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

JANUARY 2005

EXAMINATION QUESTIONS WITH EXAMINERS' KEY POINTS AND COMMENTS

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

PAPER IA

QUESTION 1

Write short notes on two of the following three:
a) Meta-analysis of observational studies
b) Measurement error in epidemiological studies
c) Community-based intervention studies

KEY POINTS IN APPROXIMATE ORDER OF IMPORTANCE

a) Meta-analyses of observational studies

- Brief definition of meta-analysis: statistical analysis of a collection of studies.
- Observational studies: non-experimental studies, mainly case control and cohort studies.
- Role of meta-analyses: could be synthetic (estimation of overall or average effect across studies) or analytic (identification and estimation of differences among study-specific effects).
- Steps of a meta-analysis: specifying study variables (exposure, outcome, confounders, intermediates, effect modifiers), systematic identification of studies, assessment of their eligibility, quantifying effects, examining heterogeneity between studies; look for evidence of publication bias, use of funnel plots.
- Pointing out that use of meta-analyses in observational studies is more controversial than in intervention studies
- Description of regression models, statistical adjustment for confounding, selection bias and misclassification
- MOOSE and other guidelines now available to assist with reporting
- Mentions statistical reanalysis of individual studies using raw data if available.

b) Measurement errors in epidemiological studies

- Some measurement or classification errors are almost inevitable in epidemiological studies.
- Could affect exposure, outcome and confounders.
- Explain differential (classification error depends on the values of other variables) vs. nondifferential (error does not depend on other variables). An example would be useful in illustrating the difference.
- The direction of bias arising from differential misclassification is often unpredictable. Can give overestimated or underestimated effects.
- Nondifferential or random misclassification generally results in bias towards the null value.
- Nondifferential misclassification of confounders results in uncontrolled or residual confounding (thus bias can be in either direction).
- Description of ways to reduce errors of measurement of exposure and outcome.
- Mentions statistical methods of dealing with measurement errors
- Mentions statistical methods of measuring observer variation

c) Community based intervention studies

- Definition: trials or experiments in which the units of study are entire communities instead of individuals

- Mentions at least two of the following reasons or need for such studies: some interventions cannot be delivered at an individual level without contaminating the control group; can be less costly per person; many behaviours are socially related; environmental changes may be more effective than trying to modify individual behaviours.
- Require control as well as intervention communities.
- Allocation to treatment groups by randomisation, pairing or matching.
- Outcome information can be collected at either individual or community level.
- Both analysis and sample size calculation need to take into account the fact that individuals within clusters are more similar.
- Description of intraclass correlation and design effect in sample size estimation
- Mentions the options of cohort sampling and repeated cross section sampling in subject selection
- Gives good examples of such studies e.g. fluoridation trial, COMMIT, North Karelia Project
- Unit of analysis is the community not individuals
- Distinctive ethical issues

COMMENTS

(a) Meta-analysis of observational studies

This question was quite well answered although quite a number of candidates simply did not cover enough aspects of meta-analysis to accrue many marks. Quite a large proportion of candidates included details of meta-analysis as it relates to randomised controlled trials when the question was about meta-analysis in the context of observational studies. A couple of people gave answers which looked like they had 'learned' the answer, perhaps from the key points of a previous paper, and demonstrated little understanding of the concepts they were writing about.

(b) Measurement error

This questions was the less frequently chosen of the three sections and when answered was generally very poorly approached. Some candidates confined their answers to type I and type II errors, which are a different issue. In addition to the inclusion of erroneous information many candidates included insufficient information to accrue many marks.

(c) Community based intervention studies

This question was generally well answered. Key items looked for were that the unit of randomisation and thus analysis is at the community rather than individual level, and that this has an effect on the precision imputed to the difference between intervention and control groups and correspondingly on the sample size required. Again many candidates wrote about too few points to accrue many marks.

Two candidates wrote answers to all three sections when answers to only two sections were required.

QUESTION 2

In a double-blind placebo-controlled trial, 109 smokers aged 18-55 years who smoked at least 15 cigarettes per day were randomised to one of two groups. The active treatment group received the drug Selegeline for 26 weeks, the control group received a matching placebo. In addition, both groups had nicotine skin patches for the first 8 weeks only.
14 of 56 subjects in the active group were continuously abstinent from smoking (did not smoke at all) for 52 weeks, compared to 6 of 53 subjects in the control group ($p=0.08$).
At 4 weeks the craving for cigarettes (strong desire for a cigarette) was lower in the active group than in the control group ($p=0.02$).

(a) Explain what is meant by the terms:

- Randomised
- Placebo controlled
- Double blind

and why these are important features of this study. (4 marks)

(b) Explain what should be inferred from the above findings, and comment on the information presented above. (6 marks)

KEY POINTS IN APPROXIMATE ORDER OF IMPORTANCE

Clear explanation of Randomised, Placebo controlled, Double blind, importance of each.
Express 14/56 and 6/53 as 25% and 11%.
Active treatment doubled abstinence rate, but difference could be accounted for by chance.
Study small to detect such a difference – prior power calculation needed.
Significant difference in craving. Although the size of the difference was not given, in both instances, confidence intervals for suitable measures of effect size e.g. difference of proportions or odds ratio should also be reported.
Abstinence from smoking throughout 52 week follow-up is binary. Analyses taking into account when as well as whether relapsed, or indeed quit at all, would be more powerful and may yield significance at the conventional 5% level. Such analyses require survival methods. Some evidence of possible benefit but chance cannot be ruled out, therefore larger study needed.
Why is craving at 4 weeks reported here? Was it measured at several times, but only significant at 4w?
Quality journals now require RCTs to be reported according to the CONSORT guidelines – key elements patient flowchart and checklist to inform referees – a lot more information than the above needs to be reported for clarity.
Need to also report drawbacks to active treatment – adverse events and drug costs – and seek to balance against apparent benefits. Even if clear cost-benefit were shown, this is still a difficult decision about whether the drug should be provided gratis, or patients should be required to pay. Ideally, economic analysis should take long-term perspective of benefits of abstinence in avoiding later healthcare demands, but this is difficult to do – conversely having to pay may help motivate continuing abstinence.

COMMENTS

This question was generally quite well answered.

(a) Explanations of randomisation, placebo controlled and double blind – and why these are important features of this study

This part of the question was worth 4 of the 10 marks yet many candidates clearly spent the balance of time on this question answering this section at the expense of the later section which was worth 6 marks.

Key features of randomisation looked for included the fact the process is specifically random and not at the researcher or participant's choice, and that every individual has an equal chance of being allocated to a particular treatment group. The importance of this is that it is an attempt to produce groups that are equal with respect to all features except the difference in treatment, and deals with known and unknown confounders, thus provided the trial is large enough acceptably balanced groups are produced.

Key features of placebo controlled were the placebo effect/Hawthorne effect and assisting with blinding by disguising whether the subject is receiving the active drug.

Key features of double blind are that blinding the prescribing doctor avoids allocation bias where the doctor gives individuals s/he thinks would benefit most (or least) the active drug; the placebo effect from the point of view of the patient who would otherwise know which treatment they have received; and the assessing clinician whose assessment will not be objective if they know which group has received which treatment.

Several candidates failed to answer the final section to indicate why these features are important in terms of the specific study i.e. that they are cardinal features of an RCT which is the best way that we have available to test the clinical effectiveness of new treatments.

(b) What should be inferred from the results

Many candidates spent a lot of time describing (often incorrectly) the technical interpretation of statistical significance and non-significance without getting to the actual meaning of the results in terms of what one might do with the information. Too little time was spent on the implications, the small study size, the lack of effect size for the craving result, the lack of information about whether these were the primary outcome measures and how they were measured or whether they were chosen for presentation in order to show best results. Many candidates failed to note that the analysis had been carried out on an intention to treat basis, leading to greater interpretability. Few failed to mention the promising yet inconclusive nature of the results and need for a further, larger study which the results of this study would greatly inform, especially in terms of the sample size calculation. Several candidates seemed to have forgotten that the study was carried out as a randomised controlled trial and mentioned features that should have been taken care of in the design e.g. the need to control for confounding rather than checking the baseline features of the two groups to ensure that in such a small study the groups were balanced.

QUESTION 3

Write short notes on the epidemiology of colorectal cancer in relation to screening for the condition.

KEY POINTS

A passing candidate would be expected to give a structured answer, e.g. based on the principles of screening of Wilson & Junger (1968) or those of the UK National Screening Committee, and to have clearly related the epidemiology of colorectal cancer, especially the presence of high risk groups, to screening for this condition.

A good candidate will give a knowledgeable and coherent answer considering different types of colorectal cancer screening, whilst an excellent candidate will demonstrate a particularly thoughtful and critical approach based on research evidence.

KEY POINTS IN APPROXIMATE ORDER OF IMPORTANCE

Definition of Screening e.g:

- Specific examination of well individuals to identify an existing illness at a pre-symptomatic stage or to identify susceptibility to a specific disease.
- The National Screening Committee uses the following definition of screening: Screening is a public health service in which members of a defined population, who do not necessarily perceive they are at risk of, or are already affected by a disease or its complications, are asked a question or offered a test, to identify those individuals who are more likely to be helped than harmed by further tests or treatment to reduce the risk of a disease or its complications.

Importance of problem:

- Colorectal cancer accounts for approx. 13% of newly diagnosed malignancies in males and females. It is third commonest cancer diagnosis in females and second in males. It is the second commonest cause of cancer-related death in UK.

Natural history and epidemiology:

- Colorectal cancer usually presents with abdominal symptoms often at relatively late stage in its course.
- Risk of colorectal cancer increases with age, diet low in vegetables (possibly high dietary total fats, or saturated fat or obesity), lack of physical exercise, cigarette smoking and bowel diseases such as ulcerative colitis.
- Some patients with colorectal cancer come from high-risk groups; about 10% have a genetic predisposition. Colorectal cancer can be inherited - familial adenomatous polyposis (FAP) or hereditary non-polyposis colorectal cancer (HNPCC).

Cost effective test:

- The sensitivity, specificity and costs of FOB tests vary depending on the test
- Screening frequency for FOB of annual or biannual have both been advocated.
- Screening by flexible sigmoidoscopy at age 55 has been evaluated as an alternative to FOB screening
- For high-risk patients – where FAP or HNPCC are diagnosed – genetic screening should be offered to first degree relatives. Screening from teens by sigmoidoscopy and colonoscopy/barium enema
- For normal risk populations, 2 trials of screening by faecal occult blood (FOB) published in 1996, (one Nottingham, one Denmark) showed earlier diagnosis and mortality benefit. Previously 5 USA trials showed earlier diagnosis, and one of the these also showed mortality benefit
- Recent UK study (BMJ July 2004) showed feasibility of colorectal screening by NHS in UK

Effectiveness of treatment:

- Screening by FOB firstly requires availability of adequate colonoscopy services to assess positives
- The main treatment option is surgery. At present about one third of patients are admitted to surgery as emergencies and clinical outcomes are poorer. Stage (severity) of disease is vital for prognosis and explains potential benefit of early diagnosis by screening

Conclusion:
Candidates should approach the question in a structured way, critically considering the potential value of colorectal screening in light of its epidemiology.

COMMENTS

A lot of candidates answers were sometimes sketchy and even scrappy in some cases.
Most candidates were aware of the main epidemiological issues but some omitted reference to genetically related high risk groups and therefore lost marks.
The main problem with the screening part of the question was that many candidates answered the question in a rather general way, relying too much on discussing the principles of screening without relating these clearly enough to the specifics of colorectal cancer.

QUESTION 4

Outline the factors leading to an increase in antimicrobial resistance in micro-organisms pathogenic to humans. Describe strategies to reduce their impact on public health in a named country.

KEY POINTS IN APPROXIMATE ORDER OF IMPORTANCE

Factors include
Large volume of antibiotic use - animals and humans - leads to selection of resistant organisms
Over prescribing in general practice especially
Delay in prescribing effective treatment because empirical choice fails
Lack of choice of antibiotic
Unable to use the correct one - e.g. contra-indications/lack of availability

For a very good answer
Reference to transfer of resistance from non-pathogenic to pathogenic organisms

Strategies for control and reduction of resistance:
Resistance appears to be linked to amount of antibiotic use and availability.

<u>Humans</u>

1. Use fewer antibiotics:

Role of guidelines (sore throats, colds etc.)
Shorter courses
Use narrow rather than broad spectrum antibiotics
Re-educate public

2. Use antibiotics better

Computer assisted choice of empirical treatment
Audit
More rapid diagnosis, new techniques not involving traditional culture
Cyclical policies on use

3. Reduce need to use (Prevent cross infection)

Education and audit
Improve cleaning
Improve building design
Decrease throughput

Is it possible to reduce pool of pathogens in community? (Role of resistant bacteria generated in animals.)

4. Develop new antimicrobial agents, may need joint government/industry investment

<u>Animals</u>

Stop the use of antibiotics as growth promoters
Makes meat more expensive
Involves long period of investment

Obviously sick animals still to be treated but some restriction may be necessary on types of antibiotic (e.g. not using novel agents soon to be used in humans).

ADDITIONAL POINTS TO IMPROVE THE SCORE

Large volume of antibiotics in use, 50% animal (33% growth promotion) 50% human (on weight basis). Many more courses prescribed in general practice than in hospital.

Problem: examples - MRSA, MDR tuberculosis, penicillin resistant pneumococci

Difficulties: transfer of resistance from non-pathogenic to pathogenic bacteria, particular role of gut bacteria in humans;
Cross resistance between veterinary/agricultural antibiotics and those used in humans;
Delay in initiating effective treatment because empirical choice fails;
Lack of choice of antibiotic;
Unable to use antibiotic.

Result: increased morbidity, increased mortality

MDR tuberculosis probably greatest threat (captain of men of death), problems with compliance with multi-drug regimens, side-effects, length of treatment, cost, link with HIV infection. Could argue for MRSA or vancomycin resistant enterococci because of possible transfer of resistance once generated.

COMMENTS

A lot of candidates answers were sketchy in places and sometimes even scrappy. The standard of answers to this question were generally reasonable.
Most candidates knew the main factors leading to an increase in antimicrobial resistance. Some however omitted reference to antibiotic usage in animals, an important issue, and subsequently lost marks.
The strategy part of the question was sometimes answered in too general a way. Better candidates offered clear and practical ideas for tackling antimicrobial resistance.

QUESTION 5

Write short notes on the relative merits of mortality data, hospital activity data and primary care data in assessing the health needs of a population in a named country.

KEY POINTS IN APPROXIMATE ORDER OF IMPORTANCE

- No single source of data can provide comprehensive picture of health needs.
- Well organised, clearly set out, well reasoned illustrated answers with evidence of clarity of thought.

Mortality data
- identifies serious health problems
- gives some indication of size of problem
- legally required
- refers to an event unlikely to be missed
- easier to diagnose that morbidity, so more likely to be accurate
- but not all disease leads to death, so only limited picture
- data can be unreliable particularly with elderly
- can calculate specific mortality rates, direct and indirect standardised rates
- can estimate avoidable deaths/years of life lost prematurely

Hospital activity data
- Indicator of morbidity
- Useful for investigating health needs which result in hospital care. Where hospital care is virtually mandatory – such as fractured neck of femur – this gives a reasonable estimate of the incidence
- Routine data source; may be variations in quality of coverage (missing data etc.) and in coding.
- Gives an indication of met need, but not unmet need

Primary care data
- High proportion of population registered and attend in a year
- Wide range of health problems dealt with including acute disease, early stages of disease, and management of chronic disease
- Useful for information on chronic diseases which make a major contribution to population morbidity, but are not well covered by either mortality or hospital data e.g. diabetes, hypertension, mental illness.
- Can provide information about prevalence of risk factors (e.g. smoking), uptake of immunisations and other preventive services, prescribing and referrals.
- Coding should be consistent between practices

- Computer systems allow selection of categories of patients
- Individual patient based
- Denominator present in system

COMMENTS

Many candidates did not answer the part of the question ".... in assessing the health needs of a population...". Most candidates had a reasonable understanding of the nature, coverage and limitations of the three types of data, but very few explained adequately how these data are useful in assessing health need. Hospital utilization data do not necessarily indicate need. It depends on hospital policy, fee structure, health financing system, etc. The same applies to primary care data.
A few candidates explained what health need means, and one discussed it in detail (which is not quite what the question asks). Sufficient to explain the uses and limitations of the three types of data, and that these data on mortality and morbidity help provide a picture of the health of the community, and can be used to identify some health needs (citing examples). For example, primary care data might be expected to give better information about some chronic diseases that do not usually cause death or necessitate hospital admission.
For certain groups, and some health conditions, these data may not be useful, and specific studies are required.

QUESTION 6

What are the strengths and weaknesses of lifestyle surveys in providing information about health-related behaviour such as smoking and dietary habits?

KEY POINTS IN APPROXIMATE ORDER OF IMPORTANCE

Strengths:
- Can collect information and estimate prevalence of risk factors where routine data not available
- Can target sub-groups of population e.g. people not in contact with services
- Can collect information and estimate prevalence of lifestyle behaviours (risk factors) such as smoking, alcohol consumption, physical activity and dietary habits in the population
- Can collect information on multiple variables from participants, once they are engaged in survey
- Can collect information on participants' views on health or health services
- Serial surveys can assess changes in lifestyle behaviours over time and monitor effectiveness of activities such as health promotion programmes
- Can be carried out in different ways e.g. face-to-face interviews, postal surveys
- There is usually no alternative source of this information

Weaknesses:
- Non-response bias.
  - Some degree of non-response is a feature of nearly all population surveys
  - The findings will not be truly representative of the whole population from which the sample was drawn.
  - There is evidence that people with different lifestyles have a different likelihood of completing a lifestyle questionnaire.

- Reporting bias
    - May arise if people underestimate or play down particular aspects of lifestyle behaviour, or conversely if they overestimate or overstate particular aspects.
    - It is difficult to assess how much this applies in a given survey unless responses can be checked against other pieces of evidence (e.g. evidence of smoking from saliva or urine tests).
- Comparability with other lifestyle surveys may be limited, unless the method and question wording are the same.
- Different approaches may have different strengths and weaknesses. e.g. Face-to-face interviews are labour-intensive. Postal surveys may allow a larger sample size, but tend to have higher levels of non-response.
- Some degree of non-response is a feature of nearly all population surveys, but attention to survey method may improve response rates. e.g. nature of initial approach, nature of letter of introduction etc.
- It may be possible to obtain some indirect evidence about the non-responders, so as to make an estimate of the kind of bias that may be introduced by their omission.

COMMENTS

Some candidates confused lifestyle surveys with studies of disease prevalence. Others confused this with a cross-sectional, analytic study to identify risk factors of a disease. Lifestyle surveys are simply prevalence studies of lifestyle risk factors, that routine health statistics cannot capture, and also serve useful public health functions - as basis for intervention.
Many candidates failed to mention that this is often the only source of information about certain health-related behaviours, and some did not mention the important weaknesses of response bias and reporting bias.
Some candidates focussed on a particular lifestyle survey rather than the general characteristics of these surveys, and a few reported results from specific surveys. A few candidates went unnecessarily into details of content, construct and other measures of validity that are more linked to psychological / mental health surveys.
Quite a few candidates misunderstood the concept of ecological studies and ecological fallacies. Lifestyle surveys are person-based and do not rely on grouped data.

PAPER IB

QUESTION 7

Define cost-benefit analysis, cost effectiveness analysis, and cost utility analysis. Describe the key steps and problems in carrying out one of them, using appropriate public health examples.

KEY POINTS

Purpose of economic appraisal is to compare two or more alternative course of action in terms of both their costs and consequences. All methods measure costs, but differ in the extent to which they measure and value consequences. Economic appraisals are dependant on having good evidence on the effectiveness of health interventions.

<u>Cost-benefit analysis.</u>
An economic evaluation where the consequences of programmes are valued in monetary terms, so as to make them commensurate with the costs. This is the broadest form of analysis in which you can judge whether the beneficial consequences of a programme justify the costs. This allows comparison of alternative programmes where the consequences are not identical.

The range of benefits that can be valued in money terms is fairly limited. Therefore, in practice, many cost benefit analyses are more limited than cost utility or cost effectiveness analysis, and are limited to only those costs and benefits that can be easily expressed in money terms. Part of this problem can be addressed by valuing benefits in health care programmes in terms of individuals' willingness to pay.

<u>Cost effectiveness analysis.</u>
An economic evaluation where the consequences are measured in the most appropriate units, such as 'years of life gained' or 'cases correctly diagnosed'. There is no attempt to value the consequences so there is an implicit assumption that the outcome is worth having. This type of study usually measures only one consequence which differs in magnitude between the alternative actions. Therefore, disparate actions with a common consequence can be compared e.g. heart surgery and mandatory cycle helmet legislation. Some analyses can present a range of outputs alongside cost and leave it to the decision maker to judge the relative values of these. (Cost minimisation analyses are a type of CEA where the consequences of alternative courses of action are equivalent.)

<u>Cost utility analysis.</u>
In this form of analysis, rather than having a financial value, the consequences are valued by their utility. This weights the consequences by the preferences of individuals or society. This allows for quality of life adjustments to be given to the outcome e.g QUALY, DALY. Such a generic outcome measure allows comparison of different programmes.

<u>Key steps</u>

1. Define the question for economic analysis (identify alternatives being compared and viewpoints from which comparison made)

2. Clear description of competing alternatives to allow consideration of generalisability.

3. Establish evidence of effectiveness of alternative programmes.

4. Identify all relevant costs and consequences for each alternative (institutional context important).

5. Consider measurement issues (units, feasibility etc)

6. Consider credibility of values ascribed.

7. Consider timing and need for discounting.

8. Consider equity issues (willingness to pay methods may be constrained by ability to pay)

COMMENTS

Most candidates defined the different sorts of economic appraisal accurately.

Very few candidates answered the key steps section well, some apparently omitting this section altogether. Many answers were poorly structured. Most answers were narrow in scope, and failed to discuss basic steps, such as defining the economic question and the interventions / alternatives. Many 'key steps' sections were limited to listing different types of costs or discussing the detail of utilities.

QUESTION 8

What opportunities are there for schools (in a named country) to contribute to reducing inequalities in the health of children and the wider community?

KEY POINTS IN APPROXIMATE ORDER OF IMPORTANCE

- Escaping the poverty trap.
- Knowledge for pupils on the wider determinants of health, health related behaviour and skills, relationships, practical skills.
- Developing community awareness, social justice.
- Reducing truancy and school exclusion.
- Reference to studies on this subject.
- Examples of evaluated projects from the literature and from practice.

The school stakeholders are pupils, teachers, governors, parents and families, local residents, and wider community residents.

Pupils - Behaviour change of pupils through healthy snacks, healthy school meals, and healthy tuck shops targeted at the disadvantaged; free fruit schemes targeted at disadvantaged; community support schemes for the school and pupils to help local disadvantaged people, e.g. painting their fencing.

School, teachers, and governors - Stop selling playing fields; devise healthy schools policies on physical activity, safe-to-school routes, healthier eating etc. Engage community in joint initiatives, use the school as a community resource for meetings, physical activity etc. Give more emphasis to tackling health inequalities in the curriculum with its structure and time allocation; ensure disadvantaged pupils participate in physical activity and healthy eating. Engage as a major stakeholder in local partnership to redress inequalities and media opportunities. Provide access to sexual health services and smoking cessation.

Families - Use the pupils to engage parents and families in health inequality issues through curriculum work.

US data suggests keeping teenagers in school reduces the teenage pregnancy rate.

COMMENTS

This question was poorly answered by many candidates. This was both surprising and disappointing given the prominence given to the topic in the recent White Paper – Choosing Health.

Many candidates adopted a narrow individualistic approach, only discussing health education for students on lifestyle factors, such as smoking, diet and exercise.

QUESTION 9

You have been asked to lead the implementation of a local strategy to reduce the prevalence of obesity in a population of 150,000. Outline the approach you would take in a named country.

KEY POINTS

(the major headings are major key points, then the points decrease in order of importance according to the degree of indent)

- Familiarise yourself with the key action points of the strategy
    - Talk with the author(s) of the strategy
    - Talk with the sponsors of the project and be clear about their expectations
    - Analyse the political base
    - Satisfy yourself with regard to the rigorousness of the strategy specifically noting:
        - Is it for children and young people, adults or both groups?
        - Is it a population or a targeted strategy or a combination of both approaches?
        - Does it cover monitoring the population, prevention and treatment?
        - Does it give a balanced approach to both health eating and promoting physical activity?
        - The content of the recommendations
        - The level of evidence for the action points recommended
        - The cost (including staff time costs) of implementing the recommendations
    - Identify the agencies and personnel that will be involved
        - e.g. Local authorities (including leisure and education departments), local health commissioning agencies, health service providers, schools, media, catering companies both small and large, retail outlets
    - Identify funding streams
    - Identify local opinion leaders

- Set up a group of key stakeholders
    - Acknowledge that members of the multidisciplinary group will have a variety of different backgrounds e.g. dietetics, leisure, primary and secondary care, school nursing, local government, local industry
    - Agree terms of reference, clarify membership, decide frequency of meetings, reporting within the team, with sponsors and with the wider constituency
    - Timescale and tasks for implementation (using for example project management methods such as Gantt charts)
    - Develop membership of the group as a team, mutual understanding of roles
    - Identify barriers to success (e.g. force field analysis)
    - Prioritise work and achieve early successes
    - Dissemination strategy e.g. newsletters, web publishing, workshops, etc; training and monitoring arrangements

- Post handover project
    - Check integration of recommendations into everyday practice
    - Analyse process of implementation and lessons learnt

Candidates should hopefully mention population v targeted approaches.

COMMENTS

Many candidates chose to focus their answer almost exclusively on the contents of an obesity strategy rather then discuss their leading the implementation of this strategy in a local setting. The practical steps of setting up a stakeholder group and identification of existing services together with resources did not feature in many answers.
A smaller proportion chose to mention specified management theories or answered using the headings of a generic strategy programme without contextualising the answer to the subject of obesity.
The need to follow through implementing a strategy with checking if practice was changing and analysing the lessons learnt from implementation were rarely recognised.
It is clear that a number of candidates had problems with spelling and some did not know the definition of the body mass index.

QUESTION 10

The commissioning body for health services in the district in which you work has been spending increasing resources on a range of complementary and alternative therapies. Local hospital clinicians protest that there is inadequate evidence of effectiveness to support this. Describe the public health specialist's role in addressing this problem.

KEY POINTS

(the major headings are major key points, then the points decrease in order of importance according to the degree of indent)

- Information gathering stage
    - What is/are the type(s) of complementary therapies being used, where is the resource being spent; who are the practitioners involved, what training have they undertaken, what is the status of their qualifications; are any audits or

patient outcome surveys available; what are the alternative conventional therapies used for the conditions in question; analyse the stakeholders involved; check the systems and policies of the local health commissioning and monitoring bodies regarding policy formulation, clinical governance and audit

- Assessing the evidence
    - The level of evidence for any benefit arising from the complementary therapy, including the quality of the evidence; summarise the findings for presentation
    - The cost effectiveness of the treatment compared with conventional therapy
    - Acknowledge any placebo effect
    - Identify which group might benefit
    - Investigate the policies of other health commissioning bodies with regard to complementary therapies

- Facilitate local discussion
    - engage all the stakeholders group at workshops especially the therapy protagonists
    - include patients, acknowledging their right to choice of treatment
    - be aware of potential for conflict and ways of resolving this
    - agree a way forward through local consensus building
    - identify any funding implications of agreed proposals
    - devise a system for assessing the outcome of service changes

- Communicate content of background investigation, summary of local discussion and way forward with local commissioning group and strategy boards, clinical governance groups, local health providers and the media

COMMENTS

Good answers conveyed the need for a balanced approach to the issue and the practical aspects of facilitating local discussion. Candidates often ignored the need to manage the issue of local professional concern. Good answers discussed power, conflict resolution, how groups function, and negotiation.
Practical suggestions for resolving the issue and improving services as a consequence and communicating the results of the conclusions of local discussion more widely were rarely included in the answers to this question.

PAPER IIA

You are part of the team seeking to reduce the use of expensive acute (hospital-based) services and move health care into the primary care/community setting, especially by improving the management of chronic diseases. There is a high rate of knee replacements locally and you have been asked to review this (electronic) BMJ paper (Arroll B, Goodyear-Smith F. Corticosteroid injections for osteoarthritis of the knee: meta-analysis. BMJ, doi:10.1136/bmj.38039.573970.7C) to assess whether a community-based service should be commissioned.

1. Write a structured abstract of no more than 250 words.
(10% of marks)

2. Write a critical appraisal, paying due regard to the task of your team.
(40% of marks)

3. What additional information would you like to have?
(15% of marks)

4. Write an outline 'service specification' for such a service in a community setting in a named country. Who might deliver such a service and how would its quality be assured through the 'service specification'?
(35% of marks)

KEY POINTS

1 Abstract: as per published abstract.

2 Important topic as this assesses the benefit of a widely used intervention for a common and disabling condition. Systematic review using meta-analysis of study data, not previously performed on this topic/data. Systematic literature search using a variety of search terms (described in paper) designed to identify all relevant publications from standard appropriate sources, but no attempt to identify unpublished trial data (other than those in the Cochrane CT Register). Systematic assessment of quality of studies using recognised technique with scoring system. Characteristics of each trial summarised and presented in a comparative fashion in a table. Sub-analysis of medium term outcomes from high quality trials.

Authors have chosen 'improvement' (defined in paper) as their preferred (dichotomous) outcome rather than what is deemed to be either an insensitive measure: a subjective pain scale (though such data is also analysed), or not patient orientated: the range of movement. The NNT for improvement is used to describe the effectiveness of the treatment. However, as they have used a dichotomous improvement outcome, some important information may have been lost, as 'improvement', measured in the minds of patients in terms of activities of daily living, may have covered a large range. Nonetheless the authors are constrained by what is in the trials.

Results are well and clearly presented. It is unlikely that candidates will notice the difference between study size and weighting in Fig 2, or indeed way the relative weightings differ between Figs 2,3 and 4, for those studies common to the Figs, but spotting that should be indicative of an good answer, an explanation evidence of excellence.

Publication bias is likely to be a problem despite the efforts of the authors (in using the Cochrane Register). None of the trials are large, though most are of a similar size, despite the common use of the intervention: the funnel plot in Fig 5 indicates an absence of small, small-effect studies but with just 6 studies plotted it would be unwise to draw conclusions. Small studies with small or no effect are unlikely to be published, both because they are unlikely to be submitted and unlikely to be published even if submitted.

Only 6 studies could be subject to meta-analysis, and with only a combined total of 160 cases and 157 controls the meta-analysis is still small. Strangely, the authors fail to give a NNT for this combination, yet give a combined NNT for those studies with statistically significant results despite these 3 studies accounting for only 50% of the combined size and 40% of the weighting.

The interventions were not all the same as some included aspiration, others did not: whilst the authors commented on the different doses (and repeated injections in one study) they did not comment on this potential confounder.

Despite the fact that this paper is published by an academic department of General Practice and Primary Health Care, and published in the Primary Care section of the BMJ, none of the 'injectors' (column 4 of table 2) were primary care practitioners (though unspecified in one trial). Its generalisability to Primary Care is therefore questionable.

3 Costs and better outcomes data. These would include comparative costs with same procedures in secondary health sector and knee replacement. Outcomes should be more related to functional abilities (daily activities) (so that QALYs and/or DALYs could be calculated) and include rates of and time to knee replacement (ie compared to no injections).

4 A service specification should include:

i) a patient pathway from diagnosis, through patient empowerment (expert patient programme, self-help programmes and training eg about exercise), medication (analgesia and NSAIDs), to knee replacement with this intervention identified (including repeat treatments).

ii) clinical guidelines describing best practice (injection techniques) and recommended type and dose of steroid as well as required skills (and training/experience), minimum activity levels (to maintain skills), and treatment thresholds along the pathway.

iii) clinical governance mechanism to monitor ii) above, especially training and clinical audit of outcomes. Practitioners must be part of a peer review process.

Candidates should address the issue of skill-mix. Choice of appropriate practitioner should be based on necessary knowledge, skills and competencies, not 'status' … "if they are able they can".
Accreditation processes should be mentioned to ensure those competencies: recognised professional qualification, formal training schemes etc.

COMMENTS

This was a straightforward and relatively simple question compared to recent past papers. Despite this, many answers were disappointing. In general, many candidates failed to answer the question as a whole, treating each of the four elements as almost separate questions rather than parts of an overall process.

Candidates should have no difficulty in writing a short and accurate abstract as the results were much clearer than in past exam papers as were all the other elements of a structured abstract: nonetheless the standard was not high.

This paper was a well reported meta-analysis carried out by standard methodology. Despite this, as usual, far too many candidates managed to convey the fact to examiners that they did not understand what they were talking about. It is not enough to say that "the statistical test was appropriate": the candidate needs to explain why it is an appropriate test. Too many candidates who tried to explain only succeeded in exposing their lack of understanding. In many cases these candidates got it completely wrong: marks are deducted for incorrect answers.

Many candidates had clearly failed to read the paper properly: there was a recurrent theme of failing to spot that one of the papers was from the USA and that there was a funnel plot on the final page of the paper; many had difficulty with the simple Fig 1 and also Table 1. Not enough candidates pointed out that a properly conducted meta-analysis lies at the pinnacle of the hierarchy of evidence. Many candidates thought that some of the papers should have been rejected because of their age: there is nothing intrinsic to the age of the paper that should require it to be rejected (the technology hasn't significantly changed over the years). In fact most of the older papers were rejected simply on the grounds of the quality of the paper. Many candidates spent quite a lot of time discussing whether foreign language journals had been identified or not. There is nothing in the search strategy described in this paper indicating that English Language papers only had been searched, the databases will have had foreign language papers in them.

Question 3 re: additional information was particularly poorly addressed: as in past papers this was looking for economic appraisal, which is almost universally absent from papers but an essential element of decision-making. There was surprisingly little in the answers re cost / outcome relationships.

In addition, many candidates answered the fourth question as if it were a different question, namely one about implementation as in past papers. The question was more specific and was essentially about describing what, by whom, to whom, where, when and to what standards (compare with the specifications of a washing machine or car). It was not about how one would implement a service. The answers to the service specification part were disappointing, and were very poor in the poor candidates. However one candidate was awarded full marks by one examiner and there was a very good answer from Hong Kong despite the apparent (but no real) problem of having a private primary care service in Hong Kong.

PAPER IIB

You are the public health advisor to an inner city primary care organisation within a large metropolitan area. In recent years, a "congestion charge" has been introduced, which requires drivers of private vehicles to pay a daily fee for entering the city centre. Traffic census data indicate that this has resulted in a 20% decline in car and light van traffic within the congestion charge zone.
You have been asked to comment on the following statistics, compiled from one of the primary care health centres within the congestion charge zone:

	Age			
	0-14	**15-64**	**65+**	**All ages**
Catchment Population at the time the congestion charge was introduced	2,243	5,801	2,054	10,098
Persons consulting for asthma (of any severity) during:				
12 months before the congestion charge was introduced	179	232	66	477
12 months after the congestion charge was introduced	190	240	70	500
Consultations for acute asthma attacks during:				
12 months before the congestion charge was introduced	68	83	24	175
12 months after the congestion charge was introduced	35	75	18	128

1. What measures of disease occurrence can be derived from the data shown in the table? (10% of marks)

2. Calculate a suitable measure for comparing the occurrence of acute asthma before and after the introduction of the congestion charge, and its 95% confidence interval. (20% of marks)

3. Describe (but do not calculate) a suitable statistical procedure to test whether the decrease in consultations for acute asthma after introduction of the congestion charge varies significantly across the three age groups. (20% of marks)

4. Comment on the results shown in the table and outline further information that you would seek in order to evaluate whether introduction of the congestion charge has reduced the occurrence of asthma locally. (20% of marks)

The local government office announces that the congestion charge zone will be doubled in size, to cover a neighbouring primary care organisation. This change will be introduced 15 months from now.

5. Prepare an outline proposal for a specially funded study to evaluate the impact of extending the congestion charge zone upon the general health of the population served by this primary care organisation. (30% of marks)

KEY POINTS

1. Persons consulting divided by catchment population is annual period prevalence. Consultations for acute asthma divided by catchment population approximates to incidence of spells/episodes of illness (although there could be more than one consultation per episode). Each of these can be derived for all ages, and for specific age groups, before and after the introduction of the congestion charge. The assumption must be made that the catchment population remained fairly constant over the period of comparison.

   Additional points - Consultations could also be viewed as a measure of workload or burden on services.
   Consultations could be compared either to all persons, or to asthmatics (i.e. consultations per patient).
   Acute asthma attacks treated in hospital may not be included in primary care statistics.

2. The comparison could be either a difference measure or ratio measure, based on rates of consultation with acute asthma. Technically, these rates are not proportions, instead the numerator may be considered as a Poisson count, with variance = count.

Difference measure = 128 – 175 = -47 per 10,098

(reduction of 4.65 acute asthma attacks per 1,000 registered patients per year)

95%CI = -47 $\pm$ 1.96*√(128 + 175) = -12.9 to -81.1 per 10,098
(reduction between 1.28 and 8.03 per 1,000 registered patients per year)

However, because the count is small compared with the practice population, similar results will be obtained by a test for difference between proportions, and credit will be given for the latter.

Ratio measure = 128/175 = 0.73 (27% relative reduction)
(this will be closely approximated by an odds ratio for which credit will be given.)

Credit will be given for providing a suitable formula for calculation of 95% confidence intervals for the ratio measure or odds ratio (eg. Woolf's approximation) although exponentiation is not expected using the calculators supplied.

Additional points – both difference and ratio measures are significant at the 5% level.

3. The age-specific counts of acute asthma consultations before and after the congestion charge could be analysed as a 2x3 table, calculating a 2df chi-square test for heterogeneity.

   Chi-square is the sum of $(O-E)^2/E$, across all 6 cells, where O=observed and E=expected counts.

Expected counts are derived from the marginal distributions of the table, e.g. top left = 103*175/303.

Additional points - Chi-square converted to p value by reference to tables of the chi-square distribution.

Degrees of freedom for a heterogeneity test = (row-1)*(col-1).
This ignores the ordering of age groups (i.e. is not a test for trend).
A 1df test for trend would be less appropriate here.

4. The prevalence of asthma (as indicated by persons consulting) appears to have remained fairly constant, increasing slightly in all age groups, whereas there has been a reduction in rates of consultation for acute asthma attacks, which is most marked in children and the elderly. The most important question is whether these changes are restricted to the congestion charge zone, or whether they reflect broader trends in asthma epidemiology or medical care. The key piece of additional information is therefore equivalent data from an area unaffected by the congestion charge policy. Then before-after comparisons should be made between the affected and control areas.

   Additional points – is there evidence of a change in air pollution exposure locally?
   - compare consultations for adults spending day time in the area v those working away from the congestion charge zone
   - compare seasonal pattern of acute asthma attacks in the two years
   - have there been changes in clinical care
   - have the diagnostic criteria changed?

5. More than one study design could be used here and credit is given for any appropriate design, although a case control study would be difficult to justify. Prospective evaluation potentially allows much greater control over the nature and scope of information to be collected. The basic design should remain a before-after comparison with control area(s), but broader:

a) Other routine health data sources, including hospital activity (and mortality, but probably too rare)
b) Wider coverage (more health centres / catchment populations) in both affected and control areas
c) Relevant health outcomes in addition to asthma (other respiratory problems, traffic-related injuries. Cardiovascular benefits of changes in activity patterns unlikely to emerge in short term)
d) Aim for data captured to be representative of study areas
e) Standardise data recording & coding as far as possible
f) Consider statistical power (mainly determined by numbers of events expected for each outcome)
g) Statistical analysis overall, and by susceptible subgroups (very young, very old, respiratory disease)
h) Undertake literature review
i) Practical issues of timetabling, resources, deliverables, ethical approval (if patients approached) etc.
j) Suitable style and structure for a funding application, making a convincing case.

Additional points - Direct effects of relieving congestion on travel times for ambulances and other health service staff could be considered but requires a different type of study. Parallel environmental studies of traffic flows, ambient air pollution exposure, personal activity patterns etc. to show that exposures have been altered. Qualitative work on perceived health (but possible bias in self-reported information). Collation of evidence from elsewhere, by systematic literature review, with or without formal meta-analysis.

COMMENTS

<u>Overall</u>

In general, this question was answered poorly. Very few candidates could be considered a "safe pair of hands" to interpret these simple data and to develop competent proposals for further work in a PCT (or similar) context.

Many candidates failed to demonstrate basic exam technique by failing to answer the question asked, making calculations when none were asked for, not making calculations asked for and apparently not allocating time in proportion to the weight given in terms of marks to each part of the question.

It was also notable that many candidates appeared to have limited knowledge and understanding of basic epidemiological and statistical measures and tests, and their appropriate use. There was a tendency to cite epidemiological and statistical concepts without demonstrating evidence of epidemiological insight – indeed, concepts were often cited inappropriately in the context of the question.

<u>Question 1</u>

Few candidates seemed to understand the nature of primary care data. Many confused patients and consultations in the third table. Very few candidates correctly noted that the second table was a measure of period prevalence. There was a tendency to cite measures of association (relative risks, population attributable risks) instead of, or in addition to, measures of disease occurrence (incidence, prevalence).

Question 2

Most candidates noted that the difference could be either a rate difference or a ratio difference. However many candidates, having noted the need to calculate a rate, then failed to make the numerical comparison. Few candidates made any attempt to calculate confidence intervals, and even fewer showed evidence that they understood the method of doing so.

Question 3

Most candidates recognised the need to use the Chi squared test. However, many candidates failed to demonstrate any knowledge of how the test was calculated. Many cited a chi-square test for trend but went on to describe a test for heterogeneity.
Only a few described the hypothetical assumptions underlying a (1df) test for trend as opposed to a [(r-1)*(c-1)]df test for heterogeneity.

Question 4

This question was answered reasonably well. Most candidates noted that there was only a very slight change in the numbers of people consulting for asthma of any severity, but there was a decrease in the number of consultations for acute asthma, particularly in the under 14s and in the 65 and over age groups. Most candidates made some useful suggestions for additional information, but many missed the essential point about the need for comparative before and after data from an area unaffected by the congestion charge.

Question 5

The standard in answering this question varied considerably. The better answers took a broad view of the possible health effects (rather than concentrating on asthma), and showed evidence of critical thinking in developing an appropriate research strategy.

Credit was given for most study designs that would address part or all of the question raised. However, some candidates merely described the study that had already been undertaken. Proposals were often vague, poorly presented and only obliquely related to the question. Many were far too thin for a section that contributed 30% of the marks for question IIB.

Others, who had chosen a health impact assessment approach, seemed to disregard completely the need for any further information collection, collation and analysis and some described a data collection exercise with no obvious study design at all. Credit was given for an "HIA" approach provided that suitable information sources and evaluation strategies were described. Simply citing a "recipe" for how to organise a health impact assessment received few marks.