Research Exam Semester 1 – 2018-2019
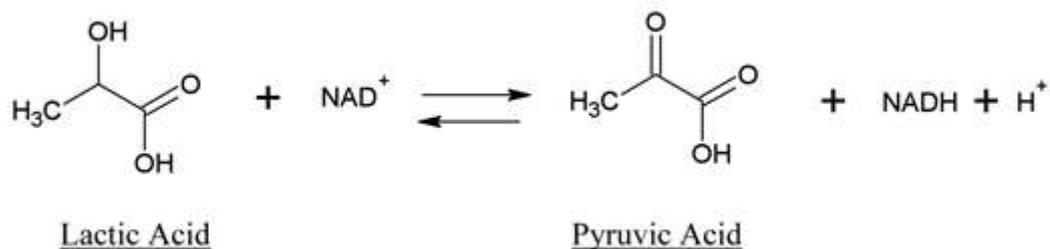# January 18, 2019

Question 1
Measuring enzymatic reactions – dr. G. Bosman
(10 points)

Lactic Acid   Pyruvic Acid

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

a. Enzymes are divided in six classes. To which class does LDH belong? Explain your answer (1 pt)

   Oxidoreductases: electron transfer is part of the reaction mechanism

b. Give a description of a laboratory method, using a spectrophotometer, for measuring the activity of LDH. (2 pt)
1. Measure the decrease of lactic acid or the decrease of NAD+ or the increase in pyruvic acid or NADH after addition of LDH (1 pt)
2. Measure the increase of NADH using the difference in absorption at 340 nm between NADH en NAD+ (2 pt)

c. Your task is the determination of the Michaelis-Menten constant and the maximal velocity of LDH. Describe the various steps that need to be taken. (5 pt)
1. Measure the conversion of various lactate concentrations at different time points bij
2. Use these data to determine initial velocities
3. Graph initial velocity versus concentration
4. Extrapolate to estimate Vmax
5. Use the graph with Vmax to determine Km

d. The concentration of LDH in the blood is an indicator for hemolysis, e.g. as a consequence of a transfusion reaction. Which concentration of lactate, <u>expressed in terms of Michaelis-Menten-constant and/or maximal velocity</u>, should be used to determine the LDH concentration in blood to check for hemolysis? Explain your answer. (2 pt)
Concentration lactate as high as possible above Km → velocity close to Vmax: velocity changes with enzyme concentration (derived as in Baynes chapter Enzymes).

Question 2
Modelling: Zero-order kinetics – dr. T. Oostendorp
(15 points)

When using single-compartment models to study pharmacodynamics it is often assumed that the efflux $\Psi(t)$ of a drug from a compartment is proportional to the concentration $c(t)$ of that drug in the compartment: $\Psi(t) = k\,c(t)$.

However, for drugs like fluoxetin, the efflux is, in good approximation, independent of the value of the concentration: $\Psi(t) = R$. This is called zero-order elimination.

a. Show that for fluoxetin the differential equation that describes the pharmacokinetics in a single compartment model is

$$V\frac{d}{dt}c(t) = D(t) - R$$

where V is the volume of the compartment, and $D(t)$ the dose velocity. (3 pts)

Change = input – output.

The input is the dose velocity, and the output is the efflux. The change is the change in *amount* of drug $= \frac{d}{dt}(Vc(t))$. The constant V can be taken out from the derivative.

b. In what units can R be expressed? There are several possibilities, just give one. (2 pt)

Anything that represents amount/time. For example: mg/s, mol/hour etc

c. Show that, according to the above model, an injected of fluoxetin at $t = 0$ will lead to a concentration as a function of time given by (4 pts)

$$c(t) = c_0 - \frac{R}{V}t$$

After the drug has been administered the dose velocity, the dose velocity is zero. We then have

$$V\frac{d}{dt}c(t) = -R$$

We have to check whether $c(t) = c_0 - \frac{R}{V}$ is a solution to this equation. For this, we have to insert $c(t) = c_0 - \frac{R}{V}t$ into the differential equation:

$$V\frac{d}{dt}c(t) = V\frac{d}{dt}\left(c_0 - \frac{R}{V}t\right) = V\left(-\frac{R}{V}\right) = -R$$

so $c(t) = c_0 - \frac{R}{V}t$ fits the differential equation.

In the introduction, it was stated that zero-order elimination is a good approximation for fluoxetin. However, this cannot be a good approximation anymore when a considerable amount of time has passed after injection with fluoxetin.

d. Explain why. (3 pts)

The model predicts that after some time ($t = Vc_0/R$ to be precise) the concentration becomes negative; that cannot happen in reality.

e. Under what circumstances will an infusion of fluoxetin with a constant dose velocity D lead to a constant concentration in the above model? Explain your answer. (3 pts)

A constant concentration occurs when $\frac{d}{dt}c(t) = 0$. We see from the differential equation in question a that this occurs if $D(t) = R$. For any other dose velocity, the concentration is not constant.

Question 3
Signals: Electromyography – dr. T. Oostendorp
(10 points)

In electromyography, the electric activity of a muscle is recorded by means of electrodes attached to the skin covering the muscle. Figure 3.1 shows a typical setup and a typical recording without filtering.
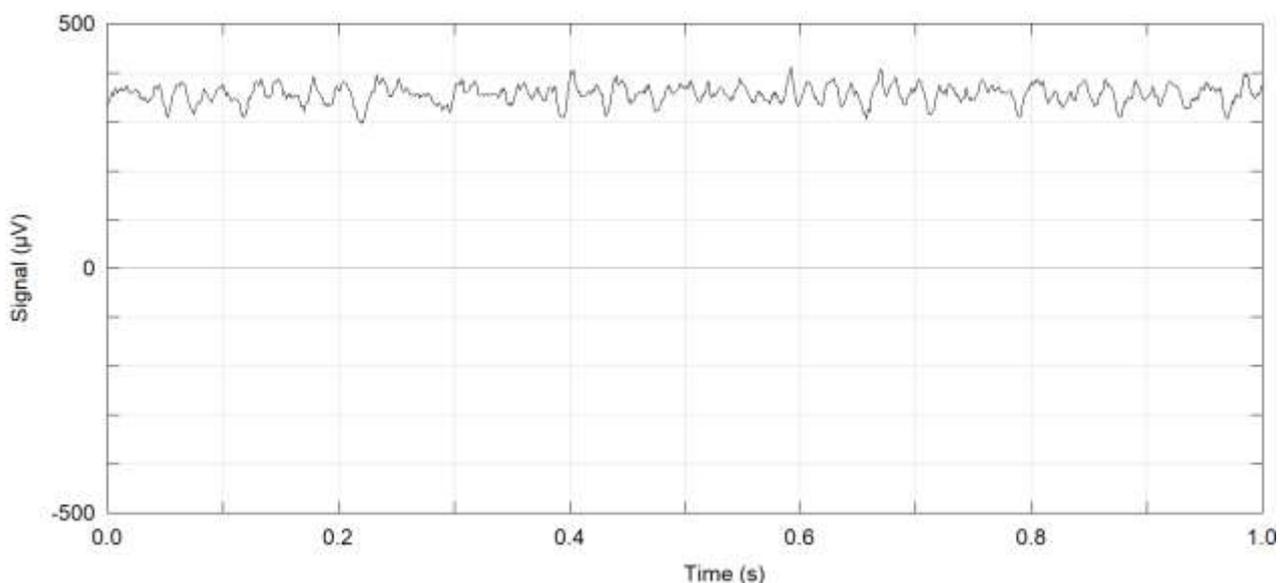

Figure 3.1   EMG recording and setup.

Figure 3.1 shows a considerable offset (vertical shift) in the EMG signal. Apparently, apart from the EMG there is constant potential of about (in this case) 350 μV. This occurs quite often when recording electrical signals from the body.

a. Explain what causes this offset (3 pt)

The electrode-skin interface functions as a battery, that generates a constant voltage of several 100 μV.

In general, amplifiers that record electro-physiological signals are equipped with a filter that eliminates this extra constant voltage.

b. Plot the frequency response of this filter. Also, add adequate label to the axes. (4 pts)
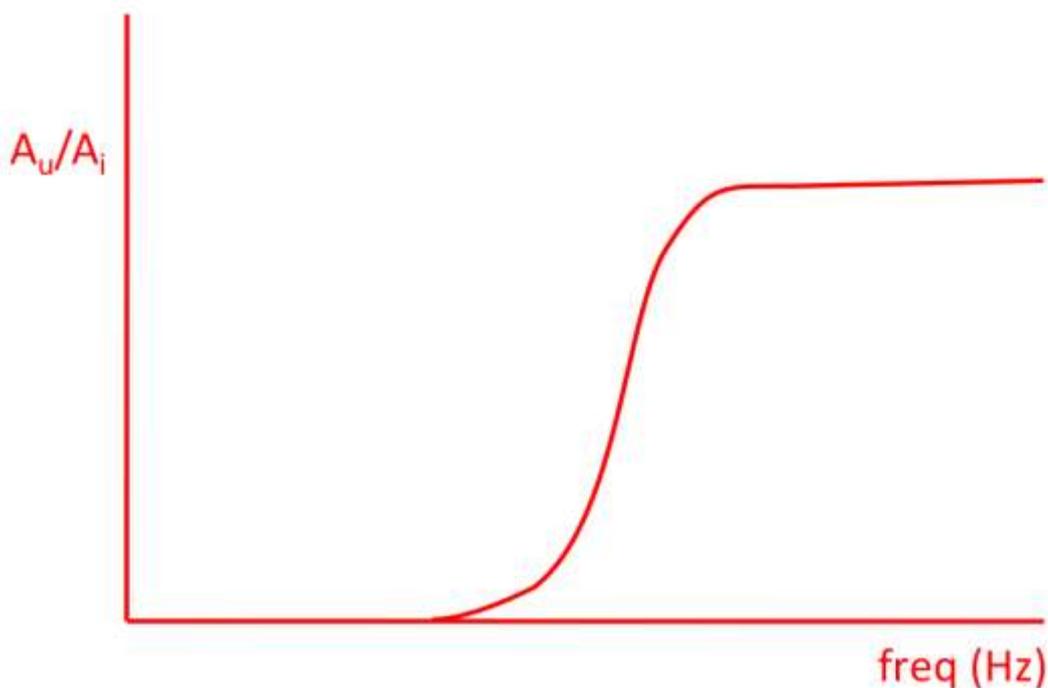

Figure 3.2 shows the recording after applying this filter. Figure 3.3 shows the spectrum of the filtered signal.
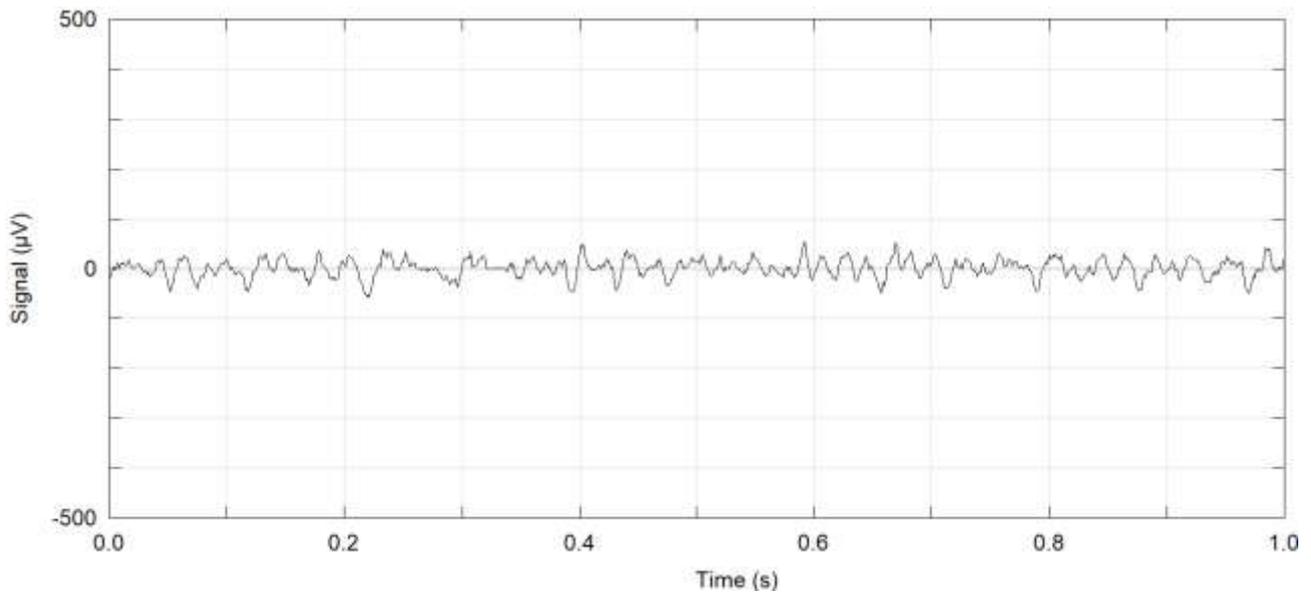
Figure 3.2   EMG recording after filtering.
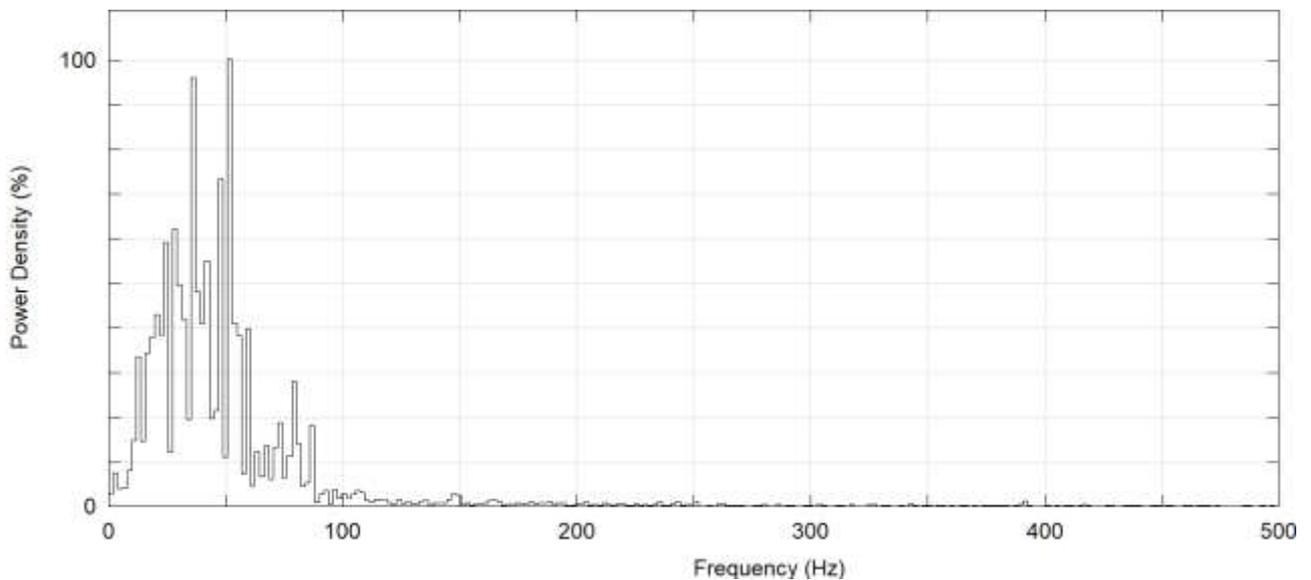
Figure 3.3   Spectrum of filtered EMG.

c. What is the minimum sample rate that can be used to record this signal while avoiding aliasing? Explain your answer. (3 pts)

Twice the maximum frequency present in the signal is needed to avoid aliasing. The maximum frequency is about 100 Hz, so a minimum sample rate of 200 Hz is needed.

(it cannot be determined very precisely what the maximum frequency in the signal is; other reasonable values are also accepted).

Question 4
Methods, measures and study designs in epidemiological studies – dr. F. de Vegt (20 points)

Use '*Brindel et al. Anthropometric factors in differentiated thyroid cancer in French Polynesia: a case–control study – abstract and tables*' for question 4 and 5.

a. In the study of Brindel et al. anthropometric factors have been measured. You can obtain information on weight and height by 1) questionnaire and by 2) measuring participants on a scale. Mention one advantage and one disadvantage for both measurement methods. (4 pts)

   1. Questionnaire, asking for body weight and height.
   + easy to obtain in a large group of people
   - selfreporting, people are reporting lower weight compared to the real body weight they have

   2. Measuring weight on a scale
   + more objective compared to questionnaire data
   - Takes more time and effort to obtain: participant needs to come to hospital / research center
   - Need equipment (scale) and researchers to measure

b. In this study, a case-control design has been used. Who were the cases and who the controls? (2 pts)

   Cases were participants with differentiated thyroid cancers (1 pnt) diagnosed between 1979 and 2004 (195 women/24 men) and controls were 359 population controls (315 women/44 men) without thyroid cancer (1 pnt).

c. Why is it not feasible to address the same research question in a randomized clinical trial? Explain your answer (3 pts)

   In a clinical trial, participants will be randomized in an intervention and control group. Here it should be a group with obesity and a group without obesity and you cannot randomize people to these categories (not feasible and not ethical) (2 pts). Then, you will have to record incident thyroid cancer in both groups. In addition, it would require large populations and/or long follow-up duration to have a number of participants with the outcome of interest – thyroid cancer (1 pnt).

d. See Table 2, the odds ratio of 1.6 (0.8-3.2). Give an interpretation of this result (4 pts).

Women with a BMI between 20.5-22.5 kg/m2 at age 18 have a 1.6 times higher risk for thyroid cancer compared to women with BMI ≤18.7 kg/m2 at age 18. (2 pts). The CI is between 0.8 and 3.2, so including the 'neutral OR' value of 1, meaning that the result is not statistical significant (2 pts).

e. See Table 2, first part, concerning **BMI before diagnosis**. Fill out the 2 × 2 table and calculate the odds ratio (4 pnt).

	cases	controls	total
BMI 20.5 – 22.9 kg/m2	35	77	112
BMI < 20.5 kg/m2	23	75	98
total	58	152	210

Table: 2 pts

Odds ratio: (35/23) / (77/75) = 1,48 of cross product: (35*75) / (77 * 23) = 1,48 (2 pts)

f. See Table 2. Why are the odds ratios adjusted for smoking? (3 pnt)

Smoking might be a confounder, disturbing the relation between BMI and thyroid cancer (2 pts), because smoking is a risk factor for thyroid cancer itself and also associated with obesity (2 pts)

Question 5
Statistics: dr. R. Donders
(15 points)

Use '*Brindel et al. Anthropometric factors in differentiated thyroid cancer in French Polynesia: a case–control study – abstract and tables*' for question 4 and 5.

In table 5.1 in the paper of Brindel et al. the histological sizes of the papillary thyroid tumours are presented. However, the percentages in that table also take the unknown values into account. In the table below, percentages are calculated on basis of all patients with a known tumour size (column '% corrected'). Note that microcarcinoma cannot be placed in any of the size categories, so these are also assumed to be of unknown size.

Table 5.1. Histological size of the papillary thyroid tumours (in mm).

Females	frequency	percentage	% corrected	Males	frequency	percentage	% corrected
≤4	30	19.7	22.4		2	11.8	13.3
5–10	50	32.9	37.3		5	29.3	33.3
11–20	22	14.5	16.4		2	11.8	13.3
21–40	25	16.4	18.7		4	23.5	26.7
>40	7	4.6	5.2		2	11.8	13.3
Microcarcinoma	6	4			0	0	
Unknown	12	7.9			2	11.8	

a. Show for one of the corrected percentages how it is calculated. (2 pts)

For one of the genders, add the number of known sizes. For example, there are 2 + 5 + 2 + 4 + 2 = 15 males with a known tumour size. For the category ≤4 the corrected percentage is (2/15)*100% = 13.3%.

b. The median tumour size is smaller for the females than for the males. Explain how you can derive the order of the sample medians. (4 pts)

If you calculate the cumulative percentages for the two genders, it becomes clear that 59.7% of the females have a tumour smaller than or equal to 10 mm. So the median size for the females will be smaller than or equal to 10 mm. However, only 46.7% of the males have a tumour smaller than or equal to 10 mm. So the median size for the males must be larger than 10 mm.

c. From Table 1 from the paper, one can derive that 6 of the 24 males were between 18 and 29 at the time of diagnosis. This is of course 25% or 0.25. Assume that by sheer coincidence this would be equal to the population proportion. What would be standard error of the sample proportion? (3 pts)

The standard error of a sample proportion is equal to , with π equal to the population proportion (0.25) and n equal to the sample size (24). So the standard error would be equal to √(0.25*0.75/24) ≈ 0.0884

In Table 2 of the paper, the relation between the presence of thyroid cancer and BMI is presented. This is done both for BMI at diagnosis and BMI at age 18. Assume that the BMI at diagnosis is calculated using measurements made in the hospital during the physical examination of the patient and that the BMI at age 18 is calculated using data about their length and weight at age 18 filled in by the patients in a questionnaire.

d. Which of the BMI measurements will have the largest random measurement error? Give a proper motivation.(3 pts)

The measurements made during the physical examination will be precise and contain little random error. Thus the derived BMI will also be precise and contain little random error, The estimates for length and weight at age 18 come from the memory of the participants and will be much less precise and contain more random measurement error. Combining these in a BMI will result is a measurement with more random measurement error.

e. For which of the BMI measurements is it more probable that a systematic measurement error will occur? Give a proper motivation.(3 pts)
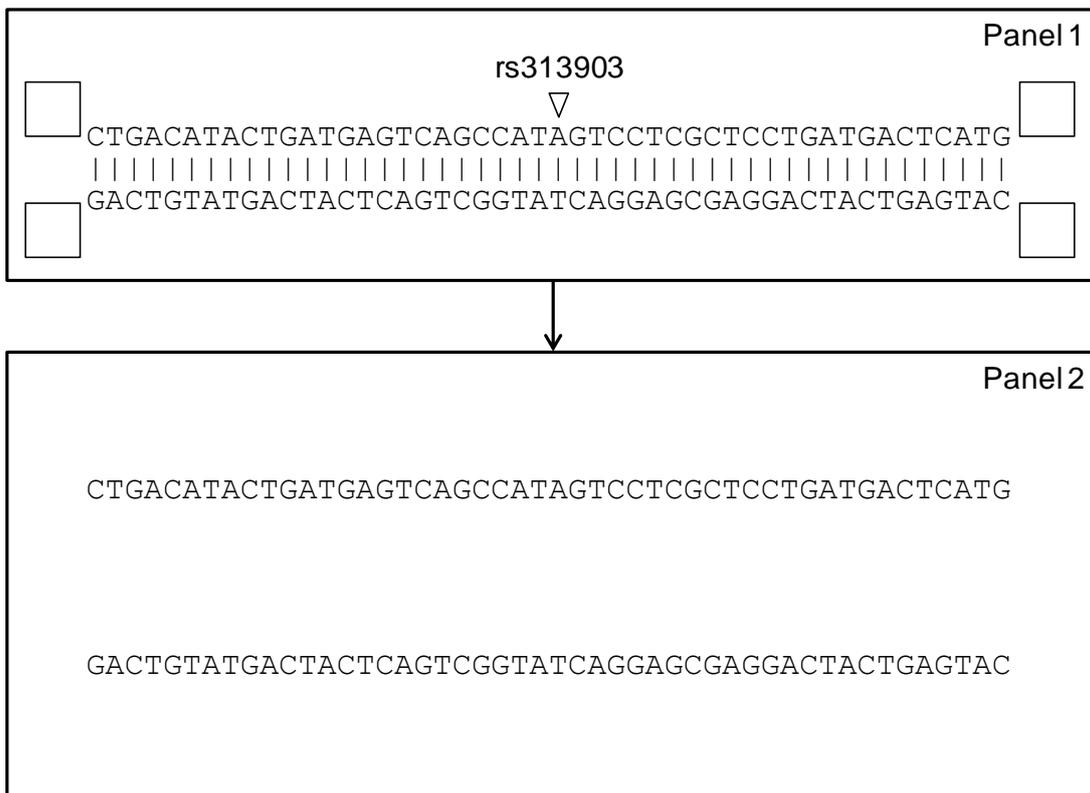
The measurements made during the physical examination will most likely contain no systematic error, but people tend to overestimate their length and underestimate their weight. So it is more likely that the measurements for BMI at age 18 will contain a systematic measurement error: these might well be underestimated.

Question 6
Genetic lab practice research: dr. D. de Bruijn
(10 points)

You are planning to perform a Sanger sequencing experiment to read a DNA fragment containing the A/C SNP rs313903. Panel 1 of the figure below displays the target sequence, a double stranded PCR product (of which only 48 base pairs are shown) with SNP rs313903 in the middle. Hydrogen bonds between the bases of both strands are depicted with single vertical lines ("|"). Panel 2 contains the same DNA sequence, but now as two single stranded DNA molecules. This figure will be used for questions a-b.

Panel 1

rs313903
▽

CTGACATACTGATGAGTCAGCCATAGTCCTCGCTCCTGATGACTCATG
||||||||||||||||||||||||||||||||||||||||||||||||
GACTGTATGACTACTCAGTCGGTATCAGGAGCGAGGACTACTGAGTAC

Panel 2

CTGACATACTGATGAGTCAGCCATAGTCCTCGCTCCTGATGACTCATG

GACTGTATGACTACTCAGTCGGTATCAGGAGCGAGGACTACTGAGTAC

```
                                                                    Panel 1
                           rs313903
                              ▽
  5'
       CTGACATACTGATGAGTCAGCCATAGTCCTCGCTCCTGATGACTCATG     3'
       ||||||||||||||||||||||||||||||||||||||||||||||||
  3'   GACTGTATGACTACTCAGTCGGTATCAGGAGCGAGGACTACTGAGTAC     5'
```

```
                                                                    Panel 2


       CTGACATACTGATGAGTCAGCCATAGTCCTCGCTCCTGATGACTCATG


           ACTGA  ─────────────────────────────────────▶
           |||||
       GACTGTATGACTACTCAGTCGGTATCAGGAGCGAGGACTACTGAGTAC
```

a. Give the correct names of the ends of all DNA strands in Panel 1. Do this in the four small boxes left and right of the double stranded DNA molecule in Panel 1. (1 pt)

b. Use Panel 2 to draw one oligonucleotide with a length of 5 bases that can be used as a primer to directly read the bases of the top strand of this DNA fragment. Allow for at least 12 bases between your oligonucleotide and the SNP location. Draw the bases of the oligonucleotide at the correct place, add the hydrogen bonds that will be formed and use an arrow to indicate the direction in which nucleotides will be built in during the Sanger sequencing reaction. (2 pts)

c. Explain what an indel is and how this type of genetic variation can be detected with Sanger sequencing. (2 pts)

An indel is a short insertion or deletion. With Sanger sequencing, you can read over this indel and determine what nucleotides have been added or deleted. Correct explanation of indel (1 pt), correct explanation of measurement (1 pt).

Below, you see a snapshot from the 1000Genomes browser with genotype information of SNP rs798766. This table contains genotype frequencies and numbers of persons per genotype (between brackets) for four different European populations (i.e. CEU=Northern and Western European ancestry, FIN=Finnish, GBR=British, IBS=Spanish).

1000GENOMES phase_3: **CEU**	**T|T**: 0.030 (3)	**C|T**: 0.354 (35)	**C|C**: 0.616 (61)
1000GENOMES phase_3: **FIN**	**T|T**: 0.040 (4)	**C|T**: 0.283 (28)	**C|C**: 0.677 (67)
1000GENOMES phase_3: **GBR**	**T|T**: 0.044 (4)	**C|T**: 0.429 (39)	**C|C**: 0.527 (48)
1000GENOMES phase_3: **IBS**	**T|T**: 0.056 (6)	**C|T**: 0.318 (34)	**C|C**: 0.626 (67)

d   Use the data in the above table to calculate the *allele frequency* of the T and C alleles in the GBR population (box 1) and calculate the expected *genotype frequencies* (TT, CT and CC) for this population under assumption of Hardy-Weinberg Equilibrium (box 2). Write down your calculations and round your results to 2 decimals. (4 pts)

Box1 (2 decimals)	T	C
Allele frequency GBR	*0.26*	*0.74*

Method 1: count the number of alleles:
T: 2x4 + 39 = 47; C: 2x48 + 39 = 135; Total: 47 + 135 = 182 (= 2x91).
Allele frequencies: T allele: 47 / 182 = 0.26; C allele: 135 / 182 = 0.74
Method 2: Add up frequencies and divide hz by 2
T: 0.044 + (0.429 / 2) = 0.26; C: 0.527 + (0.418 / 2) = 0.74

Errors:
1: Wrong method: square root of TT and/or CC frequencies (0 pt)
2: Wrong method = (freq TT)^2 + freq CT (0 pt)
3: No calculation shown, but right answer (1 pt)
4: No calculation shown, wrong answer (0 pt)
5: Possibly right method 2, but not divided by 2 (1 pt)
6: Number of total alleles = number of persons (1 pt)
7: Minor calculation or rounding error (- 0.5 pt)
8: Any other calculation error (- 1 pt)

Box 2 (2 decimals)	TT	CT	CC
Genotype frequency GBR	*0.07*	*0.38*	*0.55*

Method: use HW equation to calculate genotype frequencies:
  TT: (0.26)^2 = 0.07
  TC: 2 x 0.26 x 0.74 = 0.38
  CC: (0.74)^2 = 0.55

Errors:
1: Right method with wrong numbers from box 1 (2 pt)
2: Calculation error (- 1 pt)
3: Partial reasoning (HWE equation) shown, answer consistent with box 1 (- 0.5 pt)
4: No calculation shown, but right answer (1 pt)
5: No calculation shown, wrong answer (0 pt)
6: Small error due to rounding (- 0.5 pt)
No penalty if errors from box 1 lead to total frequency unequal to 1 in this answer

e. Use the genotype frequencies from the previous question to calculate the expected number of people (use Box 3) and compare this with the actual numbers of people for each genotype in the GBR population. What is your evaluation of the differences between these genotype distributions? (1 pt)

Box 3 (whole numbers)	TT	CT	CC
Actual number of people	*4*	*39*	*48*
Expected number of people	6	35	50

Some differences between the actual and calculated genotype distributions, but still consistent with HWE (1 pt). No mentioning of any differences or no mentioning of HWE (-0.5 pt). Alternative answers: max. 0.5 pt for an adequate evaluation of absolutely false results due to errors in previous questions.