Appendix Five

Second Exploratory Study Scoring Rubrics and Responses to Questions

A5.1 Map Question Rubric

PROBLEM VALIDITY

Present

Problem appears to be a perceived pattern in some data. As such the problem is ill-defined and it is not clear what question is being asked.

Partial

Indicate problem definition is not clear.

Absent

No reference to problem formulation.

MEASUREMENT VALIDITY

Present

Queries whether the measure is valid. State that a limb missing could be an arm or leg and that the measure is not fully explained because it is not stated whether partial limbs such as a missing forearm bone were included. Query whether the same biological mechanism is operating for different limbs, or query whether a missing limb could have several biological explanations.

Partial

Indicate that would like to investigate the classification procedure for a missing limb.

Absent

No mention or querying of the measure used.

DATA CAPTURE VALIDITY

Present

State data capture is invalid since only one year given. Would like to see data over a number of years presented.

Partial

Indicate there are not enough data.

Absent

Not mentioned.

ANALYSIS VALIDITY

An analysis validity check cannot be done on this problem.

MY OWN INTERNAL CHECK

Present

Use context knowledge to model the population distribution of New Zealand in the marked areas (ie one-third for top area, one-sixth for each of the other areas). Check whether the sample data distribution reflects this population distribution and determine that it does not. Checks plausibility of the information against own context knowledge and recalls a similar clustering incident.

Partial

Indicate more people living in an area yet no abnormalities recorded.

Absent

No mention of population distribution on the country. Describes the graph - there is 3 in this area etc.

VARIATION (statistical perspective)

Present

State that sample is small and therefore expect lots of variation even from a stable process. There would need to be something in the pattern of variation over and above the normal levels expected in a stable "nothing is happening process", before accepting anything other than random variation was operating.

Partial

Indicate in another year could see a different spread of numbers in the country or total numbers could be different.

Absent

Not mentioned.

VARIATION (context perspective)

Present

Give particular system and individual reasons, factors or suggestions for what could be causing the abnormalities in the middle of the country.

Partial

Indicate that there may be factors causing the abnormalities but do not give particular examples or give individual reasons only.

Absent

No possible causes given.

RETHINKING THE ANALYSIS

Present

State that geography of birth may not be the appropriate variable for looking at the data. Could look at the data from family tree variable or even geography of conception perspective etc.

Partial

Indicate there could be other ways to look at the data but do not give specific examples.

Absent

No reference to this aspect.

CAUSAL VALIDITY

Present

Implicit in the presentation of the data is that there is a geographical cause for this clustering of data. Another explanation could be random variation. Furthermore an alleged geographical cause cannot be altered and at the best be only a pointer to a deeper cause.

Partial

Give one of the above reasons.

Absent

No reference to implied causality.

COMMUNICATION OF DATA

Present

State communication of data could be misleading. Would like to know the exact location of each birth and the city.

Partial

Indicate presentation of data could be misleading.

Absent

No mention of how the data are presented.

EXTERNAL CHECKS

An external check cannot be done for this problem.

CONCLUSION VALIDITY

Correct

The problem suggests that there is a pattern in the data but as such it is ill-defined. The measure of missing limbs could depend on classification, and depending on the type of limb, may not be subject to the same biological mechanism, in which case the measure may not address the perceived problem. One would expect the data roughly to be in proportion to the population data if nothing special was happening. These data are not in proportion, but this is a small sample, and in small samples one obtains lots of variation in a stable process. The item may be wanting me to believe that something other than random variation is present (such as environmental or genetic factors), since it appears there is a clustering in the middle of the country. This communication of the data could be misleading as the exact location of the births could be widely spread. Also the pattern presented suggests that geography is a factor, whereas if the data were looked at from the perspective of another variable there may be another pattern. I am not willing to believe, on one year's data with such small counts, that there is a pattern in the data, and that anything other than random variation is operating. I would hope that the investigator would collate the data over a longer period of time and specifically pinpoint the towns where the births occurred.

Give at least two of the above explanations including random variation.

Partially correct

Indicate data have no meaning.

Incorrect

No conclusion made or attempt to base conclusion on context knowledge only.

A5.2 Error Rate Question Rubric

PROBLEM VALIDITY

Present

The problem is well-defined in that there is a concern about the number of errors being made by office staff and therefore the investigation is centred around finding out why. It has been defined in terms of who in the office staff is making the errors which may or may not be a useful question to address in this system.

Partial

Problem is well-defined by stating that too many errors are being made by a group of workers.

Absent

No reference to the whether the problem is well-defined.

MEASUREMENT VALIDITY

Present

State the length, type and difficulty level of the transaction may affect the number of errors. Thus the measure of the number of errors per day broken down by person may be valid for addressing differences among people but it is restrictive in discovering why those differences occur.

Partial

Indicates the type of transaction may be a factor in the error rate of the workers.

Absent

No mention of measure used.

DATA CAPTURE VALIDITY

Present

State that sampling over one month may or may not be valid for this type of situation as patterns of work over the year or during different months may be quite different. Mention that the same auditor should have been used for all four workers, that some randomisation process should be used for checking the transactions, and that the workers should be anonymous for an unbiased audit.

Partial

Mentions some facet of data capture such as one month's data may not be sufficient or may be sufficient but doesn't state why.

Absent

No mention of data capture

ANALYSIS VALIDITY

Present

The graph used is appropriate for the data.

This criterion is not marked.

MY OWN INTERNAL CHECK

Present

Reads and comprehends the graph fully. The information presented does not conflict with my own world knowledge about such situations.

Partial

Misunderstands or misreads some aspect of the graph or checks the meaning of an aspect of the graph.

Absent

Is not able to comprehend the graph.

VARIATION (from a statistical perspective)

Present

Mentions the spread of the boxplots and states that workers A, B and C are similar and that D is different. Notes that A is the most variable and D is the least variable. Also notes that there may be no difference among all four workers since the apparent difference of D may be due to sampling error (variation), as the sample size of 30 for each worker is small.

Partial

Indicates workers A, B and C are similar and that D is different.

Absent

Gives a rank order based on median only.

VARIATION (from a context perspective)

Present

Suggests system and individual explanations for variation amongst the workers such as working conditions, types of transactions etc.

Partial

Indicates other factors should be considered as to why the workers are different but does not give any suggestions, OR gives individual explanations only for the variation, OR suggests D is the best worker and the others are poor workers.

Absent

Describes the data in terms of the variables such as A has the highest number of errors per day and D has the lowest median.

RETHINKING THE ANALYSIS

Present

Suggest graph on other variables (not pertaining to explaining the difference among the workers) such as number of errors versus transaction type, or suggest graph workers in another way such as a times series plot for each worker.

Partial

Suggest looking at the worker variable may not be revealing the situation.

Absent

No suggestions for another way of looking at the data.

COMMUNICATION OF DATA

Present

States that box-and-whisker plots are a suitable method for comparing data among workers.

Partial

Indicates the graph used is suitable but doesn't say why.

Absent

No mention of how the data are communicated.

EXTERNAL CHECKS
An external check cannot be done for this problem.

CAUSAL VALIDITY
The reader is expected to do the analysis. Therefore the causal validity of a statement cannot be judged.

CONCLUSION VALIDITY
Correct

The problem is defined and useful for the system under study. The question seeks to find out who is making the errors which is a valid initial analysis. Sampling over one month is sufficient as long as the data collection procedure is reliable and randomised. I am not prepared to believe that there is systematic variation and D is making fewer errors than the other three workers, since there is not enough information to determine whether real or random variation is operating. Also I am not prepared to say that D is better than the others, as there could be many competing system and individual explanations, such as working conditions, job requirements for each worker, organisation of work, personal, the type, length and number of transactions. Also, consideration should be given to the validity of measuring on number of errors per day, when the complexity of the transaction may be a factor. I would hope that the investigator would: observe the workers in action to check out working conditions; talk to the workers; check the type of transactions done by each worker; look for other sources of variation; and reanalyse the current data - not on worker but perhaps on type of transaction - and then, if necessary, look at the third variable of worker.

Gives two or more of above reasons why no conclusion about the workers can be made.

Partially Correct

Gives one of the above reasons why no conclusion about the workers can be made.

Incorrect

On the basis of the data gives a conclusion about the workers.

A5.3 Die Question Rubric

MY OWN INTERNAL CHECK

Present

Give probability reasons such as equal chance for each number.

Partial

Indicate chance is operating.

Absent

No mention of probability.

VARIATION (statistical perspective)

Present

State that in small samples expect lots of variation. Ideas of long run relative frequency.

Partial

Ideas of long run relative frequency.

Absent

Not mentioned.

VARIATION (context perspective)

Present

Gives system and individual explanations for the numbers falling in this pattern such as the die is weighted or it is because of the way the die is tossed.

Partial

Gives individual explanations only.

Absent

No mention of causes.

CONCLUSION VALIDITY

Correct

State no reason yet to suspect the die is biased. Perhaps would toss more times to check.

Partially correct

Indicate the data have no meaning.

Incorrect

Die is biased or die is fair but the pattern is due to the way the die is tossed.

A5.4 Prison Newspaper Questions 1 & 2 Rubric

PROBLEM VALIDITY

Present

The level of prison suicides from year to year is a valid problem to be addressed. The question of finding out why there was a sudden rise in 1994 is useful for the prison system. The problem is of interest to many people as inmates should be 'safe' in penal institution and it could be considered a measure of how democratic or civilised our society is.

Give at least one reason why problem should be addressed.

Partial

Mention question should be asked but give no specific reason.

Absent

No reference to problem validity.

MEASUREMENT VALIDITY

Present

State that for the initial part of the investigation it is valid to look at the measures of number of suicides and number of inmates. Indicate that to obtain an understanding of the system need to measure or obtain a number of other possible explanatory individual (e.g. demographics of the inmates) and system (e.g. prison conditions) variables.

Partial

Give one of the above reasons.

Absent

Do not mention any measurement issues.

DATA CAPTURE VALIDITY

Present

State valid to collect all the data available over ten year period. This a census collection of data from all the prisons.
Would assume that the data-missingness rate and the data collection procedures and record keeping for the two factors of number of suicides and number of prison inmates are reasonably reliable though this may not be the case.

Partial

Mention one aspect only on data capture validity.

Absent

No mention of data capture issues.

ANALYSIS VALIDITY

Present

State analysis is not appropriate as not comparing equal intervals of years.

Partial

Query inappropriate comparisons of average numbers and single numbers.

Absent

No mention of validity of analysis.

MY OWN INTERNAL CHECK

Present

A quick two standard error check, employing the Poisson model at the upper bound, gives a value of 9.6 suicides. A mean rate of 5.1 suicides per year was used, which was calculated from data in article from 1985 to 1993. The quick rule of thumb for rates is: (mean rate) +/- 2x square-root of the (mean rate). In this case one is interested in the upper bound, where data are higher than expected. The Poisson distribution assumes independence and there is evidence that suicides may not be independent. If suicides are not independent then I would have expected a cluster to occur in 1994, which it has not, as the value 10 is just over the bounds of the two standard error limits.

Partial

State ten is not way off the average.

Absent

No attempt to check the plausibility of the information.

VARIATION (statistical perspective)

Present

Small sample (numbers) which implies lots of variation from year to year. Visualise a trend line with an envelope of limits. A quick check with the Poisson model further reinforces the notion that random variation is operating. Thus this would appear to be common cause variation rather than special cause variation.

Partial

Visualise a time series plot with a trend and numbers varying each year around the trend line. May not visualise the envelope of limits. No mention that small sample numbers produce a lot of variation.

Absent

No consideration given to variation.

VARIATION (context perspective)

Present

Give system and individual explanations such as prison conditions, age etc for the sudden increase in numbers.

Partial

Mention there could be some factors involved but do not specifically say what, OR give individual explanations only.

Absent

Give no reasons, OR describe the fact that the numbers went up in 1994, were only one in 1993 etc.

RETHINKING THE ANALYSIS

Present

The data have been analysed with averages of raw numbers. Perhaps the data should be analysed by the measures rate per 100,000 inmates, or rate per person-years at risk to get a better understanding of the situation.

(I could now do my own internal check on the claim in this article using the rate measure. I will check on the rates/100,000 inmates. In 1985 there were 8 suicides giving a rate of 288/100,000 (quoted in the article). In 1994 10 suicides gives a rate of 204/100,000 assuming a constant rate of increase for prison inmate numbers. This suggests the suicide rate has gone down. Over the nine year period 1985 - 1993 the average number of

suicides per 100,000 inmates is 136. A two standard error check using the Poisson distribution gives 159. With this measure the 1994 rate is outside the boundary and therefore appears to be special cause variation. However this is an incorrect measure for capturing the suicide rate because the prison inmate numbers are increasing over this time. The correct measure is average number of suicides per person-years at risk. Using the prison population numbers stated in the article, gives a result of 46/33750 suicides per person-years at risk assuming a constant increase in prison population. A binomial two standard error check gives a value of 11.86 revealing that the 1994 suicide numbers are well within the bounds of random variation.)

Partial

Indicate that the suicide rate seems to have gone down but do not suggest that there may be another way of looking at the data.

Absent

No mention that there could be another way of looking at the data.

CAUSAL VALIDITY

Present

Although the article claims that there appears to be no explanation for the sudden increase in suicides, the article is implying that imprisonment is a factor. The sudden increase may not be special to the prison population but may be reflecting the situation outside prison, or the situation applying to the general population. I believe that imprisonment is a manipulable cause, in that other alternatives to imprisonment in New Zealand penal institutions should be considered (which may be the reason for the article heading 'private prisons will help stop suicide'). Thus this factor can be altered but it could be also a pointer to a deeper cause. Also a plausible explanation for the increase in suicides may be random variation.

Partial

Gives one of the above reasons/explanations.

Absent

No indication that other factors should be considered.

COMMUNICATION OF DATA

Present

State communication of data is hard to follow, is misleading and inconsistent. Unable to obtain a clear picture of the trend in prison suicides. Graph would be better for communication.

Partial

Communication of data is inconsistent.

Absent

No mention of communication of data.

EXTERNAL CHECKS

Present

Checks or refers to the source/s of the article (e.g. written by a university lecturer with a doctorate in prison sociology and written for a well known mainstream newspaper), and approaches the article with a certain degree of belief but is willing to be critical.

Partial

Regards the author as a reputable researcher in this field and therefore dispensed towards accepting his judgement and statistically-based information.

Absent

No reference to the authority of the author or newspaper in which published.

CONCLUSION VALIDITY

Correct

This problem, on level of suicides, is a valid question to address in the prison system, since detained people under supervision should be safe from such acts of violence. I could believe that such data are reasonably reliable. It is valid to collect all the data over a ten year period and to initially use the measures of number of suicides and number of inmates. Although, to obtain a better understanding of why suicides are occurring, system and individual variables need to be part of the analysis. I would not agree with this reputable author that the trend threatens to continue. The high number of suicides can be explained by random variation using a rough two standard error Poisson distribution check. The measure of number of suicides per year is in question since more appropriate measures for comparison of years may be: suicide rate /100,000 prisoners, in which case a quick check reveals the suicide rate has gone down; or suicides per person-years at risk, in which case a Binomial distribution check reveals the suicide rates are within the bounds of random variation. The communication of the data is inconsistent, possibly misleading, with the averages for unequal time intervals quoted and not easily understood. A time series graph may portray the data more clearly and accurately. Although the article claims that there appears to be no explanation for the sudden increase in suicides, the article is implying that imprisonment is a factor. This sudden increase may not be special to the prison population but may be reflecting the situation outside prison, or the situation applying to the general population. I believe that imprisonment is a manipulable cause in that other alternatives to imprisonment in New Zealand penal institutions should be considered (which may be the reason for the article heading 'private prisons will help stop suicide'). Thus this factor can be altered but it could be also a pointer to a deeper cause. Despite the fact that I think the high number could be explained by random variation, I would hope that an investigation would proceed into finding possible common causes, as a high standard of prisoner safety is paramount and such variation or number is unacceptable.

Give three or more of the above reasons including at least one statistical reason why they do or do not support the claim of the author and/or how they would proceed in the enquiry.

Partially correct

Give one or two of the above reasons including at least one statistical reason why they do or do not support the claim of the author and/or how they would proceed in the enquiry.

Incorrect

Base their conclusion only on context knowledge.

A5.5 Prison Newspaper Question 4 Rubric

PROBLEM VALIDITY

Present

As an initial question it is valid to address the issue of Maori versus Non-Maori rate but other variables need to be looked at also. Particularly the information might not be useful for solving the problem, but may give an indication of what sort of data to look at.

Partial

Indicates the question is valid but does not mention that other variables need to be looked at in conjunction with this information.

Absent

No reference to problem validity

MEASUREMENT VALIDITY

Present

Queries the classification of inmates into Maori and Non-Maori categories. Who determined the classification? Admitting officer? Inmate? What definition of a Maori was used? Census definition? Prison definition? Such classifications may be arbitrary and could be biased.

Partial

Indicates that classification may be an issue but doesn't expand on why or how.

Absent

No suggestion that the measure captured may be flawed.

DATA CAPTURE VALIDITY

Present

Valid to take a census for these data over a period of ten years. Would be wary that the data are reliable since so many people would be involved in the data collection process - that is the recording of such information when the inmate is admitted to prison. There is bound to be subjectiveness, inaccuracies and data missing.

Partial

Mention the sampling mechanism only. Do not query the reliability of the data collection management and procedures.

Absent

No reference to data capture validity.

ANALYSIS VALIDITY

Present

Queries how the 1.5 was calculated and whether the different population sizes were taken into account. (Note: someone who knows about rates wouldn't normally ask this question)

Partial

Queries the accuracy of 1.5 and how calculated but does not mention the different population sizes.

Absent

No reference to analysis.

MY OWN INTERNAL CHECK

Present

Comprehends and understands how the ratio was calculated with different population sizes. The information presented sounds plausible and does not contradict any of my internal beliefs.

Partial

Comprehends the notion of a ratio but does not understand that the population size must be taken into account, OR comprehends the notion of a ratio but does not know how to compute a ratio with different population sizes.

Absent

Does not comprehend the model.

VARIATION (statistical perspective)

Present

Small number of suicides (50) over ten years so variation will be a factor in the estimate of 1.5. A small change in numbers could produce a change in the ratios (e.g. Could random variation be a factor since such small numbers are being dealt with? If 80% were Maori and 20% were Non-Maori and the suicide numbers over 10 years were 40 and 10 respectively then the ratio is 1. However if it had been 42 and 8 respectively then the ratio is 1.3 or the 43 and 7 ratio is 1.53.). Also there is no confidence interval given with this estimate which I would like to know in order to evaluate this statement.

Partial

Indicates the ratio 1.5 may not be accurate but does not say why, OR considers variation in a small sample may affect the accuracy but does not fully articulate. No mention of a confidence interval.

Absent

No consideration given to variation for the ratio estimate.

VARIATION (context perspective)

Present

Gives prison system and individual/societal reasons why Maori are more likely to commit suicide.

Partial

Gives individual/societal reasons only, OR mentions that other factors should be looked at as to why Maori are more likely to commit suicide but gives no specific factors.

Absent

No reasons given.

RETHINKING THE ANALYSIS

Present

Suggests stratifying on other variables to see if factors other than Maori reveal more. Factors such as age group and prison type (maximum, medium or minimum security).

Partial

Suggests other factors should be looked at but doesn't give any suggestions.

Absent

No suggestions of stratifying on other variables, that is removing the factor Maori entirely from the way the data are looked at.

CAUSAL VALIDITY

Present

Other plausible explanations should be offered such as age and prison type before I could evaluate this information. The proffered cause cannot be manipulated thus a deeper cause needs to be ascertained. There may be deeper, more meaningful reasons (the idea of finding a 'root' cause) why Maori and Non-Maori rates are different. In particular there may be cultural, social and environmental aspects. Race cannot be viewed as a cause because it cannot be changed or altered. Also I do not know whether this cause is special to the prison Maori population or whether it is reflecting the Maori population in general.

Partial

Gives one of the above reasons/explanations.

Absent

No indication that factors other than Maori should be considered.

COMMUNICATION OF DATA

Present

Valid way to present the data for populations of different sizes as a ratio. Common usage for comparison of data. Would like to see a confidence interval quoted for the ratio.

Partial

Valid for comparison of data - no mention of confidence intervals.

Absent

No reference to how the data are communicated.

EXTERNAL CHECKS

Present

Checks or refers to the source/s of the article (e.g. written by a university lecturer with a doctorate in prison sociology and written for a well known mainstream newspaper).

Partial

Regards the author as a reputable researcher in this field and therefore dispensed towards accepting his judgement and statistically-based information.

Absent

No reference to the authority of the author or newspaper in which published.

CONCLUSION VALIDITY

Correct

As an initial question it is valid for the author, who is an expert in this field, to address the issue of Maori versus Non-Maori rate, to use data gathered over ten years and to present the data as a ratio. The author is trying to suggest that prison is a factor in the greater number of suicides. However I would like to see the rate outside prison before considering this variable. There may be reasons why more Maori people commit suicide such as psychological reaction to prison or the attitude of wardens to Maori inmates. Other reasons should be looked at that do not pertain to the factor Maori, such as age and prison type. Variation needs to be accounted for in

the ratio in order to determine how big a difference there is between the two groups. Also I could query how the ratio was calculated and whether the different population sizes were taken into account. I would also have concerns about the reliability of the data in terms of the classification definitions for the two groups, since it would be collected by a wide range of people in the admittance procedures. I believe the data would be reliable in recording the event of a suicide. This statement at best can only be an indicator of where to look for a deeper cause that can be acted upon. At this stage I am not willing to believe that being Maori is a factor in prison suicides.

Gives two or more of the above reasons why not willing to agree with the author yet.

Partially Correct

Gives one of the above reasons why not willing to agree with the author yet.

Incorrect

Agrees with author and/or does not challenge one conclusion validity possibility.

A5.6 Fitness Newspaper Question Rubric

PROBLEM VALIDITY

Present

For this problem to be addressed there must be a belief, or some previous evidence, or assumption, that level of fitness is a risk factor for health and lowers the risk of death due to natural causes. The question as to whether exercise can lower the death rate is useful for humans, particularly if it means that the quality of life could be improved. I think another useful question would be whether exercise can lower the incidence of heart attacks rather than concentrate only on deaths. The study is being carried out under the auspices of the Cooper Institute for Aerobics Research and thus there may be a vested interest in promoting exercise.

Mention at least the first concern.

Partial

Indicates that it is important to address such a problem as the current lifestyle of people is perceived to be becoming increasingly sedentary or give some other reason.

Absent

No reference to whether the problem should be addressed.

MEASUREMENT VALIDITY

Present

Measurement issues arise as to what constitutes fitness in different age groups. Assume different criteria used. I would like to know what the procedure was if a person moved across the age group barrier in the intervening years between the first and second measure. It is probably valid to measure fitness level on a treadmill which could, for example, measure heart-rate and uptake of oxygen. I would like to know whether heart-rate is a 'good' measure of fitness. Are there other measures such as flexibility, muscle tone, walking distance, endurance? Is the treadmill test a rough measure of fitness? Are there better measures or several measures that should be taken? I would also like to know how they classified the type of death. I think another valid measure, considering their claims of preventing heart attacks, would be, for example, to measure how many had heart attacks and lived. Also there is an assumption that fit people at both measurement times continued to be fit in the following 5 years when the number of deaths was counted.

Mention at least two of the above measurement issues.

Partial

Mention one of the above measurement issues that they would like to know more about.

Absent

Do not consider any measurement issues that might arise.

DATA CAPTURE VALIDITY

Present

Would hope that a random sample of men from different occupations, ethnic groups etc., that was representative of the USA male population, was taken into account in the study. If they did the study in some large organisations, as is common practice, then hope they checked for representativeness, although the study could be biased in terms of work and living environment. The number in the study and the use of one sex are valid for the detection of any differences among the groups, since the

death rate is small. Would like to know if they had roughly equal numbers in each age group.
Would hope that the same technology, same treadmills and apparatus were used in all the tests and that the measurers had strict guidelines for obtaining reliable measurements.
Would suspect that the data-missingness rate was moderate but would like to know whether the initial measurements of fitness for this group fitted the same profile as those people who presented for both measurements.
The study was on men yet the claims in the article are for all people. Conclusions can only be applied to the group from which data were obtained which was USA men.

Mention at least two of the above concerns.

Partial

Mention one of the above concerns.

Absent

No mention of data capture or design of studies issues.

ANALYSIS VALIDITY

Present

The analysis is on deaths per 10,000 men. Working out proportions for each category and comparing them to the same base number is valid.

Partial

Mention the analysis used is valid but do not explain that the four categories will have different numbers of men and hence it is necessary to manipulate the data to the same base rate for comparison of categories.

Absent

No reference to the analysis method.

MY OWN INTERNAL CHECK

Present

From the statistics presented I can comprehend that there is an upward trend in the death rates which is explained by whether one is judged as fit, becoming-fit and unfit. Those becoming-fit had half the death rate of the unfit while those judged fit had one-third the death rate.

Partial

Recognise the upward trend only.

Absent

No description or summary of the given statistical summary.

VARIATION (statistical perspective)

Present

Rates given as per 10,000 men. Since only 10,000 men participated in the study, consideration should be given to the raw numbers in each of the four categories. A margin of error or confidence interval should be given for the rates. A very small number in one category could give rise to questions of data interpretation. The article specifically mentions fit/fit men at age 60 had a 50% lower death rate. The number of fit/fit men aged 60 may be small in which case this category may stand out simply on the grounds of random variation alone.

Partial

Recognise that sampling variation will mean that the figures quoted are only approximations (but possibly the figures do show a trend).

Absent

No reference to variation

VARIATION (context perspective)

Present

Suggest many individual and system explanations, or sources of variation that should be accounted for in the design, and in the measures such as nutrition levels, smoking status, medical conditions, stress levels, attitude, as well as the differing death rate and fitness levels of older men compared to younger men.

Partial

Suggest individual reasons only or mentions other factors should be considered but do not specify any.

Absent

No suggestions for what variables should be taken into consideration to explain the differences among the groups.

RETHINKING THE ANALYSIS

Present

Do these death rate findings apply to the 20-30 age group? Would like to have some breakdown into age group strata such as 20-40, 40-60, 60-80 to ascertain the variation and risk that would apply to each age group. This would enable me to assess my risk in a more reliable way.

Would like to see the data categorised on life threatening illnesses incurred during the study, such as a heart attack or stroke, where the men survived, rather than only on deaths, in order to properly assess the benefits of exercise.

Partial

Reference to one aspect such as age and death rate and that age must be a factor but do not explain further.

Absent

No reference to regrouping the data in another way.

CAUSAL VALIDITY

Present

I believe that factors such as nutrition levels, smoking status, medical conditions, stress levels, as well as death rate and fitness levels of older men compared to younger men, were or should be taken into account. There is an assumption that heart attacks are related to fitness whereas there may be other plausible explanations such as high blood pressure.

The cause of fitness is useful as it is something that can be changed in people to improve their chances of survival.

I believe or hope that some judgement criteria such as biological mechanism explanations and controlled laboratory experiments and possibly replications of the study may have been performed.

Give at least two of the above explanations.

Partial

Give one of the above explanations.

Absent

No reasons given to support causal validity.

COMMUNICATION OF DATA

Present

The category fit to unfit was not reported which makes one wonder if the estimated value went against the trend. This category should have been reported even if few or no people were in it. The data are easily understood but a diagram or graph would improve communication and make the findings more accessible to the reader.

Partial

Mention that a graph would help communication. Do not realise one category is missing.

Absent

No comment on communication of findings.

EXTERNAL CHECKS

Present

This is a large study published in a reputable journal and therefore scrutinised by experts. Based on current knowledge I would agree that fitness improves the chances of not having a fatal heart attack in American men.

Partial

Mention that based on their own knowledge of fitness and health they could believe these findings. (i.e. these findings fit in with commonly held community beliefs and assumptions based on wide media coverage and the setting up of the Hillary Commission etc.)

Absent

No references to their own knowledge base or reputable journal.

CONCLUSION VALIDITY

Correct

Based on current knowledge I would be in general agreement with the claim in the article that exercise would improve the risk of preventing fatality in American men, but there are aspects of the study that I would like to know about before being fully convinced. This is a large study in a reputable journal and I believe the design of the study would be valid and reliable and that factors such as nutrition levels, smoking status, medical conditions, stress levels, as well as death rate and fitness levels of older men compared to younger men, were taken into account. I would like to know whether the treadmill test truly or only partially captures the level of fitness. Also the findings are reported for the entire age cohort and for deaths only. Therefore I would like to know the relative risks among the different age groups, and to see data on survival of life threatening illnesses incurred during the study, in order to assess the claimed benefits of exercise. The study found trends for American men, yet the implication in the article is that the research applies to everyone. The question is raised whether the same trend applies to women and people in other countries and cultures. The article also claims that exercise improves your risk of preventing a heart attack, but this is not backed up by the data in the article which is only on heart attack deaths. No data are presented on the number of men who had a heart attack and survived. Sampling variation should be recognised in that the quoted figures should have had a confidence interval, and the findings should have reported on all four categories not just three categories. The recommendation that it is never too late to get fit is logical and consistent with the findings presented.

Give three or more of the above reasons why they agree or disagree with the conclusion in the article.

Partially Correct

Mention one or two aspects and reasons about why they agree or disagree with the conclusion in the article.

Incorrect

No mention of the above factors. Base their conclusion on their own knowledge.

A5.7 Analysis of Student Responses

Present or Correct is scored 2 marks, Partial or Partially Correct 1 mark and Absent or Incorrect 0 marks.

Table A5.1 Student Responses to Map Question

MAP	Prob Valid	Measu rement	Data Capt	Basic Analy	Own Check	Variat (Stat)	Variat (Cont)	Rethnk Analy	Causal Valid	Comm Data	Ext Check	Corr Concl
ISA	0	0	0		1	0	2	0	0	0		0
TEP	0	0	0		2	0	2	0	0	0		0
EAGLE	0	0	2		1	1	2	0	0	0		1
MORTA	0	0	1		2	0	2	0	0	0		1
NORS	0	0	0		2	0	2	0	0	0		0
JOY	0	0	0		0	1	2	2	0	1		0
MEAN	0	0	0.5		1.3	0.3	2	0.3	0	0.2		0.3

Table A5.2 Student Responses to Error Rate Question

ERROR	Prob Valid	Measu rement	Data Capt	Basic Analy	Own Check	Variat (Stat)	Variat (Cont)	Rethnk Analy	Causal Valid	Comm Data	Ext Check	Corr Concl
ISA	0	0	0		2	1	1	0		0		0
TEP	0	0	0		1	1	1	0		0		0
EAGLE	0	0	0		2	1	1	0		0		0
MORTA	0	0	0		2	1	0	0		0		0
NORS	0	2	0		2	1	2	2		0		2
JOY	0	0	0		2	1	1	0		0		0
MEAN	0	0.3	0		1.8	1	1	0.3		0		0.3

Table A5.3 Student Responses to Die Question

DIE	Prob Valid	Measu rement	Data Capt	Basic Analy	Own Check	Variat (Stat)	Variat (Cont)	Rethnk Analy	Causal Valid	Comm Data	Ext Check	Corr Concl
ISA					2	0	0					1
TEP					1	0	0					2
EAGLE					0	0	2					0
MORTA					2	0	1					0
NORS					1	0	0					0
JOY					0	1	2					2
MEAN					1	0.2	0.8					0.8

Table A5.4 Student Responses to Prison Newspaper Questions 1 & 2

PRS1|2	Prob Valid	Measu rement	Data Capt	Basic Analy	Own Check	Variat (Stat)	Variat (Cont)	Rethnk Analy	Causal Valid	Comm Data	Ext Check	Corr Concl
ISA	0	0	0	0	0	0	2	0	0	0	0	0
TEP	0	0	0	0	1	1	2	1	1	0	0	1
EAGLE	0	0	0	0	1	2	1	0	1	0	0	1
MORTA	0	1	0	0	0	1	2	1	0	0	0	0
NORS	0	0	0	2	1	2	2	0	1	2	1	1
JOY	0	0	0	2	1	1	2	0	1	1	2	0
MEAN	0	0.2	0	0.7	0.7	1.2	1.8	0.3	0.7	0.5	0.5	0.5

Table A5.5 Student Responses to Prison Newspaper Question 4

PRS4	Prob Valid	Measu rement	Data Capt	Basic Analy	Own Check	Variat (Stat)	Variat (Cont)	Rethnk Analy	Causal Valid	Comm Data	Ext Check	Corr Concl
ISA	0	0	0	2	1	0	1	0	0	0	0	0
TEP	0	0	0	2	1	0	2	0	0	0	0	0
EAGLE	0	0	0	0		1	1	0	0	0	1	0
MORTA	0	0	0	0		0	2	0	1	0	0	1
NORS	0	0	1	2	2	1		2	2	0	0	2
JOY	0	0	0	2	1	0	1	0	0	0	0	1
MEAN	0	0	0.2	1.3	1.3	0.3	1.4	0.3	0.5	0	0.2	0.7

Table A5.6 Student Responses to Fitness Newspaper Question

FIT	Prob Valid	Measu rement	Data Capt	Basic Analy	Own Check	Variat (Stat)	Variat (Cont)	Rethnk Analy	Causal Valid	Comm Data	Ext Check	Corr Concl
ISA	0	0	0	0	1	2	2	1	1	0	1	1
TEP												
EAGLE	2	0	2	0	0	1	2	0	2	0	0	1
MORTA	2	0	0	0	1	0	1	0	2	0	0	1
NORS												
JOY	0	2	1	0	2	0	2	0	1	0	0	1
MEAN	1	0.5	0.8	0	1	0.8	1.8	0.3	1.5	0	0.3	1