Faculty of Public Health

of the Royal Colleges of Physicians of the United Kingdom
Education Department

Working to improve the public's health

DIPLOMA & PART I EXAMINATION FOR MEMBERSHIP OF THE FACULTY OF PUBLIC HEALTH

Of the Royal Colleges of Physicians of the United Kingdom

JUNE 2004

EXAMINATION QUESTIONS WITH EXAMINERS' KEY POINTS AND COMMENTS

N.B. Please note that these are key points, not model answers

PAPER IA

Question 1

Cluster randomised trials are increasingly popular in the evaluation of certain types of health care intervention.

a) Describe circumstances where the use of cluster randomised trials is appropriate.

(4 marks)

b) Describe the principles involved in choosing an appropriate sample size.

(3 marks)

c) Describe approaches to analysing data that arise from a cluster randomised trial.

(3 marks)

Key Points

a) Some interventions are implemented at the level of an area or organisation. In this situation the effect of the intervention may be greater than when applied just to individuals. For example vaccination may result in herd immunity protecting those unvaccinated. Social networks may reinforce health promotion messages. If an individually randomised trial is used this effect on the population as a whole is not measured and individuals who are randomised not to receive the intervention may receive it (e.g. health promotion through the mass media). Interventions such as this are best evaluated by cluster randomised trial. Cluster randomised trials are also used when randomisation at individual level is either impractical or impossible (e.g. an intervention dependent on a structural facility such as a drop in clinic)

b) The ordinary sample size calculation, which involves estimating the difference between the intervention and the control group, may be carried out but it needs to be inflated appropriately, allowing for the interdependence of outcome data between individuals within a cluster. This involves quantifying the homogeneity of response between individuals within a cluster. The usual statistic required for this is the “intra class correlation coefficient” (ICC).

c) Data arising from the individuals should be aggregated to the level of the cluster, so that the notion of statistical independence is retained. The unit of analysis is the cluster, not the individual subjects within it.

The following are additional points which might improve the answer to “good” or “excellent”:

- Good example of a cluster randomised trial, with justification for the cluster design
- The Design Effect = 1 + (m-1) x ρ, where ρ is the ICC and m is the average number of individuals within a cluster. This should be multiplied by the number required for an individual based trial, where the same benefit of the new intervention is assumed.
- Aggregated data (even if derived from binary outcomes measured on individuals) may be analysed using a two-sample t test.
- If outcome data are available at the individual level, and covariates are available at both the individual and the cluster level, the benefit of the intervention may be analysed using a multilevel model with appropriate adjustments for the covariates.
- In addition, analyses are often performed on individual level data but incorporating adjustment for ICC (using methods largely developed by Allan Donner).
- If the clusters vary by size considerably (for example if the clusters are families), it is appropriate to consider weighting for cluster size.

Comments

The question was answered quite well, though few candidates achieved high marks. Most could describe circumstances where use of cluster randomised trials was appropriate, though few mentioned that the effects of interventions applied at cluster level might be greater than the sum of effects on individuals (for example social networks reinforcing health promotion messages or herd immunity in immunisation programmes). Some candidates only described general principles of choosing the sample size, without specifying the additional considerations needed in cluster RCTs. A number of candidates did not make explicit that analysis was by clusters. A few candidates talked about other study designs: studies of clusters of cases in relation to putative exposures, multi-centre trials, and multi-stage sampling procedures.

Question 2

Define how you would calculate for a cohort study:

a) the risk of disease in exposed and unexposed; (3 marks)

b) the rate of disease in exposed and unexposed; (3 marks)

and

c) Describe the information you would require to calculate the population attributable risk fraction including the assumptions you make in doing so. (4 marks)

Key Points

a) Risk is the probability of an event. Therefore, risk of disease in exposed is given by the number of exposed individuals developing the outcome during follow up divided by the total number of individuals exposed at the beginning of follow up.

Risk of disease in the unexposed is given by the number of unexposed individuals developing the outcome during follow up divided by the total number of unexposed individuals at the beginning of follow up.

b) The rate of disease in exposed and unexposed takes account of the amount of time each individual has been at risk of the outcome – this may vary from individual to individual depending on when they are lost to follow up, die or develop the outcome of interest. The rate is defined as the number of events divided by the person-years at risk. This is calculated separately for the exposed and unexposed. Bonus point if say that a multiplier is often used to make the rate a whole number.

c) The population attributable risk fraction is the proportion of all cases in the whole study population (exposed and unexposed) that may be attributed to the exposure. It is also called the etiologic fraction, the percentage population attributable risk and the attributable fraction.

It is estimated by dividing the population attributable risk by the risk in the total population. The population attributable risk is the risk in the total population (exposed and unexposed) minus the risk in the unexposed group.

If r is the risk in exposed and unexposed combined

And r0 is the risk in the unexposed

Then PAR = r-r0

PAF = PAR/r

Or (r-r0)/r

The assumptions in this calculation are:

1. The relationship between exposure and outcome are causal

2. That the measurements made in the cohort accurately represent the true situation (i.e. reflect true exposure)

3. That removing the exposure actually removes the risk

4. That the cohort studied is representative of the population to whom you wish to apply the findings.

5. Causes other than the one under investigation have had equal effects on the exposed and unexposed.

Comments

This question on risks, rates and PAF was not answered well in general. Many candidates did not define rates properly. In addition a lot of information was offered that was not asked for – risk ratios, excess risk etc. In terms of PAF some candidates gave formulae for case control studies rather than cohort. Many confused the risk difference i.e. between exposed and unexposed with the population attributable risk.

Question 3

Several weeks ago a traffic accident in a small town located in a rural region involving a tanker lorry containing chemical waste led to the spillage of a large quantity of noxious chemicals. A plume of gas hung over the town for several hours.

Since the accident, the local haematologist has reported an apparent increase in bleeding disorders in local residents. Outline a study to investigate the relationship between the event and reported health problems.

Key Points

Candidates will be expected to outline an approach that is practical and feasible and better candidates will be likely to discuss the possible barriers to undertaking the study proposed, such as obtaining funding for a long/complex study

Association and causation

Association - a statistical dependence between two or more variables. May be artefactual or spurious; true but not causal; causal

Causation - concepts of necessary and sufficient; Bradford Hill criteria: consistency; dose response; biological plausibility; temporal relationship; specificity; coherence; strength of relationship; removal of causal factor.

Designing a study: general points with specific reference to problem described

- Define problem and aim of the study
- Define type of study appropriate to problem - hypothesis testing
- Describe existing data and how you would assemble evidence to support or refute the hypothesis
- Discuss ethical considerations that might be encountered
- Define inclusion and exclusion criteria for cases and controls: who would make the diagnosis? Are the cases real or artefactual?
- Define the exposure: to the gas, chemicals or local water sources?
- Determine whether the chemicals are known to be a cause of bleeding disorders
- Define outcome measures of interest and how you would measure them
- Define your study strategy to ensure external validity - sampling frame, sample size, sample method, randomisation
- Consider a pilot study
- Estimate how much the study might cost
- Discuss analysis, potential sources of bias and interpretation of results

New suggestions:

- Health survey of population at risk with similar survey of control unexposed populations (difficulties of selecting these)
- Determination of a true excess of health events should precede looking for an environmental cause - use of routine data over time to examine temporal/spatial clustering of disease events
- Specific studies of bleeding disorders – likely to be rare so needs a case control study with examination of temporal, spatial distribution in relation to plume – behavioural variables in relation to exposure to plume.

Comments

Many candidates chose cohort study whereas case control study was most appropriate. A number of candidates discussed different methods whereas question wanted a study outline.

Question 4

Describe the essential components of a local programme for reducing unintended pregnancies, particularly among young people.

Key Points

- Assess the problem for the defined population using appropriate data on conception rates by age and outcome.

- Assess the availability and adequacy of health education and services, particularly family planning, including emergency contraception, counselling and pregnancy termination.

- Set out a framework for evaluation which should be agreed by the implementers, to include responsibilities for carrying it out and further action to be taken.

- Identification of individuals to lead, manage and monitor the programme. Reference to current evidence of what works including HDA review, poor quality of many studies.

The following are additional points which might improve the answer to “good” or “excellent”:

Relevant examples from the literature and/or experience, reference to national/government guidance and to organisation theory.

Reference to Dicenso et al (2002) review of RCT evidence showing little support for many currently used interventions and discussion of the significance of this review.

Reference to social epidemiology including international and cultural comparative data.

Comments

Needs assessment was not adequately covered, many candidates listed a range of interventions without discussing assessing the appropriateness in the local setting.

Question 5

a) Describe briefly the following measures:

i) Disability adjusted life years (DALYs) (2 marks)

ii) Potential Years of Life Lost (PYLL) (2 marks)

b) Using low back pain and lung cancer as examples, how might an assessment of the burden of disease in a population appear to differ using these health measures? (6 marks)

Key Points

- DALYs. The DALY (disability-adjusted life year) measure combines the (estimated) number of life years lost due to premature death and the number of years lived with disability using a set of disease specific empirical weights to value the level of disability following standardised methods. These were derived from the Global Burden of Disease (GBD) Study which provided quantitative, internally consistent estimates of the burden of disease, including non-fatal outcomes, attributable to 107 causes, per sex, for different age groups and per region in the world for 1990.

- PYLL. This measure attempts to quantify the potential years of life lost by looking at average age of death from conditions compared to average life expectancy. PYLL can be expressed absolutely or as a rate relative to the population at risk.

b) Low back pain has a low mortality but high prevalence with variable duration of disability. PYLL will be low; DALYs will be moderate (made up primarily from the number of years living with disability).

Lung cancer has a high mortality and a short duration. PYLL will be high; DALYs will be high (mostly from premature mortality rather than years lived with disability).

The following are additional points which might improve the answer to “good” or “excellent”:

- A well-structured answer illustrating the importance of understanding how different measures of health status are derived.
- Using mortality measures such as PYLL, to measure burden of disease takes no account of diseases and disorders that have a low mortality. Many of the most common disorders seen in primary care including low back pain, mental illness and minor infections would be invisible.
- If priority and budget setting is based on mortality measures these important conditions will be consigned to poorly resourced services.
- Using combination measures such as DALYs can, in theory, weight disability. In practice there are many methodological problems with estimating burden of disease in this way.

Comments

Most candidates were able to give a reasonably satisfactory answer, although many candidates did not know or state that DALYs are a sum of both disease specific mortality and disability adjusted years, rather than just disability adjusted years. This led some candidates in part b saying that DALYs for lung cancer would be low because people die quickly, but as DALYs include mortality as well as disability they reflect the disease burden reasonably well. A few candidates did not read the question and wrote about the use of DALYs and PYLL in economic allocation or analysis.

Question 6

What would be the public health implications for a region with:

a) Increasing fertility;

b) Decreasing fertility?

Key Points

Increasing fertility

Measured by - increasing crude birth rate, increasing general fertility rate, and increasing total period fertility rate. Leads over time to increased number and proportion of children, young people, middle-aged and then elderly people.

The effects on public health will depend on where in the world the region is, and at which point in the process the region is in. Effects can be on individual, societies, countries and regions of the world.

Individual – Increased demand for food, clean water, sanitation, and suitable housing. Potential for overcrowding in poor circumstances - risk of poverty, epidemics of illness, declining living standards etc.

Society - Increased demand for maternity and child health services; increased demand for food, clean water, sanitation, suitable housing; increasing demand for basic health services; later demand for jobs.

Decreasing fertility

Measured by - decreasing crude birth rate, decreasing general fertility rate, and decreasing total period fertility rate. Leads over time to decreasing number of children, young people, middle-aged and then elderly people.

Individual – Puts pressure on working people to support the elderly; more leisure time for exercise.

Society – Elderly people have higher dependency levels and make a greater demand on health and social services. As the proportion of working age population declines relative to the elderly there will be fewer informal carers for the elderly infirm. There will be fewer health and social services workers (unless incentives are given to work in the caring industry); less income tax revenues to put into services and pensions; more demand for specialist housing; less income to spend on fuel and food. Isolation of the elderly.

The following are additional points which might improve the answer to “good” or “excellent”:

There may also be implications for social policy if the trend needs to be countered; for example there may be an increase or decrease in family-centred or family-encouraging policies; might increase or decrease access to contraception, terminations etc. Also might have knock on effect on wider policies, e.g. immigration.

Comments

Some of the poorer candidates answered this question from a very narrow view of public health, often only considering the impact on illness and health services. Very few candidates considered how any environmental impact, such as increasing need for clean water, adequate food, housing or transport would affect health. Good answers tended to be structured in some way. A few candidates did not read the question properly and wrote about the causes of increasing or decreasing fertility rather than the implications/consequences.

PAPER IB

Question 7

What social and psychological factors determine the success of childhood immunisation programmes and what steps can be taken to improve uptake? Illustrate your answer with an example in a named country.

Key Points

1) Lay perspectives regarding immunisations and in particular views that immunisations cause harm.

2) Commitment to alternative medicine.

3) Ethnicity - lack of awareness of immunisations on offer, especially due to language barriers.

4) Lack of awareness of the potential for epidemics.

5) Indifference to disease prevention where the diseases are seen as a thing of the past.

6) Social class - educational differences and influence on positive health behaviours among parents, difficulties faced by single parents in taking children to the GP to be immunised.

7) Effects of media scares.

Steps to improve uptake: PCT immunisation coordinators to gain a better understanding of the local factors limiting uptake and targeting efforts to improve uptake accordingly, accessibility to publicity and educational materials which are easy to understand and in appropriate languages, seeking the support of community leaders and alternative medicine practitioners, integrated efforts by community staff (health visitors, practice nurses, etc) to support families with specific needs, working with the media, education at all ages including school children to influence healthy choices among the next generation of parents.

The following are additional points which might improve the answer to “good” or “excellent”:

Awareness of previous observations of a preponderance of rubella babies born to Asian mothers as evidence of the effects of language barriers in ethnic minority communities which are generally compliant with health advice if offered in languages familiar to them, religious barriers to immunisations in some communities and the need to seek the support of religious leaders, community mothers schemes aimed at improving mothers' self esteem and parenting skills which have observed improvements in immunisation rates.

New suggestions:

Accessibility of services needs to be considered. In addition, need to consider the cost and other financial incentives or disincentives to immunisation.

The Peckham study of vaccine coverage also identified two strong social factors – birth order (the later born children in families are less likely to be vaccinated) and having a sibling with a chronic illness (less likely to be vaccinated).

Comments - no examiner comments

Question 8

Write short notes on two of the following:

a) marginal cost analysis

b) opportunity cost

c) the distinction between need and demand.

Key Points

a) (Marginal cost analysis) – Makes distinction from average costs. Looks at the cost of the next unit of activity e.g. in an expanding orthopaedic demand situation, the cost of the next few hip replacements may be at marginal cost (mostly consumables) because the fixed costs are already covered. But there comes a point when the next hip replacement triggers a 'stepped' cost (steep increase) because a new ward might have to be opened or a new member of staff taken on. Another good example is in deciding optimum frequency of repeat tests in a screening programme – what is the added cost per additional screen and what is the added yield in terms of cases detected? The marginal cost per case usually climbs dramatically with each successive test (whereas the average cost may not – marginal cost is the key indicator but often overlooked by uninitiated policy makers!)

b) (Opportunity cost) – An economic term indicating that all resources have an alternative use, being the cost of what is given up by taking one particular action; opportunity cost is more than money costs, and an assessment of opportunity costs includes an examination of choice, output and scarcity value and should lead to a formal economic evaluation. It is particularly important if resources are limited. Values differ between groups – politics, ethics – and stated preferences may be different from those observed. A macro or micro economic example should be given.

c) (Need/Demand) – Demand is what people request in respect of health AND health care and is affected by many personal and societal factors in addition to ill health, e.g. education, income, pressure groups, locality, media. Need (an intervention exists which will produce a health benefit) is determined from an assessment by third parties, whose views may differ, and tends to be focussed on need for services rather than directly on need for health; normative need/felt need/expressed need and comparative need are useful distinctions OR the 'iceberg of disease' and either the opposite relation between social class and need, cf demand, or some long term trends, should be mentioned.

Comments - no examiner comments

Question 9

The improvement of health requires better collaboration between organisations that influence health. In a named country of your choice, briefly describe the barriers to effective partnerships between organisations and ways to overcome these.

Key Points

Most or all of the following would be required for a pass:

- Description of barriers:

- Cultural
- Location eg lack of coterminosity in terms of multi-agency boundaries
- Structural/organisational
- People – stakeholder capacity & personalities, culture or agency
- Resource eg staff resource, opportunity cost

- Lack of focus and/or shared objectives ie varying agendas

• Description of mechanisms to overcome barriers:

- Define the aims of a partnership on specific issue(s) eg community action partnership
- Similar service sector
- Agreement to collaborate for mutually agreed goals
- Participation by people and organisations in collaborative activities on multiple fronts
- Public involvement (could impact either way depending on the context!)
- Contribution to common assets in multiple areas
- Relevance to corporate agenda and/or national programme areas
- Leadership
- Incentives
- `Champion`
- Chief Executive Commitment
- Legitimising eg joint appointment of DPH by PCT & Local Authority
- Plentiful resource
- Creation of umbrella organisations
- Explicit process with momentum
- Success in terms of `early wins`

Outstanding answer would also be able to:

- Identify key components of a partnership as a relationship involving:
  - Power
  - Control
  - Dialogue
  - Negotiation
  - Development of shared objectives
  - Motivation
  - Context

- articulate at macro and micro levels all these points within the context of management theory and include practical examples

Comments

Not very well answered as reflected in the marking of both examiners.

Question 10

Write short notes on 3 of the following methods of funding health services. Compare their merits and disadvantages.

General taxation

Social insurance

Private insurance

User charges

Key Points

General taxation

This refers to both direct and indirect tax receipts collected by government. From a macro economic perspective, general taxation is regarded as being highly efficient, delivering strong cost containment and forcing prioritisation through what are typically overall cash-limited health care budgets set by the government. Under tax financing, the government has both a strong incentive and the capacity to control costs. From a micro economic perspective, it is also an efficient way of funding health care, typically involving low administrative costs. Because it draws revenue from a wide base, it helps to minimize distortions in particular sectors of the economy.

A reliance on general tax financing can leave a health system vulnerable in times of economic and fiscal difficulties. However, this is not a feature which is specific to tax financing, and tax financing can help individuals in difficult times when they are less able to afford out-of-pocket payments or private insurance.

General taxation ensures universal access to services irrespective of ability to pay. However, the degree of individual choice tends to be relatively limited.

Social insurance

In this system, employer and/or employee earnings-related contributions are usually paid to and managed by social insurance or 'sickness' funds. Social insurance contributions are raised from a narrower base than general taxation, with the costs falling mainly on employers and employees rather than the wider group of taxpayers. This may lead to economic distortions and disincentives as the revenue base is more concentrated, on employment.

Social insurance systems can also be vulnerable to periods of economic downturn which can result in reduced revenues into the sickness funds. As with tax-financed systems, access to health services is typically universal or near universal and not based on ability to pay. There is generally little scope for expression of individual choice.

Traditional social insurance systems have little incentive to seek to contain the payments they make to health care providers because of their ability to raise contribution rates. As a result, cost control has been weak and resulted in inefficient use of resource. The administration and transaction costs can also be high.

Private insurance

Refers to private medical insurance taken out by individuals or, by employers on their behalf. The nature and coverage of private insurance differ significantly across countries. In the US, private insurance is relied on by a majority of the population as their sole means of cover. In other countries, private insurance is largely taken out by higher income groups, either in place of social insurance (e.g., Germany) or in addition to cover provided by the government (e.g., the UK). Still, in other countries (e.g., France) private insurance is taken out widely by the population for the specific purpose of covering their liabilities for user charges within the health system.

Systems relying on private medical insurance tend to exhibit poor cost control, with spending largely demand led, an absence of global budgets and the fragmented commissioning of health care services. Management and administration costs are also high, including the costs which are required in assessing risk, setting premiums and assessing claims.

The level of access to health services is determined by the level of insurance cover, and contributions are based not on ability to pay but on an individual's health risk rating as assessed by the insurer. It will usually be the poorer, older and less healthy in society who are considered to have the greatest health risk and therefore face the highest insurance premiums (issue of adverse selection). Access depends on ability to pay and the lowest income earners tend to pay more. This may lead to a significant proportion of the population with no access to health services other than the last resort of the emergency room, either because they cannot afford the insurance premiums or because private insurers refuse to insure them because they are deemed too risky. However, expression of choice is greater with private insurance; the insured can select both the insurer and the type of care package and range of benefits.

User charges

Refers to fees and charges paid by patients for the use of particular health services, in either the public or private sector. Patients may be required to pay for all or part of the cost of a particular publicly-provided service.

Imposition of user charges can help to encourage a more responsible use of resources by limiting wasteful and unnecessary activity and contain the total amount of health expenditure which the government has to finance publicly. However, user charges may discourage people from seeking treatment including preventive care, or direct them to other areas of a health system where charges are not levied, such as emergency care.

User charges relate access much more directly to ability to pay than either general taxation or social insurance. Where user charges are levied on what are clearly regarded as clinically necessary services, adequate exemptions must be put in place to make the system equitable so that those with a clinical need are not discouraged or prevented from receiving treatment.

Comments - no examiner comments

PAPER IIA

You are acting as advisor to a national board charged with monitoring and improving the quality of care in your health service. Two recent high profile events in surgical practice have resulted in a government minister demanding that the quality of care be the subject of routine monitoring through mortality league tables (ranked mortality statistics). You have been asked to advise the board and are using the attached BMJ paper as a focus:-

Tekkis PP, McCulloch P, Steger AC, Benjamin IS, Poloniecki JD. "Mortality control charts for comparing performance of surgical units: validation study using hospital mortality data." BMJ 2003;326:786-788.

1. Complete a structured abstract of this paper in no more than 250 words.

   [Objectives provided.] (15 %)

2. Provide a critical appraisal of the paper against its objectives. (35 %)

3. The paper makes the following statement:

   'Underperforming units will lie above the upper control limits, while units with unusually good results will lie below the lower control limits. Units lying within the 95% control limits have an operative mortality that is statistically consistent with the group mean.'

   Please comment. (10 %)

4. Write a structured report to the national board, comparing the merits of mortality league tables and mortality control charts (as both point plots and running means) which will allow them to respond to the minister. Include in your report suggestions to address any problems you have noted. (40%)

Key Points

This is a difficult question because many people will be unfamiliar with the control chart methods. We should give appropriate credit for straightforward good critical appraisal. We should also give appropriate credit for people who grasp the essential elements of the control chart method:

- The group average is assumed to be best measure of expected performance

- The chart in this example is a funnel plot (people may be familiar with this from meta-analysis)
- Control limits are similar to confidence intervals around the expected performance NOT the observed rates as is usually the case
- The funnel plot shows these limits for “sample size” – hence the funnel – another interpretation of it is the precision of the mortality given the sample size – low volumes give wider confidence intervals
- It is a screening tool designed to separate variation into that which might be expected and that which is outside expected variation and which might have “special cause”
- These tools don’t tell you why, only whether, a unit is outlying
- They CAN show performance if what is being displayed is an agreed measure of performance – if surgeons agree that unit mortality is a performance measure then outlying can be regarded as good or bad
- They are an essential part of continuous quality improvement initiatives

Note:

I suspect some candidates will “rubbish” the method and say that intervals should be put around the mortality of each unit NOT around the group average in this way – how should we deal with these?

1. Abstract

- No more than 250 words
- Headings required
  - (Objective provided)
  - Design/Method
  - Data Sources
  - Statistical Analysis
  - Results
  - Conclusions

2. Critical Appraisal

a. Objectives/ design

- Study is a database/ ecological study of in-hospital mortality following major upper GI surgery for upper GI cancer based on two prospective databases
- Aim of study was to investigate use of control charts in presenting and analyzing mortality as a means of assessing unit performance
- Unit of analysis was surgical unit
- Improves on existing established technique for ranking performance of organizations e.g. hospital mortality league table
- Mortality control charts – good for comparing and monitoring outcomes of surgery
- Candidates should include some description of what contrl chart is e.g
- Control chart compare each observation with and expected observation based on the group mean – the control limits are set around the expected not the observed as is usually the case
- If all points within control limits then system is said to be under control and variability in outcomes is said to be due to "common cause variation" – outliers exhibit "special cause variation" – they are due to external or extra-process factors.

b. Analysis/ statistical methods

- Adjusted for case mix, volume using univariate analysis and logistic regression models
- Databases validated against external sources and literature in estimates of in-hospital mortality
- Mortality plotted on "funnel plot" against unit volume, control limits for 90%, 95% and 99% calculated using exact binomial distribution (correct since data is binomial - dead or alive) both as aggregated figures for all units, and run chart for two units
- Run chart plots the outcome for each consecutive patient in each unit – numerator is number of deaths upto patient of interest, denominator total patients e.g. if 4 of first 5 patients died mortality would be 4/5=80%. If none of next 5 patients died, mortality now 4/10=40%

c. Results

- Based on 1042 valid cases in 29 units
- Over half were oesophagectomies
- Key risk factors from table were POSSUM score, metastatic disease and age. Emergency surgery said to be important but model data shows odds ratio CI include 1=> non statistically significant.
- Model said to fit well
- Plotted both adjusted and unadjusted mortality against volume. 90% control limits used as "alert" and 95% as "alarm"
- 4 units lay outside 90% limits – effect of case-mix adjustment is clear – it "shrinks" units 17, 8 and 31 towards the group mean and brings them inside the control

limits showing that for 17 and 8, their high mortality is partly due to small numbers and case mix.

- The effect of small number is clearly shown and there seems to be a volume outcome relationship – the high mortality of units with low volumes is within acceptable statistical limits – this fact would be missed in a simple ranking method
- Unit 33 and possibly unit 3 have lower than expected mortality worthy of investigation
- The eagle eyed will also spot that some of the units with zero mortality fall outside the control limits and some within, but none of the zero morality units is below the 99% limit – no deaths does not mean exceptional performance
- The run chart shows clearly that unit 33 has consistently better than average performance

d. Comment

- Provides a “buffer zone” for indicating divergence from the mean mortality
- Tool to screen for possible “abnormal” results
- Data quality
  - representativeness – data source not random
  - 2 sources – validation problem, duplication of risk factors
- Primary end-point – in-hospital mortality
  - bias and confounding factors e.g. discharge policy
  - Could use other outcomes e.g. 30-day mortality, morbidity, infection, length o stay etc.
- Cannot draw conclusion because inability to discriminate and attribute specific outcomes to individual surgeon
- Could stratify by type of surgery or group by surgical volume and repeat chart.
- Choice of limits arbitrary – SPC methods often insist on 3SD~99.9% limits so these are more lenient – more “false” positives – lower limits give more sensitivity but are less specific – depends on nature of outcome and variation where limits are set.
- Overall method gives better impression of good or poor performance that simple ranking unless includes uncertainty measures – it would have been helpful to see both side by side

3. Comment on Statement

- Use of word underperforming emotive – probably better to say units with unexpectedly high or simply, high, mortality lie above control limits. Essential point is that these units are unlikely to be drawn from the same distribution as the rest and therefore differ in some way that is worthy of investigation

- Fundamental assumption is that group mean is best estimate of expected mortality for each unit. The hypothesis being tested is that observed mortality in each unit does not differ from the group mean

- Choice of 95% control limits – arbitrary and debatable - as above usual SPC methods require much wider limits but depends on balance between sensitivity and specificity required

- Can generate false positives and miss true positives

- It is true to say that Units 8 and 17 have higher than expected mortality although 17 is probably a small number/ case mix issue – unit 8 is still appears high. Unit 31 is low. All the other units fall within the expected distribution for a controlled system with the group average as best measure of performance. In any performance system it would be worth looking more closely at Unit 8 and perhaps 17 and also at unit 31 – what can be learned from it.

- Application to more complex systems questionable – why – these methods help us better understand process – can only be seen as preliminary steps

4. Report to National Board

- These methods improve on league tables in a number of ways – explicit measure of uncertainty, relationship with volume, open debate about performance, where to draw control lines etc.

- Essential tool in quality improvement programmes and processes

- Also show if "worst is worse than the worse can be" – in this example the unit with the highest adjusted mortality is shown to be consistent with average expected performance

- Mortality league tables – implication on "best" or "worst" – no measure of uncertainty; only one outcome measure of hospital performance

- Merits – allow openness about clinical performance

- Mortality control charts – screening tool to identify outliers and to stimulate further investigation

- limitations – cannot use it to label units as under (or over) performing [I disagree here – it can if you have an agreed measure of performance]

- Follow-up studies using time trends CUSUM analysis

- Sequential monitoring of mortality –can be used to give early warning of failing units – examples include SPRT, VLAD

- Prospective monitoring of surgical performance by other data e.g. post-operative morbidity, quality of life measurement

- Investigation includes excluding data problems, examining the process for abnormalities, quality of care investigations etc

Comments

This was a difficult question not helped by the confusion caused by the phrase “objectives provided”, although in practice this probably made little difference. On the whole the abstract was well done although some candidates did not appear to have noticed the word limit and rehashed the paper. Others seemed not to know standard abstract headings. Very few candidates correctly identified the study design (a validation study is not a study design per se) as an ecological study – a database study or cross-sectional study would have been acceptable.

The critical appraisal question was not well done. Many candidates criticised but were not critical – phrases such as “this method seemed appropriate”, “there should have been a table” without providing justification were common. Few candidates had a well structured approach to the appraisal, most confused the risk adjustment model with the generation of control charts, and although the statistical analysis was complex, few stated that its purpose was to control for the confounding effects of case-mix, severity, age etc. Many candidates criticised the lack of sample size estimations which was actually not necessary. What limits the generalisability is the risk adjustment model, not the control chart method. Only a handful noted that the effect of the risk adjustment was to shrink the outliers but it had little impact on the majority of surgical unit estimates. Only a few commented at all on the run chart and fewer still showed they understood it. Some candidates appear to have run out of time.

Many candidates were not familiar with control chart methods although some clearly were and mentioned Shewhart, Juran, Deming and others. On the whole part 3 was not well answered. Many candidates appreciated that the expected performance was the group average and that limits were constructed around this average but many then got confused about how to interpret the limits. Candidates did not appreciate the arbitrary nature of the choice of control limits nor how this choice affected false positive rates and the balance between missing true poor performers on the one hand or over investigating average performers on the other. Also no candidate commented, in this part of the paper, on the measure of surgical performance used.

In part 4 the main faults were not to write a structured report as the question asked, nor to answer the question. Very few candidates stated that the main problems with ranking methods are that the bottom ranked units may not be worse than the worse can be – and that ranks are unstable – that is they have wide confidence intervals. The actual rank of any one unit may be difficult to ascertain. Control charts avoid ranking and also show if units truly exceed limits of expected variation.

PAPER IIB

Hepatitis C virus (HCV) infection is usually diagnosed by testing for HCV antibodies in serum, using an enzyme immunoassay (EIA). The EIA in use by your local blood bank has been evaluated by comparing it with a "gold standard" test.

The EIA test was performed on 160 people known to be HCV-infected according to the gold standard test. 152 of these people were positive according to the EIA test.

The EIA test was also performed on 1000 people known to be HCV-negative according to the gold standard test. 20 of these HCV-negative people were positive according to the EIA test.

(a) Make a 2 x 2 table to show these results. (5 marks)

(b) What is the sensitivity and specificity of your region's EIA test? (5 marks)

Using these values for sensitivity and specificity, construct 2 x 2 tables to show the results you would get if you compared this EIA test with the gold standard among:

(c) 10,000 university student blood donors – assuming the true prevalence of HCV infection in this sub-group is 0.4%. (5 marks)

(d) 1,000 consecutive patients in the local haemodialysis unit – assuming the true prevalence of HCV infection in this sub-group is 33.3%. (5 marks)

(e) What would be the positive predictive value and the negative predictive value of the EIA test in each of these two sub-groups of people? (5 marks)

After testing the two sub-groups with the EIA test, the director of the blood bank expresses concern to you about the prevalence of HCV infection in the region. The director suggests that university students who tested HCV-positive should all be notified of their EIA test result, and that the local haemodialysis unit should be temporarily "closed down".

(f) Outline the approaches you would use to clarify the epidemiology of HCV infection in the area served by the blood bank. (15 marks)

(g) What advice would you give the director of the blood bank, giving justification?

(10 marks)

Key Points

(a)

		Gold standard		
		positive	negative	Totals
EIA	positive	152	20	172
	negative	8	980	988
	Totals	160	1000	1160

(b)

$$\text{Sensitivity} = \frac{152}{152 + 8} = 95\%$$

$$\text{Specificity} = \frac{980}{20 + 980} = 98\%$$

(c)

		Gold standard		
		positive	negative	Totals
EIA	positive	38	199	237
	negative	2	9761	9763
	Totals	40	9960	10 000

(d)

		Gold standard		
		positive	negative	Totals
EIA	positive	316	13	329
	negative	17	654	671
	Totals	333	667	1000

(e) Blood donors:

$$PPV = \frac{a}{a + b} = \frac{38}{38 + 199} = 16\%$$

$$NPV = \frac{d}{c + d} = \frac{9761}{2 + 9761} = 99.99\%$$

Haemodialysis patients:

$$PPV = \frac{a}{a + b} = \frac{316}{316 + 13} = 96\%$$

$$NPV = \frac{d}{c + d} = \frac{654}{17 + 654} = 97\%$$

f) Identify some of the key issues in relation to known epidemiology

- The HCV is a global public health problem. The 3% of the world population have been infected.

- The prevalence of the virus varies around the world. From 0.5% in UK to much higher prevalence in the Middle East, Asia and Africa.

- Blood has been identified as the main vehicle of infection

- About 80% of people who acquire the infection become chronically infected. The rest appear to clear the infection spontaneously. Around 20% of chronically infected people may develop serious liver disease such as cirrhosis after 20 years, 1–2% per year will develop liver cancer.

- Risk groups eg UK

  - The major route of HCV transmission by sharing equipment for injecting drug use.
  - There is a risk to recipients of blood transfusions (before September 1991), or blood products (before 1986) in the UK. For example, there is a high prevalence of HCV in people with haemophilia who received untreated clotting factors before 1986. All blood and blood products in the UK are now screened.
  - Mother to baby transmission does occur, but appears to be uncommon, with upper estimates of 6%. However, this is increased to around 15– 20% when there is co-infection with HIV. There is no association proven between breastfeeding
  - Sexual transmission of hepatitis C is uncommon - less than 5% in regular sexual partners.
  - Transmission can occur through medical and dental procedures where infection control may be inadequate.
  - Health care workers (and, to a lesser extent, other workers, such as police, prison staff and social workers) may be at risk of HCV infection from occupational injuries, for example needle-stick injuries.
  - There is a risk from tattooing, ear piercing, body piercing and acupuncture with non-sterile equipment.
  - There is some evidence that transmission may occur through the sharing of toothbrushes, razors and other personal toiletry items that could be contaminated with blood.

3 approaches possible

1. Assume known epidemiology applies to local population; describe epidemiology in terms of local knowledge of at risk groups

2. Undertake cross sectional study identifying at risk groups and testing for HVC or test for HVC in high risk group eg drug misusers undergoing Hepatitis B immunisation

3. Undertake case control study of those identified as being HCV infected

Good/ excellent response could include

Challenges :

The difficulty in detecting newly acquired HCV infection. Only a very small proportion of cases appear to manifest a clinical illness

The difficulty in gaining access to populations at high risk of infection. The major route of transmission is an activity that is illegal, socially marginalised and often undertaken by people who are not in close contact with the health system. In addition, people who have injected drugs but no longer do so may not be keen to acknowledge this aspect of their past. Ongoing challenge to enumerate these subpopulations.

Investigation of the natural history of HCV infection needs to take into account the long and variable course of disease progression.

Needs to take account the complex and evolving diagnostic methodology for HCV infection. Since the virus was identified, there has been a rapid development of both immunological and molecular techniques for detecting HCV antibodies and genetic material. Difficulty in gaining access to adequate diagnostic testing, much of which is expensive and requires specialised laboratory skills to carry out.

g) Give clear advice plus supporting reasons

In the low prevalence population of university student blood donors, to inform students that they were HCV-infected on the basis of the region's EIA test alone would be unwise. The PPV of 16% means that only 16% of students who test positive will genuinely be HCV-infected, and thus 84% of students with a positive test would be told that they were HCV-infected in error.

In testing for a disease in a low prevalence population needs a two-stage testing procedure:

- First, a screening test with a very high sensitivity is used to identify as many true positives as possible.

- Next, a second test with a very high specificity (often called a “confirmatory test”) is performed on individuals who had a positive initial screening test.

In this situation inform university students of their HCV positive status only if they test positive to the confirmatory test as well.

Discussions should take place with University on approach include risk lowering and education

Consideration should be given to media involvement

There is no requirement to take any specific action against the regional haemodialysis unit, since the sensitivity and specificity of the EIA test in this high prevalence population are both good, and the HCV prevalence as shown by the test is within the normal range for patients on haemodialysis.

Approach to risk reduction and prevention should be discussed with renal unit

Good/ excellent response could include

In low-prevalence populations most screening tests are likely to have a low positive predictive value.

When applied to a high-prevalence population (for example, a sub-group of injecting drug users) the same screening test will have a high positive predictive value.

Antibodies to HCV are normally confirmed by a recombinant immunoblot assay (RIBA) test.

HCV is common in people with other blood borne infections, such as HBV and HIV.

Comments

The purpose of the question was to test candidates understanding of a basic epidemiological concept and to apply that understanding to a given problem.

Most candidates demonstrated a good understanding of sensitivity, specificity and predictive values and could produce 2x2 tables.

Candidates who did not fully complete the tables lost marks. Where candidates had made calculation errors, marks were awarded if the workings showed an understanding of the concepts.

On the whole the response to the questions on outlining the epidemiology and advising the director of the blood bank were poor.

Candidates often failed to set a context in terms of the Public Health importance of HCV and some could not describe the basic epidemiological features.

A number of candidates failed to use the answers in the 2x2 tables to illustrate their answers, in particular candidates failed to comment on the variation of predictive values with prevalence in the target groups. A number of candidates chose to focus on repeating the results in the 2x2 tables rather than outlining the approaches they would use to clarify the epidemiology.

Many candidates focused on just one approach when clarifying the epidemiology rather than considering a range of approaches and identifying strengths and weaknesses of each. Too many candidates placed over reliance on routinely collected data as a means of clarifying the epidemiology. Few candidates pointed out the difficulty of working with HCV infected groups.

When giving advice a number of candidates failed to emphasise the implication of a low positive predictive value in low prevalence populations and failed to discuss the role of confirmatory testing. In addition a number of candidates did not consider wider implications such as anonymised testing, advice about prevention and protection and managing the media.

In terms of the haemodialysis unit it did not matter what the advice was so long as the justification could support it. However, few candidates were able to provide adequate justification, particularly when suggesting the unit should be closed.