Protocol for analysis of factors associated with outbreaks of SARS-CoV-2 in nursing homes in Ireland

17 February 2021

Purpose and aim

In May 2020, the National Public Health Emergency Team (NPHET) recommended the establishment of an expert panel on nursing homes to examine the complex issues surrounding the management of COVID-19 among this particularly vulnerable cohort. The COVID-19 Nursing Homes Expert Panel was appointed by the Minister for Health in May 2020 and reported back with recommendations in light of the expected ongoing impact of COVID-19 over the subsequent 12-18 months.

One recommendation relating to data analysis (recommendation 6.7) was outlined as:

> "HPSC, HSE and HIQA should produce a detailed epidemiological analysis comparing both risk and protection factors associated with having an outbreak or not at all in HIQA regulated facilities."

This protocol details the process to be undertaken by the Health Information and Quality Authority (HIQA) and the Health Protection Surveillance Centre (HPSC) to analyse both the potential risk and protective factors associated with outbreaks of SARS-CoV-2, which have occurred in HIQA-regulated designated centres for older persons (nursing homes).

1.0 Protocol for research question

Five distinct steps in the process have been identified for this analysis. These are listed below and described in more detail in sections 1.1-1.5.

1. Identify potential risk and protective factors from international literature.
2. Assess the available data in the Irish context to support an analysis.
3. Collect the data.
4. Analyse the data.
5. Summarise the findings.

The scope for the analysis is limited to a quantitative analysis using readily available data on outbreaks and nursing homes. The intention is to estimate the relative importance of different risk factors in relation to the occurrence and extent of nursing home outbreaks of COVID-19. A qualitative analysis of the context for outbreaks in individual homes is beyond the scope of this work, as the necessary data to support such analyses are not available at present.

1.1 Identify potential risk and protective factors

In line with the terms of reference of the COVID-19 Nursing Homes Expert Panel, a rapid systematic review was undertaken by a research team from the University

College Dublin, under the direction of the Panel, to investigate measures implemented in long-term residential care facilities to reduce transmission of, morbidity and mortality resulting from SARS-CoV-2. The review has subsequently been updated and the updated review will be used to inform the understanding of potential risk and protective factors.

1.2 Assess available data

Relevant data sources that may provide information on the identified risk and protective factors will be identified and assessed. The main source of data on COVID-19 outbreaks in nursing homes will be the HPSC's Computerised Infectious Disease Reporting (CIDR) database. Under the Infectious Diseases (Amendment) Regulations 2003. SI No. 707 of 2003, outbreaks of infectious disease are notifiable to the medical officer of health.[1] Data on outbreaks are captured alongside case-based notification data in the CIDR. Also relevant will be the data on nursing home characteristics managed by HIQA as part of its regulation function. The analysis will also seek to identify whether there are any other additional sources of potentially applicable data for inclusion in the analysis. For the purposes of this analysis, a nursing homes is any designated centre for older people in Ireland registered with HIQA.

1.3 Collect data

Based on the identified risk factors, it will be necessary to gather data relevant to nursing homes and outbreaks to support an analysis. It is likely that direct measures of risk may not be available, but that proxies for that risk may be estimable from the data. For example, if it is not possible to determine occupancy for a nursing home, it may be acceptable to use bed capacity as a proxy for the number of residents in a home. For the purposes of this analysis, an outbreak is defined as two or more cases of laboratory confirmed COVID-19 regardless of symptomatic presentation, or two or more cases of illness with symptoms consistent with COVID-19 infection with at least one person confirmed as a case of COVID-19, as per the HSPC definition.[2] Suspected outbreaks may also be included as part of a sensitivity analysis.

The types of data required for the analysis include:

- Outcome – whether or not an outbreak occurred in a nursing home, and the number of cases included in the outbreak
- Size of nursing home (expressed as occupancy or bed capacity)

---

[1] https://www.hpsc.ie/notifiablediseases/notifyinginfectiousdiseases/
[2] https://www.hpsc.ie/a-z/respiratory/coronavirus/novelcoronavirus/casedefinitions/covid-19outbreakcasedefinitionforireland/

- Nursing home classification (that is to say publicly provided [HSE] home, privately provided [non-HSE])
- Incidence in locality of nursing home (expressed as 14 day incidence in the administrative area the home is in, or within a radius of the home)
- Dates of when public health infection control measures were introduced or lifted (for example, embargo on nursing home visits)
- Proximity to other nursing homes (for example, homes within 5km)
- Proximity to hospitals (for example, distance to nearest public acute hospital)
- Urban-rural status and or population density of local area
- Socio-economic status of local area (for example, average deprivation score).

If available, data on the physical structure of the home (for example, the number of single-occupancy rooms) and resources (for example, staffing levels) will also be considered. It must be recognised that data on occupancy and staffing are generally reported at a point in time and may not be applicable at other points in time. Data on compliance with standards will also be considered.

As there is an important temporal component to some of the variables, such as the occurrence of an outbreak and localised incidence, the data will be required by nursing home for each day during the epidemic.

Data on COVID-19 outbreaks from the HPSC will be provided to HIQA on the basis of a data sharing agreement that stipulates the secure transfer, conditions of use and disposal of data on completion of the analysis. The sharing of data will comply with data protection legislation.

1.4 Data analysis

The approach to data analysis will be influenced by the available data. The data may include both a temporal component (that is, when the outbreak occurred) and a spatial component (that is, where the outbreak occurred). The temporal component is important for understanding the conditions at the time of the outbreak, in particular the local incidence of COVID-19. The spatial component relates to the location of the nursing homes as homes close to each other will experience similar conditions (such as, local incidence of COVID-19, population density) that may be conducive to outbreaks occurring. The spatial component may be addressed by considering the exact geolocations of nursing homes or through an area-level identifier, such as county.

Due to the changing nature of the epidemic in terms of the demography of notified cases and the public health response, the data will be analysed as two distinct waves of COVID-19 in 2020: February to July, and August to November.

Potential analytical approaches

An initial analysis will consider the presence or absence of at least one outbreak in each home over a fixed time period (such as, February to June 2020). The analysis, in the form of a logistic regression, will provide an estimate of which covariates are associated with an increased probability of an outbreak over the time period. An associated analysis will include a zero-inflated Poisson or negative binomial model that estimates the number of cases in a home. A two stage model will be used that estimates the probability of an outbreak and, conditional on an outbreak being present, the size of the outbreak. These analyses will give an indication of which covariates are important and whether additional data collection or refinement of the selected covariates may be necessary. The weakness of the analysis will be that local conditions (for example, local incidence) will be expressed as a single summary figure or average over time, which may ignore the conditions at the time of the outbreaks occurring.

A more detailed analysis will model daily outbreak data using a multi-level modelling approach. Daily outbreak data will be nested in nursing homes which may then be nested within areas or incorporated as exact locations within a spatio-temporal model. Again, both logistic and zero-inflated negative binomial models will be considered to separately model the probability of outbreaks and the extent of outbreaks. A challenge in modelling daily data is how to handle additional outbreak cases once the outbreak has started, which may suggest the use of a time-series approach to account for autocorrelation in observations.

Suitable approaches for a spatial analysis may include geographically-weighted regression and conditional autoregressive (CAR) models. Alternative approaches to estimating neighbourhood matrices will be considered given the high density of nursing homes in certain town and city areas.

In the event that more detailed information on resources within homes (that is, staffing levels, types of rooms, dependency of residents) is available for a subset of homes, further analysis will be considered. An assessment of bias will be conducted to determine if the subset of homes are representative. If they are considered representative, then a matched case-control type analysis will be used where homes with outbreaks are matched to homes without outbreaks based on propensity score matching. This form of analysis will be used to determine if the additional data provide potentially important risk factors and whether recommendations for additional data collection are justified.

Types of outputs

The outputs of a logistic model will inform an understanding of the factors that influence the probability of a home having an outbreak. The output of a zero-inflated negative binomial model will consider both the probability of an outbreak as well as the extent of the outbreak conditional on it occurring.

A temporal model will also provide information on the importance of time-specific covariates, such as local incidence at the time of the outbreak occurring and the restrictive measures in place at the time.

Additional outputs will include information on the goodness-of-fit of various models and unexplained variance in the model.

In the event that more detailed information is available on a subset of nursing homes, sensitivity analyses will be considered to determine if the additional information might improve the models.

1.5 Summarise findings

A descriptive report of the data and analyses will be prepared by HIQA and the HPSC. The report will be presented to HIQA's COVID-19 Expert Advisory Group (EAG) for clinical input and interpretation. Feedback from the EAG will be incorporated into a final report which will be shared with the COVID-19 Nursing Homes Expert Panel. The final report will provide advice to the Minister for Health based on the analysis of factors associated with COVID-19 outbreaks in nursing homes.

Published by the Health Information and Quality Authority.

For further information please contact:

Health Information and Quality Authority

Dublin Regional Office

George's Court

George's Lane

Smithfield

Dublin 7

D07 E98Y

Phone: +353 (0) 1 814 7400

Email: info@hiqa.ie

URL: www.hiqa.ie

© Health Information and Quality Authority 2021