EE1 ORC6
FACT SHEET

How to conduct a behavioural study to test chart modifications

"We have made changes to the ORCs that are not on the approved list. How can we test these modifications to assess potential risks?"

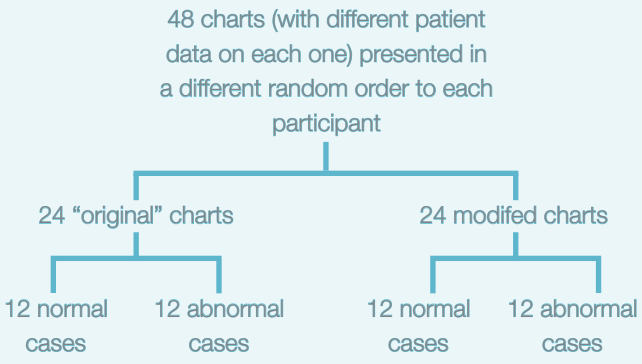
While there are many different ways to carry out an empirical study to investigate user performance with different chart designs, in the simple example outlined here, a modified chart is compared with an original observation and response chart (ORC). The aim is to determine whether the modification significantly increases error frequency and/or decision times when users carry out a task relevant to patient safety. Note that the expertise associated with conducting this type of study is not trivial and you are advised to seek specialist help (this example is just intended to provide a basic understanding of what might be involved).

This fact sheet is one of a series that provides specific information about the process of selecting, implementing and using an observation and response chart. The other fact sheets in this series are:

- EE1 ORC1 Introducing an observation and response chart
- EE1 ORC2 Modifying the observation and response chart for local use
- EE1 ORC3 Potential practice changes associated with implementing an observation and response chart
- EE1 ORC4 Training clinicians to use the observation and response charts
- EE1 ORC5 Why is it crucial to test any non-approved ORC modifications?

Figure 1: The testing process for each study participant

This figure relates to Step 4 of the study process which is outlined overleaf.



Each participant would see half of the 48 sets of patient data recorded on one chart, and the other half recorded on the other chart (that is, they would never see the same data more than once). The 24 cases presented on each chart would comprise 12 "normal" and 12 "abnormal" cases. Participants should be randomly assigned to two groups. The cases assigned to each chart for the first group should be assigned to the alternative chart for the second group.

STEP 1: Determine an appropriate task that best encapsulates the type of errors likely to be the most dangerous

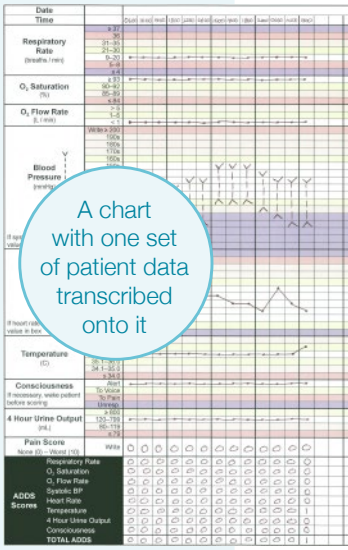
In this example, the task will be to detect whether any vital sign reading is in the abnormal range on a completed chart. Ideally, you would repeat this exercise simulating all potentially risky tasks associated with using the chart (e.g. whether data can be recorded onto the chart with minimal errors or whether appropriate actions are taken when vital signs are abnormal).

STEP 2: Create the materials needed for your task

In this example, this would involve collecting anonymous patient vital sign data (over sufficient time points to fill a chart) to be transcribed onto both the original and modified charts.

Half of these cases would have at least one vital sign outside the normal range and half would have all vital signs inside the normal range.

To make the evaluation as sensitive as possible, abnormal cases should not be dramatic (e.g. if abnormal data points equate to emergency situations then this is more likely to be noticed even on a poorly designed chart). The aim is to select cases that are subtle enough to give the best opportunity to differentiate the charts based on their design.



You will need enough data to fill sufficient charts to give your study enough sensitivity to be able to detect clinically meaningful differences in performance between the two charts (a typical number might be 48 complete sets of patient data).

STEP 3: Recruit participants to take part in your study

You could choose these to be either experienced chart users (nurses and doctors) or novices (for example, student nurses or medical students), or groups of each.

The number of people you recruit will determine how sensitive your study will be at discriminating between the two charts (a typical number might be 50 people per group).

FACT SHEET

How to conduct a behavioural study to test chart modifications

STEP 4: Test all your participants according to a set procedure

In our example, each participant in the study would be seated at a desk next to a researcher.

- Both charts would be explained to them (this could be done via video presentation).
- They would be asked to memorise relevant vital sign normal ranges (given that these often vary between and within organisations).
- Their knowledge of the normal ranges would be tested using a multiple choice test to ensure they had memorised these successfully (they would not be allowed to continue until they had obtained 100% in this test).
- They would have the experimental task explained to them.
- Typically each participant might receive 48 trials (more trials = greater study sensitivity).
- In each trial, the researcher would present the participant with completed charts and ask them to indicate whether any vital signs were abnormal.
- If applicable, the participant would then be asked to specify the abnormal vital sign(s).
- For each trial, the researcher would record their response (later to be scored as “correct” or “incorrect”) and also how long they took to respond (this could recorded using a stopwatch or computer software).

See Figure 1 for an illustration of the breakdown of charts viewed by each participant in the study.

STEP 5: Enter and analyse your data

- The data for all participants across all trials needs to be entered into a spreadsheet and then grouped by chart (modified versus original).
- The number of incorrect responses would need to be totalled for each chart and divided by the total number of trials with that chart (24 in our example, with 50 participants) in order to obtain the error rate for that chart. Then you will be able to compare average error rates across both the modified and original charts.
- You need to determine whether any difference in error rates between the two charts is statistically significant (that is, whether it is unlikely to be due to chance). You may need to consult a statistician but, typically, a two-tailed Paired T-test would be used, if the assumptions for this test were met. If the “p-value” generated by this test is less than 0.05 then we would regard any difference between the error rates of the two charts to be unlikely to be due to chance. That is, the performance on one chart is probably genuinely superior to performance on the other.
- You would repeat this process for the response time data. This would tell you which chart design results in the fastest judgements (which is also a key indicator of performance).

In addition to the type of behavioural study described here, it might also be valuable to carry out clinical assessments. This could involve, for example, conducting prospective and retrospective trials using real patient data to determine the extent to which adverse events would have been flagged by the track and trigger system and also to determine the false alarm rate of the system.

References

1. Christofidis MJ, Horswill MS, Hill A, Watson MO. Human factors design and observation charts. 7th International Conference on Rapid Response Systems and Medical Emergency Teams. Sydney, 2012.
2. Preece MHW, Hill A, Horswill MS, & Watson MO. Supporting the detection of patient deterioration: Observation chart design affects the recognition of abnormal vital signs. Resuscitation, 83(9);1111-1118.

This fact sheet was prepared by Mark Horswill, Andrew Hill, Melany Christofidis and Marcus Watson of The University of Queensland: chartdesign@psy.uq.edu.au

Further information

Further information about implementing recognition and response systems can be found in the Australian Commission on Safety and Quality in Health Care publication *A Guide to Implementation of the National Consensus Statement: Essential Elements for Recognising and Responding to Clinical Deterioration* (2012).

This can be downloaded from: www.safetyandquality.gov.au

Australian Commission on Safety and Quality in Health Care

GPO Box 5480

Sydney NSW 2001

Telephone: (02) 91263600

Email: mail@safetyandquality.gov.au