



ELSEVIER

Contents lists available at ScienceDirect

MethodsX

journal homepage: [www.elsevier.com/locate/mex](http://www.elsevier.com/locate/mex)

## Method Article

# Calculating incidence of Influenza-like and COVID-like symptoms from Flutracking participatory survey data



Emily P. Harvey<sup>a,b,c,d,\*</sup>, Joel A. Trent<sup>a,d,e</sup>, Frank Mackenzie<sup>a,d</sup>,  
Steven M. Turnbull<sup>a,b,d</sup>, Dion R.J. O'Neale<sup>a,b,d</sup>

<sup>a</sup> COVID Modelling Aotearoa, The University of Auckland, 38 Princes Street, Auckland CBD, Auckland 1010, New Zealand

<sup>b</sup> Te Pūnaha Matatini, The University of Auckland, 38 Princes Street, Auckland CBD, Auckland 1010, New Zealand

<sup>c</sup> M.E. Research, Takapuna, Auckland 0622, New Zealand

<sup>d</sup> Department of Physics, The University of Auckland, 38 Princes Street, Auckland CBD, Auckland 1010, New Zealand

<sup>e</sup> Department of Engineering Science, The University of Auckland, 70 Symonds Street, Grafton, Auckland 1010, New Zealand

## A B S T R A C T

This article describes a new method for estimating weekly incidence (new onset) of symptoms consistent with Influenza and COVID-19, using data from the Flutracking survey. The method mitigates some of the known self-selection and symptom-reporting biases present in existing approaches to this type of participatory longitudinal survey data.

The key novel steps in the analysis are:

- 1) Identifying new onset of symptoms for three different Symptom Groupings: COVID-like illness (CLI1+, CLI2+), and Influenza-like illness (ILI), for responses reported in the Flutracking survey.
- 2) Adjusting for *symptom reporting bias* by restricting the analysis to a sub-set of responses from those participants who have consistently responded for a number of weeks prior to the analysis week.
- 3) Weighting responses by age to adjust for *self-selection bias* in order to account for the under- and over-representation of different age groups amongst the survey participants. This uses the `survey` package [22] in R [30].
- 4) Constructing 95% point-wise confidence bands for incidence estimates using weighted logistic regression from the `survey` package [21] in R [28].

In addition to describing these steps, the article demonstrates an application of this method to Flutracking data for the 12 months from 27th April 2020 until 25th April 2021.

© 2022 The Author(s). Published by Elsevier B.V.

This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>)

\* Corresponding author at: COVID Modelling Aotearoa, The University of Auckland, 38 Princes Street, Auckland CBD, Auckland 1010, New Zealand.

E-mail address: [ehar061@uoa.auckland.ac.nz](mailto:ehar061@uoa.auckland.ac.nz) (E.P. Harvey).

<https://doi.org/10.1016/j.mex.2022.101820>

2215-0161/© 2022 The Author(s). Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>)

## ARTICLE INFO

*Method name:* Calculating incidence of Influenza-like and COVID-like symptoms from FluTracking participatory survey data

*Keywords:* Flutracking, Syndromic surveillance, Participatory epidemiology, Incidence, Influenza-like illness, COVID-19, Survey analysis, Survey re-weighting, Reporting bias, Logistic regression

*Article history:* Received 4 April 2022; Accepted 7 August 2022; Available online 17 August 2022

## Specifications table

Subject Area:	Medicine and Dentistry
More specific subject area:	Public health, disease surveillance.
Method name:	Calculating incidence of influenza and COVID-like symptoms from FluTracking participatory survey data
Name and reference of original method:	NA
Resource availability:	<a href="https://gitlab.com/cma-public-projects/flutracking-methods-article">https://gitlab.com/cma-public-projects/flutracking-methods-article</a>

## Method details

This article describes a method that can be used to estimate the onset of new incidence of Influenza-like (ILI) and COVID-like (CLI) symptoms from longitudinal participatory health survey data. The method mitigates some of the known biases in existing approaches for this type of data and provides more robust estimates of symptom incidence over time. In this article we also demonstrate an application of the methodology to Flutracking data for the 12 months from 27th April 2020 until 25th April 2021 in Aotearoa New Zealand.

As part of this work, we define a range of *Symptom Groupings* that can be used to estimate the incidence of symptoms consistent with COVID-19, and demonstrate the impact of the chosen Symptom Groupings on resulting estimates. Because multiple illnesses can lead to overlapping sets of symptoms, it is important to note that this method does not attempt to diagnose new onset of a specific disease (e.g. COVID-19); rather it estimates the onset of new (sets of) symptoms that are consistent with certain diseases, including COVID-19.

The method and associated code presented here, and available at [1], have been developed to analyse data from the Flutracking survey [13] in Aotearoa New Zealand, but can be easily adapted to similar longitudinal data, where self-selection biases and reporting biases are known issues.

## Background

Flutracking [13] is a participatory health surveillance system for Australia and Aotearoa New Zealand that seeks to estimate the proportion of the population with new onset of symptoms consistent with influenza [9], and more recently symptoms consistent with COVID-19. The survey was developed by Hunter Population Health in collaboration with The University of Newcastle. The survey is administered online to volunteer participants from the public. Registered participants receive a weekly email asking them to report any cold, flu, or COVID-like symptoms experienced in the previous week. The survey also allows for participants to report on behalf of family members in their household, such as young children.

Participatory health surveillance systems like Flutracking are a valuable resource for tracking outbreaks of infectious disease. They allow for analysis and reporting on how new onset of symptoms differ over time and between seasons, as well as across different geographic regions. However, there are several typical limitations that can reduce the utility of the results that they provide.

Previous studies have shown that participatory health surveillance systems can be biased in terms of survey participants and response rates over time. Firstly, some participants may be more likely to respond to a survey if they have experienced symptoms in a given week [4,20] (a form of *reporting bias*), and secondly, participatory samples rely on volunteers and are not representative of the general population [10,14] (a form of *self-selection bias*). In particular, people from lower socio-economic groups and minority ethnicities are found to be under-represented [23]. Within Aotearoa New Zealand,

we find that Flutracking respondents are disproportionately of Pākehā/NZ European ethnicity and older, with Māori and Pacific Islanders, and people under 30 years old, being particularly under-represented (see Section 4.1). Both of these biases should be accounted for where possible to improve the reliability of incidence estimates [5,6].

Participatory health surveillance studies have commonly focused on influenza tracking and have relied on strict symptom definitions that only describe one experience of influenza-like illness (ILI). For example, previous studies, including current weekly Flutracking reports [12], focus solely on experiences of ILI defined as new onset of both cough and fever in the same week [7]. While these symptoms are commonly viewed as core components of ILI [9], this criteria may be insufficient for investigating the rates of a wider range of respiratory illnesses, and in particular, would miss the majority of COVID-19 cases [19,27,29].

In the context of Aotearoa New Zealand, during the year of data considered in this article (27th April 2020 to 25th April 2021), the COVID-19 Elimination Strategy advised anyone who experienced symptoms consistent with COVID-19 to seek a test, so as to promptly detect and stamp-out any new community outbreak of SARS-CoV-2. Monitoring the performance of such an approach requires estimating the incidence of symptoms that are consistent with COVID-19, including those that do not meet the strict ‘cough and fever’ ILI definition, in order to determine whether testing rates in a given week are commensurate with incidence of new COVID-like symptoms in the same week.

In this article, we describe a methodology that seeks to address some of the limitations described above for data from the Flutracking survey from Aotearoa New Zealand. Despite this specific application, the method can be applied equally to Flutracking data from Australia, or to similar longitudinal data collected elsewhere (e.g., the “Flu Near You” survey in the United States [24], or “Influenzanet” in Europe [22]). Our approach for determining a subset of responses that are classified as ‘consistent’ in any given week, in order to reduce reporting bias and to maximise response numbers retained, is also applicable to other datasets with similar issues. For example, it could be easily applied to prevalence estimates calculated from ‘test-to-stay’ type surveillance data COVID-19 from Rapid Antigen Test results by filtering to only consider results from individuals who were regularly reporting (negatives) in the weeks leading up to a positive test.

### Flutracking data

Raw data are obtained from Flutracking New Zealand [13] through the New Zealand Ministry of Health. Flutracking symptom data consists of a presence/absence indication for each respondent for each of six<sup>1</sup> symptoms: cough, fever, sore throat, shortness of breath, runny nose, and loss of taste or smell. All of these symptoms are part of the New Zealand Ministry of Health’s definition of ‘COVID-like symptoms’ [16].

When respondents initially enrol in the Flutracking survey, they are asked to give demographic information, including age, ethnicity, location (postcode), and gender. Participants are also asked about their vaccination status for COVID-19 and the annual influenza vaccine, and whether they have been tested for COVID-19 in the previous week. An example of the weekly survey is shown in Fig. 1. If a respondent indicates that they are experiencing symptoms, they are then asked whether they sought healthcare or missed work/usual activities due to the illness, and whether they were tested for COVID-19 or Influenza.

### Method outline

For each survey week in a period of interest, the following steps are applied:

- For each response that matches the *Symptom Grouping* of interest, classify whether that week is a *new onset* of those symptoms.

---

<sup>1</sup> A seventh symptom – headache – was added to the survey in the second half of 2021, to capture better the reported symptoms for the Delta variant of SARS-CoV-2. We do not include Headache in our analysis, even after it was added, in order to maintain consistency.

---

## Symptoms

Did **you** have:

- Fever?  Yes |  No |  Don't Know
- Cough?  Yes |  No |  Don't Know
- Sore throat?  Yes |  No |  Don't Know
- Runny nose?  Yes |  No |  Don't Know
- Shortness of breath?  Yes |  No |  Don't Know
- Any change in sense of taste or smell?  Yes |  No |  Don't Know
- NEW!** Headache?  Yes |  No |  Don't Know
- None of the above

---

## Swab Test

- Did you have a nose or throat swab during the week ending Sunday 12 September?  Yes |  No |  Don't Know

---

## COVID-19 Vaccination

- Have you received **dose 1** of a COVID-19 vaccine?  Yes |  No |  Don't Know

---

## Flu Vaccination

- Have you received the **Annual Flu** vaccine in **2021**?  Yes |  No |  Don't Know

**Fig. 1.** Example of the Flutracking weekly survey, sent to participants via email. Retrieved from <https://www.flutracking.net/Demo/NZ> [13].

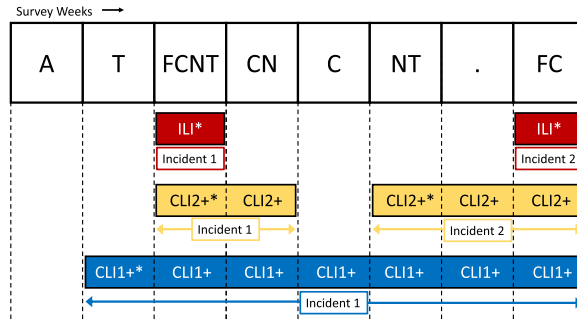
- Determine which responses each week count as *consistent responses*, and remove those that do not.
- Calculate an *age weighting factor* for each age group within the survey population, to match the age distributions of the consistent responses each week with a known population distribution. This is done using the `rake` function within R's `survey` package [21]. The age weighting factor is calculated for each week, and can be calculated at either a whole-country level or for sub-populations specified as sets of locations and ages.
- Construct *95% pointwise confidence bands* around mean incidence estimates using the `svyciprop` function in R's `survey` package [21] for specified sets of factors, including age group, Symptom Grouping, location, and survey week.

The following sections detail each of the steps of the method in more depth.

### Define Symptom Grouping criteria

We define three different Symptom Groupings:

- **CLI1+:** Responses indicate any one or more of the above *COVID-like* symptoms. CLI1+ meets the Aotearoa New Zealand Ministry of Health advice to seek a COVID-19 test [16].



**Fig. 2.** Example of symptoms reported each week for one hypothetical respondent. This illustration shows a period of 8 survey weeks, with reported symptoms (or non-response) indicated for each week. The symptoms reported include **A** Asymptomatic (i.e. a response indicating no symptoms), **C** Cough, **F** Fever, **N** runny Nose, and **T** sore Throat. A dot indicates the respondent did not answer the survey that week. In this hypothetical case, we see how the responses given can fall under each Symptom Grouping and that these *Incidents* cover different survey weeks depending on the patterns of symptoms. As CLI1+ is the most inclusive Symptom Grouping, the criteria for this are met in most weeks shown here and it is considered a single *Incident*, while the Symptom Groupings ILI and CLI2+ would be split into two separate *Incidents* during the same time period. We can also see that CLI1+ and CLI2+ (but not ILI) span a period including one week where no response was given. This is because we assume that the missing week would be a continuation of the same *Incident*, provided it meets the same Symptom Grouping criteria. The \* indicates which week will be included as the onset of the *Incident*, and will be included in the 'Incidence' calculations.

- **CLI2+:** Responses indicate two or more of the above *COVID-like* symptoms. CLI2+ allows us to be slightly more discerning, given that many non-infectious illnesses such as allergies or asthma can cause a new onset of the *COVID-like* symptoms in the survey.
- **ILI:** Responses indicate at least both *cough* and *fever*, which are symptoms of *influenza-like illness*. This Symptom Grouping is used for public Flutracking reporting [13].

### Incidence classification

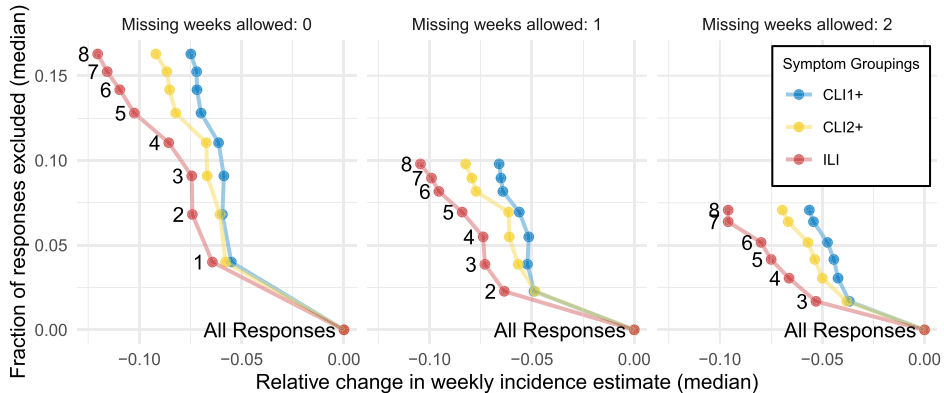
We wish to estimate 'incidence', not 'prevalence' from the Flutracking data. Specifically, we wish to identify any *new* onset of symptoms that meet the threshold for an 'incident' according to a chosen Symptom Grouping. The method we use follows the same method employed by the Aotearoa New Zealand Ministry of Health and Flutracking Australia for estimating the incidence of ILI (cough and fever), but extends to the Symptom Groupings CLI1+ and CLI2+ above, in addition to the Symptom Grouping for ILI.

Flutracking data contains each respondent's reported symptoms for a given survey week. For a selected Symptom Grouping, we first check whether an individual's symptoms in the given week meet the criteria for that Symptom Grouping. If so, we mark it as an Incident. For example, if we wanted to investigate the number of CLI1+ Incidents, we would record all weeks where an individual recorded one or more of the six *COVID-like* symptoms as an Incident.

Fig. 2 provides a hypothetical example of a respondent who reported different symptoms over the course of 8 weeks.

To distinguish new onset of illness from the continuation of a previously reported Incident, each Incident is assigned a unique ID, with this same Incident ID given to all consecutive weeks that meet the Symptom Grouping. If there is no response for one week between two weeks in which the Symptom Grouping was met, it is assumed that the interstitial week is a continuation of the previous week's Incident and it is allocated the same Incident ID. For any gaps in responses between Incidents longer than one week we allocate the second Incident a new Incident ID. This procedure is consistent with the existing methodology used by Flutracking Australia and Flutracking New Zealand.

The incidence calculation is applied independently for each Symptom Grouping. This means that if a respondent reported one symptom in their first week of illness and two or more symptoms in the second week, the second week would be recorded as an 'Incident' of CLI2+, as though it were the onset of a new illness. This means that if a participant met the criteria for CLI1+ for three consecutive



**Fig. 3.** Curves showing the fraction of responses excluded and the relative change in weekly incidence rate estimates for different choices of window size (text labels on the plot) and missing weeks allowed (panels left to right), for each of the three Symptom Groupings (colour of points and lines).

weeks, but met the criteria for CLI2+ only in the first and last weeks, this would be recorded as two Incidents of CLI2+, as shown in Fig. 2. This is capturing useful information, as the advice is that someone with new or worsening symptoms should seek a COVID-19 test.

While our approach allows for flexible Symptom Grouping criteria to be implemented within the methodology, future work may seek to apply more sophisticated, unsupervised learning approaches when defining Incidents. While not in the scope of the current research, previous work has demonstrated that unsupervised learning approaches provide a useful technique for modelling heterogeneous symptom experiences among respondents [18].

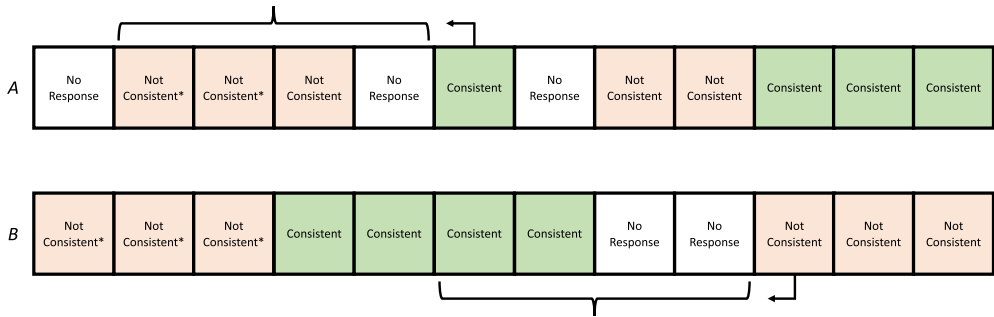
#### Determining the set of consistent responses

Flutracking participants are more likely to respond to the survey when they are experiencing symptoms than when they are well [20]. This *symptom reporting bias* can lead to overestimating symptom incidence rates if it is not accounted for. Consequently, it is important to define a subset of *consistent responses* to use for incidence rate calculations. These are defined as being the responses, in a given week, from participants who are deemed to have responded consistently to the survey, as assessed at that week.

Previous studies have determined consistent respondents by only considering responses from those participants who have completed more than some minimum fraction of all surveys (e.g., [7,10]). We improve on this approach by determining a subset of consistent responses for each week of the analysis, based on the response history of each survey participant for a defined window prior to the week under consideration. This allows a respondent's 'consistent' status to change from week to week. We implement this by looking back over a specified number of weeks (the 'window size') and requiring the participant to have responded in all those weeks, with an allowance for a certain number of 'missing weeks' within the window. This makes it possible to adjust for the known reporting bias at any given point in time, while still maximising the sample size at the corresponding time. Our approach also means that analysis can be performed on a weekly basis throughout the year, rather than only at the end of a year (or season) of data collection.

In Fig. 3 we plot trade-off curves showing the median fraction of responses excluded and the median relative change in the weekly incidence estimates as a function of 'Window size' and 'Missing weeks allowed'. This helps us to determine suitable parameter values for the selection criteria of consistent responses. For more information on how these effects vary throughout the year, see Figs. 3 and 4 in section B of the Supplementary Material.

Increasing the number of prior weeks for which a participant must have previously responded (the response window size) reduces the number of participants included in the 'consistent response'



**Fig. 4.** Hypothetical examples for the selected window size of 'the previous 4 weeks', and allowing 1 missing response in that window. The illustration above shows survey response series for hypothetical respondents A and B. Any weeks with 'no response' are left uncoloured. For weeks with a response, that response is defined as 'not consistent' (pink), or 'consistent' (green). We can see that for respondent A, the highlighted survey week is considered 'consistent' as 3 of the previous 4 survey weeks contain a response (even if for each of those previous survey weeks, a response was considered not consistent). For respondent B, we can see that the highlighted survey week is considered 'not consistent' because 2 out of the previous 4 survey weeks did not contain a response. \* indicate that the survey weeks covering the respondents first three weeks of participation are considered not consistent, due to the fact that they have not yet answered enough surveys to meet the threshold of consistency.

subset for any given week – that is, the fraction of responses excluded increases with increasing window size. Allowing respondents to have skipped some of the weeks in the consistent response window means that the responses of some participants are re-included in the consistent response subset.

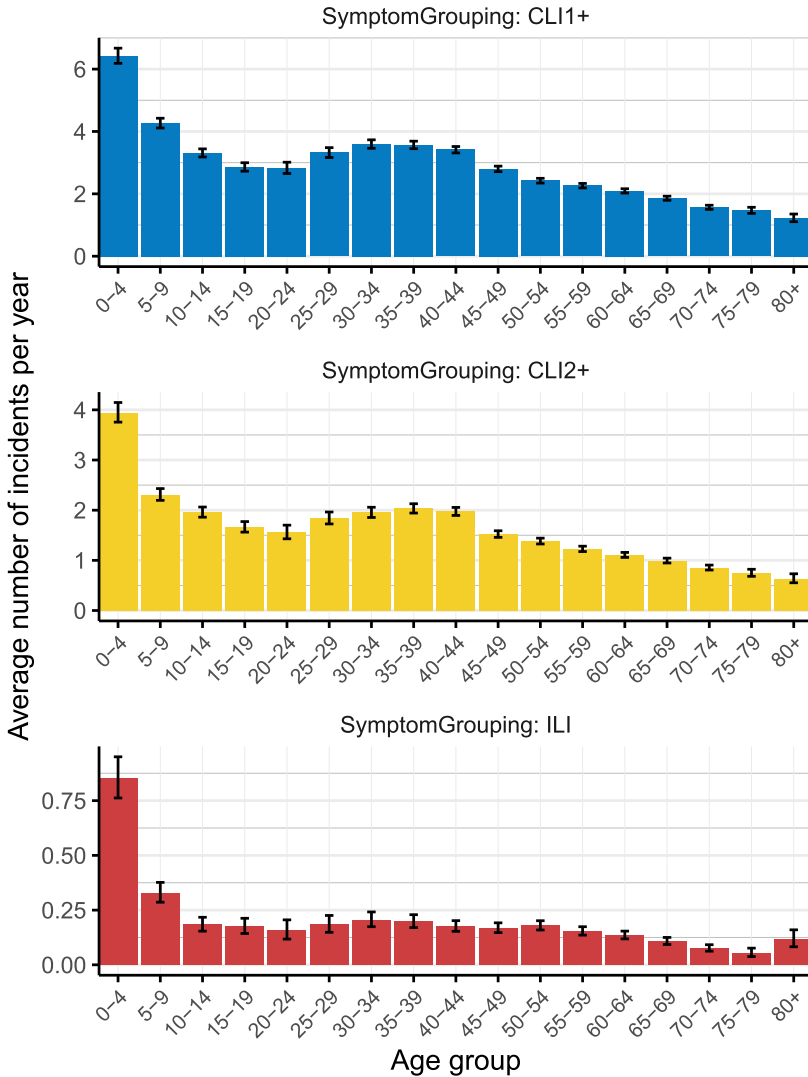
When we consider the impact of window size and missing weeks allowed on the incidence estimate we find that a more stringent (i.e. larger) window size reduces the rate of symptom incidence calculated for any given week. This is expected from previous work looking at the impact of reporting bias [20]. The biggest jump is seen when requiring *any* response in the window prior to reporting symptom onset (i.e. a one-week window, or a two-week window with one missing week allowed, or a three-week window with two missing weeks allowed). Imposing this minimal consistent response window results in an immediate decrease of approximately 5% in the incidence estimate for all Symptom Groupings. We attribute this initial decrease to a correction for the known symptom reporting bias – if participants were equally likely to respond in a given week, independent of symptom incidence, then applying this criterion would not affect the estimated incidence rate.

Increasing the consistent response window, beyond the minimum, initially causes a decrease in the number of consistent responses but does not significantly change incidence estimates. However, further increasing the window size beyond about four weeks causes both the number of consistent responses and the estimated incidence rate to decrease. We attribute this to the more stringent consistent response window introducing a different form of bias by increasing the over-representation of some groups of participants in the sample population – specifically older Pākehā/NZ European cohorts who tend to have lower incidence rates – see for example Fig. 5.

While our method allows for the user to choose any consistent respondent window size and any number of missed weeks within that, we have chosen in our subsequent analysis to use a window size of four weeks with up to one missing week. This defines a consistent response subset where small changes in the window size do not result in significant changes in the resulting incidence rate and where the fraction of responses included in the sample population for a given week is not too severely reduced, in part due to the fact that the Flutracking survey allows respondents to retrospectively enter results for up to four weeks into the past. We provide two examples of the selected values for window size and allowed missing responses using hypothetical response series illustrated in Fig. 4.

#### Accounting for self-selection biases: age weighting

There are several demographic biases in the Flutracking cohort with *registration rates* differing by age, ethnicity and location of respondents, see section A of the Supplementary Material for more



**Fig. 5.** Average number of new incidents per person for the year between 27th April 2020 and 25th April 2021, by five-year age groups (error-bars represent 95% confidence intervals). This only considers respondents who submitted at least 50 surveys over the year. In terms of Symptom Groupings, overall we see that incidents of CLI1+ are more common than CLI2+ and ILI. We can also see that the highest level of incidence, across all three Symptom Groupings, is experienced by those in the 0-4 year old age group.

details. If any of these self-selection biases also align with a corresponding variation in *incidence rates* along the same demographic factor, then it will contribute to an under- or over-estimate of incidence for the total population.

The strongest factor affecting incidence estimates is age (see Fig. 5), with younger ages, particularly 0-4 years, tending to have higher incidence across CLI1+, CLI2+, and ILI Symptom Groupings. Therefore, we adjust our incidence estimates to account for the relative number of respondents in the survey population according to age each week.



Although there is a relationship between ethnicity and incidence of several respiratory illnesses in Aotearoa New Zealand [2,3], it can be unhelpful to treat ethnicity as a predictive factor for these illnesses. Research has found that factors including socioeconomic deprivation, household crowding, and housing quality, are all associated with increased incidence of respiratory infections [3,17,31], and that these are highly correlated with ethnicity [17]. These same underlying factors are also likely to contribute to a participant selection bias within the Flutracking survey. Under-representation of lower socio-economic groups has also been noted within similar participatory surveys for Influenza. [23]

The low response rate in Aotearoa New Zealand Flutracking data for ethnic groups other than Pākehā/NZ European, combined with the under-representation of lower socio-economic groups, mean that it is unlikely that the sampled population of non-Pākehā ethnicities in Flutracking is truly representative of those groups in the wider population. Therefore, weighting based on ethnicity alone may actually amplify selection biases rather than mitigating them. For this reason, although it is possible using our method, we do not calculate or re-weight symptom incidence by ethnicity. We note that the current re-weighting methodology is easily able to be applied to ethnicity or socio-economic status, should Flutracking enrolment be sufficiently broad that responses are representative of these groups. In order to avoid this limitation in future studies, we point to previous research on best practice for participatory surveillance data collection [8,25].

The reference population used for the age weightings presented in this method is the 2021 Estimated Residential Population, calculated by Statistics NZ with 5-year age groups and spatial units of District Health Board (DHB) of residence [26]. However, any other source of reference population data can be used, provided that it can be mapped on to the age (1-year age bands) and location (Aotearoa New Zealand postcode) groupings used in the Flutracking data. In general, it is better to use higher levels of aggregation (e.g. 5-year age groups) and DHB. This is in order to avoid the low response numbers per analysis unit, and consequent lower statistical power and increased uncertainty, that a more granular aggregation will lead to.

In order to weight survey responses according to participant age, we use the rake function provided in R's survey package [21] to assign a weighting coefficient to each response. The process of raking matches the marginal distribution of ages from the survey respondents with those of a reference population.<sup>2</sup> The raking can be applied at any level of aggregation appropriate to the reported results, with the caveat that fine-grained aggregation can lead to small cell sizes which may limit the statistical power of results in some cases.

When analysing responses for sub-national spatial units (e.g. DHBs), we calculate the age weighting factor for each age group with respect to the age distributions for each of the corresponding sub-national units. However, for analysis at a national level, we calculate only national level age weightings, rather than combining sub-national weighting factors. This avoids the pit-fall of national level weighting factors inheriting lower statistical power and increased uncertainty from small cell sizes for some location-age combinations.

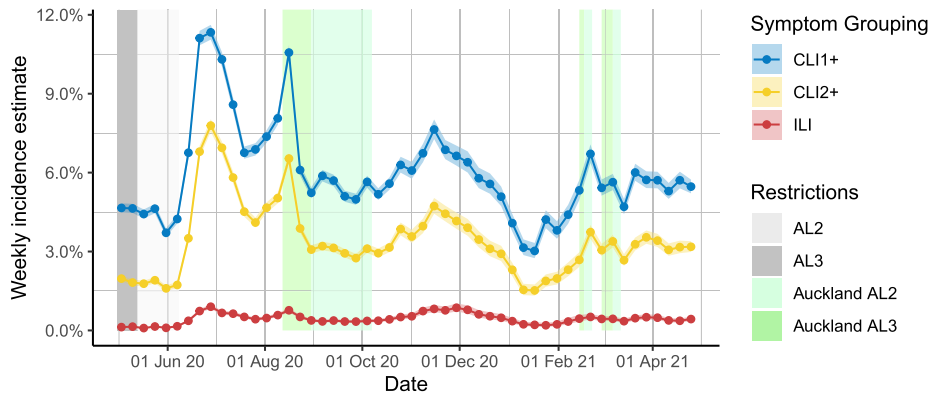
### *Constructing mean incidence estimates with 95% confidence intervals*

Constructing confidence intervals for weighted survey data is non-trivial. We construct estimates of mean incidence with 95% confidence intervals for each survey week for a given set of demographic factors and Symptom Groupings using the survey package in R [21], incorporating the age weighting previously computed in Section 1. The 95% confidence intervals for each survey week can be turned into 95% *pointwise* confidence bands across survey weeks.

We first create a 'survey design' object using our set of consistent responses, their incidence classification and weighting, and other relevant factors such as their survey week, age group, and location. This survey design object is then passed to the `svyby` function, which independently calls the `svyciprop` function for each subset in a set of factors and outputs the estimated mean incidence and 95% confidence intervals for each subset. `svyciprop` has several methods of producing confidence intervals; we use its logistic regression method (`logit`). An example set of factors could be Symptom Grouping and survey week, as in Fig. 6. One independent subset used in these

---

<sup>2</sup> Since this is specified in 5-year age groups, sets of ages must also consist of combinations of 5-year age groups.



**Fig. 6.** Weekly incidence estimates for the whole of Aotearoa NZ, for three different Symptom Groupings (CLI1+, CLI2+, and ILI) for the year between 27th April 2020 and 25th April 2021 (shaded areas represent 95% pointwise confidence bands). Alert levels [15] (AL2 and AL3 – AL3 is stronger restrictions than AL2) are shaded. We observe periods of reduced incidence during these elevated Alert Levels periods, as well as some spikes in incidence when the Alert Level is set at 1 (no restrictions), including in July 2020 – during the winter in Aotearoa New Zealand).

calculations is the Symptom Grouping ‘CLI1+’ and the Survey Week ending on Sunday the 31st of May 2020.

The process used by `svyby` and `svyciprop` is equivalent to fitting a weighted logistic regression model, where every subset in a set of factors is fitted independently such that the fitted value of any subset in the set of factors is equal to the weighted mean incidence of that subset. Thus the same model could have been fitted using the `svyglm` function from the `survey` package and then fitting interaction effects between all factors. In order to do this, a quasi-binomial model family needs to be used in `svyglm` [21] because the use of survey weights may make the number of ‘Incidents’ non-integer. However, as we were mainly concerned about the fitted means and confidence intervals for each subset rather than fitted values of regressors and their significance, we did not do this in general. Additionally, our process makes it very simple to extract confidence intervals for our estimates, which, while not too difficult (e.g. using a non-parametric bootstrap [11]), is not nearly as straightforward to do with the fitted model returned by `svyglm`.

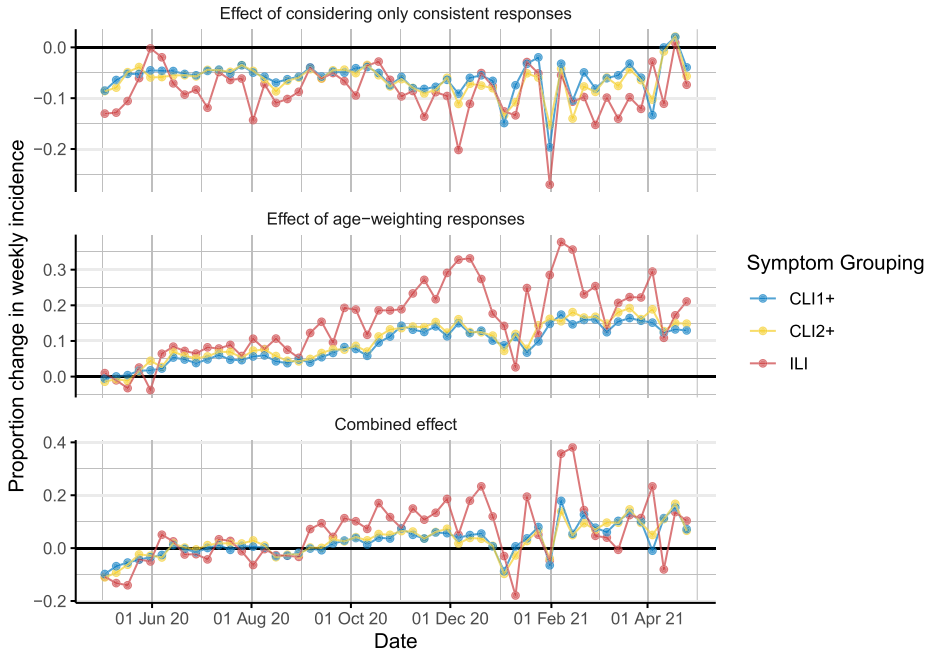
The only exception to this, is that in order to consider the significance of differences in incidence for a given Symptom Grouping, as seen in Fig. 9, we did need to use `svyglm` to test whether the difference in the estimated mean incidence between two locations on a given week was statistically significant (i.e. if the  $p$ -value for the Auckland Metro DHBs coefficient was less than 0.05).

### Method demonstration

We demonstrate the method outlined above by applying it to Flutracking data from Aotearoa New Zealand for the period from the survey week ending Sunday 3rd May 2020 to the survey week ending Sunday 25th April 2021. During this period there was no widespread community transmission of SARS-CoV-2 (the virus that causes COVID-19 disease) in Aotearoa New Zealand. The Elimination Strategy the country employed during this period meant that any detection of SARS-CoV-2 in the community was immediately followed by a system of transmission reduction measures known as Alert Levels [15], which included school closures and stay-at-home orders.

R code and supporting example data that demonstrates the application of this methodology can be found at <https://gitlab.com/cma-public-projects/flutracking-methods-article> [1]. All plots are generated using the `ggplot2` package [30] in R [28], with the code used to generate the figures available at [1].

Fig. 6 presents the weekly incidence estimates for the whole of Aotearoa New Zealand and for the CLI1+, CLI2+, and ILI Symptom Groupings after adjusting for consistent responses and age weighting.



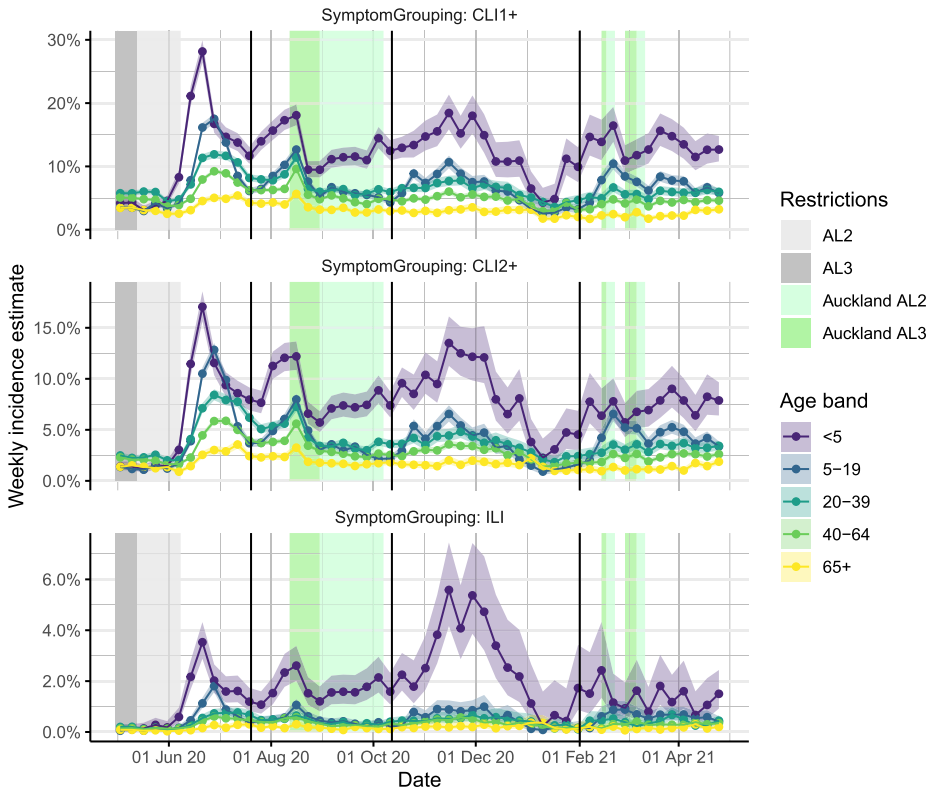
**Fig. 7.** Change in weekly incidence estimates, due to adjusting for symptom reporting bias by considering only 'consistent responses' (top), adjusting using age re-weighting (middle), and adjusting for both (bottom), relative to baseline (naive) weekly incidence estimates where all responses are used and there is no age re-weighting.

Consistent responses were weighted by age at a national level, as per Section 1. Survey week was the only variable used to explain incidence for each Symptom Grouping. The effect of elevated Alert Level interventions (indicated on the chart as shaded regions) are clearly seen in the reduction of symptom incidence for all Symptom Groupings. That is, the transmission reduction measure put in place to prevent the spread of SARS-CoV-2 also reduced the incidence of other symptoms for respiratory illnesses in general.

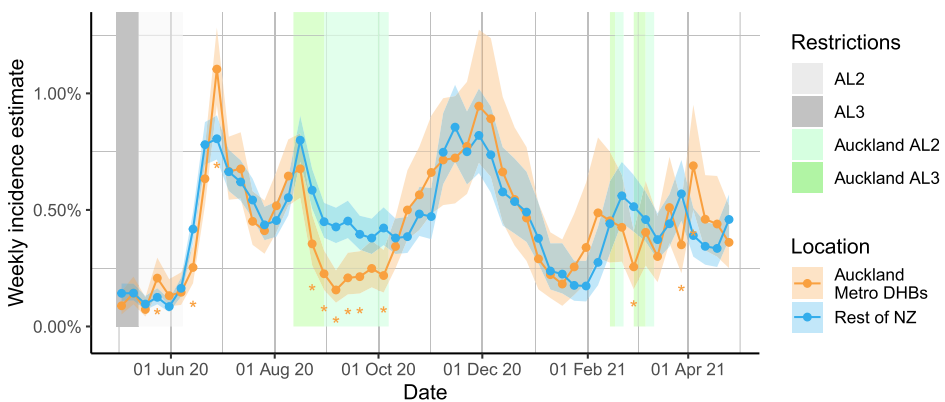
Fig. 7 shows the effect of adjustments to the incidence estimates in Fig. 6 for consistent responses and age weighting. Adjusting incidence estimates to include only the population of consistent responses shows an almost universal reduction in incidence rates, by around 5%. This is due to the removal of respondents who are preferentially (or exclusively) completing the survey when symptomatic, as has been observed in previous analysis [20]. This effect decreased in February–April of 2021, and coincides with a very stable population of Flutracking survey respondents in this period.

When (national level) age-weightings are applied to the incidence estimates, we see an increase in the calculated incidence rate. This is due to younger participants being under-represented in Flutracking responses, relative to the reference population. Younger participants also tend to have higher rates of symptom incidence (Fig. 8), hence up-weighting their responses to follow those of the underlying population increases the overall symptom incidence estimates. The size of the effect over the survey period is shown in Fig. 7. The size of this effect is not constant over time. The initial half of the survey period saw a higher number of responses from younger participants. When responses from younger participants fell in the latter half of the survey period, the effect due to age-weighted adjustment increased. This is most pronounced in the relative change in incidence rate for ILI symptoms in the age-weighted effect shown in Fig. 7 (middle panel), due to the much higher incidence of ILI symptoms in younger age groups.

Finally, we note that while the relative adjustments for the CLI1+ and CLI2+ incidence rates are mostly smoothly varying over time, the weighting factors applicable to the ILI Symptom Grouping



**Fig. 8.** Weekly incidence estimates by selected age bands, for the whole of Aotearoa New Zealand, for the CLI1+, CLI2+, and ILI Symptom Groupings (shaded areas represent 95% pointwise confidence bands). Alert Levels (AL2 and AL3) are indicated by shaded regions, where grey shaded regions are when the restrictions were applied nationwide, and green shaded regions are where they were applied only to the Auckland region. Black lines indicate the beginning of school term periods.



**Fig. 9.** Weekly incidence estimates for the ILI Symptom Grouping split by location (Auckland Metro DHBs and Rest of New Zealand, shaded areas represent 95% pointwise confidence bands). Stars indicate weeks when the difference in incidence estimate is statistically significant for the Auckland Metro DHBs, relative to the Rest of New Zealand. Alert levels (AL2 and AL3) are indicated by shaded regions.

are more volatile due to low total rates of ILI symptoms in Aotearoa New Zealand from early 2020 onward.

Fig. 8 shows the incidence of CLI1+, CLI2+, and ILI symptoms, broken down by age. Responses from consistent responses were weighted by age at a national level, per Section 1. Survey week, age band and the interaction effects between the two were used to explain incidence for each Symptom Grouping. The higher rates of incidence for younger age bands is strongest for ILI symptoms. Under-fives have 1.5–2 times the incidence rate of the next highest incidence age band for the CLI2+ Symptom Grouping, and 3–4 times higher for ILI symptoms.

Fig. 9 shows weekly incidence estimates for ILI split by location for Auckland Metro DHBs and the Rest of New Zealand. Consistent responses were weighted by age using the age distributions in the Auckland Metro DHBs and the Rest of New Zealand separately, as per Section 1. Survey week, location and the interaction effects between the two were used to explain incidence for ILI. Alert Level interventions designed to prevent the spread of SARS-CoV-2 [15], including those that were principally applied to Auckland alone, are indicated by shaded green regions. The effect of these regional interventions is clearly seen in the difference in incidence estimates. The Auckland Metro DHBs have much lower incidence numbers during periods of elevated Alert Levels (stronger restrictions) relative to the rest of the country.

## Conclusion

In this article, we have described an approach for calculating incidence estimates and corresponding confidence intervals for longitudinal participatory health survey data, such as that produced by *Flutracking*. Our method allows for analysis of arbitrary Symptom Groupings – including two Symptom Groupings developed here for relevance to the COVID-19 response in Aotearoa New Zealand. Our method has several advantages over existing methods for analysis of similar data; namely that it defines a subset of consistent responses for use in calculating incidence estimates; and that it uses re-weighting of responses by participant sub-group in order to account for registration bias. Both of these adjustments increase the rigour and accuracy of the resulting incidence calculations, compared with methods in current use. We have demonstrated our method and quantified the effect of our methodological improvements by applying it to *Flutracking* data for Aotearoa New Zealand over a period where symptom incidence was affected by a range of COVID-related behavioural changes and government interventions.

## Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Acknowledgements

We wish to thank New Zealand *Flutracking*, a collaborative initiative between the Ministry of Health (NZ) and Hunter New England Health (NSW, Australia), for the provision of *Flutracking* data. Professor Thomas Lumley provided advice on use of the survey package and the COVID Modelling Aotearoa internal review panel, lead by Dr Matthew Parry, provided valuable feedback on this article.

## Supplementary material

R code and supporting example data that demonstrates the application of this methodology can be found at <https://gitlab.com/cma-public-projects/flutracking-methods-article> [1]. The Rmarkdown document contained within this repository, 'FluTrackingMethods\_MethodsX.Rmd', was used to generate the figures found in this paper. This document can be used with a simplified masked dataset to generate the figures in this paper for a 4 month survey period and for responses from Auckland Metro DHBs and Wellington Region DHBs. To obtain the full Aotearoa New Zealand dataset for the *Flutracking* survey please contact the Ministry of Health (NZ) at [nzmoh\\_flutracking@health.govt.nz](mailto:nzmoh_flutracking@health.govt.nz).

Supplementary material associated with this article can be found, in the online version, at doi:[10.1016/j.mex.2022.101820](https://doi.org/10.1016/j.mex.2022.101820)

## References

- [1] Flutracking Methods Article COVID-19 Modelling Aotearoa public projects. <https://gitlab.com/cma-public-projects/flutracking-methods-article>.
- [2] M.G. Baker, L.T. Barnard, A. Kvalsvig, A. Verrall, J. Zhang, M. Keall, N. Wilson, T. Wall, P. Howden-Chapman, Increasing incidence of serious infectious diseases and inequalities in New Zealand: a national epidemiological study, *Lancet* 379 (9821) (2012) 1112–1119, doi:[10.1016/S0140-6736\(11\)61780-7](https://doi.org/10.1016/S0140-6736(11)61780-7).
- [3] M.G. Baker, A. McDonald, J. Zhang, P. Howden-Chapman, Infectious diseases attributable to household crowding in New Zealand: a systematic review and burden of disease estimate, 2013. <http://www.healthyhousing.org.nz/wp-content/uploads/2010/01/HH-Crowding-ID-Burden-25-May-2013.pdf>.
- [4] K. Baltrusaitis, K. Noddin, C. Nguyen, A. Crawley, J. Brownstein, L. White, Evaluation of approaches that adjust for biases in participatory surveillance systems, *Online J. Public Health Inform.* 10 (2018), doi:[10.5210/0jphi.v10i1.8908](https://doi.org/10.5210/0jphi.v10i1.8908).
- [5] J.S. Brownstein, S. Chu, A. Marathe, M.V. Marathe, A.T. Nguyen, D. Paolotti, N. Perra, D. Perrotta, M. Santillana, S. Swarup, M. Tizzoni, A. Vespignani, A.K.S. Vullikanti, M.L. Wilson, Q. Zhang, Combining participatory influenza surveillance with modeling and forecasting: Three alternative approaches, *JMIR Public Health Surv.* 3 (4) (2017), doi:[10.2196/publichealth.7344](https://doi.org/10.2196/publichealth.7344).
- [6] P. Cantarelli, M. Debin, C. Turbelin, C. Poletto, T. Blanchon, A. Falchi, T. Hanslik, I. Bonmarin, D. Levy-Bruhl, A. Micheletti, D. Paolotti, A. Vespignani, J. Edmunds, K. Eames, R. Smallenburg, C. Koppeschaar, A.O. Franco, V. Faustino, A. Carnahan, M. Rehn, V. Colizza, The representativeness of a European multi-center network for influenza-like-illness participatory surveillance, *BMC Public Health* 14 (2014), doi:[10.1186/1471-2458-14-984](https://doi.org/10.1186/1471-2458-14-984). 984–984.
- [7] R. Chunara, E. Goldstein, O. Patterson-Lomba, J.S. Brownstein, Estimating influenza attack rates in the United States using a participatory cohort, *Sci. Rep.* 5 (1) (2015) 9540, doi:[10.1038/srep09540](https://doi.org/10.1038/srep09540).
- [8] C. Dalton, S. Carlson, M. Butler, D. Cassano, S. Clarke, J. Fejsa, D. Durrheim, Insights from flutracking: thirteen tips to growing a web-based participatory surveillance system, *JMIR Public Health Surv.* 3 (3) (2017) e7333.
- [9] C. Dalton, D. Durrheim, J. Fejsa, L. Francis, S. Carlson, E.T. d'Espaignet, F. Tuyl, Flutracking: a weekly Australian community online survey of influenza-like illness in 2006, 2007 and 2008, *Commun. Dis. Intell. Q. Rep.* 33 (3) (2009) 316–322.
- [10] M. Debin, C. Turbelin, T. Blanchon, I. Bonmarin, A. Falchi, T. Hanslik, D. Levy-Bruhl, C. Poletto, V. Colizza, Evaluating the feasibility and participants' representativeness of an online nationwide surveillance system for influenza in France, *PLoS One* 8 (9) (2013), doi:[10.1371/journal.pone.0073675](https://doi.org/10.1371/journal.pone.0073675). e73675–e73675
- [11] B. Efron, Better bootstrap confidence intervals: rejoinder, *J. Am. Stat. Assoc.* 82 (397) (1987) 198–200, doi:[10.2307/2289153](https://doi.org/10.2307/2289153). Publisher: [American Statistical Association, Taylor & Francis, Ltd.].
- [12] Flutracking (2022). Weekly interim report New Zealand - week ending 29 May 2022. <https://www.flutracking.net/Info/Report/202222/NZ>, accessed: 06-06-2022.
- [13] Flutracking.net website <https://info.flutracking.net/about/> accessed: 06-06-2022.
- [14] I. Friesema, C. Koppeschaar, G. Donker, F. Dijkstra, S. van Noort, R. Smallenburg, W. van der Hoek, M. van der Sande, Internet-based monitoring of influenza-like illness in the general population: experience of five influenza seasons in the Netherlands, *Vaccine* 27 (45) (2009) 6353–6357, doi:[10.1016/j.vaccine.2009.05.042](https://doi.org/10.1016/j.vaccine.2009.05.042).
- [15] NZ Government (2020). COVID-19 alert system in New Zealand. <https://covid19.govt.nz/alert-system/alert-system-overview/>, accessed 2020-08-29.
- [16] Ministry of Health NZ (2021). About COVID-19. <https://www.health.govt.nz/covid-19-novel-coronavirus/covid-19-health-advice-public/about-covid-19>, accessed 2021-12-01.
- [17] T. Ingham, M. Keall, B. Jones, D.R.T. Aldridge, A.C. Dowell, C. Davies, J. Crane, J.B. Draper, L.O. Bailey, H. Viggers, T.V. Stanley, P. Leadbitter, M. Latimer, P. Howden-Chapman, Damp mouldy housing and early childhood hospital admissions for acute respiratory infection: a case control study, *Thorax* 74 (9) (2019) 849–857, doi:[10.1136/thoraxjnl-2018-212979](https://doi.org/10.1136/thoraxjnl-2018-212979).
- [18] K. Kalimeri, M. Delfino, C. Cattuto, D. Perrotta, V. Colizza, C. Guerrisi, C. Turbelin, J. Duggan, J. Edmunds, C. Obi, R. Pebody, Unsupervised extraction of epidemic syndromes from participatory influenza surveillance self-reported symptoms, *PLoS Comput. Biol.* 15 (4) (2019) e1006173.
- [19] E. Lavezzo, E. Franchin, C. Ciavarella, G. Cuomo-Dannenburg, L. Barzon, C.D. Vecchio, L. Rossi, R. Manganelli, A. Lorigian, N. Navarin, D. Abate, M. Sciro, S. Merigliano, E. De Canale, M.C. Vanuzzo, V. Besutti, F. Saluzzo, F. Onelia, M. Pacenti, S.G. Parisi, G. Carretta, D. Donato, L. Flor, S. Cocchio, G. Masi, A. Sperduti, L. Cattarino, R. Salvador, M. Nicoletti, F. Caldart, G. Castelli, E. Nieddu, B. Labella, L. Fava, M. Drigo, K.A.M. Gaythorpe, K.E.C. Ainslie, M. Baguelin, S. Bhatt, A. Boonyasiri, O. Boyd, L. Cattarino, C. Ciavarella, H.L. Coupland, Z. Cucunubá, G. Cuomo-Dannenburg, B.A. Djafaara, C.A. Donnelly, I. Dorigatti, S.L. van Elsland, R. FitzJohn, S. Flaxman, K.A.M. Gaythorpe, W.D. Green, T. Hallett, A. Hamlet, D. Haw, N. Imai, B. Jeffrey, E. Knock, D.J. Laydon, T. Mellan, S. Mishra, G. Nedjati-Gilani, P. Nouvellet, L.C. Okell, K.V. Parag, S. Riley, H.A. Thompson, H.J.T. Unwin, R. Verity, M.A.C. Vollmer, P.G.T. Walker, C.E. Walters, H. Wang, Y. Wang, O.J. Watson, C. Whittaker, L.K. Whittles, X. Xi, N.M. Ferguson, A.R. Brazzale, S. Toppo, M. Trevisan, V. Baldo, C.A. Donnelly, N.M. Ferguson, I. Dorigatti, A. Crisanti, Imperial College COVID-19 Response Team, Suppression of a SARS-CoV-2 outbreak in the Italian municipality of Vo', *Nature* 584 (7821) (2020) 425–429, doi:[10.1038/s41586-020-2488-1](https://doi.org/10.1038/s41586-020-2488-1).
- [20] D. Liu, L. Mitchell, R.C. Cope, S.J. Carlson, J.V. Ross, Elucidating user behaviours in a digital health surveillance system to correct prevalence estimates, *Epidemics* 33 (2020) 100404, doi:[10.1016/j.epidem.2020.100404](https://doi.org/10.1016/j.epidem.2020.100404).
- [21] Lumley, T. (2020). survey: analysis of complex survey samples, R package version 4.0.
- [22] D. Paolotti, A. Carnahan, V. Colizza, K. Eames, J. Edmunds, G. Gomes, C. Koppeschaar, M. Rehn, R. Smallenburg, C. Turbelin, S. Van Noort, A. Vespignani, Web-based participatory surveillance of infectious diseases: the Influenzanet participatory surveillance experience, *Clin. Microbiol. Infect.* 20 (1) (2014) 17–21, doi:[10.1111/1469-0691.12477](https://doi.org/10.1111/1469-0691.12477).
- [23] S.V. Scarpino, J.G. Scott, R.M. Eggo, B. Clements, N.B. Dimitrov, L.A. Meyers, Socioeconomic bias in influenza surveillance, *PLoS Comput. Biol.* 16 (7) (2020) e1007941, doi:[10.1371/journal.pcbi.1007941](https://doi.org/10.1371/journal.pcbi.1007941).

- [24] M.S. Smolinski, A.W. Crawley, K. Baltrusaitis, R. Chunara, J.M. Olsen, O. Wójcik, M. Santillana, A. Nguyen, J.S. Brownstein, Flu near you: crowdsourced symptom reporting spanning 2 influenza seasons, *Am. J. Public Health* 105 (10) (2015) 2124–2130.
- [25] M.S. Smolinski, A.W. Crawley, J.M. Olsen, T. Jayaraman, M. Libel, Participatory disease surveillance: engaging communities directly in reporting, monitoring, and responding to health threats, *JMIR Public Health Surv.* 3 (4) (2017) e62, doi:[10.2196/publichealth.7540](https://doi.org/10.2196/publichealth.7540).
- [26] Statistics NZ (2021). Technical notes DHB ethnic projections (2021 update).
- [27] C.H. Sudre, K. Lee, M.N. Lochlainn, T. Varsavsky, B. Murray, M.S. Graham, C. Menni, M. Modat, R.C. Bowyer, L.H. Nguyen, D.A. Drew, A.D. Joshi, W. Ma, C.G. Guo, C.H. Lo, S. Ganesh, A. Buwe, J.C. Pujol, J.L.d. Cadet, A. Visconti, M. Freydin, J.S.E.S. Moustafa, M. Falchi, R. Davies, M.F. Gomez, T. Fall, M.J. Cardoso, J. Wolf, P.W. Franks, A.T. Chan, T.D. Spector, C.J. Steves, S. Ourselin, Symptom clusters in Covid19: a potential clinical prediction tool from the COVID symptom study app, *medRxiv* (2020), doi:[10.1101/2020.06.12.20129056](https://doi.org/10.1101/2020.06.12.20129056).
- [28] R.C. Team, R: A language and environment for statistical computing, R Foundation for Statistical Computing, Vienna, Austria, 2021. <https://www.R-project.org/>.
- [29] P.M. Wells, K.M. Doores, S. Couvreur, R.M. Martinez, J. Seow, C. Graham, S. Acors, N. Kouphou, S. Neil, R. Tedder, P. Matos, K. Poulton, M.J. Lista, R. Dickenson, H. Sertkaya, T. Maguire, E. Scourfield, R. Bowyer, D. Hart, A. O'Byrne, K. Steele, O. Hemmings, C. Rosadas, M. McClure, J. Capedevila-Pujol, J. wolf, S. Ourseilin, M. Brown, M. Malim, T. Spector, C. Steves, Estimates of the rate of infection and asymptomatic COVID-19 disease in a population sample from SE England, *Journal of Infection* 81 (6) (2020) 931–936, doi:[10.1016/j.jinf.2020.10.011](https://doi.org/10.1016/j.jinf.2020.10.011).
- [30] H. Wickham, *ggplot2*, *Wiley Interdiscip.Rev.* 3 (2) (2011) 180–185.
- [31] Zhang, J., & Barnard, L. T. (2021). The impact of respiratory disease in New Zealand: 2020 update. <https://www.asthmafoundation.org.nz/assets/documents/Respiratory-Impact-report-final-2021Aug11.pdf>.