Office for National Statistics

Article

Coronavirus and the estimated impact on hospital episodes involving falls and fractures – sources and methods, England

The sources and methods used to estimate the impact of coronavirus (COVID-19) on hospital episodes involving falls and fractures associated with new-onset frailty and disability.

Contact:
Kathryn Littleboy
health.data@ons.gov.uk
+44 1633 651602

Release date:
6 June 2023

Next release:
To be announced

Table of contents

1 . Overview

This methodology provides details of the data sources and methods used in our Coronavirus and estimating the impact on hospital episodes involving falls and fractures, England: 2013 to 2021 article.

This research examined hospital episode records taken from the Admitted Patient Care (APC) pillar of the Hospital Episode Statistics database, focusing on records where specific International Classification of Diseases and Related Health Problems 2010 (ICD-10) codes relating to fractures, falls and frailty were present in primary or secondary diagnosis fields.

Autoregressive Integrated Moving Average (ARIMA) models were used to establish whether the number of hospital episodes involving falls and fractures in England varied significantly from their forecasted trajectory over the course of the coronavirus (COVID-19) pandemic, and to consider what this might tell us about the incidence of disability and frailty. The models were optimised individually for national, age-stratified, sex-stratified and region-stratified projections. We also considered whether geographical differences in social restrictions during the pandemic, such as tiered local lockdowns, show any correlation with fall and fractures as markers of new-onset disability and frailty.

2 . Data sources and variables

We extracted data from NHS England's Hospital Episode Statistics (HES) Admitted Patient Care (APC) database, based on International Classification of Diseases and Related Health Problems 2010 (ICD-10) codes of interest for pre-coronavirus (COVID-19) pandemic (1 January 2013 to 23 January 2020) and during pandemic years (24 January 2020 to 31 December 2021). A total of 144,148,915 hospital records were available for extraction for the pre-pandemic period, compared with 42,267,318.

A hospital episode is a time period within a hospitalisation where the patient is in the continuous care of one consultant or healthcare provider. A patient may be transferred from one consultant to another during their stay, which would mean there are two or more episode records for the hospitalisation. Therefore, it is possible that a single fall or fracture could be counted more than once in this analysis.

Hospital episodes were selected based on the presence of specific ICD-10 (2019) codes relating to fractures, falls and frailty in both primary and secondary diagnosis fields of APC records.

Fractures were defined using codes from Chapter 19 of the ICD-10, focused on the following commonly-experienced fractures by older people:

- S22.3 – Fracture of rib
- S22.4 – Multiple fractures of ribs
- S52.5 – Fracture of lower end of radius
- S72.0 – Fracture of femur

Falls were defined using the following ICD-10 codes:

- W01.0 – Fall on same level from slipping, tripping and stumbling
- W03.0 – Other fall on same level because of collision with, or pushing by, another person
- W05.0 – Fall involving wheelchair
- W06.0 – Fall involving bed
- W08.0 – Fall involving other furniture
- W18.0 – Other fall on same level
- W19.0 – Unspecified fall

Frailty was defined as any code that was also present in a predefined Hospital Frailty Risk Score index. For an episode to qualify as a frail fall or fracture, an individual must have a HES record containing a code from the Hospital Frailty Risk Score index, and a concurrent fall and fracture recorded on the same record. Data were binned to monthly frequency for time series modelling.

Episodes from 1 January 2013 to 23 January 2020 comprise the pre-pandemic sample. Episodes from 24 January 2020 to 31 December 2021 were used to represent actual episodes for pandemic years. Demographic information from the HES dataset was used to examine differences across age, sex and region.

3 . Method used for analysis

Pre-coronavirus (COVID-19) pandemic episode frequency was used to predict expected hospital episodes during the pandemic years using time series modelling, assuming that coronavirus had not occurred. Those predicted episode figures were directly compared with the actual number of hospital episodes. This was to assess changes because of the public health measures in place as part of the pandemic response, and any other changes because of the pandemic.

Hospital episodes in pre-pandemic years were stratified by age and geographical characteristics, and then averaged. They were then compared with pandemic year episodes to assess more granular changes to expected services.

Changes in patterns of hospital episodes involving falls and fractures were assessed using Autoregressive Integrated Moving Average (ARIMA) forecast models. ARIMA is a commonly-used statistical technique for time series modelling. Such models have shown strong predictive capability in the field of public health, with the ability to create short-term forecasts. ARIMA models are made up of three components: *p*, *d*,and *q.*

ARIMA model equation:

$$(1 - \emptyset_1 B - \ldots \emptyset_p B^p) \qquad (1-B)^d y_t \qquad = c + (1 + \theta_1 B + \cdots + \theta_q B^q)\,\varepsilon_t$$

AR(p) *d differences* *MA(q)*

Where:

- *p* describes the number of autoregressive (AR) terms
- *d* the number of difference terms required to render the series stationary
- *q* the number of moving average (MA) terms within the model

AR terms incorporate past values of a series, where MA terms incorporate past errors. Stationarity of underlying data were assessed by augmented Dickey-Fuller tests to determine the number of difference orders. ARIMA-estimated future values are therefore dependent on both past values of the same variable and past errors when predicting that variable.

Seasonal ARIMA models, as featured here, use differencing at a lag equal to the number of season(s) to remove additive season effects. Additional *P*, *D* and *Q* hyperparameters specify the AR, differencing and MA orders for the seasonal component of the series.

Seasonal ARIMA model equation:
$\varphi(B)\,\Phi(B^s)\,(1-B)^{(d)}(1-B^s)^D z_t = \theta(B)\,\Theta(B^s)\,a_t$

Where:

- z_t is a time series
- *B* is the backshift operator $(B_{zt} = z_{t-1})$
- *s* is the seasonal period
- $(B) = (1 - 1B \cdots {}_pB^p)$ is the nonseasonal AR operator
- $(B_s) = (1 - 1B_s \cdots P\ BP\ s)$ is the seasonal AR operator
- $(B) = (1 - 1B \cdots {}_qB^q)$ is the nonseasonal MA operator
- $(Bs) = (1 - 1Bs \cdots QBQs)$ is the seasonal MA operator

We used autocorrelation and partial autocorrelation functions to identify the autoregressive orders (p, P) and moving average orders (q, Q) in each model.

Time series modelling accuracy is typically assessed in a stepwise manner, validating the predicted value against known observations. The start of the pandemic stopped forecasted values from such validation techniques. Model accuracy was therefore assessed before the start of the pandemic, using a validation period of 2018 to 2019. Accuracy was determined across a five-year training period, using data back to 2013, in a monthly stepwise manner using root mean squared error and mean absolute percentage error metrics.

Models that minimised these error metrics were considered the most suitable for generating pandemic year forecasts. Seasonal decompositions (additive and multiplicative) were assessed by comparing non-transformed and log transformed episode frequencies. The Akaike information criterion (AIC) was smallest in log transformed models, suggesting a multiplicative decomposition model would produce a better forecast than an additive decomposition model.

Models were optimised individually for national, age-stratified, sex-stratified and region-stratified forecasts for each respective episode group. Age groups were split into the following four categories and enumerated at hospital episode age:

- children (aged 18 years and under)
- working age adults (aged 19 to 64 years)
- pension age adults (aged 65 to 79 years)
- the elderly (aged 80 years and over)

Episodes were filtered to International Territorial Level (ITL) regions (for example, North West) through completed fields in Hospital Episode Statistics (HES) records. Some 69,944 records were lacking regional information and were excluded from regional stratifications.

Aggregated models were considered accurate because of low mean absolute percentage error (MAPE) metrics (Table 1). National models were subject to a six-year training period extending backwards from December 2018. Accuracy was assessed against observed episodes for 2019.

Table 1: Model accuracy parameters for fractures, falls, and frail fall and fracture episodes, England, 2013 to 2018

Cause	Model Parameters	AICc	Mean Absolute Percentage Error (MAPE)
Fractures	ARIMA(1, 0, 1), (2, 1, 0) [12]	1478.55	2.56%
Falls	ARIMA(1, 0, 1), (1, 1, 0) [12]	1070.02	1.41%
Frail Fall and Fracture	ARIMA(0, 1, 1), (1, 1, 1) [12]	860.41	2.29%

Source: Coronavirus and the estimated impact on hospital episodes involving falls and fractures from the Office for National Statistics

Strengths and limitations of the methodology used

The strengths of the methodology include:

- the large and representative national datasets used
- the length of historical data available pre-pandemic (1 January 2013 to 23 January 2020) enabled robust modelling of predicted rates for adverse outcomes

The limitations of the methodology include:

- the extent that falls and fractures can be relied on as proxy indicators of frailty and disability

- the reliability that the hospital-wide frailty index determines premorbid frailty (although it has been well validated and applied carefully, following the protocols laid out in the Development and validation of a Hospital Frailty Risk Score focusing on older people in acute care settings using electronic hospital records: an observational study, published by the National Library of Medicine)

- the ICD-10 code listed on HES records, either in primary or secondary diagnosis fields, may not be the same reason that a patient was admitted to hospital; during the pandemic, incidental COVID-19 may have often been recorded as the primary diagnosis for hospital attendance, even if the reason for attendance was a fall or fracture

- it is possible that fewer episodes involving falls or fractures occurred during the pandemic because of reluctance to attend a hospital facility

- episodes data are inherently observational, so causal inference cannot be drawn

- HES Admitted Patient Care (APC) records account for the most serious cases, but there may be many A&E attendances that are not accounted for; most hospital activity related to falls may be captured in the Emergency Care Data Set (ECDS), however relevant condition fields in ECDS records are not mandatory and are only completed in a quarter of cases where attendance was injury-related, so any A&E data would be incomplete -- focusing on APC records means a more complete picture of serious cases likely to result in new-onset disability can be captured

- forecasting models used for this analysis make assumptions that all background conditions remain at a constant rate of change, which is not a realistic assumption

- if a patient is transferred between consultants or healthcare providers during their hospitalisation, this may be recorded as more than one hospital episode; therefore, it is possible that a single fall or fracture could be counted more than once in this analysis

4 . Glossary

Hospital episode

A hospital episode is a time period within a hospitalisation where the patient is in the continuous care of one consultant or healthcare provider. A patient may be transferred from one consultant to another during their stay, which would result in there being two or more episode records for the hospitalisation.

Pandemic period

For the purposes of this research, the pandemic period was defined as 24 January 2020 to 31 December 2021.

Pre-pandemic period

For the purposes of this research, the pre-pandemic period was defined as 1 January 2013 to 23 January 2020.

Autoregressive Integrated Moving Average (ARIMA) forecast models

ARIMA is a commonly used statistical technique for time series modelling.

5 . Future developments

Further research is required to understand how coronavirus (COVID-19) infection, social isolation, immobility, and physical and cognitive deconditioning each contributed to observed variations of hospital episodes in England involving falls and fractures from their expected trajectory. We hope to use risk modelling to assess the risk of experiencing a fall, fracture, or frail fall and fracture, if a patient reported a positive case of COVID-19.

6 . Related links

Coronavirus and the estimated impact on hospital episodes involving falls and fractures, England: 2013 to 2021
Article | Released 28 April 2023
An experimental analysis estimating the impact of Coronavirus (COVID-19) on the number of hospital episodes involving falls and fractures associated with new onset frailty and disability.

7 . Cite this methodology

Office for National Statistics (ONS), released 28 April 2023, ONS website, methodology, Coronavirus and the estimated impact on hospital episodes involving falls and fractures – sources and methods, England