**Diabetes**

Source: Pfannkuch, M., Seber, G., & Wild, C.J. (2002). Probability with less pain.

*Teaching Statistics,* 24(1) 24-30

Student worksheet

Describe what you know about diabetes and why it is of concern in New Zealand.

A standard test for diabetes is based on glucose levels in the blood after fasting for a prescribed period. For ‘healthy’ people, the mean fasting glucose level is 5.31 mmol/L and the standard deviation is 0.58 mmol/L. For untreated diabetes the mean is 11.74 and the standard deviation is 3.50. In both groups the levels appear approximately Normal.

Graph these distributions on the axes below:

Describe the two distributions.

1. ‘healthy’ people’s blood glucose levels
2. blood glucose levels for people with diabetes

Task 1

You are Doctor Diab, who has to decide the cut-off point, C, (in mmol/L) using this test. If a person’s fasting glucose level is above *C* they ‘test positive’ and we say that they have diabetes, while if it is below *C* they ‘test negative’ and we say that they do not have diabetes.

What are the implications for some people who are close to the cut-off point?

Shade regions on the graph to indicate proportions of people who would be wrongly diagnosed and indicate the disease status and test result for these groups.

The *sensitivity* of such a test is the probability of the test giving a positive result when the person has the condition. (true positive)

The *specificity* of such a test is the probability of the test giving a negative result when the person does not have the condition. (true negative)

Assume that the cut-off point is 6.5mmol glucose/L blood.

Calculate:

P(test is negative | person does not have diabetes)

P(test is positive | person has diabetes)

In 2012, 225 686 people in New Zealand had been diagnosed with diabetes out of an estimated total population of 4 433 000.

Calculate the base rate (proportion of the population with diabetes)

Base rate =

Suppose there was a screening programme introduced where the entire population of New Zealand was tested for diabetes using this test and the cut-off point was taken as 6.5mmol/L

Set up a Tinkerplots simulation\* for this base rate and test cut-off value.

From your simulation fill out the table of counts:

Base rate = 5%; cut-off =6.5mmol/L

|  |  |  |  |
| --- | --- | --- | --- |
|  |  | Test result | |
| Disease status |  | Positive | Negative |
| Has diabetes |  |  |
| Does not have diabetes |  |  |

Use your table to estimate:

P(does not have diabetes | test positive) =

P(has diabetes | test negative)

Use the model you have created in Tinkerplots to investigate the effect of different cut-off values, keeping the base rate the same.

1. Base rate = 5% cut-off = ……..mmol/L

P(test is negative | person does not have diabetes) =

P(test is positive | person has diabetes) =

|  |  |  |  |
| --- | --- | --- | --- |
|  |  | Test result | |
| Disease status |  | Positive | Negative |
| Has diabetes |  |  |
| Does not have diabetes |  |  |

Use your table to estimate:

P(does not have diabetes | test positive) =

P(has diabetes | test negative)

b). Base rate = 5% cut-off = ……… mmol/L

P(test is negative | person does not have diabetes) =

P(test is positive | person has diabetes) =

|  |  |  |  |
| --- | --- | --- | --- |
|  |  | Test result | |
| Disease status |  | Positive | Negative |
| Has diabetes |  |  |
| Does not have diabetes |  |  |

Use your table to estimate:

P(does not have diabetes | test positive) =

P(has diabetes | test negative)

c). Base rate = 5%; cut-off = ……… mmol/L

P(test is negative | person does not have diabetes) =

P(test is positive | person has diabetes) =

|  |  |  |  |
| --- | --- | --- | --- |
|  |  | Test result | |
| Disease status |  | Positive | Negative |
| Has diabetes |  |  |
| Does not have diabetes |  |  |

Use your table to estimate:

P(does not have diabetes | test positive) =

* 1. P(has diabetes | test negative) =

What cut-off point, C, would you, Dr Diab, recommend as suitable for this test for making the decision about whether a person in a screening programme is or is not diabetic. Explain why you would recommend this cut-off value.

Task 2

Use your Tinkerplots model to investigate what happens if the base rate is different than 5%. In New Zealand, the base rate for diabetes varies for different ethnic groups and ranges between about 2 and 11%.

The 2002–2003 New Zealand Health Survey of people over fifteen years of age estimated that the self-identified prevalence of diagnosed diabetes within each ethnicity was as follows: Maori 8.0%, Pacific 10.1%, Asian 8.4% and European (including ‘Other’) 2.9%, so disparities in prevalence exist [5].Ethnic disparities have also existed in the management and complications of diabetes in New Zealand [6].

<http://www.diabetes.org.nz/news/nz_news/trends_in_the_management_of_risk_of_diabetes_complications>

Use the cut-off value, C, which you recommended in Task 1 for each simulation.

1. Base rate = ……%; cut-off = ……..mmol/L

P(test is negative | person does not have diabetes) =

P(test is positive | person has diabetes) =

|  |  |  |  |
| --- | --- | --- | --- |
|  |  | Test result | |
| Disease status |  | Positive | Negative |
| Has diabetes |  |  |
| Does not have diabetes |  |  |

Use your table to estimate:

P(does not have diabetes | test positive) =

P(has diabetes | test negative)

b). Base rate = …….. % cut-off = ……… mmol/L

P(test is negative | person does not have diabetes) =

P(test is positive | person has diabetes) =

|  |  |  |  |
| --- | --- | --- | --- |
|  |  | Test result | |
| Disease status |  | Positive | Negative |
| Has diabetes |  |  |
| Does not have diabetes |  |  |

Use your table to estimate:

P(does not have diabetes | test positive) =

P(has diabetes | test negative)

Describe the effect of different base rates on the percentage of people who would get an incorrect diagnosis.

Compare the probabilities that a person having a positive result actually had diabetes if the person was a Pacific Islander or a European.

Task 3 (Extension): Estimating the base rate for diabetes

We would like to use probability to estimate the base rate, that is, the proportion of the population of New Zealand who do have diabetes.

We want to estimate ***d*** = P(a person in New Zealand has diabetes)

To find a good estimate of the proportion of people in the population who test positive for diabetes, we collect data from a random sample of people.

Suppose that 10.3% of such a sample test positive.

Use the cut-off value, C =6.5 mmol/L.

Base rate = ***d*** cut-off =6.5mmol/L

P(test is negative | person does not have diabetes) = [from task 1]

P(test is positive | person has diabetes) =

Fill in the table:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  |  | Test result | | |
| Disease status |  | Positive | Negative | Total |
| Has diabetes |  |  | d |
| Does not have diabetes |  |  |  |
| Total | 0.103 |  | 1 |

Now solve an equation to calculate ***d.***

Conclusion:

If 10.3% of a random sample of New Zealanders tested positive for diabetes then about …………..% of the population are estimated to have the disease.



This work is supported by:

The New Zealand Science, Mathematics and Technology Teacher Fellowship Scheme which is funded by the New Zealand Government and administered by the Royal Society of New Zealand

and



Department of Statistics

The University of Auckland