

Research and Critical Skills, 2002-3

Specimen Answer: Assessment 1

Some of these questions do not have a single right answer. This is an example of the sort of thing we wanted.

1. *The study is described as "single-blind" in the abstract, though no detail is given as to how this was done. What is meant by "single-blind"? What would the authors have to do to make this study single-blind? Why could the study not have been double-blind?* "Double-blind" means that treatment is concealed from both subjects and investigators, "single-blind" means that treatment is concealed from only one of these, usually the investigators. To make this single-blind, the authors would ensure that the investigator who made the measurements did not know what treatment the subjects had received. The study could not be double-blind because the subjects know whether they are exercising or not. Some measurements (quality of life) are self-reported and are therefore not blind at all.
2. *In Tables 2 and 3, means are given " $\pm SD$ ". What does "SD" mean here and what does it tell us about the data? Roughly how many observations will lie within the interval defined by mean $\pm SD$? "SD" means "standard deviation". Standard deviation is a measure of variability, the square root of the variance, in the same units as the observations. About 2/3 of observations lie within one SD from the mean. (Anything between 60% and 70% would be acceptable here.)*
3. *In Table 2, SDs are given for VO_2 . What features of the data suggest that there is a mistake in the SDs in the "before" column? There are three reasons to think the SDs are wrong. First, they are very small, less than 1% of the mean. This is very unusual for a biological measurement. Second, as 95% of observations should be within 2SD from the mean, they suggest no overlap between the three groups. These were randomised, so the only difference between them should be due to chance. Third, the SDs in the "after" column are 10 times those in the before column. They should be similar.*
4. *In Table 4, "median" and "range" are quoted. What do these terms mean? What do they suggest about the shape of the distribution of the Sickness Impact Profile score? The median is the central value when observations are arranged in numerical order, 50% of observations exceed it and 50% are less. The range is the difference between the maximum and minimum values, though here the maximum and minimum are both given. All the medians are closer to the minimum than to the maximum, suggesting that the distributions are positively skew.*
5. *In Table 3, for distance walked by the physical training group compared to the non-training group a P value < 0.003 is quoted. What does this mean? What can we conclude from it about whether physical training increases 6 minute walking distance in women with Coronary Syndrome X? The probability that we would get a difference as large as or larger than this in the sample if there were no difference in the population (the null hypothesis) is less than 0.003. This is small, less than 0.01, so we can conclude that there is strong evidence that women given training have a greater increase in their walking distance than women who are not given training.*

6. *In Table 4, for difference in Sense of Coherence score within the physical training group "NS" is quoted. What does this mean? What can we conclude from it about whether physical training increases Sense of Coherence score in women with Coronary Syndrome X? "NS" means "not significant". The probability that we would get a difference as big as or bigger than that observed in the sample if there were no difference in the population (the null hypothesis) would be greater than 0.05. We conclude that there is little or no evidence that Sense of Coherence score increases following training. However, we cannot conclude that it does not increase; the sample may be too small to detect the increase which occurs.*

Research and Critical Skills, 2002-3

Specimen Answer: Assessment 2

This is a possible answer to the exercise. It does not necessarily cover all possible points in the answers, but it is the sort of thing we were looking for.

Your script has been retained so that the external examiners can see it if they wish.

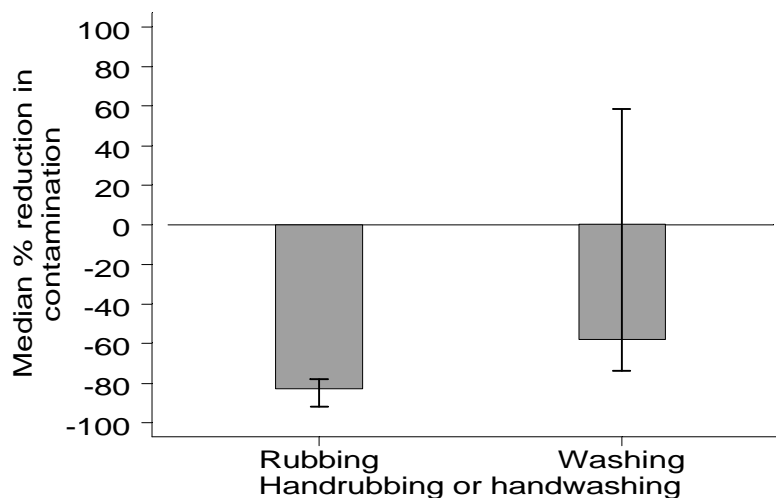
- (a) *"A total of 23 healthcare workers were included in the study and analysed; 12 were randomised to handrubbing and 11 to handwashing." What does "randomised" mean and why were the participants randomised to handrubbing or to handwashing?* "randomised" means that the subjects were put into the two groups by chance, so that the probability that each would be allocated to handrubbing was the same and unrelated to the subject's own characteristics. This was done to obtain comparable groups, which would be similar in all respects except the intervention.
- (b) *"The microbiologist who examined the culture plates and reported the microbiological results was unaware of the hand hygiene method used." Why is this important?* This aspect of the trial is single-blind. If the microbiologist had known which group the culture plate came from, it is possible that the count might have been influenced by this knowledge and tend to report in favour of one treatment or the other. Blinding prevents this possibility.
- (c) *In table 2, "Median (IQR)" for bacterial count in the handrubbing group before hand cleaning is "101 (29-380)". What is meant by the terms "median" and "IQR". What can we deduce about the shape of the distribution of bacterial count, and why?* The median is the central value when observations are arranged in numerical order, so that 50% of observations are less than the median and 50% exceed it. "IQR" is the inter-quartile range, the difference between the 25th and 75th centile, or first and third quartiles. These are numbers such that 25% of observations are less than the first quartile and 75% exceed it, and 75% of observations are less than the third quartile and 25% exceed it. Here both quartiles are given. As the median is much closer to the first quartile than to the third, the distribution is positively skew.
- (d) *In the legend for figure 2, the authors report "difference between two groups significant ($P=0.012$)". What does this mean? What can we conclude about the effect of handrubbing and handwashing?* The probability that the difference in a sample could be as large as or larger than this, if in the whole population there were no difference (the null hypothesis), would be 0.012. This is a small probability, <0.05 , and we can conclude that there is evidence that handrubbing reduces contamination more than does handwashing.
- (e) *The authors say that for the reduction in bacterial contamination "The difference in the percentage reduction between the two groups was 26% (95% confidence interval 8% to 44%)". What is a confidence interval and what does this tell us about the effect of handrubbing and handwashing?* A confidence interval is used to estimate (alternatively to assess the accuracy of an estimate of) a quantity in a population from a sample. It is a set of values within which we estimate the population quantity to lie. It is chosen so that 95% of confidence intervals will contain their population value. Here, we estimate that

handrubbing reduces contamination by between 8 and 44 percentage points more than does handwashing.

- (f) *The authors conclude that "We have shown that handrubbing with an alcohol based solution is more effective than handwashing with an antiseptic soap in reducing bacterial contamination of healthcare workers' hands during routine patient care. This was due in part to the inadequate time spent washing hands conventionally." In what way do the results of the study support these conclusions?* Bacterial counts were reduced following both methods, but more so by handrubbing. This difference was statistically significant. It is estimated to be between 8 and 44 percentage points, so it could be quite small or fairly substantial. A larger study would be needed to estimate the difference better. As the trial is randomised and the assessment blind, the difference should be the result of the hand-cleaning technique. Whether the poorer outcome with hand-washing was due to the inadequate time spent washing hands is supported by the observation that most washes were shorter than recommended, but we cannot be sure that this is the cause.

SPECIMEN ANSWER EXTRA.

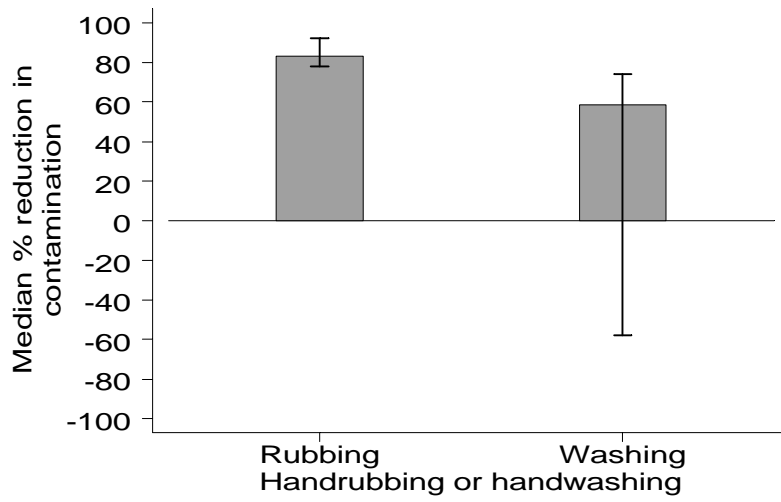
There was no question about Figure 2 in the paper, but some students were confused by it. This is not surprising. First, it did not photocopy well, one of the lines representing the interquartile range was obscured. Here is a different version.



The height of the bar represents the median difference, the line with cross-pieces represents the interquartile range. According to the text you were given, the numbers this represents are given by "Figure 2 shows that for each participant the median reduction of bacterial contamination achieved by handrubbing was significantly higher than the reduction achieved by handwashing (83% (interquartile range 78-92%) v 58% (58-74%), respectively, $P=0.012$)". In fact, there should be a minus sign here, in front of the second 58, which did not copy properly from the BMJ website: "Figure 2 shows that for each participant the median reduction of bacterial contamination achieved by handrubbing was significantly higher than the reduction achieved by handwashing (83% (interquartile range 78-92%) v 58% (58-74%), respectively, $P=0.012$)."

(This typographical error did not affect any of the questions

in the assignment.) The reduction should only be negative if there is an increase, so the graph should look like this:



It would make more sense to show the individual % reductions as separate points. There are only 23 of them.