MODELANTWOORDEN

Research Exam Semester 1 – 2020-2021

January 15, 2021 - 9.00 – 11.30 hr

Question 1 – Wet lab research - Dr. H. Pluk (10 points)

The enzyme lactate dehydrogenase (LDH) catalyses the conversion of lactic acid into pyruvic acid and vice versa.

Lactic Acid   Pyruvic Acid

Different isoforms of LDH consist in humans. In an experiment the Michaelis-Menten characteristics were determined as follows: LDH muscle isoform: V_{max} = 12 arbitrary units; K_m = 0.5 mM lactate; LDH liver isoform: V_{max} = 12 arbitrary units; K_m = 0.1 mM lactate.

A. Which of the isoforms of LDH will have the largest rate of catalysis at a lactate concentration of 0.1 mM? Explain your answer. (2 pts)

The liver isoform (0.5): this isoform has the lowest Km (0.5). At 0.1 mM lactate the speed of conversion is already half maximal; muscle will reach the same speed at a higher concentration of 0.5 mM (0.5); Vmax is the same, therefore you may conclude this, at high substrate concentrations they both have the same maximum conversion rate (0.5).

In the experiments to determine the Michaelis-Menten characteristics of the liver enzyme the following items were pipetted in a cuvette, and subsequently the absorbance at 340 nm was determined over time using a spectrophotometer.

- 0.450 ml distilled water
- 0.250 ml phosphate buffer pH 8.5
- 0.100 ml 6 mM NAD^+
- 0.100 ml lactate (in different concentrations ranging from 0.01M to 3M)
- 0.100 ml LDH liver enzyme

B. Explain why the spectrophotometer needs to be set at 340 nm for these experiments. (2 pts)

You want to measure product formation or substrate decrease in order to get an idea of the speed of the enzymatic conversion (1). At 340 nm NADH is the substance that absorbs light much more than any of the other reactants (1), including NAD^+. So you measure the formation of the product specifically.

C. If you accidently pipet 0.200 ml LDH enzyme into the cuvette and subsequently determine the Vmax and Km, what would be the expected effect on the values of Vmax and Km? (2 pts)

Vmax would <u>double/increase</u> since Vmax is dependent on the amount of enzyme used (1). Km would stay the same, it is a <u>constant of each enzyme</u> (1) and is not dependent on the amount of enzyme

D. In a follow-up experiment you want to check if compound A is influencing the reaction. You have a bottle of compound A in water with a concentration of 0.002 M compound A. What do you need to pipet in your cuvette, and how much of each solution, if you want to test the effect of 50 μM of compound A on the conversion of lactate by LDH? (4 pts)

- <u>0.425</u> ml distilled water
- 0.250 ml phosphate buffer pH 8.5
- 0.100 ml 6 mM NAD^+
- 0.100 ml lactate (in different concentrations ranging from 0.01M to 3M)
- 0.100 ml LDH liver enzyme

- X ml of compound A => 0.025 ml

1 pt for listing all components as before; 1 pt for decreasing amount of H2O to keep equal volume; 2 pt for calculation of x ml compound A

0,002 M = 2 mM compound A = 2000 μM as stock solution (1000x1000); 50 μM needed in cuvette so dilute compound A 2000/50=40 times. Final volume in cuvet 1 ml => 1000/40 = 25 ul needed = 0.025 ml

Question 2 - Modelling a biomarker for pancreas tumour growth – Dr. Thom Oostendorp (15 points)

In a healthy pancreas an equilibrium exists between the production of new cells (by cell division) and the death of old cells. Hence, the total volume of pancreatic cells remains constant. In a pancreatic tumour, the production of new cells exceeds the death of old cells, which results in tumour growth.

The differential equation that describes the dynamics of tumour size is

$$\frac{d}{dt}V_t(t) = p \cdot V_t(t) - r \cdot V_t(t)$$

with $V_t(t)$ the volume of the tumor, $p \cdot V_t(t)$ the rate at which new tumor cells are produced, and $r \cdot V_t(t)$ the rate at which tumor cells die.

A. Show that the volume of the tumour over time is given by

$$V_t(t) = V_0 e^{at}$$

with $a = p - r$, and V_0 the tumor volume at $t = 0$. You don't have to spend time in making your equations look nice, just as long as your answer is clear (4 pts).

If $V_t(t) = V_0 e^{at}$, then $\frac{d}{dt}V_t(t) = aV_0 e^{at} = aV_t(t)$.
This fits the different equation if $a = p - r$.

B. Chemotherapy increases the rate by which tumour cells die. Targeted tumour therapy, on the other hand, decreases the rate by which new tumour cells are produced. The inhibition f is the fraction by which the tumour cell growth is reduced by targeted therapy (i.e: the growth is reduced to $f \cdot p \cdot V_t(t)$). How large does f have to be in order to stop tumour growth? (3 pts)

$\frac{d}{dt}V_t(t)$ has to be negative, i.e. $fp - r$ needs to be less than zero. This happens when $f > r/p$.

Pancreatic tumours are commonly only detected in a very late stage. Early detection would benefit from a biomarker for pancreatic tumours, i.e. a substance that is produced by tumour cells and can be detected in the body. One candidate for such a biomarker is sialyl Lewis antigen A (SLA-A), which is produced by both normal pancreatic cells and pancreatic ductal adenocarcinoma (PDAC) cells.

Figure 1 shows the diagram of a model for SLA-A production. In this model, V_n and $V_t(t)$ are the total volume of the normal pancreatic cells and tumour cells respectively. $s \cdot V_n$ and $s \cdot V_t(t)$ are the rates (in mol/day) by which SLA-A is produced by the normal and the tumor cells. $B(t)$ is the concentration of SLA-A in the body (in mol/l), and $k \cdot B(t)$ is the rate by which SLA-A is removed from the body. The body is modelled as a single compartment with volume V_b.
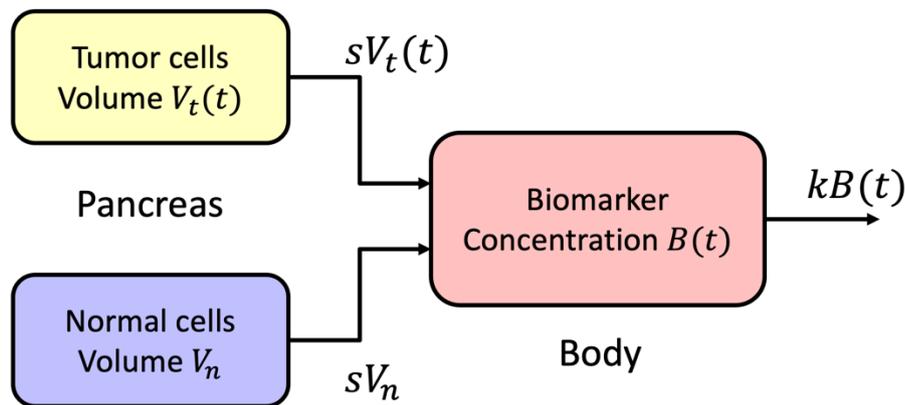
Figure 2.1   Diagram of the model for SLA-A production.

C. What is the unit of s? (3 pts)

sV_n is in mol/day. V_n is in liters, so s is in mol/day/l.

D. Write down the differential equation for $B(t)$. Again, you don't have to spend time in making your equations look nice, just as long as your answer is clear (5 pts).

Change = in - out.
in = the production of SLA-A, out = the conversion rate of SLA-A
change = the change in *amount* (not in concentration) of SLA-A in the body.
This yields:

$$V_b \frac{d}{dt}B(t) = sV_t(t) + sV_n - kB(t)$$

Question 3 - Signals – Dr. Thom Oostendorp - (10 points)

By inserting a catheter into a major artery, the arterial blood pressure (ABP) can be recorded continuously. Figure 3.1 shows an example.
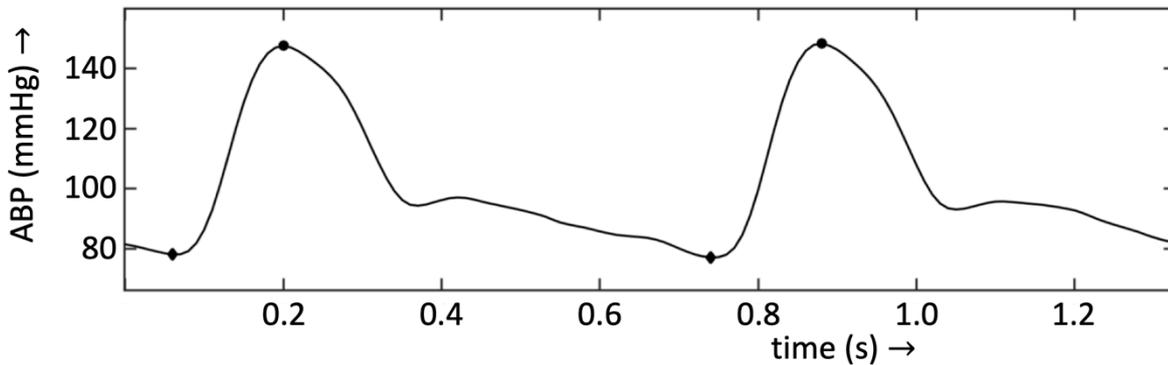

Figure 3.1 *Recording of arterial blood pressure. The markers indicate the diastolic and systolic blood pressure.*

A. What was the heart rate, in beats/min, of the subject whose ABP is displayed in figure 3.1? Show your calculations (1 pt).

The period of the signal is (0.9-0.2)=0.7 s. This corresponds to a frequency of 1/0.75 = 1.33 Hz = 80 beats/min.

B. Explain how you can deduct from figure 3.1 that the spectrum of the ABP contains more than one peak. (2 pts)

It is a repetitive signal, but the shape is not sinusoidal. If it were, the spectrum would contain only one peak, but as it is not, the spectrum will consist of several lines, [representing the base frequency and higher harmonics] (the part between the square brackets does not have to be included).

Figure 3.2 shows the ABP and its spectrum of another subject.
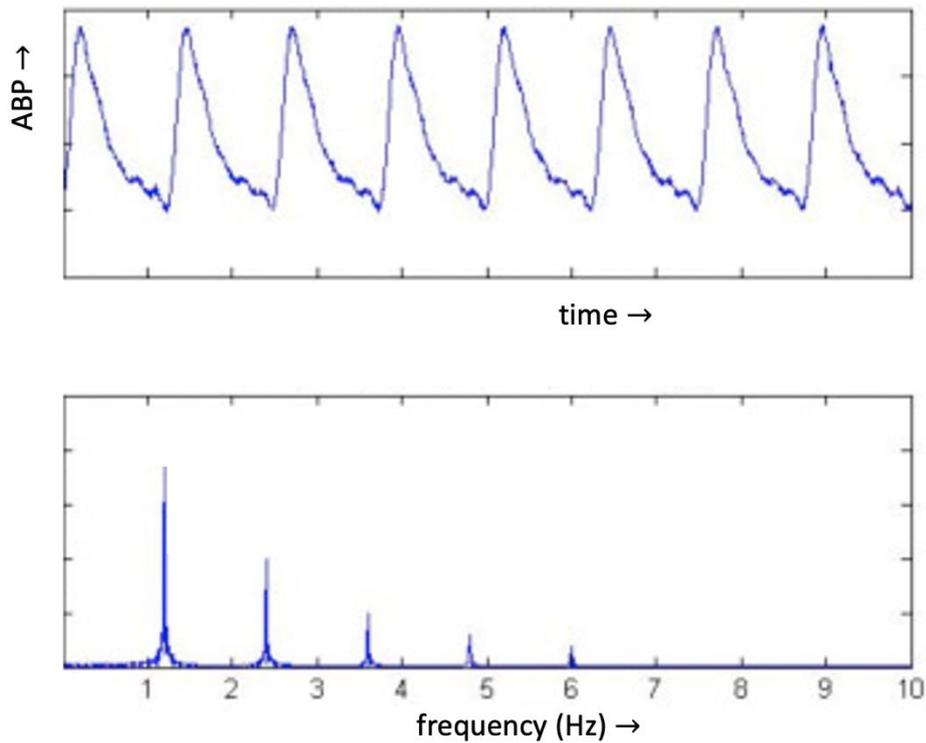

Figure 3.2   ABP and its spectrum.

C. What is the minimum sample rate needed to retain the waveform of the ABP of figure 3.2? Explain your answer (3 pts).

In order to retain the waveform you need a sample rate of 10x the highest frequency component in the signal, so 10 x 6 = 60 Hz.

D. The recording of figure 3.2 contains some high frequency noise. Sketch the amplitude response of a filter that reduces this noise. Also add appropriate labels along the x-axis and y-axis.

A rough sketch, for instance using the "draw" option in Word (*tekenen* in the Dutch version), will do. You can also sketch on paper and take a photo (4 pts).

Watch it: this is the answer!
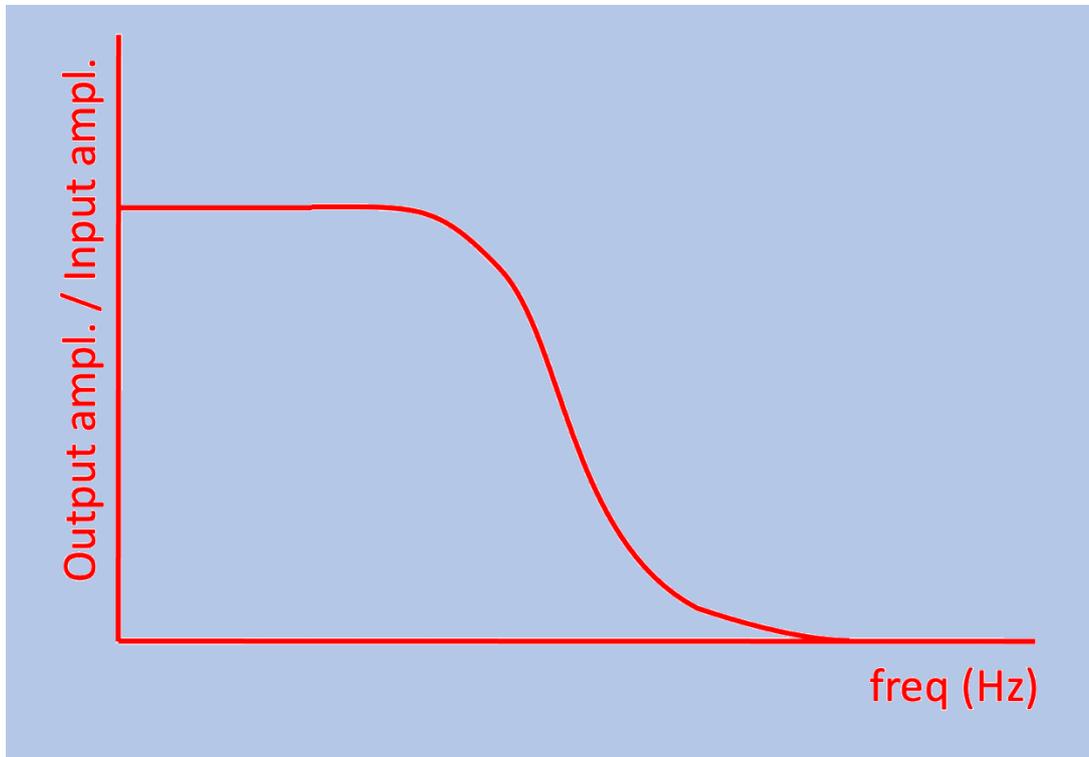
Question 4 Epidemiology and applied medical research
Dr. Femmie de Vegt (20 points)

Use the abstract, methods and tables of 'Naudin et al. Healthy lifestyle and the risk of pancreatic cancer in the EPIC study. European Journal of Epidemiology (2020) 35:975–986.

NOTE: you may interpret the hazard ratio as a relative risk

A. In their abstract, Naudin and colleagues state that "*Adherence to a healthy lifestyle was inversely associated with pancreatic cancer risk'.* Explain the meaning of '*inversely associated*' in this example. (3 pts)

The risk of developing PC in the healthy lifestyle group is lower than in the participants who have a less healthy lifestyle. The relative risk of odds ratio will be lower than 1.

B. The researchers designed an observational study. Name three different observational study designs (1 pt).

Answer – Case-control study, cross-sectional study and (prospective) cohort study

C. Which study design was used by Naudin *et al*? Motivate your answer. (2 pts)

The Epic study is a prospective cohort study. From 1992 to 2000, 521,324 participants aged from 35 to 70 years were recruited across 10 European countries, mostly from the general population. Incident Cancer cases were identified during follow-up based on population cancer registries

D. Suppose other researchers design a randomized clinical trial (RCT) for the same research question as is studied by Naudin et al. Describe and motivate the study population, intervention, time relation and health outcome measures. Is this a feasible study design? (5 pts)

Answer – Assemble a group of elderly (>50 yrs) people not having pancreatic cancer. Make two subgroups and randomise the intervention: one group will be instructed to keep a healthy lifestyle (defined in more detail) for e.g. 5 years and the other subgroup will have no instructions concerning lifestyle. Each 3 months participants have to fill out a questionnaire concerning lifestyle. The incidence of pancreatic cancer (health outcome) will be known from cancer registers.

This is not a feasible study design, as the study population has to be very large or the follow-up has to be very long to have enough cases of PC. In addition, an intervention of several years is very difficult to apply and adhere to.

E. See Table 2. Calculate the cumulative incidence (CI) of pancreatic cancer for all study participants (2 pts).

1075 incident pancreatic cancers occurred during the follow-up period in the study population of 355,472 persons. So the CI is 0.3% or 302,4 per 100,000 persons.

F. See Table 2. Calculate the incidence density (ID) of pancreatic cancer for all study participants (2 pts).

1075 incident pancreatic cancers occurred during the follow-up period of 4,968,127 person years. So the ID is 21,6 per 100,000 person years.

G. In which situations is it better to calculate an ID instead of an CI? Explain (2 pts)

In dynamic populations, because participants are not in the study for an equal time period, so you sum the follow-up time of every individual participant.
In populations with a high number of lost to follow-up.

H. See Table 3 the result 0.64 (0.52–0.78). Give an interpretation of this result (3 pts).

Participants with a HLI score of 15-20 have a lower risk for pancreatic cancer compared to participants with a HLI score of 5-9 (reference). The HR (RR) is 0.64, meaning a 0.64 times lower risk. The CI is between 0.52 and 0.78, so not including the neutral value of 1, meaning that the result is statistical significant.

Question 5 Statistics - Dr. Rogier Donders (15 points)

Use the abstract, methods and tables of 'Naudin et al. Healthy lifestyle and the risk of pancreatic cancer in the EPIC study. European Journal of Epidemiology (2020) 35:975–986.

A. This study is based on a rather big cohort (400,577 participants). According to the abstract, in this cohort 1113 subjects with pancreatic cancer (PC) were identified. Show that the sample proportion of PC cases equals approximately 0.0028. (1 pt)

P = 1113/400577 ≈ 0.0028

B. Assume the population proportion is exactly equal to the sample proportion. What would be the standard error of the sample proportion? (3 pts)

SE(p) = √{π*(1 – π)/n} = √(0.0028*0.9972/400577) ≈ 0.000084

C. The cohort consists for 75% of women. This is clearly not a representative sample from the European population. Give a clear motivation why this is not a representative sample using only the sex ratio in your answer. (3 pts)

In Europe there are approximately equal numbers of males and females. In this sample this is clearly not the case. And given the size of the cohort (and the corresponding small standard errors) this is a clear indication that men are underrepresented.

D. Of the 1113 PC cases, 57% were women. Who have a larger probability of developing PC: males or females? (4 pts)

Males have a larger probability. If males and females would have equal probabilities of developing PC the sex ratio among the PC cases would be identical to the sex ratio in the complete sample (75%). Since there are relatively fewer women among the PC cases (57%), women are relatively underrepresented amongst the PC case and appear to have a smaller probability of developing PC.

E. In Question 2 a possible biomarker for PC tumour grow is presented. This is based on SLA-A concentration. This of course has to be measured. In order to see whether it is possible to measure this correctly the following procedure is used. Blood is sampled from a single patient with PC. This is split into 10 portions and the SLA-A concentration in each portion is measured. Discuss what aspects of measurement error (systematic measurement error, random measurement error, absolute measurement error, relative measurement error) can be studied in this set up? (4 pts)

Question 6 Genetic lab practice - Dr. Diederik de Bruijn (10 points)

A. The Polymerase Chain Reaction (PCR) and the Sanger sequencing reaction are very similar at the molecular level, especially with regard to the contents of the reaction mix. Name two ingredients that are used in both reaction mixes and name one ingredient that is only used in the Sanger sequencing reaction mix. (3 pts)

Ingredients in both reaction mixes:
1. Nucleotides
2. Primer (oligonucleotide);
3. (Taq) polymerase.

Ingredient unique for Sanger Sequencing reaction mix
1. Fluorescently labeled dideoxy nucleotide
2. PCR product

1 point per good answer

Below, you see a snapshot from the 1000Genomes browser with genotype information of SNP rs243865. This snapshot contains genotype frequencies and numbers of persons per genotype (between brackets) for four different populations from across the world. These are people from: Africa (AFR), America (AMR), Europe (CEU) and South Asia (SAS).

1000GENOMES phase_3: **AFR**	C|C: 0,944 (624)	C|T: 0,054 (36)	T|T: 0,002 (1)
1000GENOMES phase_3: **AMR**	C|C: 0,634 (208)	C|T: 0,29 (95)	T|T: 0,076 (25)
1000GENOMES phase_3: **CEU**	C|C: 0,547 (275)	C|T: 0,388 (195)	T|T: 0,066 (33)
1000GENOMES phase_3: **SAS**	C|C: 0,761 (373)	C|T: 0,227 (111)	T|T: 0,012 (6)

B. Use the data from the above snapshot to calculate the *allele frequency* of the C and T alleles in the AMR and SAS populations (use box 1, **below**). Next calculate the expected *genotype frequencies* (CC, CT and TT) for both populations under assumption of Hardy-Weinberg Equilibrium (use box 2, **below**). Show your calculations (4 pts)

Box 1 (3 decimals)	C	T
Allele frequency AMR	*0.779*	*0.221*
Allele frequency SAS	*0.874*	*0.126*

***Method 1**: count the number of alleles:*
AMR-C: 2x208 + 95 = 511; AMR-T: 2x25 + 95 = 145; Total: 511 + 145 = 656 (= 2 x 328).
Allele frequencies: AMR-C allele: 511 / 656 = 0.779; JPT-T allele: 145 / 656 = 0.221
***Method 2**: Add up frequency homozygotes with half frequency heterozygotes*
AMR-C: 0.634 + 0.290 / 2 = 0.779; AMR-T: 0.076 + 0.290 / 2 = 0.221

Points awarded <u>per population</u>:
1: Right answer and calculations shown (1 pt)

Box 2 (3 decimals)	CC	CT	TT
Genotype frequency AMR	*0.607*	*0.344*	*0.049*
Genotype frequency SAS	*0.764*	*0.220*	*0.016*

***Method**: use HW equation to calculate genotype frequencies:*
  AMR-CC: (0.779)² = 0.607
  AMR-CT: 2 x 0.779 x 0.221 = 0.344
  AMR-TT: (0.221)² = 0.049

C. Use the genotype frequencies from the previous question to calculate the expected number of people (use Box 3A and 3B) and compare this with the observed numbers of people for each genotype in the AMR and SAS populations. Evaluate whether the Hardy-Weinberg equilibrium seems to hold for these populations. (3 pts)

Box 3A (whole numbers)	CC	CT	TT
Observed (AMR)	*208*	*95*	*25*
Expected (AMR)	*199*	*113*	*16*

Box 3B (whole numbers)	CC	CT	TT
Observed (SAS)	*373*	*111*	*6*
Expected (SAS)	*374*	*108*	*8*

__Calculation of expected number of people (1 pt)__: Multiply genotype frequencies with total number of persons per population
AMR population size = 328:
  AA: 0.607 x 328 = 199.1
  AG: 0.344 x 328 = 112.8
  GG: 0.049 x 328 = 16.1

Evaluation AMR (1 pt)*: <u>Major</u> differences (>10% deviation) between observed and expected (calculated) genotype distributions (0.5 pt). Not consistent with HWE (0.5 pt).*

Evaluation SAS (1 pt)*: <u>Minor</u> differences (<10% deviation) between observed and expected (calculated) genotype distributions (0.5 pt). Consistent with HWE (0.5 pt).*