). • Able to detect gait disturbances in MS compared to healthy volunteers.
[23]	Parkinson's Disease	Performance and balance tests: quiet standing, multidirectional reaching, stepping and walking on the spot, hand clasping, finger tapping, foot and leg agility, chair rising and hand pronation.	9 + 10 [e]	Vicon 3D motion analysis system	• Strong construct validity for temporal accuracy of all tests (ICCs: 0.940-0.999). • Weaker construct validity for spatial accuracy for many tests (ICCs: 0.009-0.989). • The Kinect can accurately measure the timing and gross spatial characteristics of clinically relevant movements but not with the same spatial accuracy for smaller movements, such as hand clasping or toe tapping.
Upper extremity movement					
[24]	Stroke (with chronic hemiparesis)	Shoulder flexion/extension, shoulder abduction/adduction, elbow flexion/extension, elbow abduction/adduction.	8	Impulse motion capture system (Phase Space, San Leandro, CA, USA), and Fugl-Meyer scale	• Strong construct validity (R^2: 0.64, p < 0.001). • Kinect measures as reliable as clinical assessment using the Fugl-Meyer scale.
[25]	Adhesive capsulitis	Shoulder flexion, abduction and rotation.	12 + 15 [e]	Goniometer recording	• Strong construct validity (ICCs: 0.864-0.942).
[26]	MS	Arm and finger movement: finger to nose, finger to finger, drawing squares, truncal ataxia test.	317 + 724 [e]	None stated	Using bespoke machine learning algorithms offers promise to the automated clinical assessment of cerebellar function scores for upper extremities.
Chest wall motion analysis					
[27] [f]	Cystic Fibrosis	Tidal volume, respiratory rate and minute ventilation.	9 + 13 [e]	MicroLab Spirometer (Micro Medical Limited, Rochester, UK)	• Strong construct validity (r > 0.8656).

[e] Included additional healthy volunteers (second figure).

[f] Used four perpendicularly arranged Kinect Sensors.

patient to hold the tongue extended for 10 seconds and rating the response on a 5-point verbal response scale. Facial tracking approaches may enable instrumentation of tests such as these, with increased precision of measurement and greater richness of information collected.

IV. DISCUSSION

Microsoft Kinect provides a low-cost approach to the tracking and measurement of movement in clinic and remote settings. The versatile capabilities of the camera and its SDK, combined with its low cost and ability to operate on the relatively ubiquitous Windows platform provides the potential for high utility in clinical trials. In particular, it may offer value in trials that currently rely upon subjective assessments of movement. Software developed using Kinect provides the potential to inexpensively distribute solutions to multiple sites with less requirement on specialist centers, meaning that movement information can be gathered from larger and more diverse patient populations and care settings.

Certain tests, however, may be more suited to the use of this approach. Potential limiting factors include the sampling rate of the camera, the resolution and the depth/field of vision. Sampling rate and resolution may affect the ability to measure spatial accuracy for finer movements such as toe tapping and for more rapid movements such as certain gait parameters in maximal speed walking tests. Addressing this concern, some researchers have used interpolation methods to approximate more frequent sampling rates and this has shown to improve accuracy in the assessment of fast movements.

Increasing the sampling rate of the camera would be a valuable addition in future commercial releases, although other similarly priced hardware is available that offers this, such as the Intel RealSense SR300. Tests that can be easily conducted in a confined area, such as range of motion tests, or those requiring less rapid movement over a relatively short distance such as one or two gait cycles, are less likely to be affected by sampling rate or resolution limitations. However, the literature we have reviewed reports some successful outcomes even in the measurement of more challenging movement tasks.

In general, performance tests that researchers' seek to instrument using Kinect, should be practically conducted in a small area within the depth (0.5 to 4.5 m) and field of vision of the sensor. To overcome depth and field of vision limitations, some researchers have experimented with the use of Kinect to measure gait parameters whilst patients are using a treadmill. Xu et al. [31], for example, identified that Kinect could accurately estimate heel strike but less accurately estimate toe-off events in subjects walking at various speeds on a treadmill in comparison to the Optotrak Certus motion tracking system (Northern Digital Inc., Ontario, Canada). In addition, Auvinet et al. [32] showed that increased accuracy of gait parameters could be achieved during treadmill tests by considering the extreme values of the distance between knee joints along the walking longitudinal axis to increase precision of estimated heel strike.

The ability of Kinect to accurately track joints can be affected by having other objects in the field of view. For example, a nearby chair leg could be confused with a patient's leg, although this may be greatly improved with later SDK versions which enable seated and standing modes to be differentiated. The distraction of joint tracking capabilities when use of walking supports and devices may be a consideration, however, in some patient groups.

While the development of appropriate and pertinent performance tests that can be fully measured using Kinect is one important area of continued research, for clinical trials acceptance it will also be important to be able to present robust validation evidence. Typically this will comprise a number of elements:

- Demonstration that the measurement device (Kinect) is providing measurements to an appropriate degree of accuracy and precision – for example understanding the accuracy of joint position measurement.

- Demonstration that any special characteristics of a particular population do not adversely impact the ability to make accurate and precise measurements. For example, if assessing gait parameters ensuring that walking with a stick does not affect the accuracy of measures, or patients with a more shuffling gait – as seen in some Parkinson's patients for example – does not affect the ability to make reliable measurements.

- Demonstrating the clinical relevance of health outcome measures derived from measurement data. For example, new endpoints may be derived with the greater information that motion tracking can provide and it will be important to show that these have clinical meaning if they are to be used in clinical trial regulatory submissions.

This latter element may include comparison of the clinical outcomes measured using Kinect to alternative and gold standard approaches. However, it should be noted that in some circumstances expecting to demonstrate equivalence to an existing clinical trial gold standard approach may not be appropriate. For example, comparing an objective measure to a subjective assessment may in fact fail to show equivalence due to the limitations of the subjective assessment as opposed to a failing in the new measure.

V. SUMMARY

Microsoft Kinect is a low-cost depth sensor operating on the Windows platform that enables developers to create engaging healthcare applications that may be used in clinic or in other settings to measure objectively the effects of treatment in relation to motion and mobility outcomes. It has already been used extensively in the area of serious games for rehabilitation. This paper has reviewed the technical capabilities of the platform, in addition to the Intel RealSense SR300, Leap Motion, CREATIVE® SENZ3D and Xtion Pro Live cameras in relation to tracking a variety of aspects of movement and motion. We have further reviewed a number of studies using Microsoft Kinect to enable objective assessment of aspects of motion and mobility that may be of value in clinical trials. These examples have shown that Kinect can provide clinical outcome measures to a high degree of precision and accuracy, particularly those relating to gross

spatial movements. Particular value may be seen in the ability to replace certain clinician subjective assessments with accurate and sensitive objective outcome measures, and to provide richer information with which to understand treatment effects.

While more research will be required to ensure that Kinect-based derived health outcomes measures have robust validity and clinical relevance, their potential to enhancing the experience of participating in a clinical trial, and as a means of collecting novel endpoint data is evidenced by the examples reviewed in this paper.

While highly regulated, the pharmaceutical industry remains actively interested in applying innovative approaches that improve the operation of clinical trials and the thorough understanding of treatment effects. Leveraging video gaming platforms such as Microsoft Kinect remain an industry interest area where more exploration, development and examples of successful application are needed to drive their potential adoption in clinical trials.

REFERENCES

[1] Tufts Centre for the Study of Drug Development, "Cost to Develop and Win Marketing Approval for a New Drug Is $2.6 Billion," http://csdd.tufts.edu/news/complete_story/pr_tufts_csdd_2014_cost_study, November 2014.

[2] Hamilton, M. (1960) A rating scale for depression. J. Neurol. Neurosurg. Psychiatry; 23: 56-62.

[3] Goetz, C.G., Tilley, B.C., Shaftman, S.R. et al. (2008). Movement Disorder Society-sponsored revision of the Unified Parkinson's Disease Rating Scale (MDS-UPDRS): scale presentation and clinimetric testing results. Mov. Disord. 23(15): 2129-2170.

[4] Fugl-Meyer, A.R., Jääskö, L., Leyman, I. (1975). The post-stroke hemiplegic patient I: A method for evaluation of physical performance. Scandinavian Journal of Rehabilitation Medicine; 7: 13–31.

[5] Tinetti, M.E., Williams, T.F., Mayewski, R. (1986). Fall Risk Index for elderly patients based on number of chronic disabilities. Am. J. Med.; 80: 429-434.

[6] Berg, K., Wood-Dauphinee, S., Williams, J.I. (1995). The Balance Scale: reliability assessment with elderly residents and patients with an acute stroke. Scand. J. Rehabil. Med.;27: 27–36.

[7] Shumway-Cook, A., Woollacott, M. (2007). Motor Control: Translating Research into Clinical Practice. Wolters Kluwer, Philadelphia, PA, USA; p. 413.

[8] Byrom, B., Mundt, J.C. (2005). The value of computer-administered self-report data in central nervous system clinical trials. Current Opinion in Drug Discovery and Development; 8(3): 365-373.

[9] Hondori, H.M., Khademi, M. (2014). A Review on Technical and Clinical Impact of Microsoft Kinect on Physical Therapy and Rehabilitation. J. Med. Eng.; Article ID 846514.

[10] Pagliari, D., Pinto, L. (2015). Calibration of Kinect for Xbox One and Comparison between the Two Generations of Microsoft Sensors. Sensors. 15(11): 27569-27589.

[11] Clark, R.A., Vernon, S., Mentiplay, B.F. et al. (2015). Instrumenting gait assessment using the Kinect in people living with stroke: reliability and association with balance tests. J. Neuroeng. Rehabil.; 12: 15-24.

[12] Bachmann, D., Weichert, F., Rinkenauer, G. (2014). Evaluation of the leap motion controller as a new contact-free pointing device. Sensors; 15(1): 214-33.

[13] http://uk.creative.com/p/web-cameras/creative-senz3d, accessed on 13-May-2016.

[14] Stryla, W., Banas, A. (2015). The Use of Virtual Reality Technologies during Physiotherapy of the Paretic Upper Limb in Patients after Ischemic Stroke. J. Neurology and Neuroscience; 6(3): 33-38.

[15] Archambault, P., Norouzi, N.G., Kairy, D., Solomon, J.M., Levin, M.F. (2014). Towards Establishing Clinical Guidelines for an Arm Rehabilitation Virtual Reality System. In (Jensen, W., Andersen, O.K., Akay, M. eds. Replace, Repair, Restore, Relieve – Bridging Clinical and Engineering Solutions in Neurorehabilitation, Proceedings of the 2nd International Conference on NeuroRehabilitation (ICNR2014), Aalborg, 24-26 June, 2014); pp.263-270.

[16] Breedon, P., Russell, A., Logan, P. et al. (2014). First for Stroke: using the Microsoft' Kinect' as a facial paralysis stroke rehabilitation tool. International Journal of Integrated Care (IJIC), 14.

[17] Sin, H., Lee, G. (2013). Additional virtual reality training using Xbox kinect in stroke survivors with hemiplegia. Am. J. Phys. Med. Rehab.; 92(10): 871–880.

[18] Galna, B., Jackson, D., Schofield, G. et al. (2014). Retraining function in people with Parkinson's disease using the Microsoft Kinect: game design and pilot testing. J. Neuroeng. Rehabil.; 11, article 60.

[19] Ortiz-Gutiérrez, R., Cano-de-la-Cuerda, R., Galán-del-Río, F. et al. (2013). A telerehabilitation program improves postural control in multiple sclerosis patients: a Spanish preliminary study. Int. J. Environ. Res. Public Health; 10(11): 5697–5710.

[20] Luna-Oliva, L. Ortiz-Gutiérrez, R., Cano-de-la-Cuerda et al. (2013). KinectXbox 360 as a therapeutic modality for children with cerebral palsy in a school environment: a preliminary study. NeuroRehabilitation; 33(4): 513–521.

[21] Byrom, B. (2015). Clinical Trials Re-spec: The Role of Games and Gamification in the Future of Clinical Trials. 2015 International Conference on Interactive Technologies and Games, IEEE; pp. 28-33. DOI: 10.1109/iTAG.2015.8.

[22] Behrens, J., Pfüller, C., Mansow-Model, S., Otte, K., Paul, F., Brandt, A.U. (2014). Using perceptive computing in multiple sclerosis - the Short Maximum Speed Walk test. J. Neuroeng. Rehabil.; 11:89-99.

[23] Galna, B., Barry, G., Jackson, G., Mhiripiri, D., Olivier, P., Rochester, L. (2014). Accuracy of the Microsoft Kinect sensor for measuring movement in people with Parkinson's disease. Gait and Posture; 39: 1062–1068.

[24] Olesh, E.V., Yakovenko, S., Gritsenko, V. (2014). Automated Assessment of Upper Extremity Movement Impairment due to Stroke. PLoS ONE 9(8): e104487. doi:10.1371/journal.pone.0104487.

[25] Lee, S.H, Yoon, C., Chung, S.G. et al. (2015). Measurement of Shoulder Range of Motion in Patients with Adhesive Capsulitis Using a Kinect. PLoS ONE 10(6): e0129398. doi:10.1371/journal.pone.0129398

[26] Kontschieder, P., Dorn, J.F., Morrison, C. et. al. (2014). Quantifying Progression of Multiple Sclerosis via Classification of Depth Videos, in (Golland, P., Hata, N., Barillot, C., Hornegger, J., Howe, R. eds. Medical Image Computing and Computer-Assisted Intervention – MICCAI 2014, Volume 8674 of the series Lecture Notes in Computer Science); pp 429-437.

[27] Harte, J.M., Golby, C.K., Acosta, J., et al. (2015). Chest wall motion analysis in healthy volunteers and adults with cystic fibrosis using a novel Kinect-based motion tracking system. Med. Biol. Eng. Comput.; DOI 10.1007/s11517-015-1433-1.

[28] Karson, C.N. (1983). Spontaneous eye-blink rates and dopaminergic systems. Brain; 106: 643-653.

[29] Grandas F., Esteban, A. (1994). Eyelid motor abnormalities in progressive supranuclear palsy. In (Tolosa, E., Duvoisin, R., Cruz-Sanchez, F.F. eds. Progressive Supranuclear Palsy: Diagnosis, Pathology, and Therapy, Springer); pp 33-41.

[30] Kieburtz K. and the Huntington Study Group. (1996) The Unified Huntington's Disease Rating Scale: Reliability and Consistency. Mov. Dis.;11: 136-142.

[31] Xu, X., McGorry, R.W., Chou, L.S., Lin, J.H., Chang, C.C. (2015). Accuracy of the Microsoft Kinect for measuring gait parameters during treadmill walking. Gait and Posture; 42: 145-151.

[32] Auvinet, E., Multon, F., Aubin, C.E., Meunier, J., Raison, M. (2015). Detection of gait cycles in treadmill walking using a Kinect. Gait and Posture; 41(2): 722-725.