Objective: To evaluate the use of the Human Reliability Analysis of Cataract Surgery tool to identify the frequency and pattern of technical errors observed during phacoemulsification cataract extraction by surgeons with varying levels of experience.

Design: Observational cohort study. Thirty-three consecutive phacoemulsification cataract operations were performed by 33 different ophthalmic surgeons with varying levels of operative experience: group 1, fewer than 50 procedures; group 2, between 50 and 250 procedures; and group 3, more than 250 procedures. Face and content validity were surveyed by a panel of senior cataract surgeons. The tool was applied to the 33 randomized and anonymous videos by 2 independent assessors trained in error identification and correct tool use. Task analysis using 10 well-defined end points and error identification using 10 external error modes were performed for each case. The main outcome measures were number of errors performed per task, nature of performed errors (executional or procedural), and surgical experience of operating surgeon.

Results: Analysis of 330 constituent steps of 33 operations identified 228 errors, of which 151 (66.2%) were executional and 77 (33.8%) were procedural. The overall highest error probability was associated with sculpting, followed by fragmentation of the nucleus; this was most evident in group 1. Surgeons in group 3 proportionally performed more errors during removal of soft lens matter than those in group 1 or 2. Surgical experience had a significant effect on the number of errors, with a statistically significant difference among the 3 groups ($P < .001$).

Conclusions: The Human Reliability Analysis of Cataract Surgery tool is useful for identifying where technical errors occur during phacoemulsification cataract surgery. The study findings, including the high executional error rate, could be used to enhance and structure resident surgical training and future assessment tools. Face, content, and construct validity of the tool were demonstrated.

Arch Ophthalmol. 2008;126(2):173-177
Video recordings were taken of consecutive phacoemulsification cataract operations performed by different ophthalmic surgeons from 3 UK National Health Service cataract surgery centers. Ethics committee approval was sought before video recording, and individual patient consent was obtained for all recorded videos. Complete cataract surgical operations were recorded through the operating microscope. Patients with ophthalmic comorbidity, poor pupil dilation, mature cataract, previous trauma, axial length beyond 22 to 26 mm, or other high surgical risk factors identified at preoperative assessment were excluded. Therefore, all included cases were deemed by us to have been suitable for the most junior surgeon in the cohort to undertake, thereby promoting comparative consistency. The videotapes were sent to an independent technician who digitized them using videoediting software (Adobe Premier Pro 1.5; FFCSoftware Co Ltd, Kingston, New York). The technician also removed any logos or other characteristics that could identify the surgeon or the training unit. The digitized videotapes were coded, made anonymous, and randomized.

The recorded cases were divided according to the cumulative phacoemulsification cataract operative experience of the surgeon. Group 1 had performed fewer than 50 procedures; group 2, between 50 and 250 procedures; and group 3, more than 250 procedures.

A panel of 16 senior phacoemulsification cataract surgeons was assembled to establish the face and content validity of the tool. The panel also set out to identify thresholds of error identification for this procedure in view of differing surgical techniques. To achieve the latter objective, the panel identified a “gold standard” cataract case, suitable for the most junior cohort of surgeons and free from concurrent pathological features, which was used to set error thresholds. The face and content validity survey was performed using a structured feedback form (Figure 1). There was no preference as to the technique of phacoemulsification performed because the generic end points of the Human Reliability Analysis of Cataract Surgery (HRACS) tool were deemed to be applicable to most phacoemulsification techniques. Following modification and refinement of the HRACS tool based on structured feedback received from the expert panel, the final version of the tool used in the study is shown in Figure 2. Two independent senior ophthalmic surgical trainers then applied the HRACS error analysis tool to the videos. Both assessors were surgical trainers, and each had more than 10 years of cataract surgical experience. They were trained in error identification, including viewing of the gold standard cataract case, and on the correct use of the HRACS tool.

**METHODS**

The definition of human error used in the present study is “something which has been done which was: (i) not intended by the actor, (ii) not desired by a set of rules or an external observer, or (iii) that led the task or system outside acceptable limits.”11(p1216) This definition was agreed upon at the Bellagio Conference on Human Error and has been previously used in earlier Observational Clinical Human Reliability Assessment surgical studies.6,11 Similar principles of error identification were applied to cataract surgery in this study, in which an error was defined as an action that was clearly not intended by the surgeon; an action that was not performed to a predetermined standard, as set by our panel of expert cataract surgeons; or an action that increased the likelihood of a negative consequence or was not within acceptable safety limits of the surgical procedure.

A manual procedural task may be erroneously performed in a number of ways, referred to as external error modes in the Systemic Human Error Reduction and Prediction Approach, the original project on human error in industry.7 The modified HRA method of error assessment is an observational technique that involves breaking down any manual procedure sequentially into distinct tasks. Ten manual external error modes are used to define and count errors enacted in each task. Using this as a platform, a categorization of human error was developed by Joice et al8 at the Surgical Skills Unit, University of Dundee, Dundee, Scotland (Table 1).

This comprehensive set of external error modes includes the types of observable errors committed when performing a manual procedure. The 10 external error modes can be divided into 2 groups: 1 to 6, which are procedural error modes, disrupting the correct sequence of procedural steps; and 7 to 10, which are executional error modes, implying incorrect manipulation of instruments and tissues.6,11,13 Grouping errors in this manner determines the nature of corrective measures: procedural error modes may be reduced by initiating a standardized operative task sequence, while executional error modes may be minimized by practicing surgical skills.13

**ERROR IDENTIFICATION**

Figure 1. The structured Human Reliability Analysis (HRA) feedback form used in the face and content validity survey.

**Figure 2.** The Human Reliability Analysis of Cataract Surgery tool. AC indicates anterior chamber; EEM, external error modes.
The procedure of cataract extraction was divided into 10 generic tasks similar to the Observational Clinical Human Reliability Assessment tool. These tasks were selected as end points because they are well-defined, observable, and found in all techniques of phacoemulsification surgery; their sequential completion is necessary for the operation to proceed.

DATA ANALYSIS

For every recorded procedure, each of the 10 components forming the operative tasks was observed to record errors performed and their external modes. Errors were analyzed for each component task and for surgical experience. The error probability for each task was calculated using the following formula: % Error Probability of Task = 100 × (No. of Times the Task Was Erroneously Performed/Total No. of Times the Task Was Performed).

Quantitative error data were expressed as mean (SD). The Kruskal-Wallis test was used to evaluate statistical significance, set at \( P < .05 \).

RESULTS

The results of the face and content validity survey are shown in Table 2. All feedback received was used to improve the tool. For instance, one panel member suggested we include wound enlargement (if appropriate) to the intraocular lens insertion step; therefore, the tool was amended accordingly. Other feedback included ensuring that the tool is always provided with the external error modes visible on the same page as the assessment components; this suggestion was put into practice during the study. The final version of the tool was applied to 50 recorded cataract videos; 33 were included in the study following the application of exclusion criteria previously mentioned (9 were in group 1, 7 were in group 2, and 17 were in group 3). Each recorded case was completed by a different operating surgeon in its entirety. None of the procedures used were abandoned or required a more senior surgeon to intervene. All cases used a divide-and-conquer, stop-and-chop, or pure chop technique of phacoemulsification.

Analysis of 330 constituent steps of the 33 operations identified 228 errors. Of these, 151 (66.2%) were executional and 77 (33.8%) were procedural. Figure 3 graphically summarizes the number of errors performed per task by each of the 3 groups. Table 3 shows the mean (SD) of executional and procedural errors performed within each group, along with error probabilities for each task. The overall analysis suggests that the highest proportion of errors occurs during the engagement, sculpting, and rotation or manipulation of the nucleus followed by fragmentation of the nucleus.

Statistical analysis of the impact of surgical experience on number of errors shows a statistically significant difference between the 3 groups (\( P < .001 \)). This trend is illustrated by the box plots shown in Figure 4.

COMMENT

The study of human error in surgery seems to provide an attractive quality assurance strategy that may be of interest to surgical trainees and supervising surgeons. Better understanding of such errors could have a positive effect on learning, especially on the acquisition of technical skills. This would enable us to refine our existing and future training and assessment systems by identifying underlying surgical performance error mechanisms among surgeons with varying levels of experience. Using the HRACS tool in this study, most errors identified were of the execution type (66.2%), suggesting that the surgeons were largely following the correct sequence of steps within the procedure most of the time, but failed to execute some of the component tasks adequately. The study found a statistically significant difference in the quantity of errors committed between the 3 groups of ophthalmologists studied, which reflects their operative experience. This
demonstrates construct validity of the HRACS tool when applied to phacoemulsification surgery.

Setting the acceptable standards for the procedure was a challenging and laborious task because it occasionally involved identifying arbitrary thresholds of error for some tasks based on the cumulative experience and consensus of the expert surgical panel. This was easier to perform for potential procedural errors, but more complex when considering executional errors. For instance, the incorrect sequence of tasks can be easily agreed on if erroneous; however, the threshold speed at which an action is classified as an “error” is not as easy to define. In this study, a gold standard cataract case was agreed on by the senior panel as a baseline on which error identification thresholds were set. It may transpire that this process may need to be emulated within each training program or research environment in which the tool is intended to be used so that appropriate standards can be set locally.

In this study, the assessors were both experienced surgeons and surgical trainers with considerable exposure to different phacoemulsification techniques. Despite this, they were trained in the process of error identification to help reduce the element of subjectivity and increase interassessor consistency. A limitation to HRACS use as an assessment tool is that implementing the system adequately can be labor intensive and requires training in error identification. Further research on the interrater variability of the tool is required.

Table 3. Data for Executional and Procedural Errors Observed for Each Generic Task Performed by Each Surgeon Group

<table>
<thead>
<tr>
<th>Task</th>
<th>Group 1 (n=9)</th>
<th>Group 2 (n=7)</th>
<th>Group 3 (n=17)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean (SD)</td>
<td>Error Probability, %</td>
<td>Mean (SD)</td>
</tr>
<tr>
<td>Wound construction</td>
<td>0.90 (0.74)</td>
<td>0.40</td>
<td>0.63 (0.74)</td>
</tr>
<tr>
<td>Viscodlastic injection</td>
<td>0.20 (0.42)</td>
<td>0.30</td>
<td>0.13 (0.35)</td>
</tr>
<tr>
<td>Capsulorrhexis</td>
<td>1.50 (0.85)</td>
<td>0.30</td>
<td>0.75 (0.35)</td>
</tr>
<tr>
<td>Hyrodisssection</td>
<td>0.60 (0.70)</td>
<td>0.10</td>
<td>0.13 (0.35)</td>
</tr>
<tr>
<td>Sculpting of the nucleus</td>
<td>1.50 (0.71)</td>
<td>0.50</td>
<td>0.88 (0.64)</td>
</tr>
<tr>
<td>Cracking of the nucleus</td>
<td>0.90 (0.88)</td>
<td>0.30</td>
<td>0.38 (0.64)</td>
</tr>
<tr>
<td>Removal of nuclear fragments</td>
<td>0.60 (0.70)</td>
<td>0.40</td>
<td>0.75 (0.64)</td>
</tr>
<tr>
<td>Soft lens matter aspiration</td>
<td>1.10 (0.88)</td>
<td>0.30</td>
<td>1.00 (1.07)</td>
</tr>
<tr>
<td>Lens insertion</td>
<td>1.60 (0.70)</td>
<td>0.30</td>
<td>0.87 (0.74)</td>
</tr>
<tr>
<td>Wound closure</td>
<td>0.30 (0.48)</td>
<td>0.40</td>
<td>0.13 (0.35)</td>
</tr>
</tbody>
</table>

Figure 3. The total number of errors performed per task for each surgeon group: group 1 performed fewer than 50 operations; group 2, between 50 and 250 operations; and group 3, more than 250 operations.
When constructing the Objective Structured Assessment tool, it can be used to guide assessment tool design, as was done in this study. Having established such associations, this information can be used to guide assessment tool design, as was done in this study. The HRACS tool used in this study provides valuable information to this effect.

Submitted for Publication: April 1, 2007; final revision received July 10, 2007; accepted July 11, 2007.

Correspondence: Vinod Gauba, FRCOphth, MSc, PAMedEd, 33 Alder Hill Ave, Leeds, West Yorkshire LS6 4JQ, England (vgauba@aol.com).

Financial Disclosure: None reported.

REFERENCES


Figure 4. The effect of surgical experience on the total number of errors observed in the 3 groups.