Faculty of Public Health

of the Royal Colleges of Physicians of the United Kingdom

Working to improve the public's health

PART A EXAMINATION FOR MEMBERSHIP OF THE FACULTY OF PUBLIC HEALTH

Of the Royal Colleges of Physicians of the United Kingdom

JUNE 2008

EXAMINATION QUESTIONS WITH KEY POINTS AND EXAMINERS' COMMENTS

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

Paper IA

Question 1

A randomised controlled trial (RCT) was organised to compare postoperative collaborative care between orthopaedic surgeons and physicians in geriatric medicine with routine orthopaedic care. Of 144 elderly women admitted sequentially to hospital with proximal fracture of the femur, 36 were excluded because of pre-defined criteria, and the remainder entered into the trial.

Patients were allocated either to the treatment or control group by pre-determined random sequence. The control group (n = 54) generally remained in the orthopaedic admission ward. Patients in the treatment group (n = 54) were transferred by ambulance to orthopaedic beds in a peripheral hospital 5 km away.

A sample size calculation based on pilot study data indicated that clinically important changes in the chosen measures of outcome would be significant at $p<0.05$ with a power of 0.80 if 100 patients were included in the trial with 50 in each arm of the trial. At discharge, significantly more patients in the treatment group were independent in terms of activities of daily living (primary outcome measure) than controls (41 versus 25) and their median length of stay was respectively 24 days (range 8-197) compared with 41 (range 9-365) (95% confidence interval for difference 2 to 25).

a) Explain the following terms and procedures referred to in this abstract, so that a colleague not familiar with trials terminology could understand them ***and*** give one reason for the use of each:

a. Allocation by pre-determined random sequence (20% of marks)
b. Primary outcome measure (20% of marks)
c. Sample size calculation (20% of marks)
d. Significant at $p<0.05$ (20% of marks)

b) From the information provided, what conclusion could you draw and what are the limitations of this conclusion? (20% of marks)

KEY POINTS

a) Explain the following terms and procedures referred to in this abstract, so that a colleague not familiar with trials terminology could understand them and give one reason for the use of each:

i. Allocation by pre-determined random sequence (2 marks)
Explanation
A sequence of treatment allocations which is such that one cannot predict what treatment the next patient will receive. The sequence is pre-determined by a random method so that the treating clinician is not involved in this decision. Random indicates that each patient has an equal probability of receiving either treatment ie is not haphazard allocation. (1 mark)
Reasons for use – only 1 mark allocated for a correct answer even if several correct answers are given
The reasons for the use of random allocation is so that treating clinicians cannot influence which patient receives which treatment and thereby cannot ensure that patients with particular prognostic characteristics receive one of the particular treatments about which they may have particular beliefs about its effectiveness (selection bias). (1 mark)

Randomisation is also performed to remove/reduce confounding for both known and unknown confounders (1 mark)

Any other reasonable explanation would attract the 1 mark of credit for this section of the question

ii.Primary outcome measure (2 marks)
Explanation
The primary outcome measure is the single most important clinical feature which can be measured that the treatment aims to affect. It should therefore form the basis of the hypothesis to be tested. (1 mark)

Reasons for use – only 1 mark allocated for a correct answer even if several correct answers are given
The reasons for having a primary outcome measure which forms the basis of the primary hypothesis is so multiple testing of many different outcomes is not carried out (so called fishing expedition) to find one that is 'statistically significant' (1 mark)

The primary outcome measure will be used in the sample size calculation (1 mark)

One must ensure that appropriate data are collected to measure the outcome during the conduct of the study (1 mark(

Any other reasonable explanation would attract the 1 mark of credit for this section of the question

iii.Sample size calculation (2 marks)
Explanation
The sample size calculation is the calculation that is carried out during the design of the RCT and is carried out so that the investigators know how large the study should be ie, how many individuals need to be recruited into the trial. (1 mark)

Reasons for use – only 1 mark allocated for a correct answer even if several correct answers are given
A sample size calculation is needed so that the study conducted is not too small, and therefore risks not being able to detect clinically important effects of treatment which are actually present ie. real effects (1 mark)

A sample size calculation is needed so that the study conducted is not too large such that resources are wasted when the answer to the research question could have been reached with fewer participants in the trial. (1 mark)
Any other reasonable explanation would attract the 1 mark of credit for this section of the question

iv.Significant at $p<0.05$ (2 marks)
Explanation
Significance refers to the concept of statistical significance and is a measure the probability of the result of the study occurring by chance. The level of statistical significance is conventionally set at $p<0.05$ which means that if the study were repeated 100 times the result of the study that actually occurred would be found by chance in only 1 in 20 of these 100 repeats. (1 mark)
OR
Significance refers to the concept of statistical significance and is a measure the probability of the result of the study occurring by chance. The level of statistical significance is conventionally set at $p<0.05$. This means that if in reality there was no difference between the intervention and control in gaining independence in ADL then results as extreme as were identified here would by found by chance less often than in 1 study in 20 (if repeated 20 times) – or would have less than a 5% probability of occurring by chance alone (1 mark)

Reasons for use – only 1 mark allocated for a correct answer even if several correct answers are given
The importance of statistical significance is that it gives researchers a 'measure' of whether the result from the study is a chance or real occurrence. If the result has a statistical significance of $p<0.05$ then

by convention we accept that the result is unlikely to have occurred by chance and is likely to be a 'real' finding. (1 mark)

Any other reasonable explanation would attract the 1 mark of credit for this section of the question

b) From the information provided, what conclusion could you draw and what are the limitations of this conclusion? (2 marks)

Evidence of differences in primary outcome between treatments given in these two hospitals, the result is statistically significant at $p<0.05$ but the difference between the two groups in terms of length of stay (which was a secondary outcome) may be as little as two days or as great as 25 days ie this is a very imprecise estimate of the outcome (1 mark)

But:

1. How were activities of daily living measured – blind to allocated group, same observer for both groups, using a validated scale?
2. How was the length of stay decided?
3. What other differences existed between treatment methods in the two hospitals apart from the type of physician who was caring for the patients?
4. No estimate of difference in outcome between the two groups is given for the primary outcome, this would be helpful along with its 95% confidence interval and an exact p-value.

Two of these or relevant other points would attract the single mark available for this part of the answer.

EXAMINERS' COMMENTS

This question was not well answered despite randomised controlled trials being a core topic in epidemiological methods.

Particular problems were found with candidates describing terms using terminology already given in the question ie swapping the words around in the question does not constitute an answer. For example, some candidates described allocation by random sequence as random allocation or randomisation without actually defining what these terms meant. We were looking for an answer along the lines of: random indicates that each patient has an equal probability of receiving either treatment.

There was also a tendency by some candidates to answer a different question than the one actually asked. For example the question did not ask for a specification of what was needed to perform a sample size calculation. For the latter we asked for an explanation of what a sample size calculation was to someone unfamiliar with trials terminology and the reason for its use. Valuable time was wasted by some candidates enumerating in great detail the various pieces of information needed to perform a sample size calculation (and not necessarily always correctly).

Question 2

Results were recently published from an analysis of data from a register of visual impairment in childhood in a developed country. The authors reported that the cumulative incidence to 10 years of age of incurable and uncorrectable severe visual impairment and blindness (SVI/BL) combined was 6.2 per 10,000 live births. The cumulative incidence was compared between children born as one of a twin pair and those born as singletons with a resulting crude cumulative incidence ratio of 2.1 (95% confidence interval 1.5 to 2.9). The incidence of SVI/BL was inversely related to gestational age at delivery in both twins and singletons. Twenty percent of the children with SVI/BL subsequently died with half of them dying before reaching their fifth birthday.

a) Describe cumulative incidence and in what circumstance it can be estimated (10% of marks)

b) Cumulative incidence and an incidence rate are both estimates of incidence, explain the difference between them (20% of marks)

c) Explain what the quoted crude cumulative incidence ratio and 95% confidence interval means (30% of marks)

d) Explain what 'inversely related to gestational age at delivery' means and how this might impact on the result for twins compared with singletons (20% of marks)

e) The results from this analysis were compared with data from a cross-sectional study of children whose visual acuity was measured when they were age 10 years. Explain why the prevalence of SVI/BL as measured in this cross-sectional study may differ from the estimate of cumulative incidence derived from the register-based study (20% of marks)

KEY POINTS

a) Describe cumulative incidence and in what circumstance it can be estimated (1 mark)

Cumulative incidence is a measure of disease risk and is usually calculated when data from a static cohort or population are available. That is, a fixed population is identified from the outset of observation and followed forward in for a given period of time to identify new cases of the disease of interest as they arise in the population. CI is calculated as the number of new cases over a defined period divided by the population at risk at the start of the period. (1 mark)

b) Cumulative incidence and an incidence rate are both estimates of incidence, explain the difference between them (2 marks)

An incidence rate is the estimate of incidence used when a dynamic population is being observed, that is the denominator for the population is changing as people arrive and leave the population during the period of observation. Person-years-at-risk need to be estimated as the denominator to calculate the rate. (1 mark)

In contrast, cumulative incidence is a measure of risk which can be estimated from static cohorts where information about all members of the population is available. An incidence rate is used to estimate risk when a dynamic population is being observed. (1 mark)

The incidence rate is a measure of the incidence density (or force of morbidity) at an instant in time. The cumulative incidence is a measure of disease risk over a specified period of time, which integrates the instantaneous incidence rates/densities/forces of morbidity with changes in the population at risk as cohort members die or develop disease over this time. (1 mark of extra credit if explained in this degree of detail)

c) Explain what the quoted crude cumulative incidence ratio and 95% confidence interval means (3 marks)

The incidence ratio for SVI/BL associated with being a twin compared with being a singleton was 2.1 (95% CI 1.5 to 2.9). This means that twins have 2.1 times (ie a two fold) increased risk of severe visual impairment or blindness compared with singletons. (1 mark)

The 95% confidence interval means that we are 95% confident that the true incidence ratio lies in the range 1.5 and 2.9, so that the actual incidence might be as high as nearly three fold or only just 1.5 fold. Alternatively: if 100 studies were conducted in 95 of them the result would lie in the range 1.5 to 2.9 will also attract credit. (1 mark)

Since the 95% confidence interval does not include the value 1 then this result is statistically significant at the $p<0.05$ level and this result is unlikely to have occurred by chance or will have occurred by chance only 1 in 20 times. (1 mark)

Candidates who correctly comment on the 'crude' ratio as being an unadjusted incidence ratio will attract extra credit.

d) Explain what the inverse relationship with gestational age at delivery means and how this might impact on the result for twins compared with singletons (2 marks)

An inverse relationship implies that as one thing gets bigger the other things gets smaller. In this case as gestational age at delivery increases then the risk of SVI/BL falls and vice versa at earlier gestational ages the risk of SVI/BL is higher. Thus infants born at earlier gestational ages are at higher risk of SVI/BL than infants born at later gestational ages. (1 mark)

The inverse relationship with gestational age (as with the general effect of age) might confound the relationship between plurality and the risk of SVI/BL since twins are generally born at earlier gestational ages than singletons and we are also told that the risk of SVI/BL is independently associated with gestational age ie the potential confounder is independently associated with both the exposure (plurality) and outcome (SVI/BL). (1 mark)

e) The results from this analysis were compared with data from a cross-sectional study of children whose visual acuity was measured when they were age 10 years. Explain why the prevalence of SVI/BL as measured in this cross-sectional study may differ from the estimate of cumulative incidence derived from the register-based study (2 marks)

One particular issue with SVI/BL is the fact that we know from the first study that overall 20% of children with SVI/BL died and half of those (ie 10% of the total) died before reaching age 5 years. (1 mark) Thus any incidence measure estimated at age 10 year will under-estimate the incidence by at least 10% and potentially up to 20% (since we are not told by what age the other 10% had died). (1 mark)

Candidates who appropriately discuss issues of: migration and its impact; the fact that the definition of SVI/BL might have changed over time; and the effect of sampling issues in the cross-sectional study will attract credit.

EXAMINERS' COMMENTS

Parts (a) and (b) of this question were anticipated to be challenging, and this proved to be the case, with relatively few candidates providing adequate answers, particularly to 2b. Equally, in 2a many candidates did not gain full marks as they failed to explicitly state what the denominator was.

Nevertheless, the remainder of this question was straightforward and most candidates handled these latter sections much better. A correct answer to (c) was worth 3 marks and thus only 2 further marks were required from the rest of the question for a pass. Part (c) asked for a very basic description of what the terms crude cumulative incidence ratio and a 95% confidence interval mean. Again this dealt with a very fundamental level of knowledge that one would anticipate all public health practitioners to have. In answer to part (e) many candidates, whilst recognising that death of SVI/BL children before age 10 would affect a cross-sectional survey finding at age 10, failed to gain maximum marks by simply misreporting the extent of this effect ie 20% died in total (unclear exactly when), of whom half (i.e. 10% in total) died before age 5.

Question 3

In a named country of your choice, describe how the epidemiology of meningococcal disease informs the prevention and control of the infection.

KEY POINTS

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

- Name of the country
- Description of the epidemiology of meningococcal disease:
    - morbidity
    - mortality
    - age/sex distribution
    - geographic/risk group distribution
    - seasonality/time trend
    - sporadic cases/outbreaks
    - asymptomatic carriage/mode of transmission/incubation period
    - communicability/susceptibility
    - clinical/laboratory diagnosis
    - infectious agent

- Prevention and control measures:
    - notifiable disease
    - vaccination
    - chemoprophylaxis
    - early treatment
    - education and training of the public and professionals
    - publicity and communication with the media
    - national and local policies
    - health service/public health service infrastructure.

The following are additional points that might improve the answer to "good" or "excellent":

- Structured answer, relating epidemiology to specific prevention and control measures
- Mention of relevant national policy/guidance
- Knowledge of relevant research findings
- Mention of the needs of special risk groups: asplenics, travellers
- Knowledge of the added value of enhanced surveillance systems

EXAMINERS' COMMENTS

Both examiners felt that on this question candidates who specified a country other than the UK scored less marks overall. This is a general observation, and we are not sure of the reasons why.

Question 4

Describe the relationship between transport policy and health.

KEY POINTS

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

- Role of transport (to include cycling and walking) as well as motor and air
  Benefits of social cohesion
  Physical activity and reduction of obesity
  Access to services (10 % of marks)

- Health inequalities – often most disadvantaged groups suffer most
  e.g. air pollution, accidents, breaking up communities and social networks with major roads, public transport differences between urban and rural areas, *etc*
  (10% of marks)

- Evidence of effects of policy – seat belt legislation, pricing and use of public transport/cars *etc*, speed and accidents, air pollution and CV disease and respiratory deaths and admissions
  (10% of marks)

- Risk perception – major media coverage of train/air accidents but also risk averse society *eg* walking to school declined *etc* (10% of marks)

- Additional positive contributions of policies – health impact assessments of new transport *eg* airport /motorway runways (10% of marks)

- Wider concerns on the effects of transport policy, pollution and climate change
  (10% of marks)

- Wider concerns (recent report) that whole of infrastructure and planning/ reliance on transport increases risk of obesity (10% of marks)

- Data sources on transport – department of transport, local authorities (accidents, planning and air pollution monitoring) (10% of marks)

- Poor provision of cycle pathways in many areas and those which exist are often badly thought out with regard to their potential users, with preference invariably given to motor vehicles.
  (10% of marks)

The following are additional points that might improve the answer to "good" or "excellent":
(10% of marks)

- Quoting specific experts, *eg* McCarthy & Davies
- Structuring the answer, *eg* type of intervention (engineering, enforcement, education *etc*), health promotion effects (access, recreation, physical activity, *etc*), health damaging effects (injuries, pollution, *etc*)
- Discussion of best investment for health gain

EXAMINERS' COMMENTS

Both examiners felt that the structure of these answers were less variable than for Question 3. The standard of answers was also less variable: candidates were therefore less likely to score very high or very low marks.

Question 5

a) Why is information about health and work of importance to a public health practitioner? (50% of marks)

b) Describe the routine sources of information about health at work in a named country and the limitations of such data. (50% of marks)

KEY POINTS

(Points marked * may indicate an excellent answer, if they are in addition to the basic answers)

<u>a) importance</u>

- Work (and unemployment) and health are closely linked, both generally and specifically. Work influences health: exposures to various classes of hazardous substances / conditions; and health influences work.
- Employment is one determinant of socio-economic status, which in turn is closely related to health.
- There are also specific hazards that relate to specific industries, occupations, or groups of occupations (e.g. asbestos exposure and mesothelioma). These hazards may be associated with the work process itself and/or the work environment.
- The workplace also provides an opportunity for promotion of health, both specifically (e.g. minimising specific occupational hazards) and generally (e.g. smoking, healthy eating, exercise).
- Workers suffering from certain conditions and who are employed in certain occupations pose a more general public health risk (e.g. typhoid carriers in food industry).
- * Mention of the results of certain specific studies linking occupation, grade of job and type of work to health outcomes.

<u>b) routine sources and limitations – named country</u>

- Denominator data:
  - o Demographic information – age, sex, proportion employed by age and sex, type of employment by age and sex.
  - o In the UK, this would be derived from the Census, General Household Survey and National Labour Force Survey. In Hong Kong, the first two sources apply.
  - o Other sources of data include trade union statistics. Problems: not comprehensive, not representative.
  - o * Official statistics would only include officially employed individuals. Those in unpaid employment, or working unofficially (and who might be at high risk) would not be included. To get information about these groups a special survey would be required.

- Numerator data:
  - o * Mortality data linked to occupation.
    - In the UK, occupation is registered at time of death. However this doesn't reflect occupation throughout life. The unreliability of occupation coding, e.g. in relation to the more fluid job market and the particular issues about students and the armed forces. (Same system in Hong Kong re mortality statistics; Problem of "last occupation", or "end-stage occupations", not representative of exposure to risk factors sustained in earlier occupations.)
    - The ONS Longitudinal Survey in the UK is a cohort study examining risks of death by occupation derived from the Census. (No such survey in Hong Kong)
  - o Routine morbidity data linked to occupation
    - GP records, hospital records, referrals to specialist services. (Problems in Hong Kong: poorly recorded hospital data; often occupations not recorded – the field often left blank. Data on occupation usually absent in GP records.)
    - * Overall days off work are a correlate of mental health of the workforce
    - * Most countries don't have routinely available morbidity data linked to the patients' occupation.
  - o Employment related data

- Sickness absence records. Sickness absence is not a good indicator of ill health. People may feign illness, people may fear to admit to illness for fear of losing their job, and people may be willing to put up with ill health in order to keep a job, particularly seasonal or casual workers.
- Pre-employment medical assessments.
- Ill-health retirement, and compensated illness pensions. Some employers would have access to these, but not all.

* In general, most of these statistics are not too helpful, because they are either not specific enough, are not comprehensive (hence are usually biased), or are simply not a good index of work and ill-health (not valid).

- Specific employment risks and hazards in the workplace.
  - * Information about most work-related hazards (e.g. stress) would not be available routinely, and would require special surveys.
  - Injuries at work and days off work (e.g. for back pain and repetitive strain injury) can be indicative of attention to workplace safety and ergonomics.
  - * Occupational diseases and injuries registers. (In Hong Kong, pneumoconiosis (silicosis and asbestosis) register and register for noise induced hearing loss – both for compensation purposes. Occupational injuries register – also a statutory reporting system.)
  - * Compensation payment registers. Registers of injuries that are compensateable are usually more reliable.
  - Incidents of release of toxic or hazardous substances may be recorded (e.g. in the UK by the Health Protection Agency). These statistics are often incomplete. Many diseases are not reported. (Same statutory notification in Hong Kong; same problem with incomplete reporting.)
  - * In the UK, responsibility for health and safety in the workplace falls under the Health and Safety Executive. (In Hong Kong, this is under the Labour Department).
  - Specific hazards that relate to specific industries, e.g. injuries within the off-shore shipping industry

* Incomplete reporting and coverage affects both numerator (occupational diseases and work-related illnesses) and denominator data (the size and demographic data of the specific occupation).

EXAMINERS' COMMENTS

Most of the candidates were able to answer 5a but in a very unstructured manner.

Good candidates used the 3 functions/roles of public health practitioners i.e. health protection, promotion and healthcare and used it as a structure to see what is then relevant to 'health & work' accordingly.

Question 6

a) Define infant mortality rate, perinatal mortality rate and neonatal mortality rate (30% of marks)
b) Comment on the interpretation of differences in these measures within and between populations. (30% of marks)
c) Outline the main influences on perinatal mortality rate in a named country (20% of marks).
d) Outline the main influences on infant mortality rate in the **same** named country (20% of marks).

KEY POINTS
(Points marked * may indicate an excellent answer, if they are in addition to the basic answers)

Definitions
i) Infant mortality rate is the number of deaths from birth to the end of the first year of life per 1,000 live births in that year.

* There is variation in the registration practices for very premature infants between countries; some register very premature births as live births even if there is little chance of sustained life.

ii) Perinatal mortality rate is the number of stillbirths and deaths in the first week of life (or days 0-6) per 1,000 live and still births.

* There is variation in the definition of a stillbirth between countries - some define this as a minimum gestation period (e.g. 24 weeks) or minimum foetal weight (e.g. 1000g).
* The WHO excludes stillbirths from the denominator

iii) Neonatal mortality rate is the number of deaths in the first 28 days of life (or days 0-27) per 1,000 live births.

* There is variation in the registration practices for very premature infants between countries; some register very premature births as live births even if there is little chance of sustained life.

Note that there are differences between the UK (ONS) definition and the WHO definition of perinatal mortality:

Numerator for the ONS definition of perinatal mortality is stillbirths + live births who die in first week (0-6 days)
Denominator is all live births + stillbirths

Numerator for the WHO definition of perinatal mortality is stillbirths + live births who die in the first week (0-6 days)
Denominator is just live births

Note the denominator for the WHO rate does not include stillbirths although they are included in the numerator

Further complications:

1. However the further complication is that in other places in the world eg Australia, perinatal mortality is stillbirths plus live births who die in the first month (ie neonatal deaths), not just the first week.

2. Also, although this won't necessarily come into the answer relating to the question the students have been asked, is that the definition of stillbirths is different in different places, ie the threshold between a spontaneous fetal loss (miscarriage) and stillbirth is different in different places. We use 24 weeks (we used to use 28 weeks), the WHO definition uses 22 weeks, Australia uses 20 weeks.

Interpretation

- Infant mortality is a good indicator of the health of the population generally.
- Variations within countries can be explained by ethnicity, relative affluence, smoking rates, age of first childbirth and educational attainment of women.
- As neonatal and perinatal death rates decrease a greater proportion will be due to congenital malformations, most of which can be viewed as unavoidable. Note, whilst most congenital malformations can be considered unavoidable some change can be expected with screening services in pregnancy and those reduced by interventions such as rubella vaccination or folic acid supplementation. Equally, some increase is likely with changing social patterns of pregnancy – e.g. trend towards delaying pregnancy to older ages.
- Variations between countries could be explained by differences in the definitions used – care needs to be taken when comparisons are made.

* Infant mortality is a hard measure of inequality within societies. There are pockets of society in rich countries where rates are still unacceptably high.
* When political and economic conditions change even in a well off country, the IMR is often among the first indicators to show the adverse effects on population health (e.g. the experience of Russia following the break up of the old Soviet Union).

Influences

- Social-economic status and conditions:
  Population social-economic factors e.g. housing, education, economic well-being, clean water, environmental hygiene
  Cultural factors e.g. status of women in the community
- Maternal factors:
  Age of mother, maternal height, parity, nutritional status, education, smoking, alcohol, additional morbidity (e.g. diabetes, hypertension, renal disease, infections).
- Neonate factors:
  Birth weight, multiple birth, congenital anomaly, sleeping position, feeding method
- Antenatal care – influences perinatal mortality
  Quality, availability and access to services, including interpreters
  Use and effectiveness of antenatal screening and subsequent abortion
- Care in labour and neonatal care – influences all measures
  Quality, availability and access to services, including intensive case
- Infant and child health services – influences infant mortality
  Quality, availability and access to services
  Immunisation
  Nutrition

EXAMINERS' COMMENTS

Many candidates failed to give the correct definitions of perinatal, neo-natal and infant mortality rates - this should have given candidates an easy 30% of marks.

A number of candidates answered questions not asked for - e.g. the interpretation of neo-natal mortality rate was not asked for but a surprising number of candidates gave an answer - and got no marks for it.

Too many answers were very general and poorly structured, some candidates gave very clinical answers and ignored public health considerations in their entirety.

Paper IB

Question 7

Define social capital and outline its relationship to health (50% of marks). Discuss the proposed underlying mechanisms for this relationship (50% of marks).

Key points:

This question has 2 elements which should be clearly addressed :
1. Define social capital and outline its relationship to health.
2. Discuss some of the proposed underlying mechanisms.

Definitions
<u>Social capital</u>
Social capital refers to connections among individuals.
Two key features :
  Structural – social networks [i.e the actual structure of relationships],
  Cognitive – norms of trust and reciprocity which are embedded within them.

Therefore it refers to features of social structure and organisation such as networks, norms of reciprocity and mutual aid, and interpersonal / social trust that act as resources for individuals and facilitate coordination and cooperation for mutual benefit (Coleman, Putnam).
It has two aspects: organisations and values/relationships.

<u>Basic premise-</u>
Social interaction enables people to build communities. It gives them a sense of belonging.
Trust between individuals becomes trust between strangers.
This develops a shared set of values, virtues and expectations
Requires face to face encounters.

Social capital describes the resources available to individuals through their membership of community networks. In contrast to financial capital, which resides in people's bank accounts, or human capital, which is embodied in individuals' investment in education and job training, social capital is inherent in the structure and quality of social relationships between individuals (Kawachi).

Social capital is "is the glue that holds societies together- what social cohesion depends on."

Dimensions of social capital
- *Bonding social capital.*

Good for supporting a sense of self identity, reciprocity and mobilizing solidarity – 'Superglue'
- *Bridging social capital*

Good for links to external groups or assets and access to information, can broaden peoples' identities – 'Oil'

<u>Social cohesion</u> is a broader concept (of which social capital is a subset) that includes;
- The absence of latent conflict (e.g. from income inequality, racial or political tension, or other polarising forces)
- The presence of strong social bonds (e.g. social capital; civil society (bridging social divisions) and institutions of conflict management (judiciary, democracy etc.)

Both social cohesion and social capital are collective dimensions of society and should be distinguished from the concepts of social networks and social supports that are properties of individuals.

There are diverging views on the definition of community cohesion, but there is general agreement about the features a cohesive community should display.
A cohesive community is one where:
• there is a common vision and a sense of belonging for all communities;
• the diversity of people's different backgrounds and circumstances are appreciatedand positively valued;
• those from different backgrounds have similar life opportunities; and
• strong and positive relationships are being developed between people from different backgrounds in the workplace, in schools and within neighbourhoods.6
Any cohesive society should also demonstrate the ability to integrate people from different ethnic backgrounds so that they can relate together in terms of where they live, their education, employment, and social/recreation spheres.

Social capital and health

Higher levels of social capital (measured differently, in different populations and settings) have been linked with:

- lower mortality
- better perceived health (although some studies show this for women but not for men)
- lower homicide rates.
- Lower suicide rates.
- Fewer behavioural and emotional problems in childhood
- Lower crime rates.
- Lower psychiatric morbidity
- Lower levels of psychosis
- Less binge drinking
- Better quality housing to mentally ill people.
- Improved Child development at family, school, peer group and community level.
- Public spaces are cleaner,
- Benefits of Economic prosperity
- Better health and well being – 50% reduction in risk of dying over next year!

However, *not all* measures / dimensions of social capital show an association with health indicators / outcomes.

In addition, there may be adverse impacts of high social capital on health, possibly mediated by intolerance of deviation. For example:

- Higher hospital admission rates for psychosis in areas with high perceived community safety
- Higher levels of psychosis, suicide and psychiatric hospital admission among ethnic minority groups living in areas with lower proportions of ethnic minorities. Hypothesis that social capital diminishes within the group as it becomes a smaller proportion of the population.
- Some measures of social capital, such as a high level of local identity, were associated with poorer health.

There is some evidence of a general decline in social capital. This could be caused by increasing pressure on time; increasing number of women in the paid labour force; greater sub-urbanisation and time spent commuting, and time spent watching television.

Increasing disparities between socio-economic groups in either the quantity or nature of social capital may explain increasing health inequalities.

An excellent answer will include a discussion of how social capial can be measured.

Mechanisms linking social capital and health

This part of the question should reflect the contribution of social capital to a healthier community and suggest how this might work

If the mechanisms include:

- Allowing people to resolve collective problems more easily
- Power to ensure compliance with collectively desirable behaviour
- Where people are trusting and trustworthy, everyday business and social transactions are less costly.
- People with active and trusting connections maintain character traits that are good for the rest of society

Then Social capital may affect health directly through:

1. Health behaviours
Greater diffusion of knowledge; healthy norms, social control of deviant behaviour.

2. Access to services
Cohesive neighbourhoods more able to resist closure of services or transport links; existence of pressure groups; 'appropriate' social organisations for secondary uses.or may be able to work together to ensure that remoted services can be reached by providing support for more vulnerable members of the community

3. Psychosocial processes
Social capital as source of affective support, self esteem and mutual respect. (Socially isolated individuals in cohesive communities have better health than those living in less cohesive communities.) Children growing up in areas with high social capital learn through witness /experience of public responsibility among people with no personal ties.

4 Creating more effective partnerships to provide services a. engaging patiants in their own careb. promoting interest in /action to address environmental conditions

Consider ecological fallacy - associations between low social capital and poor health may be mediated by individual-level relationships, such as social isolation and health (as opposed to societal level relationships).

An excellent answer will include a discussion of competing interpretations of how social capital affects health, e.g. neomaterialists (Davey Smith, Lynch et al) versus psychobehavioural (Kawachi, Putnam et al).

EXAMINERS' COMMENTS

Candidates performed better on 7a. However, there was a lot of confusion with many candidates focusing on social networks.

7b in general was poorly answered. Many candidates did not address the question and explored issues such as the relationship between deprivation and health.

Question 8

Describe the process for allocating resources in a health care system with which you are familiar (40% of marks). Critically appraise the extent to which this process matches service provision to healthcare needs (60% of marks).

Key Points

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

* Recognition that there are different levels of allocating e.g. between districts and within districts
* Some historical perspective of RAWP or an analogous system
* Description of the basic elements of the formula e.g. age weights, SMR, social variables, market forces factor (in England).†
* Critical discussion on the main issues which have dominated debate mentioning several of:

  * debate about age weights
  * use of SMR as a proxy for morbidity
  * size of populations and reliability of formula
  * use of data from decennial census
  * weighting of different components
  * use of regression methods to approximate need.

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

Two or more of the following:

* knowledge of local approaches to sub-district allocation
* recognition that the method of movement of districts to target is as important as the target itself
* discussion of the underlying assumption that equity of inputs is an appropriate model
* differences within the UK

† e.g. Scottish perspective:

- Arbuthnott Formula (published 2000) used since 2001/02 to calculate central funds for Hospital and Community Health Services (HCHS) and GP Prescribing to 15 and now 14 Health Boards in Scotland.
- Formula uses population and three further separate adjustments (age-sex profile, morbidity and life circumstances, and excess cost due to remoteness).
- Additional point would be that the Formula replaced the SHARE (Scottish Health authorities Revenue Equalisation) formula.
- Delivering Fair Shares for Health in Scotland (the Report of the NHSScotland Resource Allocation Committee - NRAC) published in Sept 2007: improves and refines Arbuthnott Formula; argues for new formula to use smaller geographical areas within Health Boards to better assess needs and for planning services, use timely population projections, take better account of higher relative need of elderly and very young, more accurately reflect increased need for healthcare in deprived areas, compensate for underutilisation of health services in deprived areas, and take better account of costs in providing services in different urban-rural areas.

EXAMINERS' COMMENTS

Many candidates failed to consider the whole health system but focused on specific parts (eg. allocation of resources at a PCT level).

Question 8b – many candidates failed to critically appraise approaches to resource allocation getting sidetracked on programme budgeting, economic appraisal and health needs assessment.

Question 9

Describe the circumstances in which conflict can occur within a project team (40% of marks) and the techniques that can be used to resolve this conflict (60% of marks).

Key Points

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

A project team typically consists of a variety of members with different knowledge, skills and positions; mutually respecting heterogeneity is preferable to competitive homogeneity.

The circumstances of conflict should be classified. An example is the Karen Jehn's typology of conflict:
(1) Task content conflict – refers to disagreements about the actual task being performed by members.
(2) Emotional conflict – which involves negative emotions and dislike of the other people involved in the conflict.
(3) Administrative conflict – refers to disagreements regarding the manner in which a goal should be reached.

Conflict in a team can occur at different levels :
1. Intrapersonal - the internal, mental struggle to select alternatives
2. Interpersonal - differences between individuals
3. Intragroup - differences between members of a group pursuing a similar goal
4. Intergroup - differences between groups with competing goals

Conflict management techniques :

Based on the Dual Concern Model by Kenneth Thomas, the conflict management techniques reflect a concern for both an individual's own and the opponent's outcome. Four techniques are described:
- accommodation
- pressing
- avoiding
- negotiation (incorporating compromise and collaboration) .

The principles in negotiation can be applied in project team management:
(1) Separate the people from the problem.
- Discussion should focus on issues, not personalities.
(2) Focus on interests, not positions.
- Behind opposed positions often lie compatible interests.
(3) Invent options
- Decision should aim at collaborations for the best possible solution for the team
(4) Use objective criteria.
- Consensus building process should focus on current, factual information, with a sense of fairness and equity.

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

Addressing that conflict can have both positive and negative effects on the team
- Conflict between personalities can have negative effects, while disagreements over task content may enhance team development.
- A lack of conflict can cause groupthink (phenomenon in which team members strive towards consensus at the expense of the right decision).

A conflict can be addressed not only as a short-term problem, but an opportunity for the team to learn how future changes can be handled more effectively in the long term.

Additional key points about team dynamics :

Team Formation – The Model of Group Development by Bruce Tuckman

- forming
- storming
- norming
- performing
- followed by adjourning and transforming, possibly

Roles within the Team as defined by Belbin:

- plant
- shaper
- team worker
- implementer
- completer finisher
- resource investigator
- monitor/evaluator

EXAMINERS' COMMENTS

Question 9 was approached with a wide range of perspectives, and related theories on motivation, communication, leadership and team building theories were credited. However, undue focus on "conflict prevention strategies" without tackling "conflict resolution techniques" was considered insufficient.

Question 10

Discuss the features of a hospital or other healthcare organisation that may contribute towards the detection of clinical errors within that organisation and minimise their impact.

Key Points

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

Detection of errors should be put into the context of clinical risk management, which aims to improve the quality of health care and create and maintain safe systems of care.

The purpose, structure and setting of the clinical or healthcare agency need to be clearly set out e.g. primary care trust, acute hospital or Government health department etc.

It is essential that the candidate indicates that the agency should have in place a written protocol/set of procedures covering the management of clinical errors.

The candidate should specify that the protocol/procedure should have the following features:

- **Strategic Commitment :** Approval, ownership and review process signed off by the agency's accountable officers i.e. Chief Executive and Board of Directors.
- **Agreed case definition** : An agency-wide agreed case definition and interpretation of the term "clinical error" which includes the types of clinical error (diagnostic, treatment, preventative, other – equipment failure, system failure, communication failure.
- **Accountability** : clear accountability mechanism for the reporting of clinical errors incorporating specific examples of good practice e.g. near miss approach, audit processes against standards, blame free culture for reporting.
- **Detection Mechanisms** : clear detection/reporting arrangements which ensure that appropriate clinical onfidentiality standards are maintained; monitoring by individuals, teams and organisations
- **Compliance** with:
    - recognised legal requirements (e.g.data protection legislation)
    - demonstrable evidence of agency compliance with their clinical error policy e.g. external report to the Board etc.
- **Integration within corporate governance** : clinical error management system developed within the corporate governance framework of the agency.
- **Organisational Culture**; no blame culture where excellence can flourish
- **Error reduction strategies** : systematic process for the implementation of lessons identified in clinical error management, acknowledgement of the potential for error and building in error reduction strategies at every stage of clinical practice.
    - Examples of specific methods:
        - provision guidelines, integrated care pathways, protocols and policies – including a whistle-blowing policy, redesign of services
        - within protocols acknowledge importance of workload, work patterns e.g. circadian rhythm and clinical supervision in relation to local patterns of clinical errors
        - audit, clinical governance procedures, risk assessment, complaints monitoring

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

- Addressing barriers to creating systems of safe care. They may be strategic, cultural, structural and technical
- reference to guidance on best practice locally or nationally
- reductions in the complexity of tasks

- optimise information processing by the use of protocols
- dissemination of lessons identified from error management
- explicit and sufficient financing of the corporate error management system
- Defining the nature of clinical errors :
    - The human factors approach focuses on the human component within complex sociotechnical systems.
    - Conceptual distinction between 'active errors', manifest within individual practice, and 'latent factors', located 'upstream' within the work setting.
    - Contributive factors to adverse events may be related to the patient attributes, task characteristics, individual background of the involved health personnel, team factors, work environment, organisational and managerial factors.
- system for analysis of adverse events e.g. Reason's model
- The following factors contribute to adverse events and should be considered in the analysis of adverse events:
    - *Patient factors*: Age, general health, complexity of the condition, language, social problems and personality.
    - *Task factors*: Task design, availability of protocols, test results, etc.
    - *Individual factors*: Training, physical and mental health, and awareness.
    - *Team factors*: Verbal communication, supervision, seeking help and team structure.
    - *Work environment*: Staffing levels, skill mix, shift patterns, equipment and administrative support.
    - *Organisational and managerial factors*: Financial resources, policy standards, goals and the safety culture.

EXAMINERS' COMMENTS

Most candidates demonstrated some knowledge on clinical errors detection, and the comprehensive and structured answers attracted higher marks.

Paper IIA

You are a member of a health protection advisory group with responsibility for communicable disease control in both hospital and community settings for a population of 250,000 people. The group has asked you to critically evaluate the following BMJ paper as part of a reassessment of local policies to prevent the spread of major epidemic viral infections:

Tom Jefferson, Ruth Foxlee, Chris Del Mar, Liz Dooley, Eliana Ferroni, Bill Hewak, Adi Prabhala, Sree Nair, and Alex Rivetti. *Physical interventions to interrupt or reduce the spread of respiratory viruses: systematic review.* BMJ, Jan 2008; 336: 77 - 80 ; doi:10.1136/bmj.39393.510347.BE

Candidates are supplied with the shorter version of this paper, as published in print in the BMJ, plus tables 5 and 6 from the longer electronic version. These two tables provide more detail about the case-control studies which are included in the figure on page 78 of the printed version, and about the results of meta-analyses based on these studies.

1. Write a critical appraisal of this paper, with particular reference to the findings from case-control studies. (40% of marks)
2. The authors "assigned risk of bias categories on the basis of the number of items judged inadequate in each study". Discuss the principal sources of bias that might affect the conclusions of case-control studies in this review. (15% of marks)
3. The authors "calculated numbers needed to treat (NNT) whenever … the data were robust enough to allow it." Explain what NNT means in this context (5% of marks), and describe how it could be estimated from case-control data (10% of marks).
4. Write a briefing paper of 500 words for consideration by the planning group. This paper should critically outline possible recommendations for local infection policies, explaining how these are informed by evidence in this paper.(30% of marks)

Key Points

1. This is a systematic review of the literature coordinated by one of the Cochrane Collaboration centres but was not commissioned.

The list of sources is provided along with a description of the search terms (provided online) to permit a repetition of the review. All languages were accepted, an important issue given the source of much of the data, ie relating to SARS.

Papers were selected first solely from the titles/abstracts if 'they seemed to meet our eligibility criteria' .. a rather vague and subjective concept without any clear statement of what the criteria were. These should have been clearly articulated. Selection from title/abstract can lead to the dismissal of relevant papers.

The full paper was then re-assessed. The quality of the randomised papers were then assessed against what the design requirement are for a good quality study which seeks avoids any confounding or bias, through randomisation and blinding. Non-randomised studies were then assessed using recognised scoring systems designed to address key issues that affect the validity of results from such non-randomised studies.

There is insufficient detail in this paper to know the context of the trials.

Studies were categorised into different epidemiology types with a view to meta analysis which could improve the accuracy of estimates by increasing the numbers. The authors do not describe how they did this (eg by using the paper's title or by assessing the study designs themselves) but did suggest that some authors of papers were in their view confused about the design. Two randomised trial was dismissed because of a failure of blinding because the placebo tissues 'stung the user's nose', a problem that would only have been real if the participants were aware of the significance of the stinging (did they think it signified an active ingredient and could unblinding have affected the measured outcome which is unlikely if the outcome were a laboratory confirmed diagnosis).

The authors found a considerable variation in quality amongst the papers along with designs so different as to precluded meta analysis even within epidemiological study categories. The exception were a series of case control all concerning the infection of healthcare workers in outbreaks of SARS. These 6 studies were deemed homogenous enough to permit meta analysis with pooled data sufficiently robust to permit detailed quantitative analyses

The description of the case-control papers as in table 5 is poor: in particular only those infected are described, not the source of controls. Tables 5 & 6 appear inconsistent: only in study w48 are N95 masks mentioned as an intervention in table 5 and in the text, whereas table 6 gives pooled estimates for N95 masks based on studies w45 and w47. Table 6 and Fig 2 are also inconsistent eg for wearing gloves.

Each of the 6 studies is weighted by combination of size and quality, though exactly how is not explained with the relationship between the size of the study and its weighting differing between the various intervention (Fig 2). The calculated odds ratio is ratio of the odds of having used the intervention if infected vs if not infected.

No description is given of the analysis: crude odds rations are calculated (candidates should check at least one figure for the odds ratio) but no attempt appears to have been made to calculate each independent effect given that any number of individual interventions may have been used by a participant. The benefits of any individual intervention may therefore be quite different to that calculated.

2. Sources of bias:

There should be a clear explanation of the importance of bias in observational, non-randomised studies. Simply listing sources of bias without relating them to this paper will not attract marks.

For each of the following types of bias, some discussion is expected of how it might have occurred and explanation of effects on results:
- Selection bias
- Information bias (systematic) leading to exaggeration or attenuation of the association
- Misclassification bias (non-systematic) leading to attentuation of the odds ratio
- Confounding

Detailed descriptions within these categories (eg recall bias, reporting bias, non-responder bias). Better answers will make reference to a type of bias that has NO impact in this study.

3. NNT in this situation should be the number of people needed to use the intervention to avoid one becoming infected. To calculate this it is necessary to know the infection rate (absolute risk of infection) without the use of the intervention and the infection rate when the intervention is used. The NNT is the inverse of the difference in rate (risk) ie 1/ARR (absolute risk reduction). None of this data is available so NNT could not have been calculated directly.

The authors do not state how NNTs were calculated. However, the absolute risk reduction can be estimated indirectly from the relative risk provided that the baseline incidence (infection rate) in the unexposed (unprotected) group is estimated from other sources. If a case-control study is nested within a cohort study, this information should be available more directly.

4. Briefing paper.

Candidates should note the need for different advice for avoidance of infection in an infectious disease ward setting and a community setting. They should also distinguish between dealing with an outbreak of a disease such as SARS with small numbers and low levels of circulation/low infectivity in the community, and a major epidemic (or pandemic) when levels of circulating virus in the community and infectivity may be high and a large proportion of the population are infected at any one time. They should also note that the means of spread may differ between diseases.
They should note the issue of practicality and therefore compliance with any intervention. They should also note the potential costs of an intervention.

All recommendations must be accompanied by citing the specific evidence as mentioned in the paper, specifically making the link, discussing the strengths and weaknesses so as to justify the recommendation.

Mention of benefits to a community after the epidemic has ceased of measures designed to improve general hygiene will be given credit.

EXAMINERS' COMMENTS

Answers were particularly disappointing this sitting with no candidates producing consistently good answers to all 4 questions, nor any outstandingly good answers to any one question.

a) Candidates tended to provide little more than a description of the paper in the critical appraisal simply repeating what the authors had written. Indeed it appeared that the critical faculties of the candidates had been dazzled by the provenance of a Cochrane Collaborative Centre.

Many candidates made the mistake of simply writing down as much as they could, for instance many references to blinding in relation to the case control study which appeared in several of the intervention analyses or bias due to socio-economic factors. They also called for a funnel plot despite a maximum of 6 studies, made no attempt to suggest what represented the size of the study for one of the axes (which n or N?) or attempted to draw one using say the weighting as that value. Many thought funnel plots had to show 'negative' results, rather than symmetrical plots about the central estimate, all of which could be 'positive'.

Candidates did not question how the summary ORs had been calculated: a look at the tables would have shown that the relative weighting between studies varied across interventions. Few questioned if or how the various interventions had been disentangled (by multivariate analysis).

Candidates failed to appreciate the difference, both mathematically and epidemiologically, between odds ratios and relative risk, and therefore failed to question what the authors were claiming (which in effect was that what they had calculated were RRs rather than ORs).

b) Bias: candidates mostly failed to answer in the context of this paper, giving instead a Paper I answer. Publication bias for instance was cited, yet if these standard infection control methods were found to be ineffective (a negative result) THAT would be worth publishing. Recall bias: few questioned the impact of this being a protective 'exposure' on recall or the fact that many participants were health care worker who would be expected professionally to use these measures, thus effecting recall.

c) NNTs: most but not all candidates managed to get the easy 5% by describing NNTs in this context. But VERY few noted that without a RR and AR it is impossible to calculate. Many tried to get half way there by citing the usual view that OR=RR if the disease is rare (some wrote 'common'), but failed to appreciate that its just about impossible to achieve NNTs so low in a rare disease. Even then the absolute risk is unknown in a case control study. The authors even invented a new outcome which was treatment effectiveness. There is some rationality if this were based on relative risk but it is meaningless with odds ratios. Furthermore there was no consistent relationship between this 'effectiveness' and NNTs as there should have been.

d) Candidates failed to read the question and write an appropriate answer for question 4 which required a briefing paper to a group that was already well informed. Too many brought MRSA and C difficile into a question on viral respiratory infections. Having mentioned generalisability, candidates failed to use the opportunity to expand in Q4. Many candidates failed to distinguish between public health measures and infection control. Will measures in a outbreak/small epidemic work in a pandemic? Many mentioned cost effectiveness and went along with the view that these are cheap interventions - when millions of items are used each day. What of the logistics of storage, supply and disposal during a pandemic when so many will be ill? Infection control beds would fill within days in a pandemic, ITUs within a couple of weeks and ALL hospital beds soon after that, so where will nursing of patients and infection control take place and by whom? Even in everyday use just 3 minutes per patient on hand washing adds up to 1 hour on a 20 bed ward and nurses might be scarce in a pandemic. What of the day to day benefits of public hygiene outside of outbreaks, epidemics and pandemics, which many candidates conflated?

Paper IIB

You are the cervical screening co-ordinator for Area A. Table 1 shows recently published statistics on the grades of cervical smears examined at the cervical cytology laboratory serving Area A compared with those for the laboratory serving the neighbouring Area B. Area B's cytology laboratory introduced liquid-based cytology (LBC) for processing cervical smears some years ago. Your laboratory (serving Area A) is still using conventional glass slide preparation methods.

Table 1: Comparison of Cervical Smear Results Between Areas A & B

			Grade of Satisfactory Cervical Smear Result					
	Total No.	**Unsatis-factory smears**	**Negative**	**Borderline**	**Mild dyskaryosis**	**Moderate dyskaryosis**	**Severe dyskaryosis**
Area A's Lab	6535	227	5808	249	155	64	32
Area B's Lab	20690	329	18806	592	726	105	132

Q1. Summarise and comment on the information shown in table 1. What factors might explain the findings? (20% of marks)

Q2. Calculate a measure of the precision of the difference in the rate of unsatisfactory smears between Area A and Area B and comment on the result. (20% of marks)

Results of a recently published randomised controlled trial of the accuracy of liquid-based cytology (LBC) versus conventional glass slide preparation cytology showed that the percentage of unsatisfactory smears was significantly reduced from 4.1% in the group randomised to the conventional cytology group compared to only 2.2% in the LBC group.

In this trial, 22,708 women were randomised to LBC and 22,466 to conventional glass slide preparation cytology. Women with mild dyskaryosis or worse on cervical screening were followed up with colposcopy in order to obtain a histological end point. In total 1337 women in the group randomised to LBC underwent colposcopy. The corresponding number in the group randomised to conventional glass slide preparation cytology was 661. Tables 2 and 3 show data from this trial comparing the histological findings in women referred to colposcopy with mildly dyskaryotic smears or worse for those undergoing LBC or conventional cytology. CIN denotes cervical intraepithelial neoplasia.

Table 2: Comparison of Cervical Cytology Findings and Histological End Points for Liquid-based versus Conventional Cervical Cytology for Women with Mildly Dyskaryotic Smears or Worse			
	Histological End Point Found at Colposcopy		
	No. with Grade 1 CIN or worse	No. with Grade 2 CIN or worse	No. with Grade 3 CIN or worse
Liquid-based cytology	313	99	45
Conventional cytology	184	84	53

Table 3: Detection Rate, Relative Sensitivity and Positive Predictive Value of Liquid-based versus Conventional Cervical Cytology for Women with Mildly Dyskaryotic Smears or Worse			
	Histological End Point Found at Colposcopy		
	Grade 1 CIN or worse	Grade 2 CIN or worse	Grade 3 CIN or worse
Detection Rate (%)			
Liquid-based cytology	1.38	0.44	0.20
Conventional cytology	0.82	0.37	0.24
Relative Sensitivity (95% C.I.)			
Liquid-based vs. conventional cytology	1.68 (1.40-2.02)	1.17 (0.87-1.56)	0.84 (0.56-1.25)
Positive Predictive Value (%)			
Liquid-based cytology	23.4	7.4	3.4
conventional cytology	27.8	12.7	8.0
Relative Positive Predictive Value (95% C.I.)			
Liquid-based vs. conventional cytology	0.84 (0.72-0.98)	0.58 (0.44-0.77)	0.42 (0.29-0.62)

Q3. Explain how the four measures in table 3 have been calculated from the information in table 2 and associated text above. Comment on these values. (20% of marks)

Q4. Explain why it is not possible from the information provided to calculate an absolute value for the sensitivity of LBC or conventional cervical cytology. What further information would you need to enable calculation of absolute sensitivities? (10% of marks)

Q5. Using your interpretation of the data in tables 1-3, write a briefing paper of up to 500 words for your director of public health/chief medical officer on the potential benefits and disbenefits of introducing liquid-based cervical cytology in your area, and the practical issues that need to be considered in implementing such changes. (30% of marks)

Key Points

Q1. Summarise and comment on the information shown in table 1. What factors might explain the findings? [20% of marks]

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

- Convert numbers to rates using the total number of smears taken as the denominator to show that the unsatisfactory smear rate for Area A is 3.5% compared with only 1.6% for Area B.

- There are also differences in the rate of reporting of borderline (3.9% vs. 2.9%), mildly (2.5% vs. 3.6%), moderately (1% vs. 0.5%) or severely dyskaryotic (0.5% vs. 0.6%) smears using the number of satisfactory smears as denominator. Use of the total number of smears taken as denominator would be acceptable, and would give very slightly different rates.

- These medically important differences may be because Area B is using LBC.

The following are additional points which might improve the answer to "good" or "excellent".

- Other explanations for the differences in grade of smears between Area A and B could be:
    - differences in the age structure or sexual behaviour of the target populations in the two areas;
    - differences in the level of training, experience or expertise of smear takers or smear examiners;
    - differences in laboratory reporting criteria.

- Better candidates would refer to tests of statistical significance e.g. a chi-squared test could be performed on a 5 x 2 table with 4 degrees of freedom to test for overall differences in the proportions of satisfactory smears in various categories.

- Better still candidates would refer to a chi-squared test of trend for grades of satisfactory smear since there is an intrinsic ordering of the categories of smear (normal through to severely dyskaryotic).

Q2. Calculate a measure of the precision of the difference in the rate of unsatisfactory smears between Area A and Area B and comment on the result. [20% of marks]

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

A suitable statistical test would be a parametric test i.e. comparison of two independent proportions p_1 and p_2. In this case p_1 = 329/20690 and p_2 = 227/6535. The stages are as follows:

1. derive difference in sample proportions $p_1 - p_2 = 0.019$ or 1.9%
2. derive standard error of the difference in proportions:
   $S.E\ (diff) = \sqrt{[(p_1(1-p_1)/n_1) + p_2(1-p_2)/n_2]} = 2.43 \times 10^{-3}$ or 0.24%
3. 95% confidence interval = p1-p2 +/- (1.96 X S.E. diff)
   $= 0.019 \pm 4.8 \times 10^{-3}$
   = 0.014 – 0.024 or 1.4% - 2.4%

In this case the 95% confidence interval for the difference does not overlap 0, which means that the difference in unsatisfactory smear rates between Areas A and B is statistically significant at the 5% level, with a 95% chance that the population difference lies between 1.4% and 2.4%.

Q3. Explain how the four measures in table 3 have been calculated from the information in table 2 and associated text above. Comment on these values. [20% of marks]

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

- Convert numbers in table 2 to rates using 22,708 as the denominator for the liquid-based cytology trial arm and 22,466 for the conventional cytology arm to give detection rates in table 3.

  - The positive predictive value (PPV) is defined as the percentage of women with abnormal cervical cytology who also have abnormal histology according to the reference standard of colposcopy, as shown in table A below.

Table A : Comparison of a Screening Test with a Reference Standard Test		
	Histology Abnormal	Histology Normal
Cytology Abnormal	a	b
Cytology Normal	c	d

PPV = a/(a+b)

  - In this study the positive predictive value is calculated by dividing the number of women with each grade of histological abnormality by the number undergoing colposcopy in each arm of the trial; 1337 in the LBC arm and 661 in the conventional cytology arm. Table B shows the data for CIN1 or worse.

Table B: CIN1 or Worse at Colposcopy for Women with Mildly Dyskaryotic Smears or Worse			
	= or > CIN1	No CIN	Total
Liquid-based cytology	313	1024	1337
Conventional cytology	184	477	661

  - For grade 1 CIN or worse:

    - %PPV (LBC) = (313/1337) x 100 = 23.4%

    - %PPV (Conventional) = (184/661) x 100 = 27.8%

  - The relative predictive value is simply the ratio of PPVs for LBC and conventional cytology (table B):

    - Relative PPV = (313 x 661)/(1337 x184) = 0.84 (95% CI 0.72-0.98)

  - The 95% confidence interval for this relative PPV does not overlap 1, but its upper value is close to 1. Thus the difference in positive predictive value for LBC and conventional cytology is statistically significant at the 5% level for these degrees of abnormalities but only just so, with LBC performing worse. For more severe degrees of histological abnormalities the differences are more significant, with LBC performing consistently worse than conventional cytology.

The following are additional points which might improve the answer to "good" or "excellent".

- Derive the relative detection rate for each grade of histology as the ratio of the detection rate for LBC and for conventional cytology. Since these data are from a randomised controlled trial with the same numbers in each trial arm, equal numbers of abnormalities would be expected in each trial arm, so this gives a measure of relative sensitivity. Table C below illustrates how this is derived for CIN1 or worse.

Table C: CIN1 or Worse at Colposcopy for Women with Mildly Dyskaryotic Smears or Worse			
	= or > CIN1	No CIN	Total
Liquid-based cytology	313	22395	22708
Conventional cytology	184	22282	22466

- % Detection rate (LBC) = (313/22708) x 100 = 1.38%
- % Detection rate (Conventional) = (184//22466) x 100 = 0.82%
- Relative sensitivity = (313 x 22466)/(22708 x184) = 1.68 (95% CI 1.40-2.02)
- The 95% confidence interval for this relative sensitivity does not overlap 1, so the difference between LBC and conventional cytology is significant at the 5% level for CIN 1 or worse, with LBC performing better than conventional cytology. For higher degrees of histological abnormalities the relative sensitivities of LBC and conventional cytology do not differ significantly but the confidence intervals suggest there may be underlying differences - which are not consistently in the same direction.

  Where X^2 = 4.8 and thus X = 2.2.

Q4. Explain why it is not possible from the information provided to calculate an absolute value for the sensitivity of LBC or conventional cervical cytology. What further information would you need to enable calculation of absolute sensitivities? [10% of marks]

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

- The sensitivity is defined as the proportion of disease-positive individuals, categorised according to some reference standard test, correctly identified by the test of interest.
- In this study the reference standard test is abnormal histology at colposcopy. The sensitivity is thus a/(a+c) in table A above.
- But since not all women in the trial have been colposcoped it is impossible to know how many women with normal cytology have abnormal histology i.e. the value of c in table A is unknown.
- To enable calculation of sensitivity one would either need to colpsocope all women in the trial with normal cervical cytology or, more practically, a representative sample of them.

The following are additional points which might improve the answer to "good" or "excellent".

- Performing colposcopy on women with normal cervical cytology would not be justifiable ethically.

Q5. Using your interpretation of the data in tables 1-3, write a briefing paper of up to 500 words for your director of public health/chief medical officer on the potential benefits and disbenefits of introducing liquid-based cervical cytology in your area, and the practical issues that need to be considered when doing this. [30% of marks]

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

Potential Benefits of LBC

- The data from Areas A and B indicating a favourable effect of LBC on unsatisfactory cervical smear rates are confirmed by the randomised controlled trial results.

- LBC reduces the number of repeat smears needed for unsatisfactory smear results which in turn:
  - reduces laboratory workload;
  - is more convenient for women;
  - lowers smear taker workload.

- For CIN 1 or worse, LBC is significantly more sensitive than conventional cytology. This means LBC is better than conventional cytology at correctly identifying women with cervical lesions of these histological grades.

- For CIN2 or worse LBC is more sensitive than conventional cytology, but not significantly so. But the confidence intervals suggest that LBC may be better than conventional cytology at correctly identifying women with cervical lesions of these histological grades.

Potential Disbenefits of LBC

- For CIN3 or worse LBC is less sensitive than conventional cytology, but not significantly so. Nevertheless the confidence interval suggests that LBC may be worse than conventional cytology at correctly identifying women with cervical lesions of these histological grades - potentially worrying since women with such lesions may be given false reassurance when screened using LBC.

- The positive predictive value of LBC is significantly worse than for conventional cytology for all histological grades of abnormality. This will result in more, probably unnecessary, referrals of women to colposcopy, which is an uncomfortable and stressful procedure.

Practical Issues

- capital and revenue costs for purchase of analytical equipment for labs and sampling kits for smear takers
- training for laboratory staff
- training and support for smear takers

The following are additional points which might improve the answer to "good" or "excellent".

Potential Benefits of LBC

- Liquid-based cytology provides the technical basis for performing human papillomavirus (HPV) testing.

- It is quicker and therefore promotes quicker laboratory sample turnaround times and results notification to women and smear takers.

Potential Disbenefits of LBC

- Increased referrals to colposcopy may also impact adversely on service capacity and waiting times. The effect of introducing LBC on the workload of colposcopy services would need to be monitored.

Practical Issues

- information for women about the new sampling method and the expected benefits it will bring
- logistical implications of implementing LBC, including transport of specimens, storage and waste disposal
- The local service would need to monitor closely CIN3 detection rates following introduction of LBC. It would be useful to seek reassurance from Area B (where LBC has been in use for some time) that detection rates for CIN 3 or worse hadn't reduced significantly following LBC introduction.

EXAMINERS' COMMENTS

Q1. This question was in the main answered satisfactorily. Almost all candidates realized that conversion of numbers to rates was the key. Having a structure for reasons for differences e.g. chance, population differences helped. Candidates tended to lose marks by failing to discuss the differences in rates of abnormal smears of various grades between Areas A and B.

Q2. This question was in general answered well. Nearly all candidates chose the correct statistical test i.e. the difference between two independent proportions, and understood the concept of a confidence interval. However, a surprisingly large number of candidates got the arithmetic wrong.

Q3. Replies to this question were of mixed quality. Candidates were able to see where the numbers came from, but their interpretation of the various measures was weaker.

Q4. This question was very poorly answered. Only a handful of candidates grasped that histological outcome at colposcopy was the reference standard in this scenario, and that all women would need to have colposcopy to allow calculation of the number of false negative (NOT false positive) cervical smears and hence cervical screening sensitivity.

Q5. This question was in general answered badly. Too many candidates quoted the terms e.g. sensitivity, positive predictive value from a general 2x2 table, but did not demonstrate an understanding of their implications in the scenario. Many candidates lost marks by failing to mention the practical benefits of LBC e.g. greater convenience for women, reduced lab workload. Structure and coherence in answers was lacking, with a significant number of candidates focusing too much on economic appraisal issues.

End of document