Development of an assessment tool for repetitive tasks of the upper limbs (ART)

Prepared by Health and Safety Laboratory for the Health and Safety Executive 2009
Development of an assessment tool for repetitive tasks of the upper limbs (ART)

Jeremy Ferreira
Mike Gray
Leanne Hunter
Matthew Birtles
David Riley

Health and Safety Laboratory
Harpur Hill
Buxton
Derbyshire
SK17 9JN

The prevention, control and management of work-related musculoskeletal disorders (WRMSD) are a priority for the Health and Safety Executive (HSE) and one of the key ways of improving occupational health in Great Britain. Health and Safety Executive (HSE) and Local Authority (LA) Inspectors play an important role in preventing WRMSD. As well as enforcing health and safety law, they provide advice on risk factors and control measures on a range of health and safety issues. To support this work, assessment tools are required that offer an intuitive and relatively quick process to screen workplaces for high-risk activities, raise awareness of risk factors, demonstrate the presence of risk, and recommend areas for improvement.

The Manual handling Assessment Charts tool (MAC, Monnington et al., 2002) has been recognised as a useful process to screen workplaces for high-risk manual handling operations (Melrose et al., 2006; Lee and Ferreira, 2003; Tapley, 2002). The MAC tool is designed to identify and help assess high-risk manual handling operations (lifting, lowering, carrying and team handling). However, its scope is limited to manual handling of heavy items, which primarily present a risk of lower back disorders. There was a demand for a similar tool that Inspectors can use to screen more frequent handling of light loads or other repetitive tasks and the common physical risk factors in work that may contribute to upper limb disorders (ULDs). This led to the development of the assessment of repetitive tasks (or ART) tool.

This report and the work it describes were funded by the Health and Safety Executive (HSE). Its contents, including any opinions and/or conclusions expressed, are those of the authors alone and do not necessarily reflect HSE policy.
ACKNOWLEDGEMENTS

The authors wish to acknowledge the support and contribution of colleagues in policy and operational directorates who assisted and participated in the user trials. We also acknowledge the contributions of colleagues in human factors and ergonomics teams who played an active part in the workshops and development of ART.

We are also grateful for the constructive contributions made by other ergonomists during the initial consultation exercise:

- Margaret Hanson, WorksOut
- Clare Lawton, Ergonomics and Safety Research Institute, Loughborough University
- John Ridd, JRP Ergonomics
- Glyn Smyth, Work Fit
# CONTENTS

1 INTRODUCTION............................................................................................................. 1
  1.1 Aims and Objectives................................................................................................. 1

2 CONCEPTUALISATION OF THE ASSESSMENT TOOL........................................... 2
  2.1 Design Criteria for the Assessment Tool................................................................. 2
  2.2 Initial Concept of the Assessment Tool................................................................. 2
  2.3 Selection of Risk Factors and Criteria........................................................................ 4
  2.4 Development of the Scoring System ......................................................................... 10
  2.5 Requirements to Use the Assessment Tool............................................................. 11

3 USER TRIALS ............................................................................................................. 14
  3.1 Overview.................................................................................................................. 14
  3.2 Simulated Trials – Methodology .............................................................................. 15
  3.3 Participants............................................................................................................... 18
  3.4 Simulated Trials – Phase One Findings ................................................................... 18
  3.5 Simulated Trials – Phase Two Findings ................................................................... 19
  3.6 Field Trials............................................................................................................... 27
  3.7 Summary.................................................................................................................. 31

4 BENCHMARKING ........................................................................................................ 32
  4.1 Approach to the Benchmarking .............................................................................. 32
  4.2 Results...................................................................................................................... 34
  4.3 Summary.................................................................................................................. 38

5 INITIAL CONSULTATION .................................................................................... 40
  5.1 Approach to the Consultation ................................................................................. 40
  5.2 Consultation Findings.............................................................................................. 40

6 CONCLUSIONS ......................................................................................................... 46
  6.1 Conceptualisation .................................................................................................. 46
  6.2 User Trials .............................................................................................................. 46
  6.3 Benchmarking ........................................................................................................ 47
  6.4 Initial Consultation ................................................................................................. 47

7 APPENDIX A – ART USABILITY QUESTIONNAIRE ............................................. 48

8 REFERENCES ............................................................................................................. 53
EXECUTIVE SUMMARY

Objectives

In January 2007, the Health and Safety Executive (HSE) presented the Health and Safety Laboratory (HSL) with a prototype of a tool for the risk assessment of repetitive tasks of the upper limbs. The tool was named ART (Assessment of Repetitive Tasks). The aims of the project were:

(1) To develop the tool and associated training material to the point where it is sufficiently credible and usable to be disseminated to Inspectors to pilot in an inspection campaign.

(2) To confirm that the tool is capable of providing reasonably reliable, accurate and meaningful results when used by Inspectors.

This report outlines the work on the tool, from its conceptualisation and early development, through user testing, refinement, benchmarking and initial expert consultation, up to its pilot release to a selection of HSE and Local Authority (LA) Inspectors during an inspection campaign.

Main Findings

The Assessment of Repetitive Tasks (ART) was developed as an Inspectors’ tool that can be used:

(1) To screen repetitive tasks of the upper limbs for common physical risk factors that contribute to the development of musculoskeletal disorders;

(2) To raise duty holders’ awareness and understanding of the risks of repetitive tasks;

(3) To demonstrate the presence of risk to duty holders;

(4) To give a broad indication of the level of risk; and

(5) To recommend areas for improvement.

The technical content of ART draws upon earlier work to develop the Occupational Repetitive Actions methods (OCRA; Colombini et al., 2002) and the Quick Exposure Check (David et al., 2008). The format of ART is based upon the Manual handling Assessment Charts (MAC; Monnington et al., 2002), and uses a logical flowchart process and ‘traffic-light’ system to grade risks. As a result, ART examines twelve risk factors that have been grouped into four stages:

(1) Frequency and repetition of movements;

(2) Force;

(3) Awkward postures (of the neck, back, shoulder/arm, wrist and hand); and

(4) Additional factors (which include aspects of task duration, recovery, perceived work pace and other object and work environment factors).
ART was developed using an iterative design process, where results and suggestions from repeated user trials and peer-review exercises were used to inform further improvements to the tool and associated training material (e.g. presentations and case studies). The main findings from each stage of the project are presented below.

**User Trials**

Thirty-two HSE and LA Inspectors participated in simulated user trials, which used video case studies to explore the concept of ART with users and to make improvements to the usability of ART. The observation-based assessment of quick hand and arm movements is inherently difficult. Nonetheless, the usability of ART was perceived to be favourable, and improved with successive prototypes of the tool. LA Inspectors reported that they would not use ART as frequently as HSE Inspectors, possibly due to the type of premises enforced by local authorities, and because ART is not intended for use in DSE assessments.

Experience suggests that an element of training and regular use is required in order for Inspectors to use ART reliably and to arrive at broadly similar results to Specialists (i.e. HSE/HSL ergonomists who contributed to the project). With training available, ART and its associated training materials were deemed to be sufficiently credible and usable to pilot with a selection of trained HSE and LA Inspectors during an inspection campaign. An evaluation of the extent to which training might improve reliability was beyond the scope of this study.

**Benchmarking**

A consensus of Specialists’ scores for ten tasks were used to benchmark ART against three other methods already available for assessing repetitive tasks: the Quick Exposure Check (David *et al.*, 2008), the OCRA checklist (Colombini *et al.*, 2002) and the Strain Index (Moore and Garg, 1995). When using ART, it was possible to prioritise tasks and arrive at risk levels that were in broad agreement to the other methods. There was greater agreement between the methods for the three tasks with the highest risk rankings.

Compared to the OCRA checklist and Strain Index, ART may underestimate the level of risk for some tasks, particularly those involving static exertions of the hand and awkward shoulder postures. Following further information from a pilot of ART, it may be appropriate to re-visit the weighting of some factors, the proposed exposure levels, and how ART can be adapted to assess static exertions of the hand. ART has not yet been validated as a predictor of ULD incidences.

**Initial Consultation**

A consultation exercise was held to seek an initial view on the suitability and usability of ART from four experts who are not employed by HSE/HSL but are often asked to apply, interpret, provide expert opinion on and train others in the use of HSE’s assessment tools and guidance material related to ergonomics. The general view on the suitability of ART was favourable, as long as a cautious approach was taken when interpreting the findings of the tool.

It was suggested that ART would be a useful instrument if released to others with responsibility for the design, inspection, assessment and organisation of repetitive work. It was recognised that the easiest way to reduce exposure scores obtained from ART is to reduce duration, and most companies will achieve this through job rotation. It will therefore be important to explain how ART can be used to assess task rotation and promote good task rotation.
1 INTRODUCTION

The prevention, control and management of work-related musculoskeletal disorders (WRMSD) are a priority for the Health and Safety Executive (HSE) and one of the key ways of improving occupational health in Great Britain. Health and Safety Executive (HSE) and Local Authority (LA) Inspectors play an important role in preventing WRMSD. As well as enforcing health and safety law, they provide advice on risk factors and control measures on a range of health and safety issues. To support this work, assessment tools are required that offer an intuitive and relatively quick process to screen workplaces for high-risk activities, raise awareness of risk factors, demonstrate the presence of risk, and recommend areas for improvement.

The Manual handling Assessment Charts tool (MAC, Monnington et al., 2002) has been recognised as a useful process to screen workplaces for high-risk manual handling operations (Melrose et al., 2006; Lee and Ferreira, 2003; Tapley, 2002). The MAC is designed to identify and help assess high-risk manual handling operations (lifting, lowering, carrying and team handling). However, its scope is limited to manual handling of heavy items, which primarily present a risk of lower back disorders. There was a demand for a similar tool that Inspectors can use to screen more frequent handling of light loads or other repetitive tasks and the common physical risk factors in work that may contribute to upper limb disorders (ULDs).

1.1 AIMS AND OBJECTIVES

The project began in January 2007, when the Health and Safety Executive (HSE) presented the Health and Safety Laboratory (HSL) with an original prototype of a tool for the assessment of repetitive tasks of the upper limbs. The tool was named ART (Assessment of Repetitive Tasks).

The project aims were:

(1) To develop the tool and associated training material to the point where it is sufficiently credible and useable to be disseminated to Inspectors to pilot in an inspection campaign.

(2) To confirm that the tool is capable of providing reasonably reliable, accurate and meaningful results when used by Inspectors.

To achieve these aims, the following objectives were proposed:

(1) To hold iterative user trials with Inspectors to test the usability of the tool and guide improvements to successive prototypes of the tool.

(2) To consult with HSE ergonomics specialists and external ergonomics practitioners and obtain feedback on the technical worth of the tool.

(3) To develop briefing material that can be used to deliver training to Inspectors on the use of the tool.

This report outlines the work on the tool, from its conceptualisation and early development, through user testing, refinement, benchmarking and initial expert consultation, up to its pilot release to a selection of HSE and LA Inspectors during an inspection campaign.
2 CONCEPTUALISATION OF THE ASSESSMENT TOOL

2.1 DESIGN CRITERIA FOR THE ASSESSMENT TOOL

Design criteria for Inspectors’ WRMSD assessment tools were originally specified for the development of the MAC tool (Monnington et al., 2002). These criteria were assumed to remain important for an assessment tool for repetitive tasks, which similarly was intended for HSE and LA Inspectors:

- The tool should be quick and easy to use (e.g. few pages and intuitive design).
- The tool must link to scientific studies and guidance on upper limb disorder risk assessment, particularly that published by HSE (e.g. HSG60, *Upper limb disorders in the workplace* (HSE, 2002)).
- The tool should intuitively indicate good practice for the design of repetitive tasks.
- The tool must be able to identify high risk repetitive tasks.

2.2 INITIAL CONCEPT OF THE ASSESSMENT TOOL

Several methods exist for the assessment of repetitive work tasks (David et al., 2008; Moore and Garg, 1995; Colombini et al., 2002; HSE, 2002; ACGIH, 2001; McAtamney and Corlett, 1993, Keyserling et al., 1993). Rather than develop an entirely new assessment method for repetitive tasks, the intention of ART was to make aspects of the existing tools accessible to health and safety Inspectors.

Figure 1 depicts the initial concept and process behind the development of ART. It illustrates the series of workshops held amongst the project team and wider HSE/HSL Ergonomics Team to develop a prototype that was suitable to trial with Inspectors. The project team workshops involved expert appraisal and allowed gains in usability to be made more quickly and efficiently than during user trials.

2.2.1 Technical Content

The initial concept was to base the technical content of the tool on the Occupational Repetitive Action (OCRA) methods, developed by Colombini et al. (2002) and incorporated into BS ISO 11228-3 (2007): *Ergonomics – Manual handling – Part 3: Handling of low loads at high frequency*. In particular, the OCRA checklist offered promise as an initial framework, as it was simpler to apply than the OCRA Index and more suitable for an initial screen of repetitive tasks (Occhipinti and Colombini, 2005). Thus, initial work focussed on careful examination and simplification of the OCRA checklist statements and ensuring the linkage to existing HSE guidance on upper limb disorders in the workplace (HSG60; HSE, 2002). An early peer-review workshop with ten HSE/HSL ergonomists indicated that the usability of the tool and its linkage to existing HSE guidance would benefit from applying elements of the Quick Exposure Check (QEC, David et al., 2008).
The result is that ART examines twelve risk factors that have been grouped into four stages:

(1) Frequency and repetition of movements;
(2) Force;
(3) Awkward postures (of the neck, back, shoulder/arm, wrist and hand); and
(4) Additional factors (which include aspects of task duration, recovery, perceived work pace and other object and work environment factors).

The selection and assessment criteria for the twelve risk factors are discussed further in Section 2.3.
2.2.2 Format

The initial concept also included developing a process that was similar in format and appearance to the MAC (Monnington et al., 2002). This was believed to offer an advantage, as most Inspectors, when learning how to use the tool, would already be familiar with the underlying process. A sequential flowchart process is used to guide users through the assessment in a logical way, while noting categories of risk for each factor in the flowchart. A ‘traffic-light’ system is used to grade risks, with green representing a lower level of risk and red representing a higher level of risk. The flowchart is supported with an assessment guide, which provides general instructions on the use of the tool and more detailed information on how to categorise each factor. There is also a double foldout worksheet containing forms (e.g. a task description form and score sheet) to record information collected as part of the assessment.

2.3 SELECTION OF RISK FACTORS AND CRITERIA

This section describes the selection of the twelve risk factors included in ART at the time of its pilot (February 2008), as well as the assessment criteria for the risk categories. Where the factor being assessed is believed to fall between two risk categories, it is acceptable for users to record an intermediate score.

2.3.1 Stage A – frequency / repetition

Frequent repetitive movements present a risk of injury and should be avoided within a task (BS ISO 11228-3, 2007). The risk of injury increases as the work cycle time decreases, although the overall level of risk can vary depending on other factors, including the movement pattern and the individual.

2.3.1.1 Frequency of arm movements

Examination of the simplified OCRA checklist indicated that there was considerable potential to misclassify the frequency of arm movements, particularly where users had difficulties identifying the sequence and number of technical actions in a task. In addition, Li and Buckle (1999) have raised the concern that, for many tasks, a fixed work cycle either may not exist or may vary periodically, making it difficult to assess the representative frequency of arm movements in a short period of observation. They suggested that arm movements should be categorised by the pattern of movement rather than the number of times the arm is moved within a given period. Thus, when developing the QEC, Li and Buckle (1999) proposed subjective categories for assessing the frequency of arm movements. These were applied to ART and are shown in Table 1.

<table>
<thead>
<tr>
<th>Good</th>
<th>Reasonable</th>
<th>Poor</th>
</tr>
</thead>
<tbody>
<tr>
<td>Infrequent (e.g. some intermittent movement)</td>
<td>Frequent (e.g. regular movement with some pauses)</td>
<td>Very frequent (e.g. almost continuous movement)</td>
</tr>
</tbody>
</table>

2.3.1.2 Repetition of movements

Highly repetitive tasks have been described as those with a work cycle time of less than 30 seconds or with more than 50% of the cycle time involved in performing the fundamental cycle or same motion pattern (Silverstein, 1986). Yet again, where the work cycle time may vary, it can be difficult to assess task repetitiveness over a short time period. Instead, repetition can be
assessed according to perceived movement rates (i.e. the number of similar motion patterns that are observed to repeat over a fixed period of time). Kilbom (1994) suggested that movement rates of the forearm and wrist of 10 times per minute increased risk when undertaken for more than 1 hour. From this guideline, Li and Buckle (1999) proposed three categories for assessing the repetition of forearm, wrist and hand movement. These were applied to ART and are shown in Table 2. ART users are expected to first identify the similar motion patterns of the forearm, wrist and hand (but, notably not the fingers) and then count the number of times that these motion patterns are repeated over a set period of time (e.g. 1 minute).

**Table 2** Risk categories for repetition of forearm, wrist and hand movements (adapted from the QEC)

<table>
<thead>
<tr>
<th>Good</th>
<th>Reasonable</th>
<th>Poor</th>
</tr>
</thead>
<tbody>
<tr>
<td>Similar motion pattern repeated</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10 times per minute or less</td>
<td>11 – 20 times per minute</td>
<td>more than 20 times per minute</td>
</tr>
</tbody>
</table>

These categories for frequency of arm movements and repetition of wrist / hand movements were previously shown to be practical and reliable (David *et al.*, 2008) and were previously included in HSG60 (Graves *et al.*, 2002). In terms of risk reduction and control, these categories are useful as they can take into account the use of job enlargement to provide increased variety of arm movements and speed of work.

2.3.2 Stage B – force

The aim was to develop an illustrative approach, which would allow users to assess the level of force exerted with one hand as well as the duration of the force exertion within a representative work cycle (Figure 2). ART offers two methods to assess the level of force exerted with one hand. The first and preferred method seeks out the subjective perceptions of workers; the second method is based upon user observations.

**Figure 2** Force graph
The first method requires users to ask workers if there are any actions that require muscle effort of the arms, hand or fingers. If workers do perceive actions to involve muscle effort, the users are directed to ask the operator to describe the level of individual hand force involved in each action as either ‘light’, ‘moderate’, ‘strong’ or ‘very strong’. Colombini et al. (2002) point out that posing the questions in this manner is important because workers may confuse muscle effort with the overall tiredness that they feel towards the end of the shift.

Where multiple workers perform the same task, users should interview as many workers as possible to obtain a more reliable assessment of the force involved. To improve the accuracy and validity of the assessment further, expert users can employ the Borg Category Ratio-10 Scale to help workers describe the degree of muscle exertion perceived in a given body segment (Borg, 1998). Table 3 proposes CR-10 ratings that correspond to each description of the level of individual hand force. However, in its current form, ART does not include a CR-10 scale.

Table 3: Borg CR-10 scale ratings and force descriptions

<table>
<thead>
<tr>
<th>Hand activity requires</th>
<th>CR-10 Ratings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Light force</td>
<td>&lt; 2.5</td>
</tr>
<tr>
<td>Moderate force</td>
<td>2.5 – 4</td>
</tr>
<tr>
<td>Strong force</td>
<td>5 – 6</td>
</tr>
<tr>
<td>Very strong force</td>
<td>7 – 10</td>
</tr>
</tbody>
</table>

In some circumstances, there may be language barriers as well as barriers in the physical, organisational and social environment, that prevent users and operators from communicating effectively together about the level of hand force involved in the tasks. Thus, the second method allows users to estimate the level of force exerted with one hand using observations; for example, with reference to the weight of objects handled, characteristics of the objects and the manner in which objects are gripped. The assessment guide provides some examples of actions that might require more than a minimal amount of muscle effort (e.g. pinching or gripping objects, moving levers and buttons, manipulating lids and components, pushing or forcing items together and using tools). For three of the four force categories, the graphical approach also includes levels of individual hand force: low – less than 1 kg; moderate force – 1 kg to 4 kg; and high force – more than 4 kg. These criteria were originally constructed during the development of the QEC to help workers assess the maximum force that they exert with one hand (Li and Buckle, 1999b). While convenient for the purposes of risk assessment, these criteria should be interpreted with caution. The criteria were based on a proposal that estimated _average_ hand force requirements of more than 4 kg be considered high force jobs, while requirements of less than 1 kg be considered low force jobs (Silverstein _et al._, 1986). In addition, Colombini _et al._ (2002) caution that the use of force can rarely be observed and where force is assessed by an external observer there is potential for error.

2.3.3 Stage C – awkward postures

The aim was to develop a process similar to the OCRA checklist that involved a check of both the presence of an awkward posture as well as the timing and duration of the posture. From the OCRA checklist, the tool includes an analysis of shoulder/arm, wrist and hand/finger grip postures. Further analysis of head/neck and back postures are also included, with the aim of expanding the scope of the tool to screen for a wider range of workstation design problems. Table 4 shows the risk categories and awkward posture criteria for posture assessment.
### Table 4 Risk categories and awkward posture criteria for posture assessment

<table>
<thead>
<tr>
<th>Body segment or joint</th>
<th>Awkward posture criteria</th>
<th>Good</th>
<th>Reasonable</th>
<th>Poor</th>
</tr>
</thead>
<tbody>
<tr>
<td>Head / neck</td>
<td>An obvious angle between the neck and back can be observed as a result of performing the task</td>
<td>The head / neck is in an almost neutral posture</td>
<td>The head / neck is bent or twisted a part of the time</td>
<td>The head / neck is bent or twisted more than half the time</td>
</tr>
<tr>
<td>Back</td>
<td>More than 20° of twisting or bending is observed</td>
<td>The back is in an almost neutral posture</td>
<td>The back is bent forward, sideways or twisted a part of the time</td>
<td>The back is bent forward, sideways or twisted more than half the time</td>
</tr>
<tr>
<td>Shoulder / arm</td>
<td>The elbow is raised to around chest height and the arm is unsupported</td>
<td>The elbow is kept close to the body or supported</td>
<td>The elbow is raised away from the body a part of the time</td>
<td>The elbow is raised away from the body more than half the time</td>
</tr>
<tr>
<td>Wrist</td>
<td>An obvious wrist angle can be observed</td>
<td>The wrist is almost straight / in a neutral position</td>
<td>The wrist is bent or deviated a part of the time</td>
<td>The wrist is bent or deviated more than half the time</td>
</tr>
<tr>
<td>Hand / finger grip</td>
<td>The hand or fingers hold objects in a pinch or wide finger grip</td>
<td>The hand or fingers hold objects in a power grip or do not grip awkwardly</td>
<td>The hand or fingers hold objects in a pinch or wide finger grip for a part of the time</td>
<td>The hand or fingers hold objects in a pinch or wide finger grip for more than half the time</td>
</tr>
</tbody>
</table>

Li and Buckle (1999) have suggested that practitioners prefer the use of descriptive words over specific postures quantified in degrees. It is also unlikely that users would have access to posture measuring instruments when in the field. Thus, for the neck, arm and wrist postures, descriptive criteria were used to define awkward postures, and these were supported with illustrations showing examples of awkward postures.

Where awkward postures are not present, the factor is coded as green. However, where awkward postures are present, users then consider the duration of the awkward posture or movement. For each body segment, the criteria for the amber category are described consistently as either moving to an awkward posture repetitively or holding an awkward posture for ‘a part of the time’. The Department for Trade and Industry’s document entitled ‘Instruction for consumer products’ (DTI, 1989) suggests that the phrase ‘a part of the time’ would mean 15 – 35% of the time to the majority of people. The criteria for the red category are described consistently as moving to an awkward posture repetitively or holding an awkward posture for ‘more than half the time’. In the event that the user remains undecided between the categories, or subsequent video analysis reveals an awkward posture to be adopted for either less than 15% or 36 – 50% of the time, then an amber coding and selection of an intermediate score would be acceptable.

#### 2.3.4 Stage D - additional factors

This stage considers further aspects important to the assessment of repetitive tasks, including the opportunity for recovery, the worker’s perceived workload and other psychosocial factors, object and environmental factors, and the duration of the repetitive task.
2.3.4.1 **Breaks**

It is important that work be organised in such a way that there is sufficient opportunity for recovery. A factor was included in ART to ensure users consider the time that workers perform the repetitive task without a break. This may require information on the frequency, timing and duration of both structured breaks in the repetitive work as well as time spent performing other tasks that do not involve similar repetitive arm movements. A timeline is provided for users to record this information.

BS ISO 11228-3 (2007) suggests a ratio of at least 1:5 between recovery and task duration. Thus, ART defines a break as a significant change or pause in arm or hand activity (of at least 5 – 10 minutes). Breaks include structured breaks, including meal breaks, and time spent performing other tasks that do not involve similar repetitive movements (e.g. a visual inspection or administration task). Breaks also include short ‘micro breaks’ that leave the muscle groups completely at rest for at least 10 consecutive seconds almost every few minutes (BS ISO 11228-3, 2007). Table 5 shows the risk categories for breaks in the repetitive task. There are two separate amber and red categories, with greater scores to reflect longer periods of continuous work with out a break.

**Table 5 Risk categories for breaks**

<table>
<thead>
<tr>
<th>Good</th>
<th>Reasonable</th>
<th>Poor</th>
</tr>
</thead>
<tbody>
<tr>
<td>The worker performs the task continuously without a break for less than 1 hour</td>
<td>The worker performs the task continuously without a break for 1 hour to less than 2 hours</td>
<td>The worker performs the task continuously without a break for 3 hours to less than 4 hours</td>
</tr>
<tr>
<td>There are frequent short breaks (of at least 10 seconds) every few minutes over the whole work period</td>
<td>The worker performs the task continuously without a break for 2 hours to less than 3 hours</td>
<td>The worker performs the task continuously without a break for more than 4 hours</td>
</tr>
</tbody>
</table>

2.3.4.2 **Workpace**

A workspace factor was included in the tool using similar categories to those constructed for the QEC (Table 6). The construction evidence for the QEC cites several studies that correlate time pressure and machine paced jobs with job satisfaction, fatigue and mental or physical health (Li and Buckle, 1999b). The aim was to help users investigate the extent to which workers perceive they have difficulties keeping up with the work and any reasons for these perceptions (properties of the task, an individual’s ability to cope with the demands, and the organisation of the work).

A further aim of this factor was to help users begin a dialogue with workers about other psychosocial and organisational aspects of the work that have an important influence on musculoskeletal health.

**Table 6 Risk categories for workspace (adapted from the QEC)**

<table>
<thead>
<tr>
<th>Good</th>
<th>Reasonable</th>
<th>Poor</th>
</tr>
</thead>
<tbody>
<tr>
<td>Not difficult to keep up with the work</td>
<td>Sometimes difficult to keep up with the work</td>
<td>Often difficult to keep up with the work</td>
</tr>
</tbody>
</table>

2.3.4.3 **Psychosocial factors**

With the exception of workspace, other psychosocial factors are not given a score. However, the assessment guide draws attention to the importance of recording psychosocial factors and cues users to several factors that should be considered through consultation with the workers. These
were selected from HSG60 and include: little control over how the work is done; frequent tight deadlines; incentives to skip breaks or finish early; lack of support from supervisors or co-workers; monotonous work; excessive work demands; high levels of attention and concentration; and insufficient training to do the job successfully. The ART score sheet contains a place to record psychosocial and individual factors that users identify as part of their assessment.

### 2.3.4.4 Other factors

The OCRA method lists an assortment of object and environmental factors for which there is evidence of a causal or aggravating relationship with work-related upper limb disorders (Colombini et al., 2002). These factors are important; however, they may not always be present in the workplace and thus are listed as ‘other factors’ to ensure that ART remains relevant for the majority of repetitive tasks. ART prompts the user to consider factors such as: exposure to hand/arm vibration; localised compression of anatomical structures; the need for fine movements of the fingers or hand, exposure to cold environments or cold surfaces; the use of gloves that restrict movements; and counter-shock actions, such as hammering with a tool. To be consistent with the MAC, where one factor is present, an amber category is selected, while a red category is selected where two or more other factors exist (Table 7).

![Table 7 Risk categories for other factors](image)

<table>
<thead>
<tr>
<th>Good</th>
<th>Reasonable</th>
<th>Poor</th>
</tr>
</thead>
<tbody>
<tr>
<td>No other factors present</td>
<td>One other factor is present</td>
<td>Two or more other factors are present</td>
</tr>
</tbody>
</table>

### 2.3.4.5 Duration

The aim of this risk factor was to ensure that users considered the total amount of time that workers spend performing the repetitive task in a day. A further aim was to take account of the benefits of job rotation to less demanding and/or non-repetitive tasks and discourage prolonged repetitive task durations that might occur with long shifts or regular overtime. There is little evidence to suggest how duration should be scored and it appears that several researchers have applied professional judgement (Colombini et al., 2002, Moore and Garg, 1995). In a similar approach, a system of duration multipliers was applied to show no effect on score if the task is performed for 4 to 8 hours per day (a typical shift), but decreased exposure with shorter durations and increased exposure where task durations that exceed eight hours (Table 8). The use of a ‘less than 2 hour’ category is also appropriate as HSG60 already uses 2 hours as a basic building block of exposure time within the risk filter.

![Table 8 Duration categories and risk multipliers](image)

<table>
<thead>
<tr>
<th>Duration categories</th>
<th>Risk multipliers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Less than 2 hours</td>
<td>0.5</td>
</tr>
<tr>
<td>2 hours to less than 4 hours</td>
<td>0.75</td>
</tr>
<tr>
<td>4 hours to 8 hours</td>
<td>1</td>
</tr>
<tr>
<td>More than 8 hours</td>
<td>1.5</td>
</tr>
</tbody>
</table>
2.4 DEVELOPMENT OF THE SCORING SYSTEM

2.4.1 Common problems of developing a scoring system

In a review of risk assessment methods for ULDs, Ashby (2006) notes that one of the key problems of looking to score risk factors is determining the dose. There is an inherent difficulty with offering a weighting to each individual factor to derive an acceptable or unacceptable exposure level and this problem then transfers to the overall level of risk determined from use of the method. While the literature is in broad agreement about the physical risk factors that can contribute to WRMSDs, there is more divergence about the importance of each factor and how they combine to influence overall risk. This is complicated further by the variable effect of each factor on the different joints. Therefore, the scoring systems proposed in some of the methods are a compromise between the ‘known’ and ‘unknown’ and can only be used as a reference (Li and Buckle, 1999b). This may imply that for the purposes of a general screening tool, a relatively basic scoring system can be of more value (or at least of equal value) than a more complex scoring system that presents practical difficulties for the user.

The other problem with scoring systems is establishing an appropriate level of sensitivity, while being concise enough for practical application in the work area (Ashby, 2006). Increased sensitivity may provide more accuracy, but also requires more time and ability on the part of the assessor. Li and Buckle (1999a) report how user groups, when asked about the problems with WRMSD assessment methods, describe both too much detail and paperwork on one hand but that they can be too simplistic and subjective on the other. Given that methods with a greater sensitivity must then be accompanied with more data to justify the various levels of risk differentiation, there is an argument for developing methods to be more basic.

2.4.2 The ART scoring system

The focus for ART was to apply a simple model that users would find easy to use to screen tasks for risk factors, and through the colours assigned to the risk factors, demonstrate the presence of risk to duty holders and recommend areas for improvement. Thus, the underlying assumption with ART is that the risk of ULDs from repetitive tasks can be quantified using an additive model. In other words, a value representing risk due to repetition can be added to the risk from hand force to give a combined risk score for the two risk factors. This was extended for eleven of the twelve risk factors in ART. The twelfth ‘duration’ risk factor introduces a multiplier into the model to give an indication of an individual’s exposure to the task. As it is not possible to take into account all risk factors, there will always be risk that is unattributed as well as any error in the model (Pinder, 2002).

The scoring system is represented as follows:

\[
\text{(1) Risk} = \text{Total exposure score} + \text{Unattributed risk} + \text{Error}
\]

\[
\text{(2) Total exposure score} = \text{Task score} \times \text{Duration}
\]

\[
\text{(3) Task score} = \text{Frequency} + \text{Repetition} + \text{Force} + \text{Neck posture} + \text{Back posture} + \text{Shoulder posture} + \text{Wrist posture} + \text{Hand grip posture} + \text{lack of Breaks} + \text{Work pace} + \text{Other factors}
\]

Symbolically, this can be written as:

\[
\text{(4) } R_{\text{TASK}} = R_F + R_R + R_F + R_N + R_B + R_S + R_W + R_H + R_{LB} + R_P + R_O
\]
While the units of risk are unspecified, suitable units might be a measure of the incidences of ULDs.

The ART scoring system provides different weightings to the risk factors. Where possible, the risk weightings were informed by the weightings of other methods (e.g. OCRA checklist), expert judgment and a desire to create a balanced assessment where no single stage of the assessment could dominate the overall score, be it frequency and repetition of movements, force, or awkward posture. Instead, tasks that ART would assess as high risk would typically involve repetitive movements, and either some application of force and/or awkward postures, combined with prolonged task durations and/or insufficient recovery periods. Three exposure levels (low, medium and high) are proposed for the total exposure scores. These levels are proposed for future discussion and/or evaluation, as they have not yet been validated with WRMSD incidence data.

As an additive model, the ART scoring system involves limited interactions between the risk factors, except for the ‘duration’ factor, which can greatly affect the overall risk score. Thus, it is assumed that there is no interaction between the risk due to the force and the risk due to the type of handgrip. While applying a high force with a ‘poor’ grip may be worse than applying a high force with a ‘good’ grip, simply adding the ART score for hand grip may not reflect an accurate escalation of the risk. With the assistance of the risk factor colours to demonstrate the presence of risk, this is not believed to be a significant limitation where ART is used for the purpose of a basic screening tool. In many cases, the ART risk scores will also produce a suitable estimation of the level of risk. However, as the ART scoring system is a simplified model, when undertaking a more detailed assessment, it is recommended that other suitable methods are applied alongside ART and, if necessary, with the support of specialist advice.

2.5 REQUIREMENTS TO USE THE ASSESSMENT TOOL

2.5.1 Specified Users

HSE and LA Inspectors were considered likely to become the primary user group for ART. However, previous experience of developing the MAC tool would suggest that the user group could be expanded in the future, to include health and safety practitioners with responsibility for risk assessments related to the prevention of musculoskeletal disorders (e.g. manual handling risk assessments and ULD risk assessments).

The specified users of ART, and thus many of their requirements, will be similar to those described by Lee (2005) during his work to refine the MAC. ART is intended for use by males and females in paid work. The vast majority of users will be between 18 – 65 years of age. Users must be able to communicate clearly to those employees performing repetitive tasks, using language that the employees will understand. In return, users must be able to listen to employees and understand their descriptions and opinions of the physical and organisational aspects of their work. In the UK, there will be occasions when employees performing the repetitive tasks do not have sufficient understanding and ability to communicate in English. In these circumstances, a person suitably skilled to act as a translator between the user and the employees may be required.

The nature of repetitive work will often require users to observe a full range of manual tasks involving quick and frequent movements of upper limbs. ART will be accessible to those users
who are able to read English at a reasonable level, read small font text and add a series of integers. Users must also be able to turn the pages of a booklet and make clear markings with a pen or pencil to record risk ratings and scores.

2.5.2 Suitable tasks for assessment

One of the first steps is to select a suitable task for assessment with ART. ART is applicable to repetitive tasks of the upper limbs. Colombini et al. (2002) describe repetitive tasks as those that:

1. Are carried out through actions of the upper limbs;
2. Are carried out consecutively for at least 1 – 2 hours per day; and
3. Involve work cycles that are similar to each other and of relatively short duration (e.g. a few minutes).

Where a task involves the use of one arm, an assessment can be made of just the one arm predominantly involved in the task. Where both arms are involved in the task, or where users are uncertain which arm is at greater risk, ART allows a separate assessment to be made of the left and right arm.

ART cannot be used to assess tasks that involve intense finger movements yet little movement of the wrist or hand (such as when using a keyboard). Thus, ART is not suitable for use in DSE assessment. In its current form, ART is used to assess single tasks and how it can be applied to job rotation needs to be considered carefully.

2.5.3 Equipment for use

ART is a paper booklet and a pen or pencil is required for recording scores, observations and employees’ descriptions and/or opinions on the work. A watch or stopwatch may also prove helpful to determine the level of risk for factors that involve an element of time as part of the criteria. A video camera, Borg CR-10 scale or force-measuring gauge are optional and may be valuable to those users with advanced skills or knowledge of assessing repetitive work tasks.

2.5.4 Environment of use

Observations and discussions may need to take place in environments that require the use of personal protective equipment such as hearing defenders, protective eyewear and gloves. Hearing defenders and/or noise in the environment may reduce the ability of users to communicate with employees carrying out the repetitive work, managers, safety representatives and others contributing important information about the repetitive task. Protective eyewear may reduce the ability of users to read ART and make observations of the repetitive work. Gloves may reduce dexterity and tactile feedback when turning pages or recording information on the booklet.

2.5.5 Training for use

Although it may be possible for some people to use ART with minimal training, it is expected that most users would benefit from prior training in the use of ART. It is expected that many HSE and LA Inspectors would be able to use ART with reasonable usability, accuracy and reliability following attendance at a training event lasting at least half a day, followed by further practice or familiarisation.
The training event should include topics such as:

- The identification of work situations where it is appropriate to use ART and where the use of ART is not appropriate;
- An explanation of the process that ART uses to assess repetitive tasks;
- An explanation of the risk factors included in ART;
- An explanation of the criteria used to categorise the levels of risk for each factor;
- A demonstration of how ART is applied to assess a repetitive task;
- Several opportunities where users can use ART to assess repetitive tasks and receive feedback on their performance; and
- Information on risk reduction measures, including the management of any new or existing incidences of musculoskeletal disorders.

The quality of the training material and delivery can have an important impact on the assessment results. However, there is scope to develop additional ‘standardised’ training material and make this accessible on the Internet. This additional training material should include a selection of practice assessments to consolidate learning.
3 USER TRIALS

3.1 OVERVIEW

Figure 3 shows an overview of the user trials process. This involved an iterative design process whereby the results from each of the user trials were interpreted and used to inform subsequent ART prototypes. For reporting purposes, the user trials have been grouped into three phases:

1. Simulated user trials – phase one
2. Simulated user trials – phase two
3. Field trials – phase three
The user trials methodology draws upon an approach set out by Lee (2005) for the development and refinement of paper-based MSD assessment tools.

To test the accuracy, effectiveness and usability of ART, it was necessary to ensure that users assessed a standardised set of tasks in a controlled manner. Unfortunately, it would not have been possible, or an efficient use of Inspector resource, to expose all users to identical repetitive tasks in a realistic field inspection setting. Thus, ‘simulated user trials’ were arranged, during which video footage was used to present repetitive tasks of the upper limbs and the associated risk factors. This methodology has the disadvantage in that a user cannot view the repetitive task from several angles or solicit further input from the workers. Thus, the use of video may have resulted in users making decisions from incomplete or ambiguous information, thereby tending to decrease reliability artificially. However, the simulated trials were felt to be appropriate for initial testing, where the primary aim was to explore the concept of ART with users and make improvements to the usability of ART.

During the simulated trials, users were asked to draw upon their previous experiences assessing repetitive tasks in the field. However, it was felt important to trial ART with a small sample of Inspectors in the field. This provided information on the practicalities of using ART at the work area and allowed a comparison of scores between Inspectors and Specialists in the field.

### 3.2 SIMULATED TRIALS – METHODOLOGY

Figure 4 shows the protocol followed during the simulated trials. The introduction and protocol briefing outlined the rationale for developing ART, its relation to HSE’s MSD programme of work, the purpose of the project and the role of the participants in the project. The ULDs risk factor briefing was based on the content contained within HSG60(rev) *Upper limb disorders in the workplace*. The briefing included information about the risk factors related to the: task (repetition, work postures, force, and duration of exposure); environment (physical and psychosocial environments); and individual. The ART briefing involved:

- An explanation of the process that ART uses to assess repetitive tasks;
- An explanation of how to assess each risk factor (including an explanation of the assessment criteria for each risk factor, as well as the practical challenges of making the assessment);
- An explanation of how to complete the task description form and the score sheet;
- A demonstration of how to apply ART to a repetitive task (ironing jackets); and
- An opportunity to practice using ART to assess a further two repetitive tasks (light assembly work and linnishing wheel braces) and receive feedback on the assessments.
Table 8 describes the video task material that was involved in the ART briefings. In addition to any information required to complete the task description form, some risk factors were provided to users when the information could not be gleaned from the video. Questions and group interaction about the use and development of ART were welcomed throughout the ART briefing.
Table 8 Description of video task material used during the ART briefings and those risk factors provided by the project team

<table>
<thead>
<tr>
<th>Task</th>
<th>Phase of simulated trials</th>
<th>Description</th>
<th>Work cycle times (s)</th>
<th>Risk factors provided</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1 and 2</td>
<td>The operator sprays water onto a suit coat and irons the coat</td>
<td>~24</td>
<td>Force, Workpace, Duration</td>
</tr>
<tr>
<td>2</td>
<td>1 and 2</td>
<td>The operator collects small metal parts, assembles them in the hand and then uses a press</td>
<td>~9</td>
<td>Force, Workpace, Duration</td>
</tr>
<tr>
<td>3</td>
<td>1 and 2</td>
<td>Two operators use a bench grinder to remove excess metal (‘flash’) from recently manufactured four-way cross wheel braces</td>
<td>~24</td>
<td>Workpace, Duration</td>
</tr>
</tbody>
</table>

Table 9 describes the video task material that was involved in the ART scoring assessments. As with the ART briefings, users were provided with information to complete the task description form and those risk factors that could not be assessed from video observation. During the ART scoring assessment, users were instructed to complete the assessments on their own and without further assistance from the project team.

Table 9 Description of video task material used during the ART scoring assessments and those risk factors provided by the project team

<table>
<thead>
<tr>
<th>Task</th>
<th>Phase of simulated trials</th>
<th>Description</th>
<th>Work cycle times (s)</th>
<th>Risk factors provided</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>1 and 2</td>
<td>The operator picks ice creams off a conveyor, places them in a box and, once the box is full, lifts the box onto a separate conveyor</td>
<td>~37</td>
<td>Workpace, Duration</td>
</tr>
<tr>
<td>5</td>
<td>1 and 2</td>
<td>The operator picks whole chicken breast from a container, cuts the chicken with a knife and splits the chicken open with the hands</td>
<td>~4</td>
<td>Workpace, Duration, Other factors</td>
</tr>
<tr>
<td>6</td>
<td>1 only</td>
<td>The operator uses a pen to apply enamel to buttons that have been placed on a grid support</td>
<td>~7 – 15</td>
<td>Duration</td>
</tr>
<tr>
<td>7</td>
<td>2 only</td>
<td>The operator assembles a box, and then picks a stack of envelopes, levels (‘knocks up’) the envelopes on the work desk and places the stack in the box. The operator places two more stacks of envelopes in the box before closing the box.</td>
<td>~40</td>
<td>Workpace, Duration</td>
</tr>
</tbody>
</table>

Following the scoring assessments, each user completed a usability questionnaire (Appendix A). The questionnaire was made up of two elements:

1. An open-ended usability questionnaire – previously designed to investigate specific usability problems with the risk factors and to seek general improvements to the format and wording of the MAC (Lee and Ferreira, 2003). The questionnaire has not been subjected to rigorous reliability or validity testing. However, its previous use has shown it to be useful to collect qualitative information.

2. A system usability scale (SUS; Brooke, 1996) – a simple ten-item Likert scale giving a global view of subjective assessment of usability. The SUS has been tested and
evaluated in industry and has been found to be reliable, robust and well correlated with other subjective measures of usability (Brooke, 1996). The SUS was used to measure changes in the overall usability of successive prototypes of ART. Some terminology within the SUS statements was adapted slightly for this project. The terms ‘the system’ and ‘this system’ were replaced with ‘ART’, ‘functions’ was replaced with ‘steps’, and the term ‘users’ was replaced with ‘Inspectors’. Within the context of this work, these terms were assumed to be synonymous and were modified to assist users’ understanding of the statements.

3.3 PARTICIPANTS

Table 10 shows the number of HSE and LA Inspectors involved in each phase of the trials. From the 32 Inspectors who took part in the simulated user trials, six Inspectors (19%) participated in subsequent field trials.

<table>
<thead>
<tr>
<th>User Trials</th>
<th>Prototype Used</th>
<th>HSE Inspectors</th>
<th>LA Inspectors</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Simulated Trials – Phase 1</td>
<td>P3</td>
<td>4</td>
<td>3</td>
<td>7</td>
</tr>
<tr>
<td>Simulated Trials – Phase 2</td>
<td>P4/P5</td>
<td>12</td>
<td>13</td>
<td>25</td>
</tr>
<tr>
<td>Field Trials</td>
<td>P5</td>
<td>4</td>
<td>2</td>
<td>6</td>
</tr>
</tbody>
</table>

Figure 5 shows the amount of experience that Inspectors had in their current roles when attending the trials. Twenty-seven Inspectors (81%) had received previous training in the use of the MAC. Three participants were Specialist Occupational Health Inspectors.

3.4 SIMULATED TRIALS – PHASE ONE FINDINGS

Overall, users supported the concept of ART and believed further development and refinement of ART was worthwhile to pursue. However, both users and the project team agreed that significant improvements to the usability of ART were required in order to progress with further user trials. In addition, it was agreed that the assessment of some risk factor categories did not provide a useful outcome or direct the users to risk reduction measures.
For these reasons, the results for ART Prototype 3 have not been included in this report. However, Table 11 provides a summary of user feedback for several factors where major improvements were required, along with details of the changes made to arrive at ART Prototype 4.

**Table 11** Summary of user feedback for ART prototype 3 and subsequent improvements following phase one of the simulated user trials

<table>
<thead>
<tr>
<th>Factor</th>
<th>Comments</th>
<th>Details of change</th>
</tr>
</thead>
<tbody>
<tr>
<td>A2</td>
<td>This factor just defines the types of tasks where it is appropriate to use ART. While variation of movement and longer work cycles (for example, of 30 seconds) are preferred to repeated, restrictive movements every few seconds, this is not reflected in the assessment.</td>
<td>Introduction of the QEC factor to assess repetitive movements. This allows a grading of the repetition factor according to the number of similar motion patterns repeated over a fixed 1 minute period.</td>
</tr>
<tr>
<td>B</td>
<td>The only stage where the user must select the highest score from the three factors (i.e. B1, B2 or B3), which is inconsistent with other stages of ART.</td>
<td>Simplification of stage B into a single factor for force. Introduction of a ‘force graph’, allowing the user to first assess the level of force and then the duration of any force exertion.</td>
</tr>
<tr>
<td>B</td>
<td>Difficult to assess force without talking to the operator and looking at the task closely.</td>
<td>Further instruction provided on how to involve the worker and make a subjective assessment of force in the field.</td>
</tr>
<tr>
<td>C5</td>
<td>Difficult to determine whether hand grip or finger grip is static, in contrast to dynamic, loose or relaxed gripping of an object.</td>
<td>Assessment of ‘static’ hand grip replaced with assessment of the presence of awkward hand grip postures (e.g. pinch, hook and wide finger span grips) and the proportion of the cycle time where awkward hand or finger grips are applied.</td>
</tr>
<tr>
<td>D2</td>
<td>Difficult to establish what is a ‘demanding’ piece rate system, unless you talk to the workers. Work pace set by a machine need not be ‘fast’ or difficult to watch.</td>
<td>Introduction of the QEC factor to assess workers’ perceived pace of work. This encourages users to involve the workers in the assessment. This factor then ‘opens the door’ for users, allowing them to investigate other aspects of work organisation that might have an impact on workers’ perceived demands and/or ability to cope with the work demands.</td>
</tr>
</tbody>
</table>

### 3.5 SIMULATED TRIALS – PHASE TWO FINDINGS

The methodology used in phase two was similar to the phase one methodology, although there were three notable differences:

- Where changes had been made to Prototypes 4 and 5 of ART, the training content was revised to reflect those changes.

- A new case study (‘stacking envelopes in a box’) was used for the ART scoring assessments (see Table 9). This replaced a phase one case study (‘application of enamel to buttons’), which did not provide a sufficient view the operator’s arm or whole body posture.
3.5.1 Completeness of Prototypes 4 and 5

During the phase two user trials, 98.7% (74/75) of assessments were completed, with a score next to each risk factor. During one of the trials, participants were asked to note the amount of time that they required to complete each assessment. Of the 13 participants, 11 provided this information, which suggested that 90% of assessments were completed in 10 to 15 minutes. The average time to complete an assessment was 13 minutes and 42 seconds.

3.5.2 Extent of scoring agreement for Prototypes 4 and 5

3.5.2.1 Risk factor scores

During the trials, each of the 25 participants was asked to assess three tasks, and for each task, this involved making a judgement on 12 risk factors. This provided 888 judgements in total, as one participant completed an assessment of only two of the tasks.

Figure 6 depicts the extent of scoring agreement for those factors that require observation of the task. Table 12 details the extent of scoring agreement for each risk factor. Overall, 69% (610/888) of judgements agreed with the expert consensus for colour code. 97% (858/888) of judgements either agreed with the expert consensus or were within one category of the expert consensus. However, information on work pace and the duration risk factors was provided to participants, as it could not be obtained from video. These risk factors scored 100% for full agreement. If this data is removed from the analysis, 740 judgements are available for comparison. In this case, 62% (462/740) of judgements agreed with the expert consensus for colour code and 96% (708/740) of judgements either agreed with or were within one category of the expert consensus. This compares to a target of 75% of colour codes in full agreement with the expert consensus and 90% of colour codes in agreement or within one category of the expert consensus.

<table>
<thead>
<tr>
<th>Risk factor</th>
<th>Full agreement with expert consensus (%)</th>
<th>Full agreement or within one category difference of expert consensus (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A1 Frequency</td>
<td>80</td>
<td>100</td>
</tr>
<tr>
<td>A2 Repetition</td>
<td>31</td>
<td>97</td>
</tr>
<tr>
<td>B Force</td>
<td>66</td>
<td>100</td>
</tr>
<tr>
<td>C1 Neck</td>
<td>64</td>
<td>96</td>
</tr>
<tr>
<td>C2 Back</td>
<td>66</td>
<td>92</td>
</tr>
<tr>
<td>C3 Shoulder</td>
<td>64</td>
<td>92</td>
</tr>
<tr>
<td>C4 Wrist</td>
<td>58</td>
<td>99</td>
</tr>
<tr>
<td>C5 Hand grip</td>
<td>41</td>
<td>71</td>
</tr>
<tr>
<td>D1 Breaks</td>
<td>77</td>
<td>89</td>
</tr>
<tr>
<td>D2 Work pace (given)</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>D3 Miscellaneous factors</td>
<td>64</td>
<td>99</td>
</tr>
<tr>
<td>E Duration (given)</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>Overall</td>
<td>69</td>
<td>97</td>
</tr>
<tr>
<td>Overall (excluding givens)</td>
<td>62</td>
<td>96</td>
</tr>
</tbody>
</table>

Table 12 Extent of scoring agreement for risk factors
Figure 6 Extent of scoring agreement between users and expert consensus for risk factors that require observation of task performance

(▁ = full agreement, ▃ = one category difference, ▄ = two categories difference)
3.5.2.2 Total exposure scores

Table 13 summarises participants’ total exposure scores for the three tasks assessed in the simulated trials.

<table>
<thead>
<tr>
<th>Task</th>
<th>N</th>
<th>Mean</th>
<th>Std deviation</th>
<th>Median</th>
<th>Range</th>
<th>Expert consensus</th>
</tr>
</thead>
<tbody>
<tr>
<td>Boxing ice cream</td>
<td>25</td>
<td>25.2</td>
<td>3.17</td>
<td>26</td>
<td>18 – 29</td>
<td>23</td>
</tr>
<tr>
<td>Chicken chopping</td>
<td>25</td>
<td>25.7</td>
<td>4.3</td>
<td>26</td>
<td>17 – 35</td>
<td>21</td>
</tr>
<tr>
<td>Packing envelopes</td>
<td>24</td>
<td>20.6</td>
<td>4.0</td>
<td>20.5</td>
<td>14 – 28</td>
<td>17</td>
</tr>
</tbody>
</table>

Figure 7 shows the extent of agreement in the total exposure scores for each task and for the trials overall. Overall, 70% of participants scored within 5 points of the expert consensus. However, there was large variation between the case studies. Inspectors tended to score the total exposures higher than the expert consensus. This was primarily due to scoring the repetition or force risk factors higher than the expert consensus.

![Figure 7](image_url)

**Figure 7** Extent of total exposure score agreement (%) compared to the expert consensus

3.5.3 Usability of Prototypes 4 and 5

Inspectors reported good support for the development and use of ART. Overall:

- 64% of respondents felt that ART improved or greatly improved their understanding of ULD risks.
- 84% of respondents felt that ART improved or greatly improved their confidence when assessing ULD risks.
- 64% of respondents felt that ART helped identify what action to take.
Table 14 shows the descriptive statistics for the System Usability Scale (SUS) data and gives an overall measure of usability and user satisfaction.

Table 14 Combined SUS descriptive statistics for Prototype 4 and 5

<table>
<thead>
<tr>
<th>Inspectors</th>
<th>Mean (%)</th>
<th>Standard deviation (%)</th>
<th>Median (%)</th>
<th>Minimum (%)</th>
<th>Maximum (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>HSE</td>
<td>74.8</td>
<td>11.5</td>
<td>77.5</td>
<td>57.5</td>
<td>87.5</td>
</tr>
<tr>
<td>LA</td>
<td>63.5</td>
<td>8.8</td>
<td>65.0</td>
<td>50</td>
<td>80</td>
</tr>
<tr>
<td>Overall</td>
<td>68.9</td>
<td>11.5</td>
<td>67.5</td>
<td>50</td>
<td>87.5</td>
</tr>
</tbody>
</table>

Figure 8 shows the extent to which the mean SUS score changed with each subsequent prototype. This shows that:

- Overall usability and satisfaction was favourable throughout the trials (i.e. more than 50%).

- The SUS measure of overall usability and satisfaction increased during the iterative design process.

- Overall, the mean SUS score did not meet the 75% target. However, the mean SUS score for P4 and P5 did meet the 75% target amongst the group of HSE Inspectors.

LA Inspectors reported significantly lower SUS scores compared to HSE Inspectors (Mann Whitney U = 33.5, p<0.05). The slightly lower usability result for LA Inspectors was due to several aspects:

- More LA Inspectors reported they “needed to learn a lot of things before they could get on and use ART quickly”. LA Inspectors similarly reported that ART improved their understanding of ULD risks to a greater extent than HSE Inspectors.

- Fewer LA Inspectors reported “they would use ART frequently”, possibly due to the type of premises enforced by Local Authorities (particularly in the London area).
More LA Inspectors reported that there was too much inconsistency in the scoring. When asked how easy it was to score each risk factor, respondents rated 10 of the 12 risk factors as favourable on average (Figure 9). However, there was considerable variation in opinion between individual respondents. The risk factors that were rated more difficult to score were frequency of arm movements, repetition, force, shoulder posture and wrist posture. Where respondents reported that a risk factor was difficult to score, they were asked to describe why it was difficult. Table 15 provides a summary of user feedback for the factors that were rated as more difficult to score. Where possible, improvements were made to the design of ART and the associated training material to address these problems. At this stage of the project, the scope to make further fundamental changes to the structure and design of ART was limited. Thus, where further clarity or information was requested for particular risk factors, the project focussed on how improvements could be made to the content and delivery of the accompanying training material.

Some of the reported difficulties were due to the limitations of assessing repetitive tasks using the video case studies. Such problems included:

- The inability to speak to operators and ask further questions where scoring was uncertain (e.g. assessment of force).
- The inability to view postures at multiple angles.
- Temporary obstruction of certain hand/wrist postures due to the position of the camera.

Many users anticipated that it would be easier to assess the level of force and awkward postures when using ART in the field.

![Figure 9 Mean usability rating and standard deviation for each risk factor](image)

Figure 9 Mean usability rating and standard deviation for each risk factor

(1 = very difficult; 5 = very easy)

Table 16 provides a summary of changes that users requested to improve the format and general content of ART.
### Table 15 Summary of user feedback for specific risk factors that were more difficult to score

<table>
<thead>
<tr>
<th>Factor</th>
<th>Comments</th>
<th>Change</th>
<th>Details of change</th>
</tr>
</thead>
<tbody>
<tr>
<td>A1</td>
<td>Confusion between scoring the frequency of arm movements and shoulder/arm posture (C3), as both factors refer to temporal aspects of the task</td>
<td>Training</td>
<td>Further instruction that Stage A refers to upper limb movements, while Stage C refers to particular upper limb postures.</td>
</tr>
<tr>
<td>A2</td>
<td>Difficult to count the number of similar motion patterns, particularly where a task involves more than one type of motion</td>
<td>Training</td>
<td>Further instructions given to first identify the similar motion pattern of the hand/wrist/forearm that occurs most frequently, before trying to count the similar motion patterns. Further strategies provided to count similar motion patterns (e.g. refer to task cycle time if possible; tasks with cycle times of less than 3 seconds would involve more than 20 similar motion patterns per minute).</td>
</tr>
<tr>
<td>A2</td>
<td>Confusion between scoring the repetition of hand/wrist movements and wrist posture (C4), as both factors refer to temporal aspects of the task</td>
<td>ART</td>
<td>Risk factor name changed from ‘wrist/hand repetition’ to ‘repetition’.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Training</td>
<td>Further instruction that Stage A refers to upper limb movements, while Stage C refers to particular upper limb postures</td>
</tr>
<tr>
<td>B</td>
<td>No indication provided of the level of force involved in ‘normal’ tasks</td>
<td>Training</td>
<td>Further emphasis provided on the importance of obtaining workers’ subjective opinion of the level of force and the potential for error when relying on observer’s own opinion of the level of force. Experience handling 1 kg and 4 kg loads during training might give users an indication of low and high force tasks.</td>
</tr>
<tr>
<td></td>
<td>Uncertain how to score where two or more workers use different techniques or exert different levels of force during the same task.</td>
<td>Training</td>
<td>Further emphasis that users must gain insight into the level of force from the perspective of all workers where this is possible. Further emphasis that it is possible to select intermediate scores where workers report different levels of force (e.g. light to moderate force or moderate to strong force) but that the assessment should focus on the reasons why particular operators might report more force than other operators.</td>
</tr>
<tr>
<td>C4</td>
<td>Force description table not as user-friendly as other tables in ART</td>
<td>ART</td>
<td>Descriptions/examples of ‘moderate force’ simplified in table on page 3.</td>
</tr>
<tr>
<td></td>
<td>Difficult to observe wrist posture from video and if wearing gloves. Users believed this would be easier in the field.</td>
<td>Training</td>
<td>Further instruction provided to view wrist posture from several angles and ask operators to demonstrate wrist/hand movements in slow motion.</td>
</tr>
</tbody>
</table>
Unsure how many people should be asked about the work pace to get a representative answer

Further emphasis that users must gain insight into workspace from the perspective of all workers where this is possible, and to focus on the reasons why workers might have difficulty keeping up with the work.

Unsure what other miscellaneous factors could be specified that ART does not already consider.

Place for user to specify other miscellaneous factors removed. Score sheet contains a box to record other risk factors not included in ART.

Difficult to collect the information needed to determine the duration of repetitive work (the simulated trials provided a lot of information not available in reality)

Where structured shift patterns with fixed periods of work and recovery are not present, users are instructed to ask workers on average how much time they spend per day doing the task with reference to the duration categories (i.e. less than 2 hours, about half a day, more than half a day, or more than 8 hours).

Table 16 Summary of changes that users requested to improve the format and general content of ART

<table>
<thead>
<tr>
<th>No.</th>
<th>Comments</th>
<th>Change</th>
<th>Details of change</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>ART does not refer to the hierarchy of control measures</td>
<td>ART</td>
<td>HSG60(rev) guidance included in ART (page 2) reminding users to look for higher-order solutions where the priority is to consider if it is reasonably practicable to eliminate the hazard through re-design of the work, substitution of tools or components, or automation of the task.</td>
</tr>
<tr>
<td>2</td>
<td>Would only need one assessment guide, but several score sheets for each inspection. It is easier to refer to the score sheet if on a different sheet</td>
<td>ART</td>
<td>Separate A4 worksheet provided for experienced users. The worksheet contains the task description form and score sheet printed on one side and the flowchart on the opposite side.</td>
</tr>
<tr>
<td>3</td>
<td>Too much flicking backwards and forwards between the assessment guide, flowchart and score sheet</td>
<td>ART</td>
<td>Perforated edge inserted between assessment guide and worksheet. This allows the task description form and score sheet to be removed easily from the assessment guide. Training</td>
</tr>
<tr>
<td>4</td>
<td>Information required on control measures before use in the field</td>
<td>Training</td>
<td>Information on control measures included in the accompanying training. Training refers to information provided in appendix two of HSG60(rev) and already available to users.</td>
</tr>
<tr>
<td>5</td>
<td>Notebook size would be preferable</td>
<td>ART</td>
<td>ART reduced from A5 size to custom size, similar to Inspectors’ field notebooks and MAC.</td>
</tr>
</tbody>
</table>
Finally, users were asked for their overall impression of ART, including those aspects that users liked and disliked. The aspects of ART that users liked included:

- ART “gives a process to work through a difficult area.”
- ART “directs your thought process and helps narrow down what [factors might be] causing problems.”
- ART is “useful to show the duty holder how [repetitive tasks] are assessed.”
- ART is “used as a tool to assist and not an absolute.”
- “Like MAC and so familiar.”
- ART “could be objective when scoring certain points” (In contrast, some users also reported that ART was too subjective and depending upon the point of view, one observer could score very differently to another).
- ART “helps identify where action is needed but not what action is needed.”

### 3.6 FIELD TRIALS

From the 32 Inspectors who took part in the simulated user trials, six Inspectors (19%) participated in subsequent field trials. Three ergonomists acted as Specialists during the field trials. Two were members of the project team involved in the development of ART. The other Specialist, although not involved in the project team directly, had received training in the use of ART at the same time as the Inspectors with whom the joint visits were conducted.

#### 3.6.1 Tasks Assessed

During the field trials, ART was applied at eight different premises and used to assess 15 different tasks in total. The sample of tasks assessed in the field trials included:

- Packaging fruit;
- Feeding linen into ironing machines;
- Assembling hose pipes;
- Icing cakes;
- Building cardboard boxes;

There were occasions when a repetitive task provided an opportunity for more than one comparison between the Inspector and Specialist scores. For example, during one visit, two trained Inspectors accompanied the Specialist on the visit. Thus, for the three tasks assessed during the visit, comparisons were made between each Inspector’s scores and the Specialist’s scores, resulting in six comparisons in total for these tasks. In addition, for seven of the tasks, the Inspector and Specialist made separate assessments for the left and right arm. In these circumstances, the comparisons between the Inspector and Specialist were included for each
arm. However, where the risk factor did not distinguish between a specific arm only one comparison was counted.

Thus, the 15 different tasks assessed provided a total of 29 comparisons for risk factors that are specific to each arm (frequency, repetition, force, shoulder posture, wrist posture, hand grip and miscellaneous factors) and 18 comparisons for risk factors that are not specific to a single arm (neck posture, back posture, breaks, workspace and duration).

### 3.6.2 Extent of scoring agreement

#### 3.6.2.1 Risk factor scores

Table 17 shows the extent of scoring agreement between the Inspectors and Specialists. Overall, 67% of Inspectors’ risk factor scores agreed with the Specialist’s scores and 95% of Inspectors’ risk factor scores either agreed with or were within one category of the Specialist’s scores. This compares to a target of 75% full agreement and 90% either full agreement or within one category difference.

The full agreement target was not met overall, nor for nine of the 12 risk factors. This demonstrates the complexity of assessing repetitive tasks with observation-based methods. The target for ‘full or within one category difference’ was met overall, and for 11 of 12 risk factors. The only risk factor that did not meet the target in this respect was hand grip and some uncertainty remained when deciding whether the observed hand grip was awkward enough to merit scoring. This is an area where further clarification and attention to training may be required. Figures 10 and 11 depict the extent of scoring agreement between Inspector and Specialist for each factor when ART was trialled in the field.

#### Table 17 Extent of scoring agreement between Inspectors and Specialists for risk factors

<table>
<thead>
<tr>
<th>Risk factor</th>
<th>Full agreement with Specialist</th>
<th>Full agreement or within one category difference of Specialist</th>
</tr>
</thead>
<tbody>
<tr>
<td>A1 Frequency</td>
<td>19/29</td>
<td>28/29</td>
</tr>
<tr>
<td>A2 Repetition</td>
<td>21/29</td>
<td>26/29</td>
</tr>
<tr>
<td>B Force</td>
<td>20/29</td>
<td>27/29</td>
</tr>
<tr>
<td>C1 Neck</td>
<td>7/18</td>
<td>17/18</td>
</tr>
<tr>
<td>C2 Back</td>
<td>14/18</td>
<td>18/18</td>
</tr>
<tr>
<td>C3 Shoulder</td>
<td>18/29</td>
<td>27/29</td>
</tr>
<tr>
<td>C4 Wrist</td>
<td>15/29</td>
<td>26/29</td>
</tr>
<tr>
<td>C5 Hand grip</td>
<td>14/29</td>
<td>25/29</td>
</tr>
<tr>
<td>D1 Breaks</td>
<td>12/18</td>
<td>18/18</td>
</tr>
<tr>
<td>D2 Work pace</td>
<td>12/18</td>
<td>18/18</td>
</tr>
<tr>
<td>D3 Miscellaneous factors</td>
<td>26/29</td>
<td>29/29</td>
</tr>
<tr>
<td>E Duration</td>
<td>18/18</td>
<td>18/18</td>
</tr>
<tr>
<td>Overall</td>
<td>196/293</td>
<td>277/293</td>
</tr>
</tbody>
</table>
Figure 10 Extent of scoring agreement in the field between Inspectors and Specialist for risk factors that require observation of task performance

(\(\square\) = full agreement, \(\square\) = one category difference, \(\square\) = two categories difference)
Figure 11 Extent of scoring agreement in the field between Inspectors and Specialist for risk factors that do not require direct observation of task performance

( □ = full agreement, △ = one category difference, ▼ = two categories difference)

3.6.2.2 Total exposure scores

Figure 12 compares the total exposure scores between the Inspector and Specialist ($R_s = 0.865$ at 0.01 level of significance). For the 29 assessments conducted during the field trials, the total exposure scores of Inspectors were within two points of the final scores of Specialists’ for 41% of assessments, within five points for 76% of assessments and within ten points for 97% of assessments.

Table 18 provides some descriptive statistics for the total exposure scores that Inspectors obtained during the field trials. Total exposure scores ranged from 1 to 42, with an average score of 15.2. This again shows a good spread and gives some indication that ART will be able to help prioritise interventions. ART Prototypes 4 and 5 allocated total exposure scores into four proposed risk levels. During the field trials, the Inspectors’ action levels agreed with the Specialist’s action levels for 59% (17/29) assessments and either agreed or were within one category difference for 97% (28/29) of assessments.

Table 18 Descriptive statistics for total exposure scores of Inspectors

<table>
<thead>
<tr>
<th>Mean</th>
<th>Median</th>
<th>Mode</th>
<th>Minimum</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>15.2</td>
<td>20</td>
<td>21</td>
<td>1</td>
<td>42</td>
</tr>
</tbody>
</table>
3.7 SUMMARY

A summary of the user trials is as follows:

(1) Thirty-two HSE and LA Inspectors participated in simulated user trials, which used video case studies to explore the concept of ART with users and make improvements to the usability of ART.

(2) The observation-based assessment of quick hand and arm movements is inherently difficult. Nonetheless, the usability of ART was perceived to be favourable, and improved with successive prototypes of the tool. Inspectors liked that ART provided a familiar process that showed how to assess repetitive tasks, helped to narrow down those factors that might be causing problems, and identified where action was needed.

(3) Compared to HSE Inspectors, fewer LA Inspectors reported that they would use ART frequently. It is possible that ART will not be applicable at many of the premises enforced by Local Authorities, which primarily involve work with display screen equipment (DSE).

(4) A limited number of field trials with ART were conducted. These suggested that trained Inspectors will arrive at reasonably similar results compared to Specialists. Experience suggests that an element of training and regular use is required to use ART reliably. However, an evaluation of the extent to which training might improve reliability was beyond the scope of this study.

(5) ART and the associated training materials are sufficiently usable to pilot with a selection of trained HSE and LA Inspectors during an inspection campaign.
4 BENCHMARKING

The purpose of the section is to compare ART with existing methods used for risk assessment of repetitive tasks of the upper limbs.

4.1 APPROACH TO THE BENCHMARKING

For the benchmarking, a small number of existing assessment methods were selected for detailed comparison. These were:

- Assessment of Repetitive Tasks (ART – Prototype 6);
- The Quick Exposure Check (QEC; David et al., 2008);
- The Occupational Repetitive Actions checklist (OCRA checklist; Colombini et al., 2002); and
- The Strain Index (SI; Moore and Garg, 1995)

4.1.1 Task assessment

Ten videos of industrial repetitive tasks were selected for the benchmarking (Table 19). The tasks were scored with a consensus of at least three ergonomists involved in either the development workshops and/or the development of the user trial material. This consensus approach was necessary to minimise variation between ergonomists, particularly when using the more complex assessment methods (Strain Index and OCRA checklist).

<table>
<thead>
<tr>
<th>Task</th>
<th>Description</th>
<th>Work cycle times (s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>The operator uses a pneumatic lifting device to turn stone flags onto a conveyor and then pushes the flags along the conveyor</td>
<td>15</td>
</tr>
<tr>
<td>B</td>
<td>The operator collects small metal parts, assembles them in the hand and then uses a press</td>
<td>9</td>
</tr>
<tr>
<td>C</td>
<td>The operator uses a pen to apply enamel to buttons that have been placed on a grid support</td>
<td>7 – 15</td>
</tr>
<tr>
<td>D</td>
<td>The operator places stockings onto an automatic ironing machine above shoulder height and pulls the stockings down</td>
<td>4</td>
</tr>
<tr>
<td>E</td>
<td>The operator picks whole chicken breast from a container, cuts the chicken with a knife and splits the chicken open with the hands</td>
<td>4</td>
</tr>
<tr>
<td>F</td>
<td>The operator picks ice creams off a conveyor, places them in a box and, once the box is full, lifts the box onto a separate conveyor</td>
<td>37</td>
</tr>
<tr>
<td>G</td>
<td>The operator picks ups a plastic cup of grapes, tips the contents into a bag and returns the empty cup</td>
<td>23</td>
</tr>
<tr>
<td>H</td>
<td>The operator uses a Stanley knife to strip yarn off stacks of cones</td>
<td>3</td>
</tr>
<tr>
<td>I</td>
<td>The operator assembles radiators by inserting a metal ‘wand’ into a tube and then inserting the tube through holes in the heat transfer sheets.</td>
<td>5</td>
</tr>
<tr>
<td>J</td>
<td>Two operators use a bench grinder to remove excess metal from recently manufactured four-way cross wheel braces</td>
<td>24</td>
</tr>
</tbody>
</table>
Where a method resulted in a separate score for each hand, the hand with the highest score was selected for comparison. Early development workshops indicated that each method assesses the exposure to repetitive tasks (i.e. the total task duration, the distribution of breaks in the shift and the performance of other repetitive tasks) in different ways, which can greatly influence the assessment outcomes. Thus, during the benchmarking, each method was applied under the following assumptions:

- The task was performed for 7 hours in total during the work period (i.e. the operators worked for 8 hours and received 1 hour of breaks during the work period);
- The breaks were evenly distributed throughout the work period (e.g. a 15 minute morning break from 1.75 – 2 hours, a 30 minute meal break from 3.75 – 4.25 hours and a 15 minute afternoon break from 6 – 6.25 hours); and
- There was no system of task rotation.

Then, to investigate the effect of reduced exposure, the benchmarking was then repeated under the following assumptions:

- The task was performed for 3.5 hours in total during the work period (i.e. less than half of a typical 8 hours working period);
- The breaks were evenly distributed throughout the work period (e.g. a 15 minute morning break from 1.75 – 2 hours, a 30 minute meal break from 3.75 – 4.25 hours and a 15 minute afternoon break from 6 – 6.25 hours); and
- The operators rotated to a non-repetitive task for the other 3.5 hours in the work period.

### 4.1.2 Analysis of task assessment results

The use of each method resulted in a ‘risk score’ (or total score) for each of the ten tasks. Along with other indicators, the risk scores can be used to help prioritise tasks for risk reduction. However, direct comparison of these scores across all methods was not possible due to the different types of data that the methods return. Therefore, the risk scores had to be converted into rank orders for comparison across all methods.

For each method, the total score can also be converted into a ‘risk level’ or exposure level, either using a three point ordinal scale (ART and Strain Index), four point ordinal scale (QEC) or five point ordinal scale (OCRA checklist). This has the effect of reducing the variability of the scores, while also proposing a level of urgency for intervention. The risk levels for each method have been proposed with various levels of validation. In several studies, the Strain Index risk levels have shown good validity in predicting incidences of ULDs in the elbow, wrist and hand (Spielholz et al., 2008; Rucker and Moore, 2002; Moore and Garg, 1995). The OCRA checklist risk levels have been calibrated to the risk levels of the OCRA Index (Occhipinti and Colombini, 2005), which itself has undergone tests of validity to predict the expected prevalence of diagnosed ULD cases and persons affected by ULD cases (Occhipinti and Colombini, 2007). With the QEC, a relationship between the risk levels and health outcomes has not been published and thus, the primary purpose of the scores remains to allow a comparison of the task before and after an ergonomics intervention. Brown and Li (2003) did propose action levels for the QEC, which were matched to the action levels of Rapid Upper Limb Assessment (RULA; McAtamney and Corlett, 1993). However, little information has been published on the relationship between RULA action levels and health outcomes.
To facilitate comparison, it was decided to convert the scores of each method into a four point ordinal scale. This was done using the ranges described in Table 20. For ART, the second exposure level (‘moderate’) was split into two categories: ‘low – moderate’ and ‘moderate – high’. The QEC total scores were converted into risk levels using the score boundaries proposed by Brown and Li (2003). The first and second OCRA checklist levels (‘acceptable’ and ‘borderline/very low’) were combined into a single level, as it was felt that each of these levels would describe very few repetitive tasks included in the benchmarking. For the Strain Index, the original risk levels (0 – 3, 3.1 – 6.9, and 7+) were retained and, to create a fourth greater risk level, a boundary of 10.1 was selected. This 10.1 boundary was selected as both Moore and Garg (1995) and Rucker and Moore (2002) found that all tasks with a SI score of 10.1 or greater were associated with incidences of distal upper extremity disorders.

### Table 20 Risk level boundaries used for benchmarking the risk assessment methods

<table>
<thead>
<tr>
<th>Risk levels</th>
<th>Risk interpretation</th>
<th>ART</th>
<th>QEC</th>
<th>OCRA</th>
<th>Strain Index</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Low</td>
<td>0 – 11</td>
<td>36 – 65</td>
<td>0 – 11</td>
<td>0 – 3</td>
</tr>
<tr>
<td>2</td>
<td>Low – Medium</td>
<td>12 – 17</td>
<td>66 – 80</td>
<td>11.1 – 14.0</td>
<td>3.1 – 6.9</td>
</tr>
<tr>
<td>3</td>
<td>Medium – High</td>
<td>18 – 23</td>
<td>81 – 113</td>
<td>14.1 – 22.5</td>
<td>7 – 10.0</td>
</tr>
<tr>
<td>4</td>
<td>High</td>
<td>24+</td>
<td>114 – 162</td>
<td>22.6+</td>
<td>10.1+</td>
</tr>
</tbody>
</table>

### 4.2 RESULTS

#### 4.2.1 Risk scores

Table 21 presents the risk scores (or total scores) that the four methods produced for the ten tasks. These scores reflect the assumption that the tasks were performed for a total duration of 7 hours. For easier comparison, the tasks have been labelled A – J and are presented in the order of lowest ART score to highest ART score.

### Table 21 Risk scores (total scores) for the benchmarking

<table>
<thead>
<tr>
<th>Task</th>
<th>ART</th>
<th>QEC</th>
<th>OCRA</th>
<th>Strain Index</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>4</td>
<td>82</td>
<td>7</td>
<td>1.5</td>
</tr>
<tr>
<td>B</td>
<td>11</td>
<td>90</td>
<td>17.5</td>
<td>6</td>
</tr>
<tr>
<td>C</td>
<td>11</td>
<td>96</td>
<td>21.5</td>
<td>13.5</td>
</tr>
<tr>
<td>D</td>
<td>22</td>
<td>104</td>
<td>43</td>
<td>8</td>
</tr>
<tr>
<td>E</td>
<td>24</td>
<td>104</td>
<td>32</td>
<td>40.5</td>
</tr>
<tr>
<td>F</td>
<td>25</td>
<td>106</td>
<td>23.8</td>
<td>27</td>
</tr>
<tr>
<td>G</td>
<td>25</td>
<td>110</td>
<td>29</td>
<td>13.5</td>
</tr>
<tr>
<td>H</td>
<td>27</td>
<td>112</td>
<td>33</td>
<td>36</td>
</tr>
<tr>
<td>I</td>
<td>28</td>
<td>110</td>
<td>53</td>
<td>36</td>
</tr>
<tr>
<td>J</td>
<td>33</td>
<td>110</td>
<td>51</td>
<td>72</td>
</tr>
</tbody>
</table>

Figure 13 compares the risk scores for ART and the OCRA checklist, which are based on a similar framework of risk factors and an additive scoring system. The boundaries for the lowest and highest risk levels are also shown for ART (solid lines) and the OCRA checklist (dotted lines). While ART scores consistently lower than the OCRA checklist, the boundary of the highest risk level for ART is positioned slightly above the OCRA checklist boundary and may need to be lowered to reflect the lower ART scores.
4.2.2 Task rankings

Figure 14 shows how the different methods ranked the ten tasks for 7 hour task durations. Even a comparison of task rankings is problematic because each method varies in its scope. The Strain Index only assesses risks to the elbow and wrist/hand. The OCRA checklist assesses risks to the shoulder, elbow, wrist/hand. In addition to assessing risks to the upper limbs, the QEC and ART include neck and back posture in their approach to assessment.

However, Figure 14 shows that there appears to be broad agreement between the four methods, particularly for those tasks ranked as having the lowest scores and the three tasks ranked as having the highest scores. For the four tasks with middle ranking scores, there is greater variation between the rankings, yet it is possible that these tasks involve a similar level of risk as many of the methods assign these tasks similar risk scores (Table 21).

With the task duration reduced to 3.5 hours, the ART, Strain Index and OCRA checklist all reduced the task scores by a factor of 0.75 and so this did not have an effect on the task rankings. However, the effect of reduced task duration on the QEC system was more complex and did alter its ranking of tasks compared to ART (Figure 15 and 16).
Figures 15 – 18 show how the ART scores compare to the task rankings for each of the other three assessment methods individually. Where a task ranking between ART and another assessment method differs by two or more positions, an explanation is offered in Table 21.

**Figure 14** Ranking of risk (total) scores of the different methods

**Figure 15** Ranking of risk scores for QEC vs. ART (7 hours task duration)

**Figure 16** Ranking of risk scores for QEC vs. ART (3.5 hours task duration)
4.2.3 Task risk levels

Figures 19 and 20 show the risk levels that each assessment method assigned to the ten tasks when performed for 7 hours and 3.5 hours respectively. Unlike the other methods, the QEC categorised all tasks to the same risk level when using the risk boundaries recommended by Brown and Li (2003). ART assessed three tasks to involve a low level of risk; however, it was only for Task A (flag turning) that this finding was supported by the OCRA checklist and Strain Index. For Tasks B and C, the OCRA checklist and Strain Index assigned a higher level of risk and possible explanations for this difference are provided in Table 21.
Figure 20 Task risk levels for 3.5 hours task duration

4.3 SUMMARY

A summary of the benchmarking findings is as follows:

1. Consensus scoring of ten tasks allowed ART to be benchmarked against three other methods already available for assessing repetitive tasks: the Quick Exposure Check (David et al., 2008), the OCRA checklist (Colombini et al., 2005) and the Strain Index (Moore and Garg, 1995)

2. Benchmarking across all four methods was problematic because of the different scoring systems adopted. In addition, the scope of each assessment method varied: the Strain Index only assesses risks to the elbow and wrist/hand; the OCRA checklist assesses risks to the shoulder, elbow, wrist/hand; the QEC and ART also include neck and back posture in their assessment framework.

3. Despite these limitations, the benchmarking exercise found that ART was able to prioritise tasks in broad agreement with the other methods. There was greater agreement between the methods for the three tasks with the highest risk rankings.

4. Compared to the OCRA checklist, ART may underestimate the level of risk for some tasks, particularly those involving static exertions of the hand and awkward shoulder postures. Following further information from a pilot of ART, it may be appropriate to re-visit the weighting of some factors, the proposed exposure levels, and how ART can be adapted to assess static exertions of the hand.

5. Unlike the Strain Index, and to some extent the OCRA checklist, ART has not yet been validated as a predictor of ULD incidences.
### Table 21: Tasks where ART task rankings or risk levels differ to the OCRA checklist or Strain Index by two or more positions

<table>
<thead>
<tr>
<th>Task</th>
<th>Method</th>
<th>Explanation</th>
<th>Comparison of risk levels</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td>Strain Index</td>
<td>Task involves static hand postures to grip a pen (duration of exertion is 100%). The Strain Index system of scoring multipliers considers these interactions to a greater extent than ART and this is reflected in the higher score for the Strain Index.</td>
<td>Strain Index assigns a much higher risk level to the task compared to ART.</td>
</tr>
<tr>
<td></td>
<td>OCRA</td>
<td>The OCRA checklist assigns higher scores to static technical actions and awkward hand postures. ART takes account of awkward hand postures but does not assign higher scores to static exertions (technical actions).</td>
<td>OCRA checklist assigns a higher risk level compared to ART.</td>
</tr>
<tr>
<td>D</td>
<td>OCRA</td>
<td>Task involves working above shoulder height for more than half the time. OCRA checklist assigns higher scores to awkward shoulder postures (e.g. 24) compared to ART (e.g. 4)</td>
<td>OCRA checklist assigns a higher risk level compared to ART.</td>
</tr>
<tr>
<td>E</td>
<td>Strain Index</td>
<td>Task involves frequent moderate intensity exertions with a knife and marked ulna deviation of the wrist. The Strain Index system of scoring multipliers considers these interactions to a greater extent than ART and this is reflected in the higher score for the Strain Index.</td>
<td>Strain Index assigns a higher risk level to the task compared to ART.</td>
</tr>
<tr>
<td>F</td>
<td>OCRA</td>
<td>Task involves sitting side on to the conveyor and twisting the neck and back to access items on the conveyor. Unlike ART, the OCRA checklist does not take account of neck and back postures and thus assigns a lower ranking score for the task.</td>
<td>ART and OCRA checklist assign an equivalent risk level.</td>
</tr>
<tr>
<td>G</td>
<td>OCRA</td>
<td>The task involves static back flexion and excessive reaching for items on a far conveyor. The task also involves static neck flexion to place items on the near conveyor. Unlike ART, the OCRA checklist and Strain Index do not take into account neck and back postures, and thus assign a lower ranking score for the task.</td>
<td>ART, OCRA checklist and Strain Index assign an equivalent risk level.</td>
</tr>
</tbody>
</table>
5 INITIAL CONSULTATION

5.1 APPROACH TO THE CONSULTATION

Following the completion of the user trials, the project team sought the informed opinion of four UK-based ergonomics consultants external to HSE. The purpose of the consultation was to seek an independent view on the suitability and usability of ART from experts who are not employed by HSE/HSL but are often asked to apply, interpret, provide expert opinion on and train others in the use of HSE’s assessment tools and guidance material related to ergonomics. All members of the consultation team had previous research experience, regular consultancy experience and, in some cases, expert witness experience in the topic of WRMSD.

The work specification required consultants:

1. To attend a 2 hour briefing on the development and use of ART (this was similar, albeit shorter, to the briefing provided to Inspectors during the user trials);

2. To apply ART to three different repetitive work tasks that they will assess in their role as an ergonomics consultant (or if such opportunities did not arise, to apply ART to video of a repetitive task that they had previously assessed);

3. To provide a description of each task, the assessment findings with ART, a comparison of the ART findings to their own personal assessment or conclusions about the task, and other assessment methods or tools that were used at the time;

4. To provide a general view on the suitability and usability of ART and what further improvements might be required; and

5. To provide a view on whether ART would be appropriate for other health and safety practitioners and what changes might be required to take this forward.

5.2 CONSULTATION FINDINGS

5.2.1 Tasks assessed

In total, ART was applied to 15 different tasks. Four tasks were assessed at the work area itself, while nine tasks were assessed away from the work area using video that was taken during previous assessments. The tasks were based in the following sectors:

- Food and drink (slicing meat products, inspecting and packing meat products, decanting frozen produce, stirring vats of prawns, transferring large and small milk bottles between roll cages).

- Manufacturing (packing phone cards into boxes, attaching buttons to a painting jig, knocking up newspaper inserts, loading hoppers with newspapers, skimming cement across a tile mould and placing the mould into a press shuttler, attaching suture thread to needles, winding suture thread around pegs and packing sutures, cutting slots into paper wallets).

- Health services (transabdominal sonography).
5.2.2 Comparison to personal assessments

For 10 of the 15 tasks (67%), consultants reported general agreement between the ART findings and their personal assessment conclusions. Explanations were provided where there was disagreement between the ART finding and experts’ personal assessments. This helped the project team to inform users about certain limitations with ART during the inspection campaign briefing:

- ART is not appropriate for assessing tasks that involve intense movements of the fingers yet no movement of the wrist or hand (for example, keyboard activity). The ‘other factors’ risk factor prompts users to consider tasks that require fine precise movements of the hand or fingers (e.g. over areas of 2 – 3 mm); however, ART does not allow users to grade the frequency/intensity of finger movements.

- When used as a tool for grading the level of risk, ART is limited to repetitive tasks and it is not appropriate for varied tasks, which involve longer work cycle times and a range of ‘sub-tasks’, but nonetheless might have a history of ULDs. When determining whether a task is repetitive and suitable for assessment, users are directed to definitions that already exist in HSG 60, and that Silverstein et al. (1986) first proposed to describe repetitive tasks:

  1. There is a cycle or sequence of movements that is repeated twice per minute or more; OR
  2. More than 50% of the task involves performing a repetitive sequence of motions.

- ART does not take account of static postures unless they meet the criteria to be classified as an awkward posture. For example, with the hand grip factor, a power grip is described in the green category, and therefore is not considered to present any risk irregardless of whether there is sufficient opportunity to relax the grip during the work cycle and irregardless of the total task duration.

For these reasons, ART should be described, as the name implies, as a tool for assessing repetitive tasks, and not an “Upper Limb Disorders Assessment Tool”. These limitations are raised during the training, however, there is little reference to these limitations in the ART document itself.

5.2.3 Qualitative feedback on usability

The general view on usability was favourable:

- “ART is quick and straightforward to complete.”
- “At the expense of the font size, the booklet is a good usable size and will be easy to use in addition to the standard notebook.”
- “The usability was good – it was relatively straightforward to use particularly for those with experience of using the MAC.”
- “Usage becomes easier with practice, but it will only be maintained with regular practice.”
• “I have enjoyed using the tool and believe that I would continue to do so, although this would partly be encouraged by clients who want to know how they measure up to HSE’s criteria.”

• “The separate ‘pull out sheets’ are a reasonable means of collecting and storing the data and they are reasonably easy to use.”

However, both at the briefing and as a result of completing some assessments, a number of improvements were suggested to improve the content and usability of ART (Table 22). Wherever the project plan allowed, these recommendations were incorporated into the version of ART piloted during the inspection campaign.

5.2.4 Qualitative feedback on suitability

The general view on suitability of ART was favourable, provided that a cautious approach was taken when interpreting the findings of the tool.

• “ART provides a good overview of the risk factors.”

• “ART provides a clear indication of the level of associated risk and a good starting point for an inspection for ULD risks.”

• “ART would be a good screening tool, providing users understood its limitations. The colour coding and risk rating for each [factor] was good for picking out the high risks, which would be useful for further [detailed] assessment and determining control measures.”

• “The subject material is more complicated to interpret for applying ART than are the issues that have to be considered for MAC assessment. Therefore, it would take a reasonable amount of educated use before anyone could produce a reliable score and recommendation. Without regular practice, the outcomes must be used only as broad indicators of risk, which is ‘all’ (as I understand it) that you require of the tool.”

5.2.5 Appropriateness of ART for use by health and safety practitioners

One consultant raised concerns about the current gap in the understanding and use of readily available methods for assessing highly repetitive tasks:

• “In a survey of 100 manufacturing workplaces, several questions probed issues relating to which methods companies use to assess ULD risks. One of the key findings was that many companies use the MAC to assess all manually intensive tasks…. There appears to be a lack of understanding when differentiating between manual tasks that require manual [handling] and those that involve low loads but high rates of repetition. This tool will be of significant benefit to health and safety representatives and supervisors in assessing ULD risks and furthering understanding of ULD risks.”

Most consultants reported that ART could be useful for health and safety practitioners, where they have the appropriate knowledge, skills and experience to apply ART and interpret the findings:

• “Those companies using MAC really like the layout and scoring system. It is highly likely that the similar format used in ART will be effective, accessible and familiar to those users.”

42
• “It would be a good screening tool for health and safety practitioners, with the provision that it isn’t a full assessment and that it might underestimate the risks associated with some tasks.”

• “I believe some consultants would find [ART] useful and they will be able to apply it appropriately. I would have reservations about health and safety professionals using this as a risk indicator, presumably following notification of a problem. I phrase it in this way because I cannot see that using ART as a routine means of assessing risk would be a particularly cost-effective use of employers time. My own use of the filter in HSG60 showed very quickly the extent of problems and possible risks associated with the tasks – to use ART for the same tasks would take considerably longer. Once an indication of risk has been identified, then the employer would be far better advised to seek specialist help, either to apply ART or the full assessment in HSG60.”

A number of suggestions were made to tailor ART to the needs of health and safety practitioners:

• “For [users] with less experience [in ULD assessment], the tool may need to be refined to include actual definitions within the [flowchart] and the [assessment guide] of the types of ‘awkward’ postures to look for and the ranges of movement, either through presenting ranges of motion in degrees or images.”

• “It may be useful to include an action sheet which records the outcome and links the assessment to the required actions.”

• “The easiest way to reduce exposure scores obtained from ART is to reduce duration, and most companies will achieve this through job rotation. It is therefore important [to explain] what is meant by ‘good’ job rotation and text [should be] inserted to explain how the ART tool could be used to: ensure [job] rotation [involves] significantly different tasks in terms of [ULD] risks and determine for what duration specific tasks can be conducted.”
<table>
<thead>
<tr>
<th>No.</th>
<th>Comments</th>
<th>Change</th>
<th>Details of change</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>May not be necessary to record start and finish times for shifts and breaks (in addition to the timeline). The key information to draw from the table is for how long an individual performs the task (a) without a break and (b) during the day (excluding breaks). Workers may know this without having to report work and rest patterns in detail. In addition, work and rest patterns may vary greatly depending upon a number of factors (e.g. individual, task demands on the day, and work organisation).</td>
<td>ART Training</td>
<td>On the task description form, the table for recording start and finish times for shifts, breaks and time spent performing other repetitive tasks removed. Separate boxes included for users to record (a) for how long an individual performs the task without a break and (b) for how long an individual performs the task during the day (excluding breaks). Training updated to instruct users to record work periods and rest periods directly onto the timeline, only if this information is needed to clarify, confirm or describe the pattern of work and recovery periods with workers or employers.</td>
</tr>
<tr>
<td>2</td>
<td>Risk criteria for the 'breaks' factor do not always describe how long an individual performs a task without a break (e.g. where there are two breaks in a 7 to 8 hour shift, an individual could perform the task for many more hours if the spacing of the breaks is not optimised)</td>
<td>ART</td>
<td>Risk criteria for the 'breaks' factor altered to reflect consistently how long an individual performs a task without a break (less than 1 hour, 1 hour to less than 2 hours, 2 hours to less than 3 hours, 3 hours to less than 4 hours, and more than 4 hours).</td>
</tr>
<tr>
<td>3</td>
<td>The layout of the score sheet may be counter-intuitive with the colour code recorded on the left half of the score sheet and the numerical score recorded on the right half of the score sheet. Mentally, users may score the colour code and numerical factor together.</td>
<td>For future consideration</td>
<td>Considered changing the score sheet so that the colour codes and numerical score for the left hand are recorded on the left half of the page and the colour codes and numerical score for the right hand are recorded on the right half of the page. This change was not implemented before the pilot, as without time to test, it was uncertain whether the revised score sheet structure would introduce additional scoring errors.</td>
</tr>
<tr>
<td>4</td>
<td>On the flowchart, there is potential to make an error in translation, as the left and right hand scoring columns for each stage of the assessment must be copied to the bottom of the page where the left hand scores are entered into the top row and the right hand scores are entered in the bottom row.</td>
<td>ART</td>
<td>At the bottom of the flowchart where scores are entered into rows, the positions to enter the left hand scores are staggered slightly left of the positions to enter the right hand scores.</td>
</tr>
<tr>
<td>No.</td>
<td>Comments</td>
<td>Change</td>
<td>Details of change</td>
</tr>
<tr>
<td>-----</td>
<td>--------------------------------------------------------------------------</td>
<td>--------</td>
<td>----------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>5</td>
<td>Use of repeated wording throughout the tables in the assessment guide (e.g. shoulder and arm movements are infrequent, shoulder and arm movements are frequent)</td>
<td>ART</td>
<td>Repeated words removed.</td>
</tr>
<tr>
<td>6</td>
<td>Action level 2 (changes required in the near future) and action level 3 (changes required soon) are too similar and need re-phrasing to make them more distinct</td>
<td>ART</td>
<td>Action levels 2 and 3 combined into a single ‘medium’ exposure level where “further investigation is required”. Within the single action level, scores can be used to assist with prioritising interventions.</td>
</tr>
<tr>
<td>7</td>
<td>On occasion it was necessary to flick through the accompanying information to check duration definitions and multipliers</td>
<td>ART</td>
<td>Duration definitions and multipliers included on flowchart.</td>
</tr>
<tr>
<td>8</td>
<td>Breaks risk factor does not consider ‘micro-breaks’ within the work cycle</td>
<td>ART</td>
<td>Short frequent breaks within the work cycle that leave the muscle groups totally at rest included in ART (e.g. micro-breaks of at least 10 seconds every few minutes over the whole work period).</td>
</tr>
<tr>
<td>9</td>
<td>List of ‘other factors’ within the scoring table is confusing, as it requires users to consider the presence, duration and total number of ‘other factors’ before selecting a category.</td>
<td>ART</td>
<td>List of ‘other factors’ placed before the scoring table. The requirement to consider the duration of ‘other factors’ within the work cycle was removed from the assessment. Users instructed to check the list for the presence of ‘other factors’ in the task, and the score the risk factor according to the total number of ‘other factors’ identified. This structure has the benefit of being similar to ‘other environmental factors’ for MAC</td>
</tr>
</tbody>
</table>
6 CONCLUSIONS

The Assessment of Repetitive Tasks (ART) was developed as an Inspectors’ tool that can be used:

(1) To screen repetitive tasks of the upper limbs for common physical risk factors that contribute to the development of upper limb disorders;

(2) To raise duty holders’ awareness and understanding of the risks of repetitive tasks;

(3) To demonstrate the presence of risk to duty holders;

(4) To give a broad indication of the level of risk; and

(5) To recommend areas for improvement.

6.1 CONCEPTUALISATION

The technical content of ART draws upon earlier work to develop the Occupational Repetitive Actions methods (OCRA; Colombini et al., 2002) and the Quick Exposure Check (David et al., 2008). The format of ART is based upon the Manual handling Assessment Charts (MAC; Monnington et al., 2002), and uses a logical flowchart process and ‘traffic-light’ system to grade risks. As a result, ART examines twelve risk factors that have been grouped into four stages:

(1) Frequency and repetition of movements;

(2) Force;

(3) Awkward postures (of the neck, back, shoulder/arm, wrist and hand); and

(4) Additional factors (which include aspects of task duration, recovery, perceived work pace and other object and work environment factors).

ART was developed using an iterative design process, where results and suggestions from repeated user trials and peer-review exercises were used to inform further improvements to the tool and associated training material (e.g. presentations and case studies).

6.2 USER TRIALS

Thirty-two HSE and LA Inspectors participated in simulated user trials, which used video case studies to explore the concept of ART with users and make improvements to the usability of ART. The observation-based assessment of quick hand and arm movements is inherently difficult. Nonetheless, the usability of ART was perceived to be favourable, and improved with successive prototypes of the tool. LA Inspectors reported that they would not use ART as frequently as HSE Inspectors, possibly due to the type of premises enforced by local authorities, and because ART is not intended for use in DSE assessments.

Experience suggests that an element of training and regular use is required in order for Inspectors to use ART reliably and arrive at broadly similar results to ‘Specialists’ (i.e. HSE/HSL ergonomists who contributed to the project). With training available, ART and its associated training materials were deemed to be sufficiently credible and usable to pilot with a
selection of trained HSE and LA Inspectors during an inspection campaign. An evaluation of the extent to which training might improve reliability was beyond the scope of this study.

6.3 BENCHMARKING

A consensus of Specialists’ scores for ten tasks were used to benchmark ART against three other methods already available for assessing repetitive tasks: the Quick Exposure Check (David et al., 2008), the OCRA checklist (Colombini et al., 2002) and the Strain Index (Moore and Garg, 1995). When using ART, it was possible was able to prioritise tasks and arrive at risk levels that were in broad agreement to the other methods. There was greater agreement between the methods for the three tasks with the highest risk rankings.

Compared to the OCRA checklist and Strain Index, ART may underestimate the level of risk for some tasks, particularly those involving static exertions of the hand and awkward shoulder postures. Following further information from a pilot of ART, it may be appropriate to re-visit the weighting of some factors, the proposed exposure levels, and how ART can be adapted to assess static exertions of the hand. ART has not yet been validated as a predictor of ULD incidences.

6.4 INITIAL CONSULTATION

A consultation exercise was held to seek an independent view on the suitability and usability of ART from four experts who are not employed by HSE/HSL but are often asked to apply, interpret, provide expert opinion on and train others in the use of HSE’s assessment tools and guidance material related to ergonomics. The general view on the suitability of ART was favourable, as long as a cautious approach was taken when interpreting the findings of the tool.

It was suggested that ART would be a useful instrument if released to others with responsibility for the design, inspection, assessment and organisation of repetitive work. It was recognised that the easiest way to reduce exposure scores obtained from ART is to reduce duration, and most companies will achieve this through job rotation. It will therefore be important to explain how ART can used be used to assess job rotation and promote ‘good’ job rotation.
Please answer the following questions to assist us in evaluating the effectiveness and usability of the ART. The information that you provide will help ensure that repetitive tasks can be assessed with accuracy, ease and confidence.

**BACKGROUND INFORMATION**

**Who is your employer?**

- [ ] HSE
- [ ] A Local Authority
- [ ] Other, please specify___________

**What is your current role in health and safety?** (Tick most appropriate)

- [ ] Inspector
- [ ] Occupational health inspector
- [ ] Health and safety awareness officer
- [ ] Other, please specify___________________________________________________

**How much experience do you have in your current role?**

- [ ] Less than 2 years
- [ ] 2 – 5 years
- [ ] 6 – 10 years
- [ ] More than 10 years

**What previous training or briefings have you received on the topic of musculoskeletal disorders?** *(Tick all that apply)*

- [ ] MAC tool
- [ ] MSD course for inspectors (3 day)
- [ ] Better Backs campaign briefing
- [ ] Display screen equipment training
- [ ] Other(s), please specify___________________________________________________

**Have you used a previous version of ART before today?** *(Tick all that apply)*

- [ ] Yes, used ART to assess tasks in the field
- [ ] Yes, used ART to assess tasks on video
- [ ] No, not used ART before today
SYSTEM USABILITY SCALE (SUS)

Please indicate to what extent you disagree or agree with the following statements

1. I think that I would like to use ART frequently

2. I found ART unnecessarily complex

3. I thought ART was easy to use

4. I think that I would need a technical person to be able to use ART

5. I found the various steps in ART were well integrated

6. I thought there was too much inconsistency in ART

7. I would imagine that most inspectors would learn to use ART quickly

8. I found ART very cumbersome to use

9. I felt very confident using ART

10. I needed to learn a lot of things before I could get going with ART

Strongly Disagree □ □ □ □ □ Strongly Agree □ □ □ □ □
<table>
<thead>
<tr>
<th>Risk Factor</th>
<th>Rating</th>
<th>If 1 or 2, please describe why it was difficult:</th>
</tr>
</thead>
<tbody>
<tr>
<td>A1. Shoulder / arm movements</td>
<td>1----2----3----4----5</td>
<td></td>
</tr>
<tr>
<td>A2. Repetition</td>
<td>1----2----3----4----5</td>
<td></td>
</tr>
<tr>
<td>B. Force</td>
<td>1----2----3----4----5</td>
<td></td>
</tr>
<tr>
<td>C1. Head / neck posture</td>
<td>1----2----3----4----5</td>
<td></td>
</tr>
<tr>
<td>C2. Back posture</td>
<td>1----2----3----4----5</td>
<td></td>
</tr>
<tr>
<td>C3. Shoulder / arm posture</td>
<td>1----2----3----4----5</td>
<td></td>
</tr>
<tr>
<td>C4. Wrist posture</td>
<td>1----2----3----4----5</td>
<td></td>
</tr>
<tr>
<td>Risk Factor</td>
<td>Rating</td>
<td>If 1 or 2, please describe why it was difficult:</td>
</tr>
<tr>
<td>-----------------------------</td>
<td>---------</td>
<td>--------------------------------------------------</td>
</tr>
<tr>
<td>C5. Hand grip</td>
<td>1-2-3-4-5</td>
<td></td>
</tr>
<tr>
<td>D1. Breaks</td>
<td>1-2-3-4-5</td>
<td></td>
</tr>
<tr>
<td>D2. Work pace</td>
<td>1-2-3-4-5</td>
<td></td>
</tr>
<tr>
<td>D3. Miscellaneous factors</td>
<td>1-2-3-4-5</td>
<td></td>
</tr>
<tr>
<td>E. Duration</td>
<td>1-2-3-4-5</td>
<td></td>
</tr>
</tbody>
</table>
If available, would you use ART for investigating upper limb disorder risks in the workplaces you visit? *(Circle most appropriate)*

<table>
<thead>
<tr>
<th>Never</th>
<th>Seldom</th>
<th>Occasionally</th>
<th>Often</th>
<th>Almost Always</th>
</tr>
</thead>
</table>

Are there any repetitive upper limb tasks that you have come across where you feel that ART is not suitable for assessing the risks? If so, please describe the situation.

To what extent do you feel that the use of ART improves your confidence when assessing upper limb disorder risks? *(Circle most appropriate)*

<table>
<thead>
<tr>
<th>Does Not Improve</th>
<th>Slightly Improves</th>
<th>Improves</th>
<th>Greatly Improves</th>
</tr>
</thead>
</table>

To what extent do you feel that the use of ART improves your understanding of upper limb disorder risks? *(Circle most appropriate)*

<table>
<thead>
<tr>
<th>Does Not Improve</th>
<th>Slightly Improves</th>
<th>Improves</th>
<th>Greatly Improves</th>
</tr>
</thead>
</table>

To what extent do you feel that the use of ART helps to identify what action to take to reduce upper limb disorder risks? *(Circle most appropriate)*

<table>
<thead>
<tr>
<th>Does Not Help</th>
<th>Slightly Helps</th>
<th>Helps</th>
<th>Greatly Helps</th>
</tr>
</thead>
</table>

Please describe any sections of ART that could be improved with additional written instruction. Be as specific as possible. *(Please include page number)*

**What do you like or not like about ART?** *(You may continue writing on the next page, if you wish)*

Many thanks again for completing this questionnaire!
8 REFERENCES


Kilbom, A. (1994)

*Selecting and applying a series of usability methodologies to MAC: an iterative design approach*. MSc Dissertation, Robens Centre for Health Ergonomics, University of Surrey.


Li, G. and Buckle, P. (1999a)

Li, G. and Buckle, P. (1999b)
Current techniques for assessing physical exposure to work related musculoskeletal risks, with emphasis on posture based methods. *Ergonomics*, 42(5), 674 – 695.

McAtamney, L. and Corlett, E.N. (1993)


Development of an assessment tool for repetitive tasks of the upper limbs (ART)

The prevention, control and management of work-related musculoskeletal disorders (WRMSD) are a priority for the Health and Safety Executive (HSE) and one of the key ways of improving occupational health in Great Britain. Health and Safety Executive (HSE) and Local Authority (LA) Inspectors play an important role in preventing WRMSD. As well as enforcing health and safety law, they provide advice on risk factors and control measures on a range of health and safety issues. To support this work, assessment tools are required that offer an intuitive and relatively quick process to screen workplaces for high-risk activities, raise awareness of risk factors, demonstrate the presence of risk, and recommend areas for improvement.

The Manual handling Assessment Charts tool (MAC, Monnigton et al., 2002) has been recognised as a useful process to screen workplaces for high-risk manual handling operations (Melrose et al., 2006; Lee and Ferreira, 2003; Tapley, 2002). The MAC tool is designed to identify and help assess high-risk manual handling operations (lifting, lowering, carrying and team handling). However, its scope is limited to manual handling of heavy items, which primarily present a risk of lower back disorders. There was a demand for a similar tool that Inspectors can use to screen more frequent handling of light loads or other repetitive tasks and the common physical risk factors in work that may contribute to upper limb disorders (ULDs). This led to the development of the assessment of repetitive tasks (or ART) tool.

This report and the work it describes were funded by the Health and Safety Executive (HSE). Its contents, including any opinions and/or conclusions expressed, are those of the authors alone and do not necessarily reflect HSE policy.