

Exam HEVAL 4200 2019 – Problem set with solution

Front page:

Department of Health Management and Health Economics
Faculty of Medicine
University of Oslo

School exam for HEVAL4200, March 29th 2019

Examination support material: Approved calculator (Citizen SR-270X, Citizen SR170X BTSPU, TI-30X IIS, or Casio FX-82MS). An online calculator is also available.

In addition, the following two attachments will be handed out:

- Formulas
- Hofer et al 2018 (with line numbers)

The exam consists of 13 questions with a maximum score of 180 points.

Introduction to solution

The exam was arranged digitally on the Inpera platform and consisted of 13 questions. Questions 9–13 were applications of a check-list from the course's main textbook, Drummond et al. (2015), on the "exam article" by Hofer et al. (2018). The exam article was made available to the students approximately 6 weeks before the exam, so the candidates were expected to be well prepared. Different solutions are possible, but in general each point on the check-list should have a clear answer with an appropriate reference to the text (e.g., page and line numbers).

The maximum total score was 180 points. The following relation between grades (A–F) and points [0– 180] was used: A: [165-180]; B: [145-164], C: [120-144], D: [99-119], E: [80-99], F: [0-79].

1. Economic evaluation and priority setting (8 points)

Oslo University Hospital recently hired you as a health economist to provide guidance on efficient resource allocation in the hospital. Your new colleagues at the hospital are new to the field of health economics and economic evaluation and ask you a few questions. However, they are busy taking care of patients, so you need to respond briefly (1-3 sentences for each question).

- a) What is economic evaluation?

Answer: Economic evaluation is the comparative analysis of alternative course of action in terms of both their costs and consequences. In other words, it deals with the choice between alternatives, and the inputs and outputs associated with these alternatives.

- b) Why do we need to prioritize – or put differently, why can't we just provide all beneficial health services to all people?

Answer: The rationale for economic evaluation is that we have limited resources, such as doctors, nurses and hospitals. At the same time, we have unlimited wants, meaning that we always wish to have "more health" or "better health". These factors combined imply that it is not possible to satisfy all health needs when resources are limited. Thus, priority setting is unavoidable.

2. Allocating the budget (20 points)

On your second day at Oslo University Hospital, you are considering how to spend the available budget to maximize gains in quality-adjusted life-years (QALYs). You are provided with the following table of alternative treatment programs and their expected total costs and benefits.

Program	Cost (NOK)	Effect (Gains in QALYs)
A (Physical exercise program)	500 000	40
B (Mental health program)	700 000	200
C (State of the art imaging)	1 800 000	400
D (Cancer treatment center)	5 000 000	625

Assume you can fund any combination of these programs, and the programs are infinitely divisible (i.e., you can fund parts of a program, at proportional costs and QALY gains).

What combination of the program(s) would you choose to implement if your operating budget was NOK 5 000 000? How many QALYs would you gain?

Answer:

Program	Cost (NOK)	Benefit (Gains in QALYs)	C/E
A (Physical exercise program)	NOK500,000	40	12,500
B (Mental health program)	NOK700,000	200	3,500
C (State of the art imaging)	NOK1,800,000	400	4,500
D (Cancer treatment center)	NOK5,000,000	625	8,000

You start by implementing the program with the lowest cost per QALY, and so on. This means you will first choose B, then C, then D and finally A. However, with a limited budget of NOK5,000,000 you will choose the following:

- First implement program B at a cost of NOK700,000 – you still have NOK4,300,000 left of your budget.
- Second, implement program C at a cost of NOK1,800,000 – you still have NOK2,500,000 left of your budget.
- The next program you want to prioritize is D – however, you cannot afford the entire program, which costs NOK5,000,000. Instead, you can spend your remaining NOK2,500,000 of your budget and implement 50% of program D.

The resulting QALY gains associated with implementing programs B, C and 50% of D, are:

- Program B = 200
- Program C = 400
- Program D = $625 * 0,50 = 312,5$
- In sum: $200 + 400 + 312,5 = 912,5$ QALYs

3. Choosing between alternative programs (12 points)

The hospital is deciding between alternative rehabilitation programs (E1, E2 and E3), which are mutually exclusive. In the following table, the QALYs gained and costs are measured against the alternative of doing nothing.

Program	Cost (NOK)	QALY gains
E1	2 500 000	87
E2	1 700 000	60
E3	900 000	62

Calculate the incremental cost-effectiveness ratios. Which program should the hospital implement if the willingness-to-pay threshold is NOK 50 000 per QALY gained?

Answer:

E3

Program	Cost (NOK)	Benefit (Gains in QALYs)	ICER
E3	900 000	62	14 516
E2	1 700 000	60	dominated
E1	2 500 000	87	64 000

Program E2 is dominated because it has higher costs and lower gain in QALYs, compared to program E3. We therefore rule out program E2 and calculate the ICER of E1 compared to E3. If the willingness-to-pay threshold is NOK 50 000, we would implement program E3, because the ICER of program E1 is higher than the threshold value.

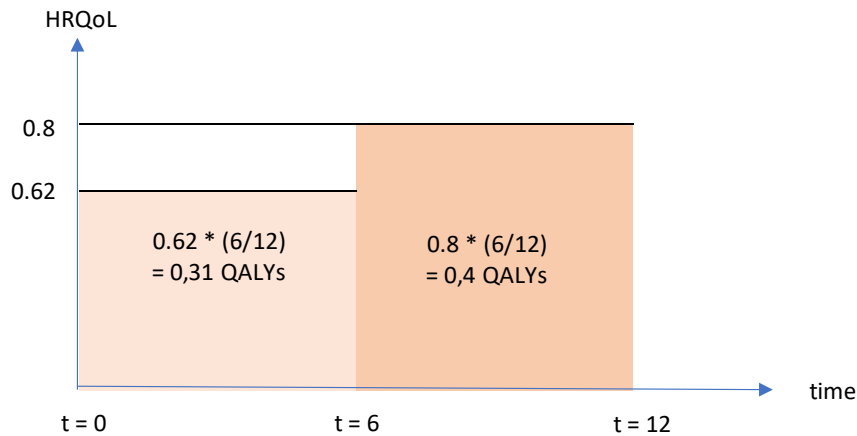
4. Valuing health (22 points)

On your third day at Oslo University Hospital, a physician comes into your office and asks for advice. The physician has received data from an ongoing clinical trial, in which case they have collected health-related quality of life (HRQoL) from patients. One patient has reported the following HRQoL:

- From baseline and until 6-months follow-up: 0.62
 - From 6 months follow-up and until 12-months follow-up: 0.8
- a) Explain what is meant by quality-adjusted life-years (QALYs)
 - b) Calculate the associated QALYs during the time from baseline and until the 12-months follow-up.
 - c) Explain how the researchers may have collected HRQoL-data in the study. That is, explain alternative approaches in terms of direct and indirect measures, and provide three examples of direct measures and three examples of indirect measures.

Answer:

- a) A quality adjusted life year (QALY) is defined as the length of life multiplied with a QALY weight. The QALY weight is usually measured on a scale between 0 and 1, where 1 represents full health and represents the state of being dead (negative numbers are sometimes used to represent states considered worse than death).
- b) The QALY diagram for this patient:



The number of QALYs for the patient from time = 0 to time = 12 is equal to: $0,62 * (6/12) + 0,8 * (6/12) = 0,31 + 0,4 = 0,71$ QALYs

c) Direct measures: VAS, SG, TTO.

Indirect measures: generic measures such as EQ-5D, SF-6D and HUI3, and disease-specific measures.

Full score is given to candidates who have provided three examples of each as well as having provided some description/details of the different measures.

5. Probabilities and rates (8 points)

After working at the hospital for a few weeks, you have started to build a Markov model to evaluate the cost-effectiveness of a new drug for the treatment of COPD. You have reviewed the literature to inform the transition probabilities in your model, which has a 1-year cycle length. A study has reported that, during the 5-year follow-up time, 432 out of 6452 patients improved their COPD condition (from severe to moderate COPD).

Assuming that the rate is constant over the period, what would be the transition probability from the health state "severe" to "moderate" to use in your model based on this study? Report 6 decimals.

Answer:

The 5-year probability is $432/6452 = 0,066956$

The corresponding 1-year rate = $-\ln(1 - 0,066956) / 5 = 0,013861$

The corresponding 1-year prob = $1 - e^{(-0,039340)} = 0,013765$

The correct answer is 0,013765

6. Diagnostic uncertainty (8 points)

The prevalence of a disease is 5%. A diagnostic test is available, with a sensitivity of 60% and specificity of 80%. What is the positive predictive value of the test?

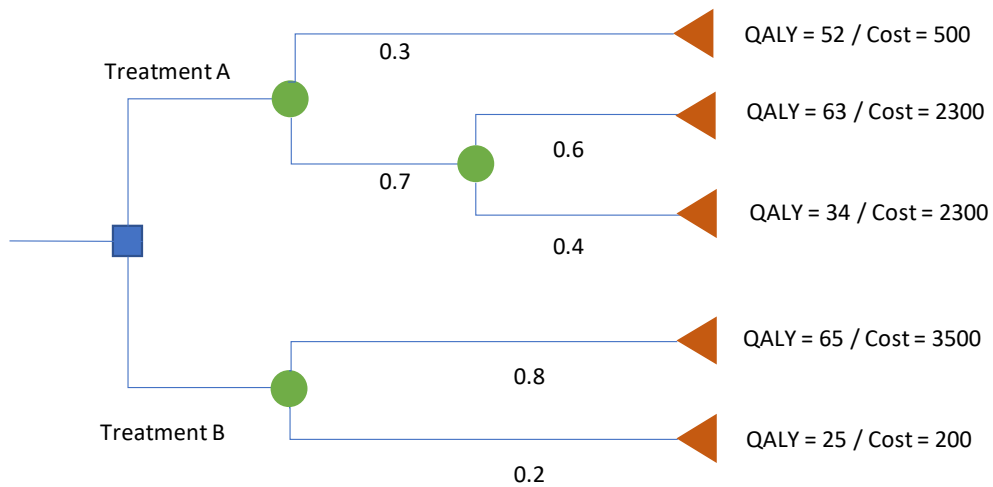
Answer:

$PPV = (sens * prev) / (sens * prev + [(1-spec)*(1-prev)]) = 0,136364$

PPV = 13,636%

7. Decision-tree model (22 points)

Consider the following decision tree model with two treatment options, Treatment A and Treatment B:



Calculate the expected QALYs and expected costs associated with each treatment option, and calculate the ICER of Treatment B compared to Treatment A. Under what willingness-to-pay threshold would you adopt Treatment B?

Answer:

Treatment A:

- Expected QALY = $0,3 \cdot 52 + 0,7 \cdot [0,6 \cdot 63 + 0,4 \cdot 34] = 51,58$
- Expected cost = $0,3 \cdot 500 + 0,7 \cdot [0,6 \cdot 2300 + 0,4 \cdot 2300] = 1760$

Treatment B

- Expected QALY = $0,8 \cdot 65 + 0,2 \cdot 25 = 57$
- Expected cost = $0,8 \cdot 3500 + 0,2 \cdot 200 = 2840$

$$\text{ICER} = (2840 - 1760) / (57 - 51,58) = 199,26$$

You would adopt Treatment B if the willingness-to-pay for one additional QALY was at least NOK199,26.

8. Markov model (sketch sheet) (32 points)

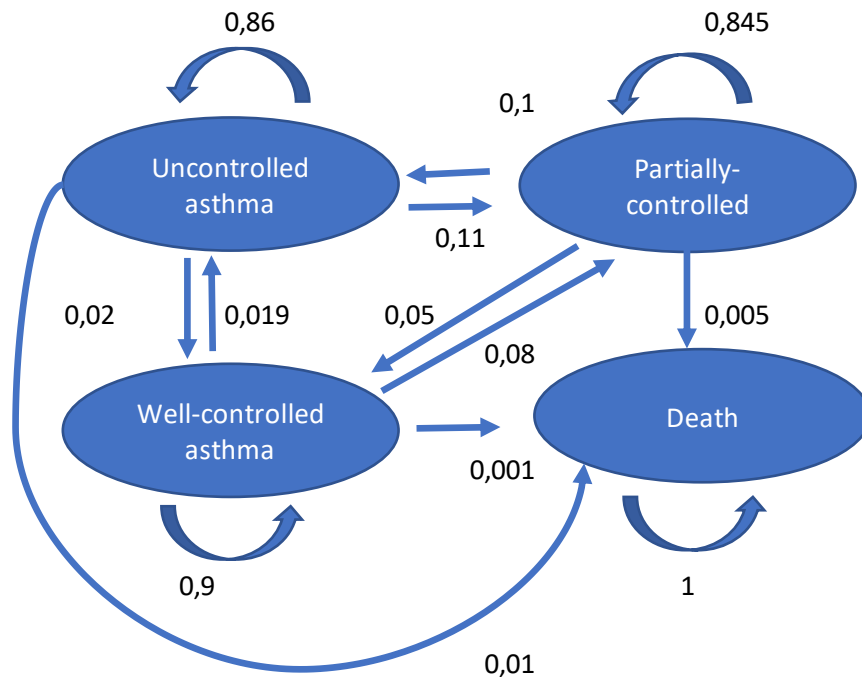
Consider a Markov model with the following health states: uncontrolled asthma, partially controlled asthma, well-controlled asthma, death. The following table presents transition matrix:

	Uncontrolled asthma	Partially controlled asthma	Well-controlled asthma	Death
Uncontrolled asthma	0.86	0.11		0.01
Partially controlled asthma	0.1		0.05	0.005
Well-controlled asthma		0.08	0.9	0.001
Death	0		0	1

a) Use the sketch sheet and draw the state transition diagram for this model. Fill in the transition probabilities that have been omitted from the table.

Answer:

	Uncontrolled asthma	Partially controlled asthma	Well-controlled asthma	Death
Uncontrolled asthma	0,86	0,11	0,02	0,01
Partially controlled asthma	0,1	0,845	0,05	0,005
Well-controlled asthma	0,019	0,08	0,9	0,001
Death	0	0	0	1



b) Assume that everyone starts out in the health-state well-controlled asthma in cycle 0. What proportions of the cohort are in the different health states in cycles 1 and 2, respectively?

For states uncontrolled, partially controlled, well-controlled and death:

Cycle 0: [0 0 1 0]

Cycle 1: [0.019 0.080 0.900 0.001]

Cycle 2: [0.041 0.142 0.814 0.0025]

ALTERNATIVE PRESENTATION

For states well-controlled, partially controlled, uncontrolled and death, respectively:

Cycle 0: [1 0 0 0]

Cycle 1: [0.9 0.08 0.019 0.001]

Cycle 2: [0.814 0.142 0.041 0.0025]

Drummonds checklist

Comment: Overall, evaluation of the candidates' responses to the checklist was based on how the candidate demonstrated independent thinking and general understanding of economic evaluation.

9. Checklist point 2 (9 points)

2. Was a comprehensive description of the competing alternatives given (i.e. can you tell who did what to whom, where, and how often?)

2.1. Were any relevant alternatives omitted?

2.2. Was (should) a 'do-nothing' alternative (be) considered?

Answer:

2. Yes, the study compares annual LDCT screening with standard clinical care among current and former heavy smokers aged 55 to 75 years (abstract l. 110-111; p.2 l. 31-33 and l. 34-45). The screening algorithm is well-described (p. 2, l. 128-131 and Figure 2). Treatment and aftercare is also well-described (p. 2, l. 90-108).

2.1. The study only includes a single screening algorithm compared with standard clinical care. Ideally, the study should have considered alternative screening strategies, for example by considering alternative screening methods, target age groups and screening frequencies. Alternative screening frequencies is however considered in sensitivity analysis.

2.2. LDCT screening was compared to standard clinical care. A do nothing alternative (in a strict sense) would not have been relevant in this context; any patient diagnosed with lung cancer would receive treatment and not doing so would be considered unethical. One could consider standard clinical care as a "do nothing" (no screening) alternative.

10. Checklist point 4 (12 points)

4. Where all the important and relevant costs and consequences for each alternative identified?

4.1. Was the range wide enough for the research question at hand?

4.2. Did it cover all relevant perspectives? (Possible perspectives include those of patients and thirdparty payers; other perspectives may also be relevant depending on the particular analysis.)

4.3. Were capital costs, as well as operating costs, included?

Answer:

4. The authors used a healthcare payer perspective and did not include indirect costs such as patient's time cost (p. 2, l. 15-17).

4.1. Given a healthcare payer analytic perspective, it appears the range was wide enough.

4.2. The study did not cover all relevant perspectives, as a societal perspective would have been relevant when considering whether to implement a screening program. For example, patient's time costs could be an important factor in determining the societal cost-effectiveness of a screening program.

4.3. The study used data from the German outpatient reimbursement catalogue to inform screening costs (p. 3, l.63-64), and treatment costs based on a previous study using administrative data from a German statutory health insurer (p.3, l. 136-150), which may or may not have included capital costs. In addition, the authors have added a lump sum per case to represent the additional costs typically incurred in structured screening programs (p. 3, l 65-67), which is a type of operating cost.

11. Checklist point 7 (10 points)

7. Were costs and consequences adjusted for differential timing?

7.1. Were costs and consequences that occur in the future "discounted" to their present values?

7.2. Was any justification given for the discount rate(s) used?

Answer:

7: Yes, costs and effects were adjusted for differential timing using discounting. It seems that costs were measured in 2016 Euros (€), although this is only mentioned one place in the paper (p. 3, l. 140).

7.1: Yes, both costs and effects were discounted to the present value at 3% per year (p. 2, l. 55).

7.2: No justification for the discount rate was given, but the discount rate was varied in the sensitivity analysis.

12. Checklist point 8 (6 points)

8. Was an incremental analysis of costs and consequences of alternatives performed?

8.1. Were the additional (incremental) costs generated by one alternative over another compared to the additional effects, benefits, or utilities generated?

Answer

8. Yes, an incremental analysis of costs and consequences was performed (p. 6, l. 25-43). The ICER is the efficiency metric and main outcome in the analysis. The authors have calculated the ICER using both life-years and quality-adjusted life-years in the denominator.

8.1. Yes, the additional costs and effects (life-years and QALYs) of LDCT screening is compared with standard clinical care.

13. Checklist point 9 (12 points)

9. Was uncertainty in the estimates of costs and consequences adequately characterized?

9.1. If patient-level data on costs or consequences were available, were appropriate statistical analysis performed?

9.2. If a sensitivity analysis was employed, was justification provided for the form(s) of sensitivity analysis employed and the ranges or distributions of values (for key study parameters)?

9.3. Were the conclusions of the study sensitive to the uncertainty in the results, as quantified by the statistical and/or sensitivity analysis?

9.4. Was heterogeneity in the patient population recognized, for example by presenting study results for relevant subgroups?

Answer

9. Overall, the uncertainty in estimates seems to have been adequately characterized. However, the study lacks justification for why the specific uncertainty analyses were undertaken.

9.1. The analysis is not based on patient-level data.

9.2. The analysis includes seven deterministic one-way sensitivity analyses and a probabilistic sensitivity analysis (PSA) which the authors refer to as “Monte Carlo simulation with 10,000 repetitions” (p. 5, l.122-130). In Table 1, the probability distributions used in the PSA have been reported. Assignment of distributions seems appropriate. The authors state that the PSA is conducted “to account for heterogeneity” (p. 6, l. 18), but it is unclear whether they have both used sampling *and* microsimulation, or whether they have just used sampling (referred to as “10,000 draws” in Figure 4). To account for heterogeneity, microsimulation or subgroup evaluation is required. The authors do not state how heterogeneity is captured in the PSA. The paper lacks justification for the parameters (and their ranges) that are explored in the sensitivity analyses.

9.3. The authors state that “model results were robust with no variation exceeding €31000 per life year gained or €48000 per QALY gained” (p. 6, l 44-47). These are relevant thresholds; in the discussion section, the authors state there is no official German cost-effectiveness threshold, although the program would be regarded as cost-effective using the WHO threshold of €48,000 per QALY (p. 6, l. 70-78). Although results were consistently below the relevant thresholds across uncertainty analyses, the ICER were sensitive to some assumptions, as discussed in the results section (p. 6, l. 44-60) and showcased in the tornado diagram in Figure 3.

9.4. No, study results were not presented for different subgroups, although the authors mention in the discussion section that this should be the focus of future research (p. 9, l. 56-59). The authors state that heterogeneity is captured in PSA, but it is not clear how the PSA captures heterogeneity.