Abstract

Background: Total knee replacement currently lacks robust indications and objective follow-up metrics. Patients and healthcare staff are under-equipped to optimise outcomes. This study aims to investigate the feasibility of using an ear-worn motion sensor (e-AR, Imperial College London) to conduct objective, home-based mobility assessments in the peri-operative setting.

Methods: Fourteen patients on the waiting list for knee replacement, and 15 healthy subjects, were recruited. Pre-operatively, and at 1, 3, 6, 12 and 24 weeks post-operatively, patients underwent functional mobility testing (Timed Up and Go), knee examination (including range of motion), and an activity protocol whilst wearing the e-AR sensor. Features extracted from sensor motion data were used to assess patient performance and predict patients' recovery phase.

Results: Sensor-derived peri-operative mobility trends correlated with clinical measures in several activities, allowing functional recovery of individual subjects to be profiled and compared, including the detection of a complication. Sensor data features enabled classification of subjects into normal, pre-operative and 24-week post-operative groups with 89% (median) accuracy. Classification accuracy was reduced to 69% when including all time intervals.

Discussion: This study demonstrates a novel, objective method of assessing peri-operative mobility, which could be used to supplement surgical decision-making and facilitate community-based follow-up.

Keywords: total knee replacement; mobility; post-operative; sensor

Introduction

The prevalence of total knee replacement (TKR) is increasing. In the UK, 90,842 knee replacements were carried out in 2012, an increase of 7.3% compared to 2011.[1] In the context of an obesity epidemic and increasingly elderly population, this trend is set to continue,[2] making it imperative that management strategies of such a significant healthcare burden are optimised.

Post-operatively, rates of functional convalescence and long-term outcomes are poorly measured and documented, with surgical revision acting as the surrogate gold standard outcome measurement, despite its many limitations,[3] perhaps best exemplified by the 20% of non-revised TKR patients continue to complain of knee pain or disability post-operatively.[4] Measures to monitor rehabilitation are similarly blunt, and whilst major post-operative complications are likely to be reported by the patient, difficulty with rehabilitation exercises due to manageable causes may not be identified until the next out-patient appointment, if at all.

Outcome following knee replacement is most often reported in the literature using patient reported outcome measures. The limitations of this approach are well documented; cultural and psychological bias on interpretation of questions,[5, 6] physical confounders reducing test specificity,[7] and the amplification of bias with repeat testing.[8] Kinematic and biomechanical assessment in clinical practice and research may provide a more objective measure of patient outcome, but is limited by cost, availability, and technical expertise. An ear-worn activity recognition sensor developed in our department (e-AR, Sensixa Ltd and Imperial College London) is capable of providing valid and reliable detection of activities of daily living,[9] estimating gait parameters,[10] and predicting energy expenditure.[11]

Studies involving objective measures of functional mobility following knee replacement are scarce. Kleijne et al. profiled the recovery curve of unilateral knee replacements at 4 time points over 2 years using a wearable sensor in the out-patient clinic,[12] whilst Mattsson et al. investigated walking speed[13] and Weidenhielm et al. used electro-goniometry[14] during walking pre- and post-operatively.

Healthcare providers such as the National Health Service (NHS) in the UK are trying to improve the cost-effectiveness of specialist services by moving appropriate care provisions into the primary care setting, and giving patients more responsibility in the management of sub-acute conditions. This dramatic reshaping of the post-operative care paradigm requires considerable planning, and the development of technologies designed for safe patient surveillance. As such, the aim of this study was to assess the feasibility of using a wearable sensor to monitor patient recovery following TKR in the community setting.

Methods

A prospective observational pilot study was designed to assess the feasibility of sensor-based mobility assessments of patients undergoing TKR in the home environment. The study was approved by the National Research Ethics Committee (07/12/139). Patients scheduled to undergo TKR at Imperial College Healthcare NHS Trust between January 2009 and January 2011 were identified and assessed for study eligibility. Patients aged less than 18 years old, those with symptomatic contralateral knee pain, another known mobility disorder or previous lower limb surgery, were excluded. Patients were informed of the research study in the outpatient clinic by their direct healthcare team. Following expression of interest and verbal consent, the patient was given a study information sheet and the option to meet with a member of the research team. Written informed consent was gained from all study participants. There was no financial incentive offered to participate.

Appointments at the patients' homes were scheduled approximately 1-week pre-operatively, and again at 1, 3, 6, 12 and 24 weeks post-operatively. Each appointment was planned to include 3 sections: completion of validated outcome measures (quality of life and knee-specific questionnaires and the timed up and go test), brief clinical examination (including bilateral knee range of motion), and an activity protocol whilst wearing a motion sensor (e-AR). A group of healthy volunteers (n=15) were assessed once using the same protocol.

The timed up and go test (TUG) is a simple clinical test used to assess the basic mobility skills required to perform activities of daily living, including standing, walking, turning, and sitting.[15, 16]

The activity protocol was based on the Dynaport Knee Test (Table 1),[17] for which construct validity has been proven by correlation with the physical therapists' ratings of disability.[18] Patients were asked to complete as many activities as they were comfortable with. The e-AR sensor (Fig. 1) contains a tri-axial Micro-Electro-Mechanical-Systems (MEMS) accelerometer (ADXL330) and wireless transmitting capabilities, allowing subject movements to be collected in a free-living environment, without the restrictions and costs associated with wired systems and gait laboratories.[19] A sampling rate of 50Hz was used, within an acceleration range of +/- 3g. A tablet computer operated by the attending researcher wirelessly received data from the e-AR sensor. To aid data analysis, activities were electronically tagged and segmented using a real-time marking system on the tablet.

Table 1.

Figure 1

Components of Feasibility

In investigating the feasibility of the sensor for peri-operative assessment, 3 components were evaluated, chosen based on their potential value in clinical application. The first was a performance score based on sensor data, which was compared against current assessment techniques (TUG, knee ROM). The second was the assessment of gait through a visual analysis of longitudinal sensor data. The third was an assessment of the similarities in the performance of patients at specific time points through the application of computational clustering techniques.

Data analysis

The raw 3-dimensional signal from the e-AR sensor was segmented into activities. A number of abstract signal features (n=36) were extracted from sensor motion data, which were used to create a kinematic impression of subject performance of each task. The healthy subject data were used to plot typical activity performance, which would act as a baseline for comparison. A computational model was used to quantify the abnormality in patient performance for each task, whereby a bigger number indicates a greater degree of abnormality.[9] This sensor score provided the primary outcome measure to compare with other assessment techniques. Artificial Neural Networks were used to classify patients into their most likely performance group, clustering subjects exhibiting similar kinematic profiles of activity.

Statistical analysis

Descriptive statistics were calculated for subject demographics and all outcome measures. Spearman's rho correlation coefficients were calculated between sensor scores and gold standard measurements (TUG and ROM) to assess concurrent validity. The construct and ability of the sensor to distinguish between patients of expectedly dissimilar states (e.g. pre-op vs. 1-week post-op) was investigated using Wilcoxon signed-rank tests. Statistical analysis was performed using the Statistical Package for Social Sciences version 19.0 (SPSS, IBM corp, NY, USA).

Results

Study participants

Between January 2009 and January 2011, 14 patients undergoing TKR and 15 healthy subjects were recruited. The patient group included 6 males and 8 females, with a mean age at the time of surgery of 69.3 ± 4.6 years (SD) (range 60 – 84). The mean Body Mass Index was 29.2 ± 2.8 kg/m^2 (SD) (range 22.1 – 33.8). Healthy subjects had a mean age of 27.1 ± 1.9 (SD) (range 23 – 49), and all had a BMI of less than 30 kg/m^2. Fifty-four out of 84 planned data points were collected (64%), with randomly distributed data loss. None of the subjects requested withdrawal from the study. Twenty-seven out of 84 (32%) questionnaires were completed and as such formal analysis was omitted.

Sensor Performance Score

Patient function estimated by the sensor, TUG and ROM consistently declined following the operation, before gradually returning to a level similar or greater than that during pre-operative testing. (Fig. 2, Table 2) Pre-operatively, sensor scores for gait, slope walking, tray carry and bag carry (4/10 activities) correlated with TUG, but the correlation between all activities combined and TUG did not reach statistical significance (R = 0.427, p = 0.167). No pre-operative activities showed significant correlation with ROM. The correlation between gold standard measurements (TUG and ROM) was not significant (R = 0.323, p = 0.306). The biggest changes in performance were seen within the first 6 weeks of the surgery (including the pre-operative time point), after which performance plateaued. Changes in patient performance estimated by the sensor followed similar trends to TUG (R = 0.245, p = 0.097) and ROM (R = 0.282, p = 0.057) yet correlations failed to reach statistical significance overall. A weak correlation was seen between TUG times and ROM (R = 0.324, p = 0.026).

Figure 2

Table 2

At 12-weeks post-operatively, 1 patient was seen to have a substantial knee effusion with accompanying pain and stiffness. This patient's recovery curve was noticeably different to the typical recovery profile exhibited by the majority of the cohort. (Fig. 3)

Figure 3.

Gait analysis

Longitudinal changes in performance of the activity protocol were visible on the raw e-AR sensor data. (Fig. 4) During walking, for example, it was clear that the walking speed and confidence (reflected by sharpness and amplitude of the steps) reduced immediately after the operation, before gradually returning to a level similar or greater than that during pre-operative testing. Sensor performance scores for walking individually showed a moderate correlation to changes in knee ROM ($R = 0.462$, $p = 0.001$), but not to TUG times.

Figure 4.

Patient Classification

At the defined time intervals patients appeared to exhibit similar kinematic traits. Subjects were clustered according to sensor motion features, which allowed blind classification into normal, pre-operative and 24 week post-operative groups with an accuracy of 89% (median). Classification accuracy was reduced to 69% when including all post-operative time intervals. (Fig. 5) Misclassification usually occurred into a chronologically adjacent group. (Table 3) Classification accuracy decreased with increasing post-operative time.

Figure 5.

Table 3.

Discussion

This study demonstrates the feasibility of a lightweight ear-worn motion sensor to objectively assess peri-operative patient mobility in the community setting. Kinematic features derived from sensor motion data were used to quantify movement abnormality as compared with a healthy cohort, producing performance scores for a selection of activities of daily living. Expected trends in patient performance were seen in the sensor data, which bared resemblance to the gold standard measurements (TUG and knee ROM). Although not all of the correlations were statistically significant, neither were some between the TUG and ROM. The sensor score was able to identify 1 episode in which the patient's recovery trajectory deviated away from that of the rest of the cohort. The motion features used to derive the score consider elements such as the magnitude of linear accelerations, consistency in repetitive tasks, and left/right symmetry. Combining such features, in theory, provides an impression of the physical manner in which a patient completes their activities of daily living. It is possible that this novel outcome measure is more representative of physical function, compared to the one-dimensional TUG and ROM scores, where coping strategies and adaptation may provide misleading results. Alternatively, the sensor score could be seen as a quality control metric, whereby a quicker TUG time is only valid upon a satisfactory physical performance (i.e. without hopping or an uncontrolled descent into the chair).

Healthy subjects and patients at the same peri-operative stage were seen to exhibit similar kinematic profiles throughout the activity protocol. This allowed subjects to be clustered and classified into particular groups at a high level of accuracy. Classification accuracy was seen to decrease at the latter time points, probably reflecting the increased variation in longer-term outcomes. One would expect a Gaussian distribution in performance at each time point during recovery from a larger patient population, taking into account confounding factors

such as age, cardiovascular fitness, and contra-lateral pathology. Once a more reliable recovery curve is established from a larger data set, the clinical utility of classification would be to identify those patients whose actual post-operative stage differs to that predicted by the sensor. This information could facilitate an early warning system, building towards the resource-sparing paradigm of appointments when needed, as opposed to arbitrarily planned.

Our results are similar to the, albeit sparse, related literature. Directly comparable studies with TKR and similar time intervals are not available, however, patients of a comparable age range undergoing uni-compartmental knee replacement displayed similar trends in recovery to our cohort, in which the most significant improvements in function were also made in the first 12 post-operative weeks.[12]

A common alternative to objective, functional assessments in research and clinical practice is the use of questionnaires for quality of life (e.g. SF-36) and/or knee function (e.g. Noble and Weiss score).[20] Although it would be feasible for patients to complete a questionnaire every day during rehabilitation, many groups have questioned the reliability of this method.[8] Repeated completion of a questionnaire is likely to result in amplification of the subjectivity in response. Furthermore, questionnaire completion is vulnerable to confounding factors including mood, culture and comorbidities.[5-7] Functional tests such as the TUG, 6-minute walk and others may also be used. However these test are somewhat unspecific, failing to identify problem areas or tasks. There is also a risk of factors that are unrelated to the knee, including behaviour (e.g. competitiveness) and level of fitness, confounding the results. The e-AR sensor provides a performance score for several different tasks, increasing the resolution of the assessment and catering for a wide spectrum of performance levels. For example, a task more physically demanding than walking may be required to distinguish between 2 patients at the latter end of the recovery curve. In addition, using a single sensor

has practical advantages over multi-sensor systems such as the DynaPort system including placement consistency, patient usability, and minimum restriction in activity.

Interpretation of the data must be performed in the context of limitations in the study design. Although the prevalence of TKR is high, recruiting patients with unilateral symptoms is challenging. Also, collection of data from multiple visits to a patient's home is socially invasive and logistically demanding, which resulted in higher than anticipated attrition rate and partial appointments where questionnaires were incomplete. The distribution of data loss was not associated with patient function, though was more prevalent during the latter post-operative stages.

To provide a reference for comparison in our study, we recruited 15 healthy subjects to complete the activity protocol wearing the e-AR sensor. However, the mean age of the group was significantly less than that of the patient cohort. Whilst this might not have been an ideal reference for direct comparison, the data provided a useful baseline with which to compare longitudinal changes in patient performance.

Collecting data in the home environment was proposed with the future in mind, based on recent strategies to relocate healthcare services into the community. However, there was variable access to walking space and stairs, and the stairs were different in all homes. Although some activities had to be omitted, so long as the environment was consistent, longitudinal follow-up results need not be affected.

Clinical impact

In the community, a remote assessment device such as the e-AR sensor has potential applications as a screening tool, or a surveillance method for patients with musculoskeletal

pathology. Having an objective measure of function would allow primary care services to titrate and monitor the effects of conservative management strategies such as physiotherapy and analgesia with more confidence. The decision to refer to a specialist could be more carefully timed, with the referral including accurate longitudinal performance data.

Post-operatively, regular patient-led assessments combined with feedback may provide a source of motivation to complete rehabilitation exercises. Frequent acquisition of performance data could be used to personalise rehabilitation and detect complications. Recovery data available to a clinician during a follow-up appointment would allow them to make more informed decisions on the continued management of the patient. Similarly, practice managers and national governing bodies may benefit from such data when evaluating quality of service and developing best practice guidelines.

For such implementation to be realised, further work must address the need to develop an effective patient user interface and communication infrastructure to facilitate unsupervised data collection and upload. It will be important to define the level of post-operative data granularity required from knee replacement patients to provide an adequate picture. Data analysis is likely to evolve from calculating the magnitude of movement abnormality to providing detailed mobility parameters. Although methods described in this paper would suffice for identifying gross complications requiring clinical contact or surgical revision, more subtle abnormalities such as asymmetrical loading or altered temporal gait features would provide more useful information to the patient and healthcare team. This level of information is more likely to surpass the use of current clinical tests, and promote a strong case for using the sensor when considering detailed assessment and personalised, dynamic rehabilitation strategies.

Acknowledgements: With thanks to Mr Dominic Spicer FRCS (Orth) who was the lead consultant surgeon to the patients recruited

Declarations and funding: e-AR Sensor was provided by Sensixa Ltd, a spin-off from Imperial College London, for which GZY is the director. Funding was provided by the Engineering and Physical Sciences Research Council (EPSRC) as part of the ESPRIT programme.

References:

1. National Joint Registry. National Joint Registry for England and Wales: 10th annual report. 2013. http://www.njrcentre.org.uk. (2013
2. Arden N, Nevitt MC. Osteoarthritis: epidemiology. Best Pract Res Clin Rheumatol 20 (2006)3-25
3. Goodfellow JW, O'Connor JJ, Murray DW. A critique of revision rate as an outcome measure: re-interpretation of knee joint registry data. J Bone Joint Surg Br 92 (2010)1628-31
4. Bourne RB, Chesworth BM, Davis AM, et al. Patient satisfaction after total knee arthroplasty: who is satisfied and who is not? Clin Orthop Relat Res 468 (2010)57-63
5. Tugwell P, Bombardier C, Buchanan WW, et al. The MACTAR Patient Preference Disability Questionnaire--an individualized functional priority approach for assessing improvement in physical disability in clinical trials in rheumatoid arthritis. J Rheumatol 14 (1987)446-51
6. Gilson BS, Gilson JS, Bergner M, et al. The sickness impact profile. Development of an outcome measure of health care. Am J Public Health 65 (1975)1304-10
7. Wolfe F. Determinants of WOMAC function, pain and stiffness scores: evidence for the role of low back pain, symptom counts, fatigue and depression in osteoarthritis, rheumatoid arthritis and fibromyalgia. Rheumatology (Oxford) 38 (1999)355-61
8. Ryd L, Karrholm J, Ahlvin P. Knee scoring systems in gonarthrosis. Evaluation of interobserver variability and the envelope of bias. Score Assessment Group. Acta Orthop Scand 68 (1997)41-5
9. Ali R, Atallah L, Lo B, et al. Detection and analysis of transitional activity in manifold space. IEEE Trans Inf Technol Biomed 16 (2012)119-28

10. Jarchi D, Wong C, Kwasnicki RM, et al. Gait parameter estimation from a miniaturized ear-worn sensor using singular spectrum analysis and longest common subsequence. IEEE Trans Biomed Eng 61 (2014)1261-73

11. Bouarfa L, Atallah L, Kwasnicki RM, et al. Predicting free-living energy expenditure using a miniaturized ear-worn sensor: an evaluation against doubly labeled water. IEEE Trans Biomed Eng 61 (2014)566-75

12. Kleijn LL, van Hemert WL, Meijers WG, et al. Functional improvement after unicompartmental knee replacement: a follow-up study with a performance based knee test. Knee Surg Sports Traumatol Arthrosc 15 (2007)1187-93

13. Mattsson E, Brostrom LA, Linnarsson D. Changes in walking ability after knee replacement. Int Orthop 14 (1990)277-80

14. Weidenhielm L, Olsson E, Brostrom LA, et al. Improvement in gait one year after surgery for knee osteoarthrosis: a comparison between high tibial osteotomy and prosthetic replacement in a prospective randomized study. Scand J Rehabil Med 25 (1993)25-31

15. Faria CD, Teixeira-Salmela LF, Nadeau S. Development and validation of an innovative tool for the assessment of biomechanical strategies: the Timed "Up and Go" - Assessment of Biomechanical Strategies (TUG-ABS) for individuals with stroke. J Rehabil Med 45 (2013)232-40

16. Shumway-Cook A, Brauer S, Woollacott M. Predicting the probability for falls in community-dwelling older adults using the Timed Up & Go Test. Phys Ther 80 (2000)896-903

17. van den Dikkenberg N, Meijer OG, van der Slikke RM, et al. Measuring functional abilities of patients with knee problems: rationale and construction of the DynaPort knee test. Knee Surg Sports Traumatol Arthrosc 10 (2002)204-12

18. Mokkink LB, Terwee CB, van Lummel RC, et al. Construct validity of the DynaPort KneeTest: a comparison with observations of physical therapists. Osteoarthritis Cartilage 13 (2005)738-43

19. Atallah L, Wiik A, Jones GG, et al. Validation of an ear-worn sensor for gait monitoring using a force-plate instrumented treadmill. Gait Posture 35 (2012)674-6

20. Kreibich DN, Vaz M, Bourne RB, et al. What is the best way of assessing outcome after total knee replacement? Clin Orthop Relat Res (1996)221-5

Table and Figure Legends

Table 1. Activity protocol (based on Dynaport Knee Test)

Table 2. Median (Interquartile range) activity performance scores and longitudinal statistical assessment for significant difference (down respective columns)(Wilcoxin signed rank)

Table 3. Aggregate of confusion matrices for classification of post-operative week (for perfect classification only diagonal elements will prevail). Off-diagonal elements signify the confusion between 2 classes of post-operative stage.

Figure 1. E-AR sensor

Figure 2. Peri-operative patient performance as measured by the sensor (all activities combined), TUG and ROM (normalised values). All data were normalised relative to the median pre-operative performance (i.e. week 3 score / median pre-op score).

Figure 3. Case comparison – Patient A exhibits a typical recovery curve whilst patient B shows late signs of deterioration (reflected by sensor and knee range of motion).

Figure 4. Raw data (arbitrary units) from e-AR sensor captured during walking. For illustrative purposes, this shows changes in the signal throughout recovery. Each axis is shown by a different colour (Blue – Medio-lateral, Red – Supero-inferior, Green – Antero-posterior). Downward spikes in the red and green axes reflect ground contact events.

Figure 5. Sensor accuracy for classifying study subjects into chronological groups. Grey columns consider correct classification into 1 of 3 possible groups, and white columns into 1 of 7 groups, with expectedly lower accuracy.