A 3-D assessment and feedback tool for Ankylosing Spondylitis from the perspective of healthcare professionals

Shijuan Li [1], Stephen Kay [2], Stuart Porter [3]

[1]Department of Information Management, Peking University, Beijing, China, [2] THISC Ltd, UK, [3]School of Health Sciences, University of Salford, Salford, UK.

Abstract

To investigate the utility of 3-D visualisation technology to augment assessment and feedback for Ankylosing Spondylitis (AS), a visualisation prototype was developed, and both subjective and objective measures of current assessment instruments were compared.

To verify and establish a base line for the prototype's effectiveness, motion data and measurement data from a healthy adult in a laboratory environment were collected. To validate the prototype, a qualitative evaluation was undertaken using multiple methods including a pilot study, focus groups and individual interviews. Research subjects comprised physiotherapists in clinical practice and academia and content analysis of their responses was used to substantiate the findings.

The prototype enhanced both assessment and feedback of AS from the physiotherapist's perspective and they believed it to be superior to the current methods used in practice for assessing the condition and in documenting variations for subsequent treatment. The physiotherapists believed that such a system had potential to encourage multidisciplinary working, and to be patient-centric, both with respect to the process of treatment and with regard to the convenience it offered to patients in managing their own condition. 3-D visualisation of AS symptoms and its treatment via exercise is a valuable technique as demonstrated by the prototype system.

Key Words: Ankylosing Spondylitis, assessment and feedback, patient-reported outcomes, 3-D visualization.

Correspondence: Shijuan Li, Department of Information Management, Peking University, Beijing, China. E-mal: shijuan.li@pku.edu.cn

1. Introduction

Ankylosing Spondylitis (AS) is a rheumatic disease that affects mainly the spine [1], and periphery [2]. It is a long term, chronic condition that requires clinical assessment to inform the

ongoing management process. An integral part of the current assessment is the range of measurements taken; these comprise a number of subjective and objective factors that are believed to affect the person's health status over time.

The current measurements try to encompass the overall impact of AS on a person at a particular time and these are compared with historical data. The assessments have been summarized by the "Assessment of SpondyloArthritis International Society" (ASAS). ASAS identified six core domains: pain; spinal stiffness; patient global assessment of health status; functional ability; spinal mobility; and fatigue. These six are all measured either by a visual analogue scale (VAS) [3,4] or a numerical rating scale (NRS): pain (the intensity is measured by a VAS or NRS); spinal stiffness (morning stiffness duration is measured by a VAS or NRS); patient global assessment of health status (VAS or NRS); functional ability (measured by Bath Ankylosing Spondylitis Functional Index (BASFI) [5] or Dougados Functional Index (DFI)) [6]; spinal mobility (chest expansion, Modified Schober, Occiput to wall, Cervical rotation and Lateral spinal flexion or Bath Ankylosing Spondylitis Metrology Index (BASMI)) [7]; and fatigue (VAS or NRS). In the UK, a typical assessment includes interview and form completion designed to capture demographic details as well as measurements provided by a range of instruments to assess these domains. These instruments include several well-respected indices; BASFI and BASMI being the most relevant for this study.

BASFI uses the VAS extensively, attempting to measure a characteristic or attitude that ranges across a continuum of values. It takes the form of a 10 cm line scale and the patient marks the line in such a way as to provide an answer to the question being asked [3,4]. Within the AS assessment the VAS captures the subjective response from the patient that the clinician then uses to calculate a numeric score by measuring where the mark is on the line. BASMI is different from BASFI and requires clinicians to use various types of measuring device to obtain more 'objective' measures. Both in the process of subjective and objective assessment, healthcare professionals play a critical role by contributing to the measuring and in helping subjects to manage their condition.

Bath indices were regarded as gold standard for AS assessment and we believe that can be improved upon in ways not previously thought possible. BASFI records sensitivity through its capacity to make use of the entire scale of the index by using the VAS, although it has been suggested that it may be improved by using a numeric rating scale for the response options [8].

It is known that clinical and laboratory indicators of disease activity are poor predictors of radiological damage [9]. The BASMI instrument does not relate well to radiology treatment but it takes account of normal physical limitation and soft tissue. BASMI is, however, judged to be important in assessing AS [10].

Given the chronic nature of the condition, the on-going assessment of an AS patient is by means of scheduled, periodic reviews. Ideally the elapsed time between such assessments is directly related to the patient's health state, but, pragmatically the appointment relates to a general schedule dictated by the particular clinic's own resource constraints rather than a specific patient's needs. The trend, due to a combination of factors, is towards longer periods between assessments and this trend is deemed to be unsatisfactory by both AS support groups and clinicians alike who believe that more infrequent assessments may result in the earlier onset of advanced disability. Furthermore, the variety of measuring devices available at each clinic and the unique characteristics of the patient often result in inconsistency and inaccuracy of assessment.

Given the problems identified above, this study was concerned with how new innovations in technology might be used to improve assessment. Furthermore, given the visual nature of the AS condition, and the visual aspects of the instruments used to assess the patient's condition, it was appropriate to consider whether the field of Virtual Reality (VR) had any value for this research. VR has been claimed to be "revolutionary" for healthcare [11]. More specifically, VR as an assessment/feedback tool has been used for cognitive performance assessment [12,13], for assessing clinician's skills through a simulator [14], and for haptic force feedback [15]. These different applications were encouraging and 3-D visualization/ displays are considered to be the core of VR, which suited our purpose very well [16].

Furthermore none of the published work in 3-D visualisation has looked at integrating different tools to support the management of multiple problems related to the same condition. Nor did there seem to be an appreciation of how evaluation of performance is required to be an integral part of the tool set. AS was an appropriate exemplar condition as it has multiple symptoms and requires evaluation (including assessment and feedback) as an integral part of its ongoing treatment process and these all rely to some extent upon visual forms of assessment.

Significant progress in the understanding of AS have been made since the initial ASAS recommendations, which even then acknowledged that they could likely change in the light of

new research evidence. We believe that the results from this study with the visualisation prototype provides potentially more effective ways to both measure and manage the assessment process and that this innovation can also contribute to the documentation and feedback of the condition, benefiting both the clinical process and patient outcomes.

2. Methods and Procedures

The traditional ways of measuring and documenting AS and its progression in a particular patient employed linear scales and instruments in a paper-rich environment. These were often inaccurate and bureaucratic with limited patient inclusion and interaction. One clear alternative is to use digital representation rather than paper; however merely substituting video for paper alone is insufficient. This study sought to go further and to investigate how new technology could add value to the content of patient measurement records by creating more useful information including improved patient utility and feedback.

To overcome the limitations of the linear scales used for representing the functional ability of complex entities, visualisation technology was used here to create a 3-D virtual person (avatar) which was envisaged as being a major component and interface to a new tool. As noted above, the current forms of representation and measurement are unsatisfactory and the use of a 3-D avatar was to be tested as an alternative to the VAS. A decision was made to make the avatar a neutral actor representing the subject of care; as a photographic or 'life-like' video was found to be unhelpful. Poor self-esteem arising from a recognisable image when faced with a deteriorating condition was found to be discouraging and potentially offensive to the patient, as the avatar would capture and display to the patient's posture over time. The following describes the avatar as a 3-D virtual person.

The 3-D visualisation system comprises a 3-D virtual person and multi-functional interfaces providing feedback information in both static and dynamic 3-D form: the steps to develop this new form of assessment were as follows:

1) A 3-D virtual person was created to represent the subject who needs to be measured and assessed.

2) Motion data were collected from a healthy volunteer in a laboratory environment using sensors and cameras.

3) These motion data are used to drive the 3-D virtual person and to test the system's limitations. The animated avatar, using these data, was then the centre-piece of the prototype.

4) Healthcare professionals, specialising in AS, drawn both from practice and academia, evaluated this prototype with respect to future assessments.

2.1 Creation of a 3-D virtual person

The process began with the possibility of "creation" of a 3-D model, followed by "characterization" into Motionbuilder®. The process is illustrated by **Figure 1** and more fully explained in detail in *A visualisation tool for augmenting assessment of ankylosing spondylitis* [17] to show the potential of implementation into other environments. Since the "creation" and "characterization" of a 3-D model assists the research focus, rather than being central to the hypothesis being considered, it is only introduced here for context. A simplified 3-D "actor" was utilised to assist rather than expend too much time on constructing a more complicated 3-D model. However, the use of the file format "FBX" which is compatible with a range of 3-D modelling environments such as 3ds Max® would enable greater sophistication when used in real world applications.
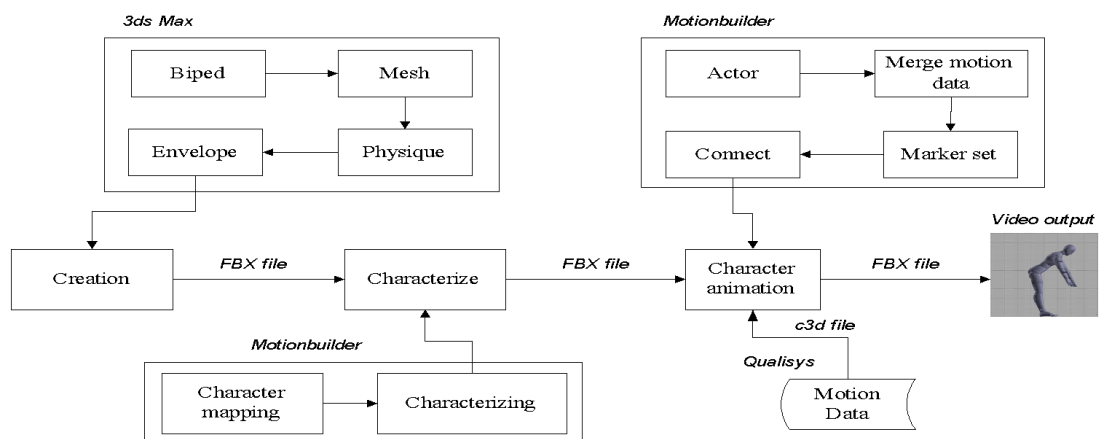

Figure 1. The process workflow. This figure shows the construction of a 3-D virtual person and where 3-D motion data can be used to drive the avatar.

2.2 Experiment collecting motion data

Mapping 3-D motion data to the virtual person required the pipelining of outputs from a set of software tools, including Qualisys® and Autodesk® Motionbuilder. Mapping motion data

5

required a healthy adult to undertake selected actions wearing markers, as shown in **Figure 2** and **Figure 3**.

Figure 2 Healthy volunteer wearing markers for capturing motion (Marker set A)

Figure 3 Image of whole body marker set (Marker set B)

2.2.1 3-D data collection

Equipment and calibration

The optical motion capture system Qualisys®, controlled by Windows-based data acquisition software (QTM), was used to collect motion data from a healthy adult. Qualisys® is a marker based technology that tracks the movement trajectories of retroreflective markers which are attached to a person and to make precise measurements and to deliver high quality 3-D data. The data was collected in a twin laboratory environment, and two sets of Qualisys® systems were used. One system comprised 12 Oqus^TM cameras in the organisation's gait lab whereas the other system used 6 Oqus^TM cameras within the usability lab. The latter setting with fewer cameras was used to test the possibility that a future development might be to collect data in a relatively limited space like a client's home. The main advantageous feature of the Oqus cameras is the ability to calculate marker positions with impressive accuracy and speed. Hundreds of markers can be measured at thousands of frames per second with ten, fifty or even

more cameras – and it can be run using an ordinary laptop. No workstation or switch is needed which makes the system easy to use in the field for example at rheumatology or physiotherapy clinics. Subjective assessment activities and physiotherapy exercises were performed in the gait lab, and objective assessment activities were collected in the usability room reflecting the different space and accessibility requirements.

The compatibility of data output formats was also considered so as to prepare for future applications. Prior to each data collection session, the cameras were calibrated using a T-shaped wand and an L-shaped reference structure in order to avoid measurement errors and incorrect integration of all measurement systems into the same timeframe.

Marker set

The data collection was conducted on a healthy adult. Whole body marker sets, including both single marker and marker clusters, were used to ensure compatibility with the following process and to provide rich information. Based on a whole body marker placement system that is used in the gait lab, two modified whole body marker sets (marker set A and marker set B) were used for data collection.

Marker set A (as shown in **Figure 2**) is for subjective assessment and exercise activity, and marker set model B (as shown in **Figure 3**) is for objective assessment. This is deliberate given that the aim of this work was not to compare efficiency between marker sets or to focus on a kinematic and kinetic study. The marker sets were calibrated by an experienced engineer who used a simple rule to capture efficient information without affecting the healthy adult's movements. Marker set model A was shared across subjective assessment and the National AS Society (NASS, http://nass.co.uk) exercises. Marker set model A includes single markers on the joints, i.e. wrist, elbow, shoulder, knee, ankle etc., and marker clusters on various segments, i.e. forearm, upper arm, thigh, leg, pelvis etc. These were appropriate points given the rheumatologic condition under study. Marker set model B was shared across objective assessment. Compared to model A, a smaller number of markers were used, as only single markers were set up. This slight change was made to fit the measurement purpose without sacrificing data quality. In this study, the impact of different marker set models on data capture is deliberately skipped because it falls off the research focus. And the quality of the data within Qualisys® manager was sufficiently accurate for our purpose.

2.2.2 3-D data process

The healthy adult was asked to complete the pre-selected activities step by step, whilst the 3-D motion data were captured and processed within QTM. Once the data were deemed satisfactory according to certain criteria, they were stored in the c3d file to be ready for output to the next step.

Previously captured 3-D motion data were imported to Autodesk® MotionBuilder® pro for further processing. This provided an environment where the imported data could be processed and used to drive the 3-D actor, therefore representing the activities of the healthy adult. Both subjective assessment data and objective assessment data in c3d format were imported, and mapped to a 3-D actor one by one, with files stored in FBX available for future output. Compatibility of 3-D data format is an important consideration especially when data transmission standards are lacking.

2.2.3 3-D data output

3-D video of a virtual person was chosen as the priority output format, because it was possible for visualisation to use low tech equipment, i.e. a standard pc, also because it provided more complete information as opposed to separate images, scales and charts that were not connected. The 3-D video of the virtual person was directly output from Motionbuilder® into "avi" or "mov" files. The virtual actor was utilised due to its neutral appearance, which was regarded as easy to view and convenient to compile.

2.3 A 3-D visualization prototype

The prototype presented a 3-D virtual person, partly as a means to assess the applicability of 3-D visualisation for AS, and partly because a 3-D virtual person presented an alternative to the current forms of representation and measurement, which were considered unsatisfactory. This had clear advantages over forms of assessment such as the simple VAS and also permitted dynamic recreation and representation of functional and exercise-related activities in a more realistic fashion; it had the added benefit that measurements for assessment could be completed automatically during the capture of activity. The Bath Indices are considered a gold standard for functional assessment and therefore became our main focus on current instruments in clinical practice. The BASFI was used to measure the functional ability; related mobility issues were measured by BASMI, which were selected as our focus. 3-D visualisation functions for a

range of activities are discussed in the following order: assessment activities, physiotherapy exercises and objective measurements.

2.3.1 Assessment activities visualization

For subjective assessments, 4 out of 10 questions in BASFI were simulated and successfully mapped to the 3-D virtual person. **Figure 4** shows snapshots of BASFI Q. 2 "bending forward to pick up a pen without bending your knees" in sequence as an example. Other questions successfully mapped to the 3-D avatar have been discussed elsewhere, including subjective assessments and objective measurements [18].
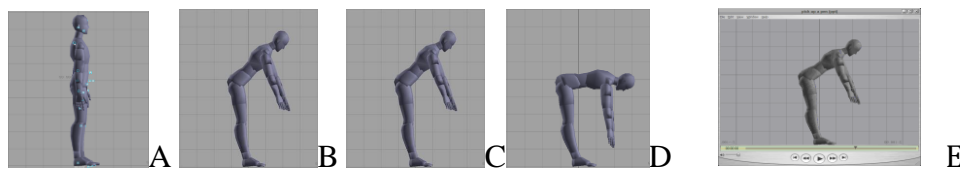

Figure 4 Snapshots of BASFI Q. 2 "bending forward to pick up a pen" in sequence (Image A with markers on; image B, C and D without Markers on; Image E video)

2.3.2 Physiotherapy exercises visualization

Q. 9 of BASFI was expanded by creating four different kinds of exercises: I. "Spinal side flexion"; II. "Spinal rotation"; III. "Neck rotation"; and IV. "Hamstring stretch". The snapshots of selected actions are shown in **Figure 5**.

| A, Q. 2 Bending forward to pick up a pen | B, Q. 4 Standing out of an armless chair | C, Q. 8 Looking over shoulder | D, Q. 9 I. Spinal side flexion |

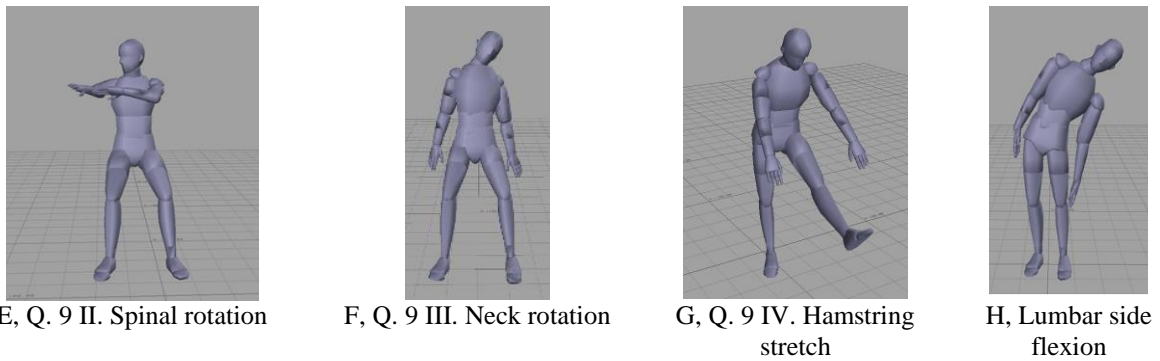| E, Q. 9 II. Spinal rotation | F, Q. 9 III. Neck rotation | G, Q. 9 IV. Hamstring stretch | H, Lumbar side flexion |
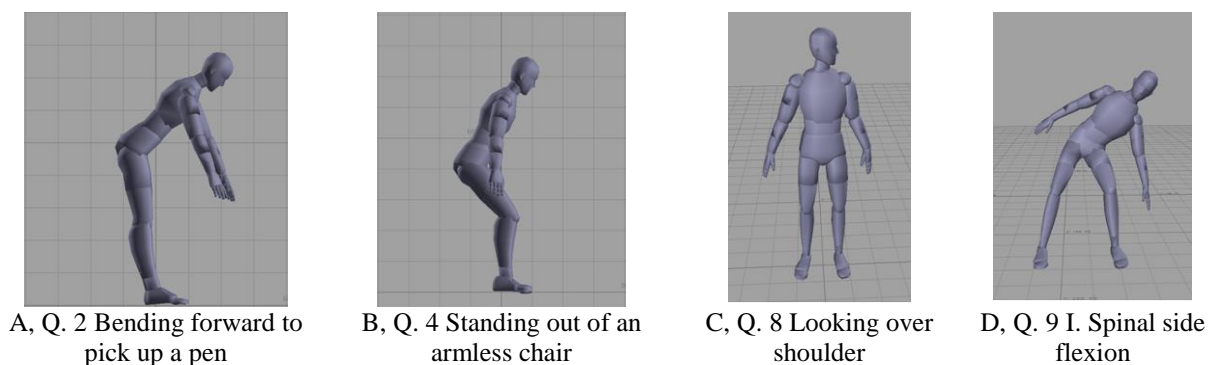
Figure 5 Summary of snapshots for selected actions. First two images (A and B) are from a side view; the others (C, D, E, F, G and H) are from a perspective view.

2.3.3 Objective assessment visualization

For the objective measurements of BASMI, two out of the five measurements were picked up and tested. They are: I. "Lumbar side flexion" (as shown in Figure 5, image H) and II. "Tragus to wall distance" measurements.

It was possible to integrate the three methods (an unadorned 3-D avatar; a 3-D avatar with markers on; and a video) into this one single tool, with the added advantage that automatic measurement generation was possible. For example, 'lumbar side flexion' measurement could be captured simultaneously at the very beginning of an action in BASFI Q.2, 'bending forward to pick up a pen'. This could be accomplished by simply applying an extra marker to each middle finger tip of the patient as the three axis value (X, Y, Z) were available through the cameras. Then the average 'lumbar side flexion' distance could be calculated automatically. The hitherto manual measurement and calculation could be removed from the treatment steps, and multiple tests could be completed by a single one via this automatic capture and generation courtesy of the tool.

2.4 Qualitative evaluation design

Focus groups were selected as the method of choice because the group's interaction could produce grounded data on the exact issues of this research.

Participants' recruitment

In order to minimize sample bias, participants were purposively recruited from a number of sources, such as different UK NHS (National Health Service) hospitals rather than from a single source since the goal was to represent a full spectrum of experiences and opinions. That having

been said the treatment of AS is a very specialist topic and the number of experts available meant that the groups were small, albeit well qualified. This homogeneity and the fact that it was a relatively close community allowed for more free-flowing conversations among participants within groups and also facilitated analyses that examined differences in perspective between groups [19].

The participants were recruited from two major groups of interested parties (the "panels"). In practice these overlap, i.e. individuals could be said to have experience in both panels. Therefore, for the most part, the data from each panel was not analyzed separately. Ten AS specialists participated and were allocated randomly into three groups for the evaluation. The first group was used as a pilot study. Only one participant attended the third group due to reasons beyond the control of the research team.

- **Panel A: Practitioners**. This panel included those from northwest UK NHS trusts, including primary care trusts (PCTs), who routinely conduct assessment and feedback to AS patients. Such individuals need to know how to assess the patient and how to interpret and feedback assessment results to the patient and other stakeholders.

- **Panel B: Academia.** This panel included those who have publications within the area of AS management. Such individuals are closely involved in the training and assessment of student physiotherapists yet also understand patients' requirements.

Evaluation execution

Groups were mixed by sex, race or age because none of the factors had been seen to affect the participants' comfort in the discussion or the analysts' ability to make useful comparisons. Participants recruited from different NHS sites had met occasionally on a Physiotherapy Networking Meeting. They were asked however not to talk to each other about the topic prior to the meetings.

The group meetings were held in a usability laboratory. An experienced moderator moderated the groups and an assistant also attended. An interview guide was used which ensured consistency between groups. However there was some heterogeneity in the order of questions between groups. Each group session lasted no longer than two hours and was recorded for subsequent analyses.

First, as part of the study, formal models of assessment and feedback developed were presented and discussed [20].

Then, the technique of considering a hypothetical patient at various stages in the assessment and feedback processes was used to design how the prototype could work, and also as a framework to elicit comments and create a focus for the discussions. The hypothetical patient (or persona) was selected as being a typical AS stereotype of a young person, who first knew about the condition in her late 20s or early 30s. The scenarios start from the first assessment at the onset of the condition and then the focus moves to later assessments made throughout the course of her life with the disease. Although feedback can be given along with the assessment process, in the model it is artificially separated from the assessment, as it is a significant area for the current research. It therefore has its own place in the four distinct scenarios. The persona and scenarios were approved in the pilot study. A number of logistical issues were raised during the pilot, and these were addressed in order to improve the internal validity of the evaluation. The details of persona and scenarios setting is as following:

Persona: Angie Stone; Female; 30 years old; Accountant

Assessment Scenarios

Assessment Scenario 1: 1st assessment by the physiotherapist

Ms Stone is diagnosed with AS and is referred to the physiotherapist. This is the first attendance, which is introductory in nature and sets a base-line.

Assessment Scenario 2: nth assessment by the physiotherapist

Ms Stone revisits the clinic and sees the same or maybe a different physiotherapist at the scheduled time. This scenario is about the subsequent visits by the patient for assessment as these differ from the very first. Subsequent assessments have one or more previous assessments to compare.

Assessment Scenario 3: prescription of exercises by the physiotherapist

A physiotherapist prescribes exercises to Ms Stone on the 1st visit or nth.

Assessment Scenario 4: documentation by the physiotherapist

After or during assessment of Ms Stone, a physiotherapist writes up her record on the 1st visit or nth.

Feedback Scenarios

Feedback Scenario 1: 1st assessment by the physiotherapist

Ms Stone visits the clinic for the first time, the assessment has been done, feedback information.

Feedback Scenario 2: nth assessment by the physiotherapist

Ms Stone revisits the clinic, the new assessment is done, feedback information again.

Feedback Scenario 3: previous and future exercises given by the physiotherapist

Ms Stone is asked about the exercises given to her.

Feedback Scenario 4: different aspects of AS from the physiotherapist

Ms Stone visits/revisits the clinic. After several assessments have been done, the physiotherapist feeds back information on all aspects of the condition and treatment.

In order to improve comparison and support the analysis process, the participants were asked to go through all the scenarios twice. Regarding the evaluation of the prototype, both the persona and scenarios were first used to capture data in procedures used in current practice. The participants in each group went through the scenarios for the first time to collect narratives of procedures that are being used in current practice. Then, the participants were shown our 3-D visualization prototype within the usability lab. The second time was for narratives of assessment and feedback with the prototype. The "funnel strategy" was conducted for the evaluation process, which meant each group began with a less structured approach that emphasized free discussion and then moved towards a more structured discussion of specific questions.

3. Evaluation Results

3.1 Data analysis

All data were recorded and transcribed, and the transcriptions have been cross-checked by another researcher prior to the analysis. The data generated from the three sessions, including from the pilot study, the focus group and the interview, were transcribed using Transana® 2.22 and analyzed using Nvivo® 8.0 as shown in **Figure 6**.
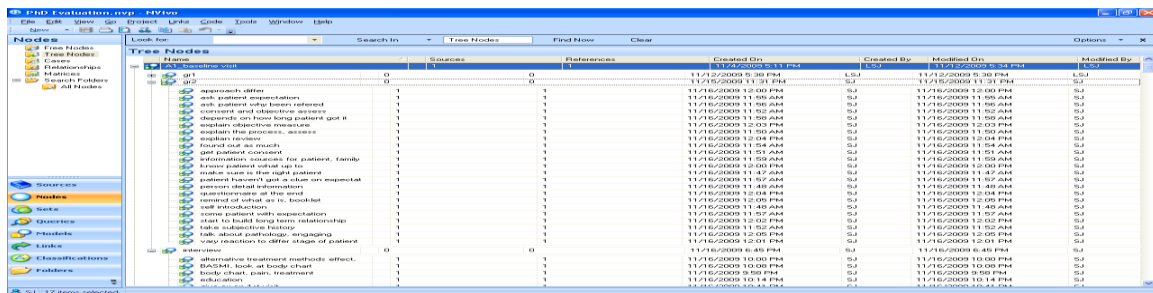
Figure 6 Snapshot of the Nvivo® 8.0 working interface with data

The interpretation and analysis of the data employed content-based thematic analysis which involved classifying and coding the content into themes [21]. The basic procedure in content analysis is to design categories that are relevant to the research purpose and to sort all occurrences of relevant words or other recording units into these categories [22]. Categories and themes are both generated from research emphasis and emerge from the data. Decisions about what counts as a category come from theory, literature, research experience and the data

13

itself. Each broad theme was then subjected to a more detailed manual analysis, which led to the formation of more specific categories within each theme. The framework of four scenarios for both assessment and feedback were used to elicit comment and create discussion for the evaluation. These became the general context of the evaluation that provides the environments of generating data and establishing themes.

3.2 Main outcomes and results

A summary of the narratives from the participants going through designed scenarios is presented as in **Figure 7**. This figure shows an overall view of current practice process depending on different sites and professionals. The process and content in **Figure 7** are within the scope of proposed models [20]. The assessment activity follows three phases: input (discover), process (learn), and output (perform). Both objective and subjective facts/results are fed back from the clinician to patients along with the process. Note the timing and frequency of the visits varied from hospital to hospital. The patient's first visit is usually to meet the physiotherapist, build the base-line from which the physiotherapy will manage the treatment. The 2nd or 3rd visit is a follow up and these happen within a relatively short time. Thereafter, regular visits will be scheduled but the time between visits will usually be longer than the earlier ones. These differences pose potential problems for tracking the progress of the condition in the individual and for comparing results across all AS patients.

Bath indices, including BASFI, BASMI and BASDAI, are mainly used for the assessment process. Exercise remains a key therapy, however, NASS exercises are recommended in all participants' hospitals except one, where there is a specific department generating exercise leaflets. There are some guidelines but what exercises are done depends on the patient but also the experience of the physiotherapist. Currently, there is no consensus about the most effective numbers of repetitions for exercises. What is shown in **Figure 7** was sent back to the participants for validation as part of the evaluation report was created from the participant discussions.
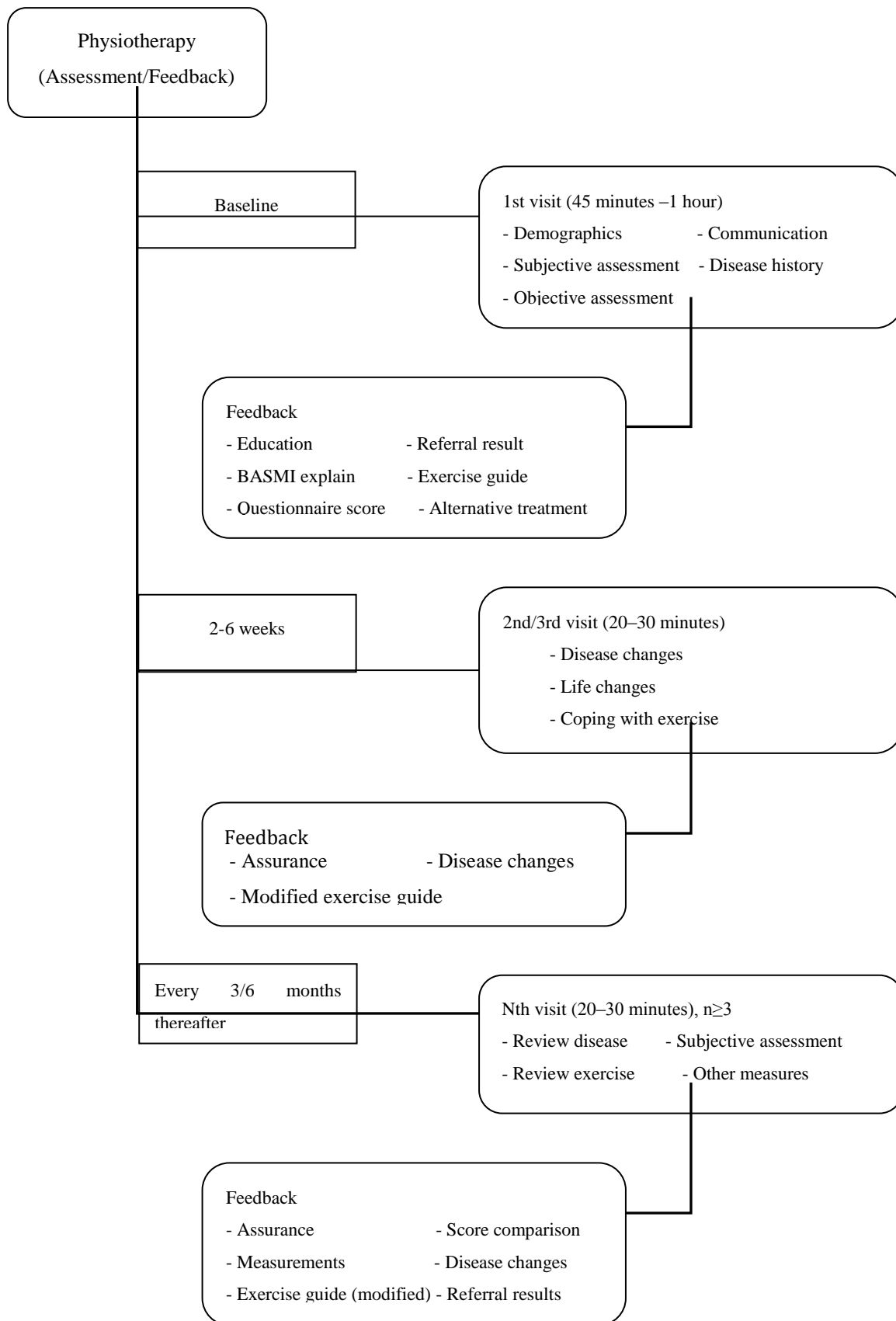
```
┌─────────────────────┐
│    Physiotherapy    │
│ (Assessment/Feedback)│
└─────────────────────┘

┌──────────────────┐        ┌─────────────────────────────────────────┐
│    Baseline      │────────│ 1st visit (45 minutes –1 hour)          │
└──────────────────┘        │ - Demographics        - Communication   │
                            │ - Subjective assessment  - Disease history│
                            │ - Objective assessment                   │
                            └─────────────────────────────────────────┘

┌─────────────────────────────────────────┐
│ Feedback                                 │
│ - Education          - Referral result   │
│ - BASMI explain      - Exercise guide    │
│ - Questionnaire score  - Alternative treatment│
└─────────────────────────────────────────┘

┌──────────────────┐        ┌─────────────────────────────────────────┐
│    2-6 weeks     │────────│ 2nd/3rd visit (20–30 minutes)           │
└──────────────────┘        │    - Disease changes                    │
                            │    - Life changes                       │
                            │    - Coping with exercise               │
                            └─────────────────────────────────────────┘

┌─────────────────────────────────────────┐
│ Feedback                                 │
│  - Assurance            - Disease changes│
│  - Modified exercise guide               │
└─────────────────────────────────────────┘

┌──────────────────┐        ┌─────────────────────────────────────────┐
│ Every  3/6 months│────────│ Nth visit (20–30 minutes), n≥3          │
│ thereafter       │        │ - Review disease    - Subjective assessment│
└──────────────────┘        │ - Review exercise   - Other measures    │
                            └─────────────────────────────────────────┘

┌─────────────────────────────────────────┐
│ Feedback                                 │
│ - Assurance            - Score comparison│
│ - Measurements         - Disease changes │
│ - Exercise guide (modified) - Referral results│
└─────────────────────────────────────────┘
```

Figure 7 Summary of current physiotherapy process in the UK

The themes provide a comprehensive overview of the interpretations of the participants' perceptions in terms of the current practice context and this prototype. For the primary research interest and a clearer path, these paralleled themes are gathered with respect to the limitations of current instruments, the perceptions in the proposed models, and the prototype. The themes focusing on the participants' views are detailed in **Table I**.

Table I Themes and categories from evaluation

Broader themes	Main categories		
The perceptions of current instruments' limitations	Limitations of the Bath indices and of VAS	Inconsistent measurement	Desire for computerized documentation
The perceived credibility on assessment for AS	Credibility coming from visualizing movements in 3-D	Potential credibility coming from visualizing pain in 3-D	Credibility of encouraging multi-disciplinary (MDT) working
The perceived credibility on feedback for AS	Credibility related to educating the patient	Credibility from capturing disease trend and status	Credibility of feedback information flow changes
The perceptions regarding physiotherapy exercises for AS	Physiotherapist prescribing exercises	Participants' anticipation of patients' compliance	Patient transforming exercises
Discussions and concerns about this prototype	Discussions on the form of this prototype	Concerns regarding cost	Concerns of communication changes

There was agreement from the participants on the limitations of current AS measurement instruments. The participants perceived that inconsistent measurement existed in the practice due to the varied measurement between individuals; the problems of the patient conceptualising VAS were confirmed according to their experience; and there was a desire from clinicians for computerised documenting of AS treatment process that the prototype would support and surpass.

"Like you might do for a patient's blood results …… you have a graphic showing this is the trend available in the last 6 years. And you can see how this has gone up and down. That

is what we hope to do with Bath scores in 6 months time. But, that is not happening. At the moment, I suppose the feedback is fairly informal …...." (Participant, Interview)

Secondly, the proposed assessment and feedback models were accepted by the participants as the assessment basis, which captured what they did in practice.

"It would be easier to show them visually, what score means, rather than 'cm' …… And that for them 'watch' for sure is easier." (Participant A, Focus Group)

"One thing comes into my mind is that it might help in that. You know we said this is a fluctuated disease, which depends on which bit of the day you catch them on. If the patient could somehow take average [measurements] over a week, do measurement everyday for a week. And then they come to see the physiotherapists, and say, 'look, this is what I've got'; downloading something into the physiotherapists' computer, that would give them indications." (Participant C, Pilot)

Thirdly, the participants perceived the prototype positively. The prototype was convincing. It has been perceived as novel and helpful in the visualisation capability of movements and even pain in a 3-D figure. This prototype was also perceived as beneficial in educating patients and capturing disease status and trend. And the information flow has been augmented within this prototype by placing the patient as the central actor within the process. Moreover, if this prototype could be used in practice, it may be possibly beneficial to standardise the objective assessment resulting in increased reliability and trust. Now exercise is used as a key physiotherapeutic tool but it is lacking evidence as to which one is better and whether or not the patient is complying. This prototype is thought to be a way to start to build the evidence base but that assumes it can be used directly by the patient in their home.

"There are three or four different exercises for 'lumbar spine rotation'. I think the patient can look at the booklet, and choose which one they want to do. So they would rather do one set …… So I think if it is a chronic condition, and long term exercise treatment is needed, I am happy the patient has initiated how they are going to perform their exercises." (Participant, Interview)

Finally, the 3-D visualisation prototype utilises a neutral figure to represent the patient; the participants believed that the neutral appearance choice is more acceptable than a simple video because the specialists believed that people do not like to see themselves especially when the progression of a condition results in a deterioration and/or deformation of their body. Although,

17

there are concerns regarding cost and communication changes, as technology is becoming less expensive over time and people are getting more engaged in technology, it is seen as something that might become readily available for patients to use at home and shared with the professionals. And it is thought of as a tool that would encourage multi-disciplinary team (MDT) communication in the feedback process.

> "……There isn't MDT communication between assessments, because we have patients to be assessed by us on a Thursday day, and the next Tuesday, they went to the nurses and the nurses would repeat the assessment. The patient is going like, 'this is done, and you've not got the information from the physios'. Maybe it is a more novel tool and MDT use would be encouraged." (Participant, Interview)

4. Discussion

We believe the prototype has shown that the current gold standards for AS assessment can be improved upon in ways not previously thought possible. Regardless, the prototype can a) provide both continuous visualisation and linkage to numeric scores, with full reflection of disease status range, and b) avoid the vagaries of any scoring system. Whatever the 'impossible' tag at the far end of the 10 cm VAS means, there are a range of visualisations for the same score of 'impossible' available in this prototype. This prototype does have the potential to take account of normal physical limitations and soft tissue. It would be necessary to manually adjust the markers' positions, although the measurement and 3-D visualization would be automatic.

4.1 Potentials of 3-D visualisation

The Bath indices are mainly patient reported outcome measures, and make heavy use of the VAS, a form of assessment that is of particular interest to this study. Given the "visual" aspects it was appropriate to consider 3-D approaches as requirements for this and future studies: 1) AS is part of the spondyloarthropathy conditions which are characterized by visible symptoms that are capable of being represented by visual means; 2) The assessment of symptoms currently relies on a VAS assessment, which is an early form of visualisation; 3) The recommended treatment of AS involves exercise as a treatment for mobility problems [23], and this too can be represented by visualisation techniques; 4) The current forms of assessment include multiple visual elements, but these are often separate and difficult to bring together in a coherent way.

This is because AS has multiple symptoms, such as stiffness, pain, fatigue that all have to be measured by different types of instruments.

There are additional advantages for focusing on 3-D visualisation, as it requires relatively low-tech, less expensive environments which make it practical for potential integration and possible implementation in clinical settings that are complex. It also makes it feasible to consider applying it to multi-problem chronic conditions in which the patient has to actively be involved in their life-long treatment. In the literature there still remains a paucity of evidence concerning application of visualisation techniques to chronic conditions. A further advantage of this approach, that was not tested, could stem from the attractiveness of similar technology to the sufferer who is often diagnosed during early adulthood [1] and is likely to be tech-savvy.

4.2 Meaningful feedback

Three types of output were available at this stage of the prototype's development: an unadorned 3-D avatar; a 3-D avatar with markers on; and accompanying video, as shown in **Figure 4**. 3-D video represented the main output method for the prototype; more detailed information regarding its 3-D functionalities, i.e. 3-D visualisation of subjective and objective assessment, its integration capability and meaningful feedback were integrated into this prototype as well.

BASFI is considered to be an example of Patient Reported Outcome Measures (PROMs). It is likely that more research will be commissioned in creating such constructs as their importance increases [24]. However, like many PROMs the information is uni-directional, i.e., flowing from the patient to the clinician, and often the end result/score is relatively abstract or perhaps requires too much clinical knowledge to have accessible meaning for the patient. Nevertheless, the objective measurement BASMI is completely executed by the clinician and is mainly for the clinician.

However, it is possible in the foreseeable future to envisage, with the avatar within this prototype being driven by movement data taken directly from patients, that what you see is what you measure (what is termed, WYSWYM). Rather than depending on patients' understanding of a scale and score, or their recall of their activities since the last assessment, the actual action would form part of a visual diary that could be captured and stored for future analysis, replay, reflection and comparisons. The avatar's performance is a more faithful representation than a VAS as it accurately shows the extent of the patient's progress, as well as the capability of presenting other facets of the condition at that time.

19

A big advantage of this approach is that the means of collecting the individual's AS health status is precisely the same one that can provide meaningful feedback directly to the patient with minimal mediation. Question 2 in BASFI, asks "Please indicate your ability with bending forward from the waist to pick up a pen from the floor without an aid". Leaving to one side whether or not this is a good thing to do without bending the knees, it is much easier and more natural for an assessor if the avatar's function enables direct assessment by simple comparison, rather than by transferring the degree of difficulty to a scale. Furthermore, this prototype adds objectivity to the measurement by removing inconsistencies that arise due to differences in subjective reporting, imprecise protocols and differences in the use of measurement devices. So awkward questions regarding the current BASFI score system such as, "Does a reduction of 20% from 8 to 6 in BASFI have as much significance as a reduction of two points from 6 to 4, or 4 to 2?" [25] can be avoided.

With this prototype, the feedback process can be modified as shown in **Figure 8**. The patient now becomes the central actor within this modified model. And the information flow is not uni-directional anymore, i.e. flowing from the patient to the clinician; rather the prototype permits a two-way partnership. The previously clinician-oriented result/scores are a by-product whereas movement is much less abstract and requires less clinical knowledge for the patient to understand their progress. The new information flow also allows more open interaction during the treatment process between members of the MDT. This model therefore supports the previous requirements for better communications throughout the evaluation process concerning both assessment and feedback.
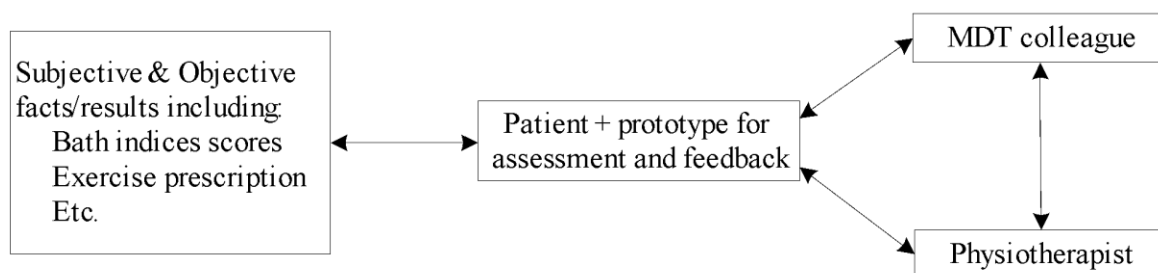

Figure 8 Meaningful feedback model

4.3 Limitations

A greater number of focus groups would have given extra confidence in the findings. The participants were recruited from the Northwest Physiotherapy Networking and most of them

showed great interest in this work. However, the requirement for participants to be experts in AS restricted their number as there are only a relatively few such experts in the country. Originally, the inclusion of patient trials was considered for the evaluation sessions. The patient's view is obviously the most critical but it would have been premature and perhaps therefore unethical to expose such patients to the different lab tests at this early stage of the prototypes development. We do think it is very likely that the dependence on markers and sophisticated laboratories will become a thing of the past for the patient's participation as advances in technology move these 'lab-based' prototypes experiments to home-friendly gaming consoles and/or combined with the increased use of sensors in the patient's home to track functions in daily, 'natural' living. However, there are lessons from this study that transcend any particular platform.

5. Conclusion

Formal assessment and feedback models for AS have been proposed and validated, and no such models have been reported in the current literature to our knowledge. These models provide concepts, which help to capture the information, albeit in a simplified form. These models provided the scope within which a prototype has been designed and developed so as to study the perceived value of 3-D visualisation technology in augmenting assessment and feedback for AS.

Although visualisation techniques have been deployed previously for individual symptoms (see [26,27]), this is the first attempt to integrate several assessment tools within a single visualisation tool. A 3-D virtual person has been designed as the main vehicle for presentation rather than relying upon abstract linear instruments. This is much clearer and believed to be consistent and robust especially with respect to presenting the actual exercise activity. The qualitative evaluation sessions consisting of participants from both academic and practice environments provided very positive feedback. The participants all stated that 3-D visualisation technology has great value in augmenting assessment and feedback of AS, and that they would readily use it should it become available in practice. Furthermore, it provided the potential means to obtain subjective assessments, provide multiple objective assessments and support exercise activity via one tool. A potential outcome of realising this research in practice is that such a tool can increase knowledge of AS by providing more accurate and comprehensive

assessments and more meaningful feedback over time. This has the potential to change the way the patient, clinicians and the service as a whole manage this chronic condition.

Acknowledgements

Disclosure statement

All authors declare no conflicts of interest. Ethical approval was granted by the Research Governance and Ethics Sub-Committee (RGEC) of the Research Committee at University of Salford, reference number RGEC08/057.

References

1. Calin A. Ankylosing Spondylitis. Medicine 2006; 34 (10): 396–400.

2. Walker J. Ankylosing Spondylitis. Nursing Standard 2006; 20 (46): 48–52.

3. Crichton N. Information point: visual analogue scale (VAS). Journal of Clinical Nursing 2001;10 (5): 706.

4. Wewers ME, Lowe NK. A critical review of visual analogue scales in the measurement of clinical phenomena. Research in Nursing and Health 1990;13:227–36.

5. Calin A, Garret SL, Jenkinson TR, Whitelock HC et al. A new approach to defining functional ability in ankylosing spondylitis: the Bath Ankylosing Spondylitis Functional Index (BASFI). J Rheumatology 1994; 2:2286-91.

6. Dougados M, Gueguen A, Nakache J et al. Evaluation of a functional index and an articular index in ankylosing spondylitis. J Rheumatology 1988; 15: 302–7.

7. Jenkinson T, Mallorie P, Whitelock H, Kennedy L, Garrett S & Calin A. Defining spinal mobility in ankylosing spondylitis (AS): the Bath AS Metrology Index. J Rheumatology 1994; 21 (9): 1694–8.

8. Eyres S, Tennant A, Kay L, Waxman R, Helliwell P. Measuring disability in ankylosing spondylitis: comparison of Bath Ankylosing Spondylitis Functional Index with Revised Leeds Disability Questionnaire. J Rheumatology. 2002; 29:979–86.

9. Taylor H et al. The relationship of clinical and laboratory measurements to radiological change in ankylosing spondylitis. Br J Rheumatology 1991; Vol. 30: 330–5.

10. Kennedy LG et al. Ankylosing spondylitis: the correlation between a new metrology score and radiology. British Journal of Rheumatology 1995; Vol. 34: 767–70.

11. Satava RM, Jones SB. Chapter 55. Medical Applications of Virtual Reality. In: VR Handbook. 1999.

12. Parsons T, Silva T, Pair J, Rizzo A. Virtual environment for assessment of neurocognitive functioning: virtual reality cognitive performance assessment test. Stud Health Technol Inform. 2008; 132: 351–356.

13. Kang YJ, Ku J, Han K, Sun IK, Yu TW, Lee JH, Park CI. Development and clinical trial of virtual reality-based cognitive assessment in people with stroke: preliminary study. CyberPsychology & Behavior 2008; 11 (3): 329-339.

14. Oostema J, Abdel M, Gould J. Time-efficient laparoscopic skills assessment using an augmented-reality simulator. Surg Endosc. 2008; 22 (12): 2621–2624.

15. Di Diodato L, Mraz R, Baker S, Graham S. A haptic force feedback device for virtual reality-fMRI experiments. IEEE Trans Neural Syst Rehabil Eng. 2007; 15 (4): 570–576.

16. Melanie Tory, Arthur E. Kirkpatrick, M. Stella Atkins, Torsten Möller. Visualization Task Performance with 2D, 3D, and Combination Displays. IEEE Transactions on Visualization and Computer Graphics. 2006; 12(1): 2-13.

17. Li S, Kay S, Porter S. A visualisation tool for augmenting assessment of ankylosing spondylitis. Proceedings of The Medical Informatics Europe (MIE): medical informatics in a united and healthy europe. Sarajevo, Bosnia and Herzegovina 2009: 821–5.

18. Li S, Kay S. Transforming assessment and feedback for ankylosing spondylitis through the use of 3-dimensional visualisation. Proceedings of IADIS International Conference: e-Health 2009. Algarve, Portugal 2009: 175–80.

19. Clifford C. Qualitative research methodology. In: Nursing and healthcare. New York: Churchill Livingstone, 1997: 61

20. Li S, Kay S, Hardicker N. Virtual reality: towards a novel treatment environment for ankylosing spondylitis. Proceedings of Medical and Care Compunetics 4. Amsterdam, Netherland 2007: 190–8

21. Tesch R. Qualitative Research: Analysis Types and Software Tools. London: Falmer, 1990.

22. Morgan D L. Designing focus group research. In: Stewart M, Tudiver F, Bass M J, Dunn E V, Norton P G, ed. Tools for primary care research. Newbury Park, CA: Sage, 1992:194–208.

23. Nass. NASS Guidebook for patients: a positive response to ankylosing spondylitis, 2009.

24. Prom Group. Chronic conditions-a structured review. In: Report to the Department of Health. Oxford: UHCE University of Oxford, 2006.

25. Calin A. Defining outcome in ankylosing spondylitis: Where have we been, where are we and where do we go from here? CPD Rheumatology 2001; 2: 77-80.

26. Hoffman H, Richards T, Coda B, Bills A, Blough D, Richards A, Sharar S. Modulation of thermal pain-related brain activity with virtual reality: evidence from fMRI. Neuroreport 2004; 15 (8): 1245–1248.

27. Krijn M, Emmelkamp P, Olafsson R, Schuemie M, Van Der Mast C. Do self-statements enhance the effectiveness of virtual reality exposure therapy? A comparative evaluation in acrophobia. Cyberpsychology & Behavior : the impact of the Internet, multimedia and virtual reality on behavior and society 2007; 10 (3): 362–370.