

Exam in Heval4200, 25 March 2021:

Suggested/outlined solution

Version intended for candidates

Knut R. Wangen,

k.r.wangen@medisin.uio.no

30 April 2021

Introduction

The exam was arranged digitally on the Inspira platform and consisted of 7 questions.

Question 7 was an application of a checklist from the course's main textbook, Drummond et al. (2015), on the "exam article" by Wilkinson et al. (2017). The exam article was made available to the students nearly 4 weeks before the exam, so the candidates were expected to be well prepared.

The maximum total score was 180 points. The following relation between grades (A–F) and points [0–180] was used: A: [155–180], B: [140–154], C: [110–139], D: [90–109], E: [70–89], F: [0–69].

Explanation of grades. After the grades are published, the candidates will have access to their individual scores by logging on to Inspira. The present suggested solution, which will be available to the candidates, explains the general principles for how scores and grades were awarded.

Appeal against grades. Appeals against grades must be submitted following the procedure described on the course webpage (that is, not submitted to the course coordinator).

Informal feedback. Candidates who wish to receive additional feedback may ask the course coordinator for an informal talk (phone call or Zoom meeting). The candidates must then email the course coordinator 2-3 suggested times. These suggested times should be within normal office hours (0900h-1700h), with at least two workdays notice, and within the period 5-14 May 2021.

Suggested solution

Question 1 (20 points)

The question is open-ended and can be answered in many different ways. The following points are considered important:

- Health related quality of life and QALY weights (concepts, scale 0-1, QALY calculation)
- Properties of EQ-5D and indirect methods
 - Generic (comparison between different patient groups)
 - Questionnaire (domains, levels)
 - Tariff (valuation methods; sample of patients or from the general population)
 - Limitations (e.g., versus disease specific methods)
- Repeated measurements for each patient

- (QALY gains: Requires some sort of comparator)

Question 2 (24 points)

a)

Patient group 1				Patient group 2			
Treatment	ΔC	ΔE	$\Delta C/\Delta E$	Treatment	ΔC	ΔE	$\Delta C/\Delta E$
A	30	1.1	27.27	D	22	1.2	18.33
B	39	-0.5	-78.00	E	43	0.2	215.00
C	6	0.3	20.00	F	45	1.1	40.91

b)

Exclude B, which is dominated by A; exclude E, which is extendedly dominated by F, then recalculate:

Patient group 1				Patient group 2			
Treatment	ΔC	ΔE	$\Delta C/\Delta E$	Treatment	ΔC	ΔE	$\Delta C/\Delta E$
A	30	1.1	27.27	D	22	1.2	18.33
C	45	-0.2	-225	F	88	1.3	67.69

Treatment C is also excluded because it is dominated by A.

$\lambda < 18.33$, the current standard of care should be chosen for both patient group.

$18.33 \leq \lambda < 27.27$, the current standard of care should be chosen for group 1, and treatment D should be chosen for group 2.

$27.27 \leq \lambda < 67.69$, treatment A should be chosen for group 1, and treatment D should be chosen for group 2.

$67.69 \leq \lambda$, treatment A should be chosen for group 1, and treatment F should be chosen for group 2.

c)

Treatment A costs EUR 30 million. Treatments D, E, and F cost EUR 22 million, EUR 65 million, and EUR 110 million, respectively. The EUR 130 million budget is not sufficient to cover both treatments A and F.

The maximum number of QALYs gained is 2500 and can be achieved by choosing either 1) treatments A and E, at total cost EUR 95 million, or 2) treatment F at cost EUR 110 million. It seems reasonable to assume that the government would prefer alternative 1) because it has lower costs.

Question 3 (21 points)

Treatment 1: $E(\text{QALY}) = 4.5$ and $E(\text{Cost}) = 1150$

Treatment 2: $E(\text{QALY}) = 2.75$ and $E(\text{Cost}) = 52.5$

$\text{ICER} = (1150 - 52.5) / (4.5 - 2.75) = 627.1$

If the willingness to pay threshold is less than 627.1 we would recommend treatment 2. If the threshold is higher, we would recommend treatment 1.¹

Question 4 (20 points)

Probabilities or relative frequencies:

$$P(T+ \cup D+) = P(T+ | D+) \cdot P(D+) = 0.985 \cdot 0.0018 = 0.001773$$

$$P(T- \cup D-) = P(T- | D-) \cdot P(D-) = 0.986 \cdot (1-0.0018) = 0.984225$$

For one million people:

	D+	D-	Sum
T+	1 773	13 975	15 748
T-	27	984 225	984 252
Sum	1 800	998 200	1 000 000

Expected number of cases:

True positive ($T+ \cup D+$): **1773**

False positive ($T+ \cup D-$): **13 975**

Question 5 (20 points)

a)

$$dV/dP = 27 \$ / 0.00012 = 225\ 000 \$.$$

b)

Possible relevant factors could include:

- Stated preferences (students) versus revealed preferences (private clinic), in general.
- Different risk profile or risk preferences, say, young students versus perhaps older patients at the clinic.
- Different levels of income and wealth (ability to pay).
- (Selection bias) In the population consisting of all regular patients at the private clinic, it is conceivable that only some individuals have high WTP ($\geq 75 \$$) while the average WTP is lower (say, 40 \$). In addition, individuals who have high WTP for the vaccine, but who are not regular patients at the private clinic, may contact the clinic for the sole purpose of getting the vaccine.

¹ The decision tree considers only two options. One of them could be the old standard of care but this is not specified.

Question 6 (25 points)

(Other solutions are possible)

- 1) Long-term symptoms can be represented by splitting the recovered state into new states. One of the new states could be “Fully recovered” while others could represent different types of recovery (the virus infection itself is cleared in all these states). The probability of reaching these new states could potentially depend on what infection state the patient experienced (Mild symptoms, Severe disease, Critical disease). For instance, it could be less likely for a patient to experience long-term lung problems if she had only mild symptoms than if she experienced critical disease.
- 2) The probabilities for developing severe or critical disease increase in age and tend to be higher for males than females. This can be handled for instance by having separate Markov models for the different sub groups (say, 10-year age groups and sex). In principle, all transition probabilities could vary between the subgroups.
- 3) Some of the new virus variants are claimed to be more contagious, but this is outside the scope of the presented model, because the model’s starting point is a newly infected individual (not a susceptible individual). Given that an infection has occurred, the model structure could be the same but all transition probabilities may in principle depend on the virus variant. For instance, some of the variants are claimed to be more likely to cause severe disease and to be more lethal. If recovering from one virus variant does not provide immunity against another variant, reinfection could be possible (the subsequent transition probabilities could be different from those who have not had any COVID-19 infection before).

Question 7 (50 points)

a)

Introduction

(Page 903, line 25-27): “*This study follows a hypothetical cohort of 1000 pregnant women experiencing one live birth over a 2-year time horizon.*” This means that the study is based on a model where the parameters are obtained from other studies.

The model is a decision tree (Figure 1), where probabilities are presented in Table 1, costs items are presented in Table 2, and utilities are presented in Table 3. Additional information can be found in an online supplement (p. 905, l. 104) but the supplement is beyond the scope of this exam.

Checklist point 2: Yes, the alternatives seems to be clearly described.

2.1: No, the screening intervention and the current standard of care seem to be the most relevant alternatives in the given context. Variants of the screening could perhaps have been considered, for instance by having nurses providing the screening instead of physicians (p. 912, l. 45-50).

2.2: No, a “do nothing” alternative is not likely to be acceptable in this context.

2.3: The article does not mention patient subgroups. Potential subgroups could for instance be women with known risk factors.

Checklist point 4: It seems like the most important costs and consequences were identified.

4.1: Yes, the range seems wide enough, given the chosen perspective.

4.2: No, the article did not cover all relevant perspectives. It uses a Medicaid payer perspective (p. 903, l. 30-31; p. 906, l. 92-94), and this perspective is considered a limitation because “many women will likely lose their Medicaid coverage before the end of our 2-year time horizon” (p. 911, l. 86-89). Costs that have to be covered by the women themselves, when the Medicaid coverage lapses, are not included (p. 912, l. 4-7). Neither are the negative implication for the women’s family members (p. 911, l. 24-29).

4.3: No, capital costs are not explicitly included. This could be appropriate assuming that the alternatives do not require expensive equipment. The cost items (Table 2) seems to focus on variable costs (e.g., doctor’s time) and overhead and fixed costs (say, training of personnel or use of office space).

Checklist point 6: Yes.

6.1: Yes, the sources of all values were identified in Tables 3 and 4. The sources seem credible.

6.2: Yes, it seems like most cost items were wages (e.g., doctor’s time) or from the Medicaid fee schedule (drug costs), which could be considered as approximate market values. We should note that the costs of postpartum psychosis (PP) treatment were obtained from two scientific articles, Guo et al. (2007) and Qiu et al. (2009), confer Table 2 (with notes).

6.3: This point does not seem to be relevant.

6.4: Yes, the consequences were valued in number of remissions and QALYs (p. 904, l. 54-55). The QALY measures takes several relevant health states into account (Table 3).

Strengths:

- The model is clearly presented (structure and inputs)

Weaknesses:

- Hypothetical model, based assumptions and not an actual intervention
- Possible demarcation problems regarding costs towards the end of the 2-year horizon (confer point 4.2 above)
- Subgroups are not considered

b)

Checklist point 9: Yes.

9.1: This point does not seem to be relevant.

9.2: Deterministic one-way sensitivity analysis and probabilistic sensitivity analysis (PSA) were performed (p. 909, l 12-63). The choices regarding the parameters’ low and high estimates and their

distribution (Table 1 and 3) seem not to be justified. The cost parameters were assumed fixed and not included in the sensitivity analysis (p. 909, l. 39-44).

9.3: Yes, to some extent. The perhaps most interesting ICER estimate was \$13,857/QALY, clearly below the commonly accepted U.S. WTP threshold of \$50,000/QALY (p. 909, l. 76-83). From the PSA, we can infer that it more than 90% likely that the screening intervention would be cost effective at the \$50,000/QALY threshold. The sensitivity for individual model parameters are presented in Figure 2.

9.4: The article does not mention patient subgroups.

Relation to part a):

- Assuming that the costs are known is problematic in general. Note also points 4.2 and 6.2.
- Subgroups could perhaps have been related to the sensitivity analysis. For instance, PP prevalence is the single most influential parameter in Figure 2, and this prevalence could perhaps vary between subgroups.