

No part of the candidate's evidence in this exemplar material may be presented in an external assessment for the purpose of gaining an NZQA qualification or award.

SUPERVISOR'S USE ONLY

S

93201



932010

Draw a cross through the box (X) if you have NOT written in this booklet

☐

+

OUTSTANDING SCHOLARSHIP EXEMPLAR



Mana Tohu Mātauranga o Aotearoa
New Zealand Qualifications Authority

Scholarship 2024 Statistics

Time allowed: Three hours
Total score: 32

Check that the National Student Number (NSN) on your admission slip is the same as the number at the top of this page.

You should answer ALL the questions in this booklet.

Pull out Formulae and Tables Booklet S–STATF from the centre of this booklet.

Show ALL working.

Check that this booklet has pages 2–24 in the correct order and that none of these pages is blank.

Do not write in any cross-hatched area (⚡). This area may be cut off when the booklet is marked.

YOU MUST HAND THIS BOOKLET TO THE SUPERVISOR AT THE END OF THE EXAMINATION.

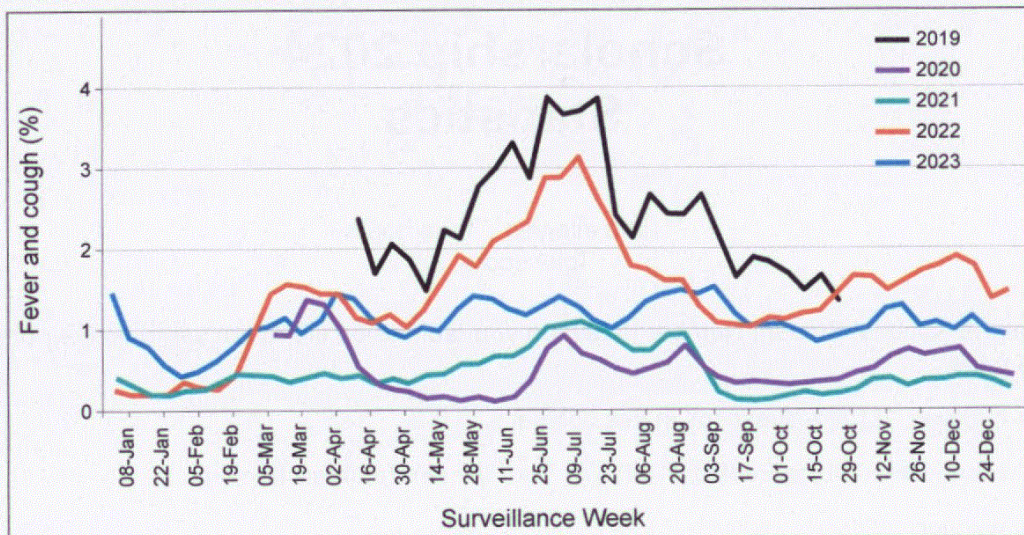
QUESTION ONE

FluTracking is a surveillance system that harnesses both the power of the internet and community spirit for monitoring influenza. Each week, over 110 000 volunteer participants in New Zealand and Australia fill in an online survey about their illness symptoms during the previous week, including whether they have experienced the following:

- fever
- cough
- sore throat
- runny nose
- shortness of breath
- any change in sense of taste or smell
- headache.

- (a) Figure 1 shows the incidence rate of participants in New Zealand who reported having both a fever and a cough during the surveyed week.

Figure 1: FluTracking New Zealand weekly fever and cough symptoms (%) 2023



Adapted from: www.flutracking.net/Info/Report/202353/NZ

- (i) Write one short paragraph describing the key features of the data for the incidence of fever and cough symptoms over the period April 2019 to December 2023.

The first year, 2019, has almost always the highest incidence of cough and fever symptoms for each ~~fortnightly~~ ^{weekly} period. It also, like all the years other than 2023, sees a noticeable spike in Jun-Aug, the winter months. In 2019 this peak reaches almost 4% in the June 25 week. 2020 has no data in Jan and Feb, and ends March with a mild increase. It troughs until the winter flu season in June and July. This trough period is lower than any other year. Other

than winter 2022, the years after 2019 have not surpassed 2% weekly cough and fever symptoms.

- (ii) Discuss THREE possible limitations of the study design used by FluTracking, including the impact of the design on any conclusions that could be made.

Because the survey is self reported, and conducted online participants may not consider their symptoms severe enough to record. This could result in the true rate of symptoms being higher than surveyed.

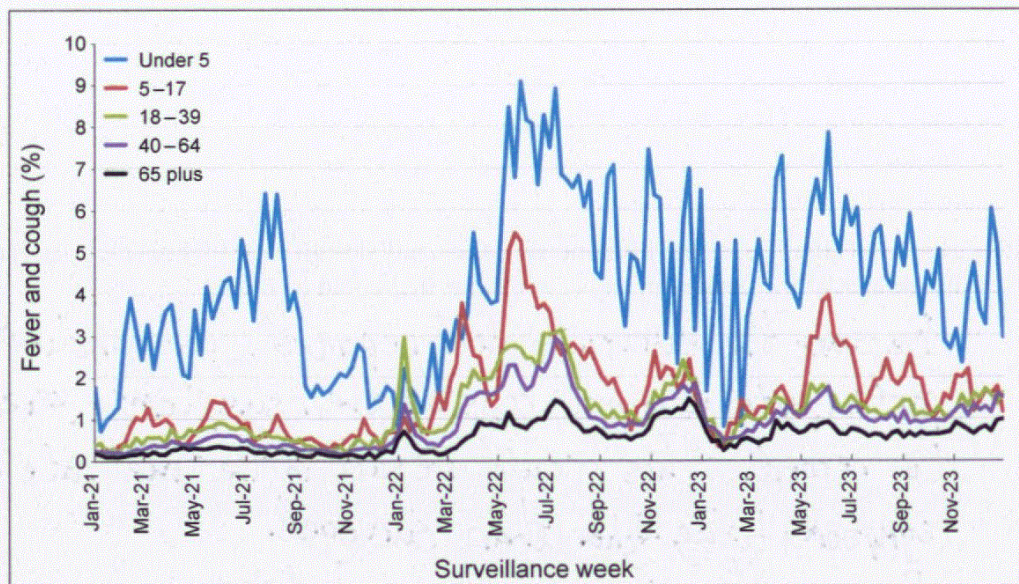
Because the survey uses volunteers, more commonly called self selection sampling, the sample may ~~be more~~ have a higher proportion of chronically ill or people with poor immune systems. This may limit the survey's findings to a subset of the population. In other words, healthier people are less likely to use the survey.

Similarly as the survey is conducted weekly, it is possible that people not experiencing symptoms may not report on the survey, due to forgetting. Online surveys have no way of ensuring that participants follow up regularly, relying on the participant's memory. This means the value may be higher than the true amount of symptoms.

The survey is also conducted online, which also ~~is~~ means that ~~for~~ affluent respondents, as well as younger respondents will be ~~prioritised~~ more common due to technological literacy or barriers to access. This affects the applicability of the results of the study on the wider population.

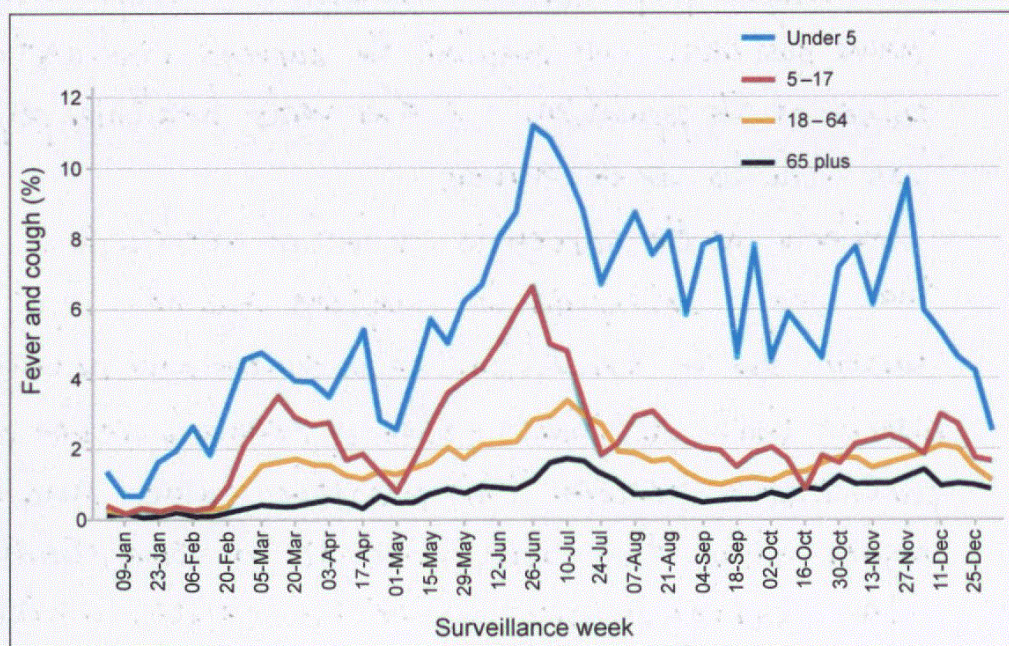
- (b) FluTracking releases weekly reports for both Australia (Figure 2) and New Zealand (Figure 3). These reports include a breakdown of symptoms by age group.

Figure 2: FluTracking Australia weekly fever and cough symptoms (%) by age group 2021–2023



Adapted from: www.flutracking.net/Info/Report/202353/AU

Figure 3: FluTracking New Zealand weekly fever and cough symptoms (%) by age group 2023



Adapted from: www.flutracking.net/Info/Report/202353/NZ

Use Figure 2 and Figure 3 to make THREE statements that compare the incidence of fever and cough in different age groups between the two countries.

Overall, the relative rates between the different age groups appear similar. Under 5s are consistently the highest, followed by 5-17, occasionally matching 18-64 (which has been split into 18-29 and 40-64 in Australia, but these are roughly similar, although 40-64 ~~is~~ typically ~~lower~~ has a lower percentage). This age range is consistently below 2% in Australia and 3% in NZ. Finally the lowest age bracket is consistently 65+. These never surpass 2% in the sample provided.

Australia ~~has~~ and NZ both share a similar ~~seasonal~~ under 5 pattern across 2023. There is a predictable spike in May-Aug due to the winter cold and flu season, which drops slightly to 5% for the remainder of the year apart from a short spike in Nov-Dec.

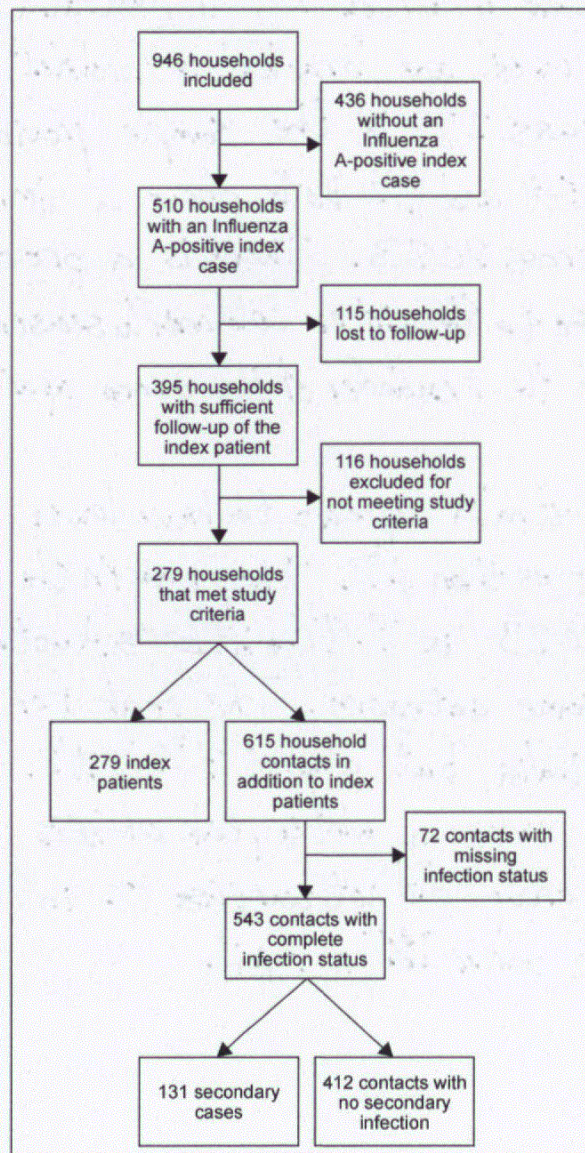
Overall, Australia appears to have lower fever and cough symptom rates than NZ. The peak in the winter season for under 5s in 2023 is 7.9% in ~~the~~ Australia, but 11% in NZ for the same category. The peak for 5-17s in 2023 is 4% in Australia, but almost 7% in NZ. 65+ ~~respondents~~ in NZ 2023 reached a winter peak of ~~2.2~~ 1.9% but in Australia the same group did not surpass 1% in 2023, with a ~~winter~~ peak of just under 1% in April.

- (c) In order to quantify the risk factors of influenza transmission in households, a prospective study was conducted during the 1999/2000 winter season in France.

946 households were enrolled in the study, where one member of the household, the index patient, had visited their doctor because of an influenza-like illness. 510 of the index patients tested positive for Influenza A. Secondary cases of influenza among household contacts of the index patients were identified using a standardised daily questionnaire over a period of 15 days.

Secondary cases were defined as household contacts who had gone on to develop clinical influenza within 5 days of the disease onset in the index patient. 395 households completed the questionnaire, of which 116 were excluded. Figure 4 describes the study design.

Figure 4: Study Design (1999–2000 Winter Influenza Season in France)



Adapted from: Viboud, C. et al (2004), Risk factors of influenza transmission in households, *The British Journal of General Practitioners*, 54(506):684-89

- (i) In the 279 households that met the study criteria, the mean number of children under 15 years of age per household was 0.71.

Calculate an estimate of the number of children under 15 who were recorded as secondary cases in this study.

$$279 \times 0.71 = 198.09 = 198 \text{ children under 15 in the study households.}$$

$$\text{Some of these could be index cases } P(\text{index}) = \frac{279}{279+615} = 0.3121$$

$$198 \times (1 - 0.3121) = 136.21 = 136 \text{ household contacts}$$

$$136 \times \frac{543}{943+72} = 120.08 = 120 \text{ contacts complete infection status}$$

$$120 \times \frac{131}{131+412} = 28.95 = 29 \text{ children under 15 recorded as secondary cases.}$$

- (ii) Discuss any issues that may influence the accuracy of your estimate in part (c)(i).

Young children's immune systems are still growing and are weak, so it is possible that some probabilities were not independent of age, such as the proportion of index patients and the proportion of contacts who contracted influenza A.

Households with one child may also be more likely to transmit the virus, as younger children require more care which may involve physical touch, which may affect the number of household contacts.

QUESTION TWO

There are over 200 different viruses that can cause colds. These viruses spread through the air when someone with a cold sneezes or coughs. Once cold symptoms begin, they will usually be resolved within a week.

- (a) A study on the duration of symptoms of the common cold virus was undertaken across Auckland. A researcher conducted an online survey of 620 volunteer participants. Each participant was asked to report the duration of their most recent cold symptoms from onset, also known as day 0. Of the 620 participants in the study, 121 experienced a fever and 612 experienced a sore throat, while 8 participants did not experience either of these symptoms.

Figure 5 shows the density distributions of the duration of fever and sore throat symptoms reported by study participants.

Figure 5: Density distributions of the duration of fever and sore throat symptoms in days

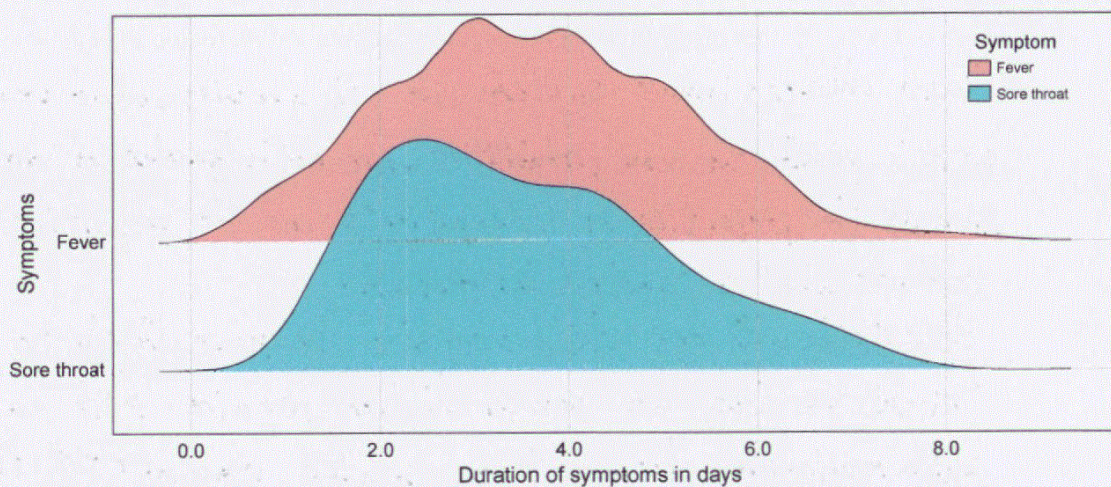
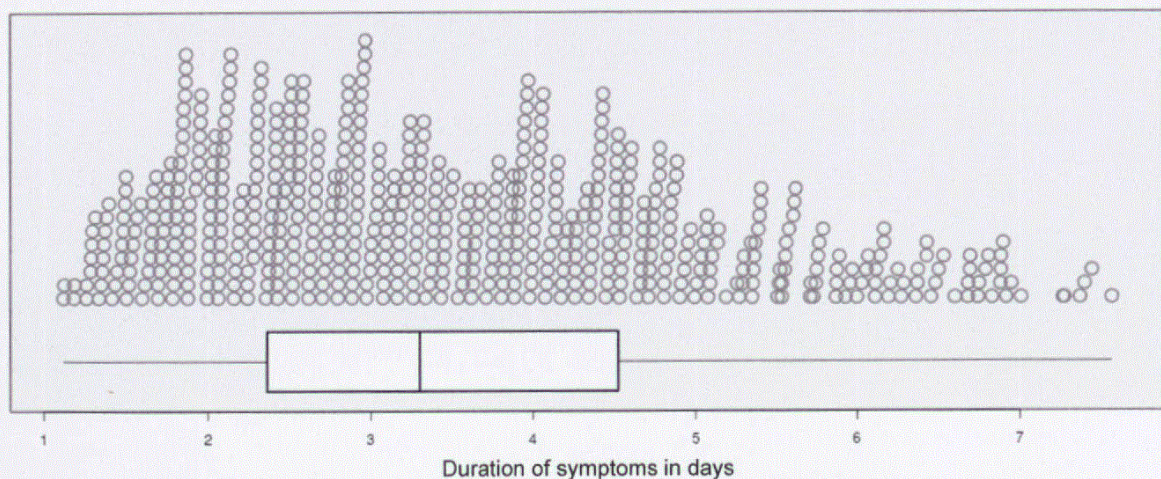


Figure 6 shows the sore throat symptoms, in days, reported by study participants.

Figure 6: Sore throat symptoms in days



Min	25%	Median	75%	Max	Mean	SD	Sample Size
1.12	2.361	3.3	4.521	7.563	3.536	1.471	612

- (i) Suggest a suitable probability distribution model, including relevant parameters, to model the duration of fever symptoms.

Comment on the limitations of the probability distribution model you have chosen.

As the fever symptom chart appears ^{somewhat} symmetrical and bell shaped, a normal distribution with $\mu = 3.8$ and $\sigma = 1.2$ could be used. However, as the normal distribution is asymptotic, values < 0 have a very small model probability, which is not actually possible. These are unlikely to be an issue for wide scope investigations. Another issue is that the distribution is not entirely bell shaped, so some details around the centre may be lost.

- (ii) It has been suggested that the duration of sore throat symptoms can be modelled by a triangular distribution with the parameters $a:1$, $b:7$, $c:2$.

Using the data and information provided in Figure 6, discuss whether the proposed probability distribution model is appropriate.

Use probability calculations to support your answer.

$$h = \frac{2}{7-1} = \frac{1}{3}$$



The data states that ~~the~~ $P(X > 4.521) = 0.25$.

The model predicts $P(X > 4.521) = \frac{1}{2} \times (7 - 4.521) \times \frac{2(7 - 4.521)}{(7-1)(7-2)} = 0.2048$.

This is somewhat close. The data also has ~~a~~ $P(X < 2.361) = 0.25$.

The model predicts $P(X < 2.361) = \frac{1}{2} \times (2.361 - 1) \times \frac{2(2.361 - 1)}{(7-1)(2-1)} = 0.3087$.

This is also slightly close.

Based on these predictions, the proposed triangular model is a mildly good fit, but should not be relied on.

- (b) Minimum sick leave entitlements in New Zealand increased from 5 days per year to 10 days per year from 24 July, 2021. Paid sick leave can be taken by employees when they, or a dependent family member, are unwell. If an employee uses all their available sick leave entitlements, they may take further sick leave as unpaid leave, annual holiday leave, or as sick leave in advance.

The Workplace Wellness Report 2023 surveyed 137 businesses employing 135 742 people across New Zealand. One of the questions in the survey asked for the number of days taken by employees as sick leave in the previous year 2022. Figures showed that in 2022, employees of these businesses took an average of 5.6 whole days sick leave per person. This was an increase of 1.2 days above the average for 2012.

Table 1: Sick leave entitlements and average sick leave taken 2012 and 2022.

	Yearly sick leave entitlement	Average sick leave taken
2012	5 days	4.4 days
2022	10 days	5.6 days

- (i) Using the Poisson distribution model and information provided in Table 1, compare an estimate for the probability that an employee exceeded their yearly sick leave entitlement in 2012, with an estimate for the probability of exceeding their entitlement in 2022.

$$2012 - \lambda = 4.4 \quad P(X > 5) = 0.4488$$

$$2022 - \lambda = 5.6 \quad P(X > 10) = 0.6891$$

$$\frac{0.0391}{0.4488} = 0.1316$$

The risk that ~~the~~ an employee exceeds their entitled sick leave ^{in 2022} is 13.16% of the same risk in 2012.

- (ii) Discuss TWO potential issues with using the Poisson distribution model with this data.

Poisson assumes the events are randomly occurring. This may not be the case as illnesses frequently persist over multiple days.

Poisson also assumes that the events always have the same chance of occurring. This is not the case as winter means more people will get sick, and therefore the risk of getting sick is higher.

- (c) A health student from Otago University wanted to investigate whether New Zealanders who were unwell were more likely to take time off from their regular duties, such as work or school, in the winter months compared with the summer months. They conducted two online surveys of randomly sampled New Zealanders in August 2023 (winter) and January 2024 (summer), and asked respondents whether they had experienced an influenza-like illness (a fever with a cough or a sore throat) during the previous month.

The respondents who had experienced an influenza-like illness were then asked an additional question about whether this illness resulted in them being absent from their regular duties. The results of the second question for both surveys is displayed in Table 2 below.

Table 2: Absence from regular duties due to influenza-like illness (ILI)

	Absent from regular duties due to ILI	Not absent from regular duties despite ILI	Total in study
July 2023	1434	670	2104
December 2023	519	275	794

Construct a confidence interval and use it to evaluate the claim that New Zealanders with an influenza-like illness are more likely to take time off from their regular duties in the winter months than in the summer months.

$$P(\text{Absent July}) = \frac{1434}{1434 + 670} = 0.6816$$

$$P(\text{Absent Dec}) = \frac{519}{519 + 275} = 0.6537$$

$$0.6816 - 0.6537 = 0.0279$$

The study states that NZers with an ILI are 2.79 percentage points more likely to take time off in the winter.

Both probabilities are $0.3 < x < 0.7$ so $\frac{1}{\sqrt{n}}$ approximation can be used

$$ME_{\text{July}} = \frac{1}{\sqrt{2104}} = 0.02180$$

$$ME_{\text{Dec}} = \frac{1}{\sqrt{794}} = 0.03549$$

$$MOE \text{ for independent studies} = 1.5 \times \frac{ME_1 + ME_2}{2}$$

$$MOE = 1.5 \times \frac{0.02180 + 0.03549}{2} = 0.0429675$$

$$0.0279 \pm 0.04297 = [-0.01507, 0.07087]$$

Because this confidence interval contains 0, the study's claim that NZers with an ILI are more likely to take time off in the ~~summer~~ ^{winter} months cannot be supported.

QUESTION THREE

- (a) Asthma is one of the most common illnesses in New Zealand, affecting an estimated 1 in 8 New Zealand adults, and 1 in 7 New Zealand children. There is no cure, but recent advances in medication may improve outcomes for asthma sufferers.

Report 1: Study a 'game-changer' for treating Kiwi asthma sufferers

Asthma affects more than 500 000 Kiwis, but a ground-breaking New Zealand-led study could be about to change the way patients with the respiratory illness are treated.

Currently, most adults with mild asthma are given two different inhaler medications: a preventer and a reliever. The preventer is an inhaled corticosteroid which is taken twice a day. The reliever is an inhaled bronchodilator which is taken whenever needed to relieve asthma symptoms, such as wheezing.

More than 670 people from New Zealand, the United Kingdom, Italy, and Australia with mild asthma took part in the study's trial – 82 per cent were from New Zealand. Participants had been taking only the reliever medication before they started the study. They were randomly allocated into three groups; one third used a reliever inhaler for 12 months whenever they had symptoms. Another third used both a preventer inhaler twice daily and a separate reliever inhaler whenever they had symptoms. The final group used a single combined preventer-reliever inhaler whenever they had symptoms.

The study showed that the risk of an asthma attack was reduced for those who used the single combined preventer-reliever inhaler. The study's results showed the combined preventer-reliever group's risk of an asthma attack was reduced by 60% compared to those who used a reliever alone. The single combined approach was also shown to have a 56% reduction compared to those who used both medications separately.

Adapted from: www.stuff.co.nz/national/health/112844094/study-a-gamechanger-for-treating-kiwi-asthma-sufferers

- (i) Write a short paragraph that identifies the key elements of the study design using appropriate statistical terminology.

The study was an experiment with 3 treatments, ~~just reliever~~, ~~separate~~ which were different combinations of types of inhaler. There was a reliever inhaler, separate preventer and inhaler, and combined reliever and preventer. The experimental units were people with ^{mild} asthma. The response variable was the risk of an asthma attack. The controlled variables included the medication the subjects were previously on, ~~the~~ and the application time and method of the inhalers.

- (ii) Discuss TWO reasons researchers randomly allocated participants into one of the three groups: reliever only, reliever and separate preventer, combined preventer-reliever.

Participants, while ~~they~~ all had mild asthma, may have different symptom types and intensities. Random allocation would evenly distribute different symptoms into the 3 groups. This would better test the efficacy of the treatments on a wide group of symptoms. Some participants may have a resistance to the treatments, so randomly allocating them minimised the variation between groups.

- (iii) The study in Report 1 was an open-label trial; there was no blinding, and participants were aware of what medication they were taking. Previous studies were double-blinded and all participants were provided with two inhalers. Participants in the double-blinded studies were randomly allocated into one of three treatment groups:

- active reliever inhaler and placebo preventer inhaler
- active reliever inhaler and separate active preventer inhaler
- active combined preventer-reliever inhaler and placebo preventer inhaler

Discuss why researchers may have chosen not to blind the study in Report 1.

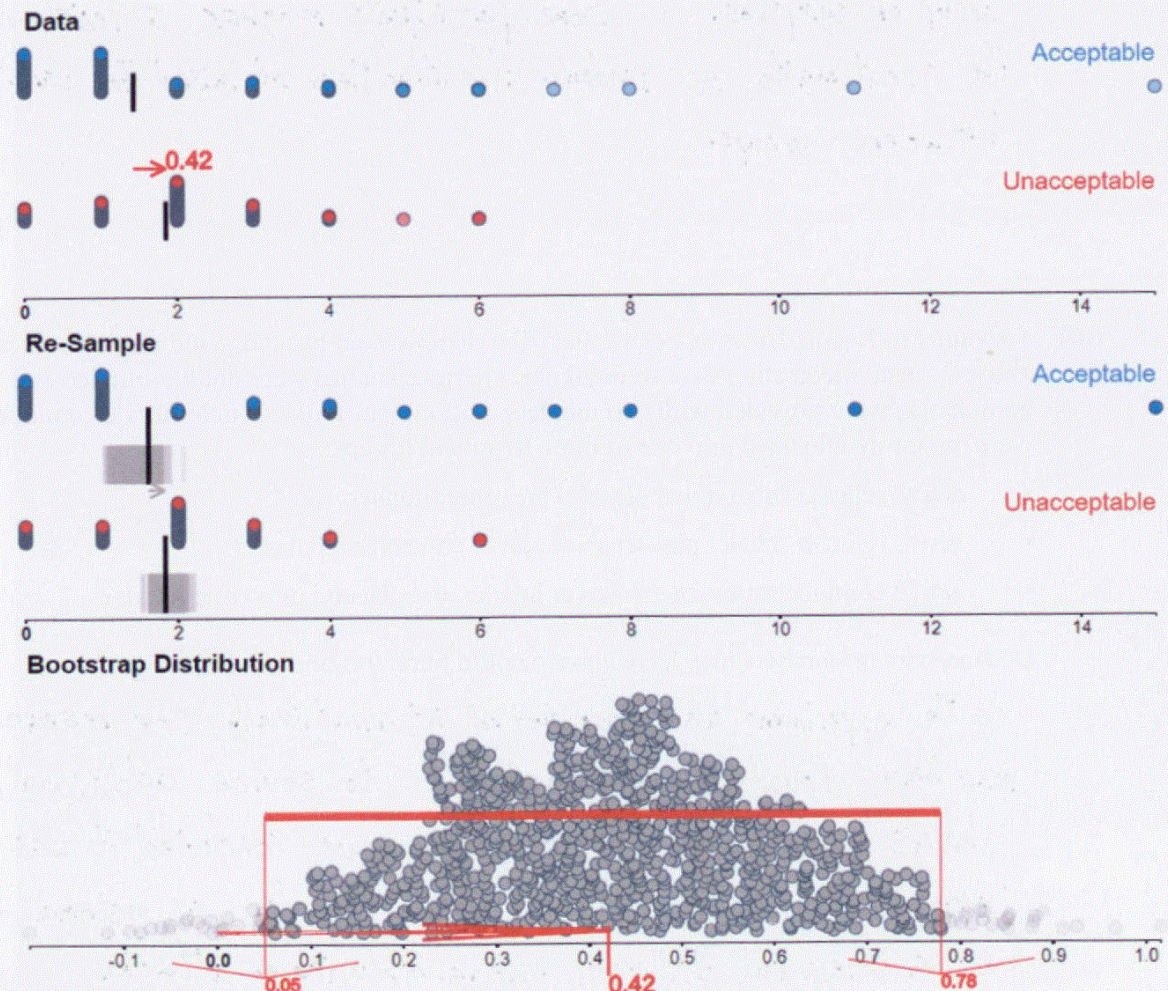
As the experiment was conducted internationally, the researchers may have found it cost-ineffective to source additional inhalers overseas. They may also have concluded it was immoral to assign a non-functional inhaler as asthma can be a stressful illness, especially during an attack.

- (b) The New Zealand Healthy Homes insulation standard requires all rental homes to have suitable ceiling and underfloor insulation by 1 July 2025. Adequate insulation results in homes that are warmer, drier, and less prone to mould. The Asthma and Respiratory Foundation of New Zealand recommends that all New Zealand homes be insulated to this standard to help keep those with asthma well, and to minimise their symptoms and potential complications with the illness.

In January 2024, results from a random sample of 2000 rental properties in Otago, Canterbury and the West Coast showed that 12.5% of the occupants in the rental properties surveyed were currently taking medication for asthma. The survey respondents were asked whether they believed their property met the current Healthy Homes insulation standard for both ceiling and underfloor insulation, as well as how many asthma-related healthcare visits each household member made during 2023.

The survey data for household occupants who were currently taking medication for asthma was analysed further. This data was used to construct a bootstrap confidence interval (Figure 7) for the difference between the mean number of healthcare visits by those who live in rental homes with acceptable insulation, and by those with unacceptable insulation.

Figure 7: Bootstrap distribution test output



Discuss what can be concluded from both the features of the sample data distributions and the confidence interval constructed using the sample data, including any reservations you may have about extending any findings from the results.

The ~~median~~ mean number of healthcare visits by those who live in rental homes with unacceptable ~~house~~ insulation on the West Coast is 0.62 more than the mean number of healthcare visits by those who live in rental homes with acceptable insulation on the west coast in my sample of asthma medication takers.

The overall spread of rental homes with acceptable insulation is much wider than those with unacceptable insulation on the west coast by those taking asthma medication. Acceptable houses range from 0-~~15~~ 6 visits, with the two outliers at 11 and 15 drawing up the mean, and unacceptable houses range from 0 to 6 visits.

The shape of acceptable houses is that of right skew.

This means acceptably insulated houses ~~are~~ tend to result in fewer hospital visits. The unacceptably insulated houses have a bell curve shape. This means that these houses ~~may~~ are ~~equally~~ similarly likely to have a number of hospital visits above or below the mean.

The greater range in acceptable houses means that inhabitants ~~may~~ possibly ~~also~~ access healthcare more than unacceptably insulated households. The shift in mean suggests that on average an unacceptably insulated rental is ~~more~~ likely to go access healthcare more than acceptably insulated homes. The bootstrap confidence interval shows that the population of rentals in Otago Canterbury and the west coast is 95% likely to ~~be~~ have a difference between the mean number of healthcare visits related to asthma ~~and the~~ by people living in unacceptably insulated homes and in acceptably insulated homes of ~~0.05~~ 0.05 and 0.78 in 2023. This does not include 0, so ~~the~~ the claim can be made that unacceptably insulated houses result in more ~~has~~ healthcare visits related to asthma by renters in Otago, Canterbury and the West Coast in 2023.

cont.

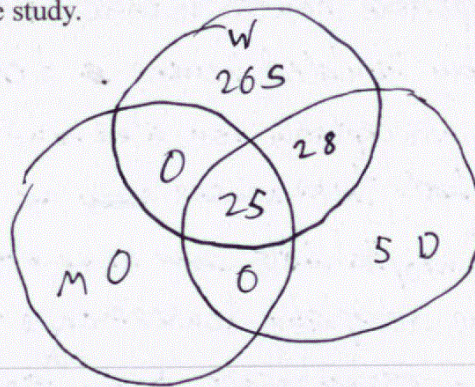
- (c) In Uganda, large-scale population data on the prevalence of asthma has been lacking. A survey was conducted among 3416 people aged 12 years or older to address this data gap. A questionnaire was used to collect participants' socio-demographics, respiratory symptoms, medical history, and known asthma risk factors.

Participants who reported either a wheeze (difficult, noisy breathing) in the past 12 months, a doctor's diagnosis of asthma, or current use of asthma medications were classified as having asthma. Any one of these three criteria was sufficient for a participant to be recorded as having asthma by the survey organisers.

Of all 323 participants recorded as having asthma, 318 reported wheeze in the past 12 months, 58 had a doctor's diagnosis of asthma, and 25 currently used asthma medications. 270 participants had only one of the three criteria, 28 patients had only two criteria, and 25 participants had all three.

Data from: Kirenga, B. (2019), Prevalence and factors associated with asthma among adolescents and adults in Uganda: A general population-based survey. *BMC Public Health* 19(1).

- (i) Calculate the proportion of participants who either reported wheeze in the past 12 months or currently used asthma medications, that had been given a diagnosis of asthma by a doctor before the study.



$$P(W \cup M | D) = \frac{P((W \cup M) \cap D)}{P(D)}$$

$$P((W \cup M) \cap D) = \frac{53}{3416} = 0.0155152$$

$$P(D) = \frac{58}{3416} = 0.0169789$$

$$P(W \cup M | D) = \frac{0.0155152}{0.0169789} = 0.9192$$

- (ii) Describe how you could construct a confidence interval for the proportion of Ugandan people aged 12 years or older meeting any of the three criteria for asthma.

Discuss any reservations or considerations you have about using this method.

As the proportion in the study that has asthma is $\sim 10\%$, the $\frac{1}{n}$ approximation cannot be used. Instead, the MoE of $1.96\sqrt{\frac{p(1-p)}{n}}$ can be used.

$$0.09455 \pm 1.96 \sqrt{\frac{0.09455(1-0.09455)}{3416}} = [0.08474, 0.1044] \quad (95\%)$$

One issue with the study is that one factor is reliant on the participants' memory of wheezing in the past 12 months. This could cause an incorrect value, as relying on memory ~~causes~~ can cause mistakes. This means the proportion of asthma sufferers may be higher than the study records, rendering the confidence interval incorrect.

QUESTION FOUR

- (a) For patients in the United Kingdom who are hospitalised with influenza, asthma is one of the most common pre-existing medical conditions. Data was collected from patients admitted to 75 hospitals in 31 United Kingdom cities or towns with a confirmed influenza infection. Some of the differences between the patients are listed in Table 3.

Table 3: Treatments and clinical outcomes for asthmatics and non-asthmatics hospitalised with influenza infection

	Asthmatics	Non-asthmatics
Total patients	385	1135
Pre-admission treatments:		
Antibiotics	78 20.2%	202 18.0%
Antivirals	44 11.4%	128 11.3%
Inhaled corticosteroids	203 52.7%	51 4.5%
Clinical outcomes:		
Pneumonia	66	188
Severe outcomes (e.g. admission to intensive care)	43 11.7%	225 19.8%

Data from: Myles, P. et al (2013). Differences between asthmatics and non-asthmatics hospitalised with influenza A infection. *European Respiratory Journal*, 41: 824-831

- (i) Compare the rates of both 'Pneumonia' and 'Severe outcomes (e.g. admission to intensive care)' for asthmatics and non-asthmatics.

$$P(\text{pneumonia} | \text{asthmatics}) = \frac{66}{385} = 0.1714$$

$$P(\text{pneumonia} | \text{non asthmatic}) = \frac{188}{1135} = 0.1656$$

$\frac{0.1714}{0.1656} = 1.035$. Asthmatics are 3.5% more likely to develop pneumonia after being hospitalized with influenza than non asthmatics.

$$P(\text{severe} | \text{asthmatic}) = \frac{43}{385} = 0.1117$$

$$P(\text{severe} | \text{non asthmatic}) = \frac{225}{1135} = 0.1982$$

$\frac{0.1117}{0.1982} = 0.5636$. Asthmatics are 56.36% less likely to develop severe outcomes after being hospitalized with influenza than non asthmatics.

- (ii) Based on the data in Table 3, describe which pre-admission treatment might best explain the observed differences in rates of severe outcomes for the two groups.

Asthmatics have a ~~207/385~~ $\frac{207}{385} = 52.7\%$ chance to have inhaled corticosteroids as a preadmission. Non asthmatics have a 4.5% chance of the same. The other pre admission treatments are all evenly distributed between groups. (see annotated Table 3)

$\frac{52.7}{4.5} = 11.7$ Asthmatics are 11.7 times as likely to have received ~~the~~ corticosteroids as non asthmatics. They are also ~~11.7~~ $\frac{11.7}{1.19} = 9.83$ 9.1% as likely to develop severe outcomes. This suggests, as all other pre admission treatments are administered at similar rates, that corticosteroids reduce severe outcome risk in asthmatics.

- (iii) 19% of all patients in the study had short hospital stays (less than two days duration), of which 71 individuals were asthmatics.

Compare the proportions of short hospital stays for asthmatics compared to non-asthmatics.

$$385 + 1135 = 1520 \text{ total influenza patients in the study}$$

$$1520 \times 0.19 = 288.9 \approx 289 \text{ short hospital stays}$$

$$71 \text{ asthmatic short stays}$$

$$289 - 71 = 218 \text{ non-asthmatic short stays}$$

$$P(\text{short} | \text{asthmatic}) = \frac{71}{385} = 0.1844$$

$$P(\text{short} | \text{nonasthmatic}) = \frac{218}{1135} = 0.1921$$

$$\frac{0.1844}{0.1921} = 0.9599$$

Asthmatic patients are 95.99% as likely to have a short hospital stay following an influenza hospitalisation than non asthmatic patients.

Question Four continues on the following page.

- (iv) For the 385 asthmatic patients in this study, the probability that two randomly chosen, consecutive patients are children is 0.0356.

Use this information to determine how many children were in this study.

$C = \text{number of children}$

$$\frac{C}{385} \times \frac{C-1}{384} = 0.0356$$

$$\frac{C^2 - C}{147840} = 0.0356$$

$$C^2 - C - 5263.104 = 0$$

$$C = 73.049, -72.049$$

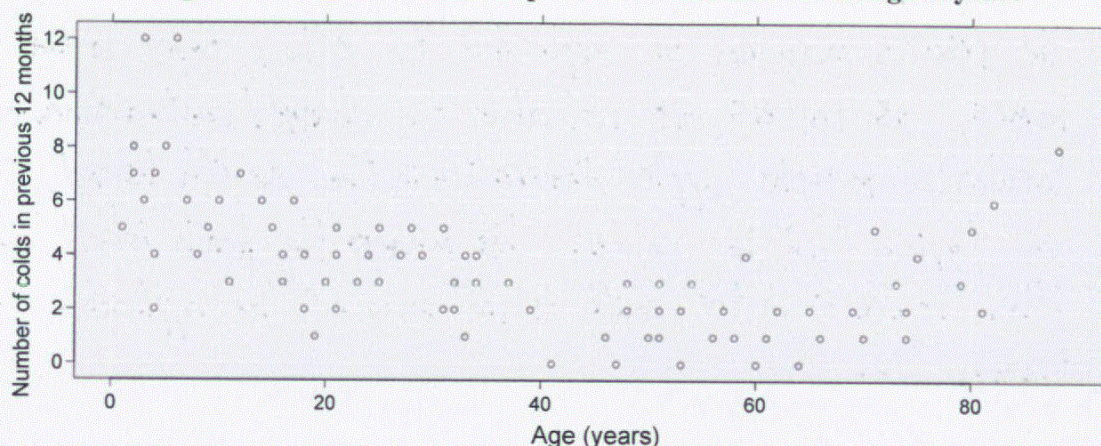
Negative children is impossible

$C = 73.049$, there were 73 children in the study.

- (b) The New Zealand Ministry of Health claims that as people age, they tend to get fewer colds due to a build up of immunity to some cold viruses. To test this theory, researchers took a random sample of 80 patients from four GP practices in Auckland. These patients were asked to recall the number of colds they had experienced in the previous 12 months. When a patient was under the age of 14, their caregiver was asked to recall the number of colds the patient had experienced in the previous 12 months.

Figure 8 shows a scatterplot displaying each person's age in years plotted against the number of colds each person experienced in the previous 12 months

Figure 8: Number of colds in previous 12 months versus age in years



Identify key features of the data shown in Figure 8, and use these to discuss the relationship between age in years and the number of colds experienced in the previous 12 months.

The graph shows a negatively associated, moderately strong relationship. The ~~association~~ trend may also be quadratic, as the values past age 60 appear to tend upwards. No major groups are visible, although there is moderate scatter between 20 and 70 years of age. Weaker scatter is outside this range. The ~~tight~~ moderate scatter shows that the relationship is quite strong between the participants' age and number of colds. Outside of the wider ranges (< 20 , > 60) the number of colds typically has a range of around 4, which ~~is rarely~~ shifts slightly depending on age. The model does appear to be quadratic, which suggests that the young and old extremes are more susceptible to colds. One issue is the memory recall of the study. Participants or their caregivers may not be able to accurately recall the number of colds across the previous year, which may affect any conclusions made.

Extra space if required.
Write the question number(s) if applicable.

QUESTION
NUMBER

- 3b) These results may not be applicable to other years as the landlords may be upgrading or ~~is~~ planning to in the years up to the new Healthy Homes requirement.
- The results may not be applicable to other types of ~~the~~ homes, as renters are not able to change/contract new insulation. Insulation may be applied differently in the south island, as typically it is colder. This means the North island, or other regions in the south island may not be applicable or transferable.

Outstanding Scholarship

Subject: Statistics

Standard: 93201

Total score: 28

Q	Score	Marker commentary
1	8	This question provides evidence for 7o because the candidate has gained at least 6s grades and 1o grade overall. The candidate discussed at least two trends or key features of time series data in context and with numerical evidence from Figure 1. They have at least three sufficient statements describing non-sampling errors in context. They have at least three sufficient comparison statements, with evidence, comparing relevant trends or features of time series data from two graphs from Figures 2 and 3. They have calculated a valid estimate for the proportion of children who were recorded as secondary cases in the study in Figure 4, and they have discussed two valid issues that may have influenced the accuracy of an estimate made from Figure 4.
2	6	This question provides evidence for 5s because the candidate has gained at least 6s grades overall. The candidate has identified a suitable probability distribution and parameters. They have identified at least one limitation of this model. The candidate has not calculated one suitable probability for evaluating the goodness-of-fit of a probability distribution model. They have not correctly calculated probabilities from a Poisson distribution. They have evaluated the Poisson model for the given situation, and have discussed two appropriate limitations. They have calculated and appropriate estimate for the margin of error for a given situation, but have not interpreted the confidence interval appropriately.
3	7	This question provides evidence for 8o because the candidate has gained at least 6s grades and 1o grade overall. The candidate described two key features of an experiment in context using statistical language. They discussed at least two reasons for random allocation in context. They identified a suitable purpose for an open label trial. They have compared relevant features of sample data from a bootstrap distribution, and have interpreted the confidence interval in context. They

		have adequately discussed at least two potential issues with the sampling strategy for this study. The candidate has not correctly calculated a conditional probability. They have described the use of a margin of error formula that will not over-estimate the confidence interval.
4	7	This question provides evidence for 7o because the candidate has gained at least 6s grades and 2o grades overall. The candidate has correctly compared appropriate rates using a relative approach. They have reasoned with evidence that the use of corticosteroids is associated with the observed difference in severe outcomes for asthmatic patients, and again have used a relative approach. They have calculated appropriate proportions for short hospital stays for asthmatics compared to non-asthmatics, and compared these using a relative approach. The candidate has calculated the number of children in a study and has considered non-replacement. They have identified the nature and strength of the data in Figure 8, but have not identified the effect of outliers or unusual groupings.