The University Of Sheffield.

The use of English Admitted Patient Care Hospital Episode Statistics to inform time-trends in the rates and costs of secondary care associated with tobacco and alcohol

Version 1.1.0

December 2020

Duncan Gillespie[1], Laura Webster[1], Colin Angus[1] & Alan Brennan[1]

[1]School of Health and Related Research (ScHARR), The University of Sheffield

Address for correspondence:
Dr Duncan Gillespie
Section of Health Economics and Decision Science,
School for Health and Related Research,
The University of Sheffield,
Regent Court, Regent Street, Sheffield, S1 4DA, UK
Email: duncan.gillespie@sheffield.ac.uk

WARNING

This is a working version that is subject to review and future versions are likely.

The code that accompanies this report can be found here in the hesr R package (https://stapm.gitlab.io/r-packages/hesr/). This report is licensed to The University of Sheffield under a CC by 4.0 license.

Summary

This methodology report sets out a description of how we estimate the rates and costs of English inpatient hospital care attributable to tobacco and/or alcohol. We review past methods and then explain our own choice of methods.

Glossary of terms

Table 1: Explanation of terms.

Term	Definition
NHS	The UK's National Health Service
NHS Digital	The national information and technology partner to the health and care system <https://digital.nhs.uk/>
Hospital Episode Statistics (HES)	A database containing details of all admissions, A and E attendances and outpatient appointments at NHS hospitals in England.
Secondary care	Planned (elective) care such as a cataract operation, or urgent and emergency care such as treatment for a fracture
Inpatient care	Care of patients whose condition requires admission to a hospital
ICD-10 code	International Statistical Classification of Diseases and Related Health Problems 10th Revision https://icd.who.int/
OPCS code	Details of any procedures or interventions performed. Office of Population, Censuses and Surveys: Classification of Interventions and Procedures, 4th revision
Diagnostic fields or positions	The list of ICD-10 disease diagnosis codes recorded for an episode of care
Finished consultant episode (FCE) or just 'episode'	The time a patient spends in the continuous care of one consultant using hospital site bed of one health care provider
Spell	The total continuous stay of a patient using a hospital bed on premises controlled by a health care provider during which medical care is the responsibility of one or more consultants. Also known as "provider spells"
Continuous inpatient spell (CIPS)	The total continuous stay of a patient using a hospital bed across premises controlled by multiple health care providers (i.e. joining spells where patients were transferred between providers). Also known as "super spells"
Population Attributable Fractions (PAFs)	The proportional reduction in population disease or mortality that would occur if exposure to a risk factor were reduced to an alternative ideal exposure scenario
Primary diagnosis	The ICD-10 disease code recorded first in the order of diagnoses given for an episode of care
Secondary diagnoses	Disease diagnoses that are recorded after the primary diagnosis
Healthcare Resource Group (HRG) code	HRGs are a measure of case mix, presenting standard groupings for clinically similar treatments, which consume a common set of health care resources
Narrow measure	A method of assigning a single diagnosis to a spell of inpatient care by looking at the primary diagnosis of the admission episode of the spell, and also looking at whether any external causes were recorded
Broad measure	A method of assigning a single diagnosis to a spell of inpatient care by looking at all diagnosis codes and selecting the diagnosis most strongly associated with the risk factor(s) being investigated
Admission episode	The episode of care that comes first among the episodes within a spell or CIP
Hospital admission	The start of a spell or CIP of hospital care
Secondary Use Services (SUS)	SUS is a secure data warehouse that stores patient-level information and applies complex derivations which support National Tariff reimbursement https://digital.nhs.uk/services/secondary-uses-service-sus

Contents

Summary **2**

Glossary of terms **3**

1 Introduction **5**

2 Background **5**

2.1 The HES Admitted Patient Care data . . . 5

2.1.1 Episodes and spells . . . 5

2.1.2 Diagnosis and procedure codes . . . 5

2.1.3 Costs . . . 6

2.2 Programme of work . . . 6

2.2.1 Objectives . . . 6

2.2.2 Diseases attributable to tobacco and/or alcohol . . . 6

3 The aim of this report **6**

4 Review of existing methods **7**

4.1 Attributing inpatient care to specific diseases . . . 7

4.1.1 Primary diagnostic position of the admission (first) episode . . . 7

4.1.2 Narrow and broad methods for alcohol . . . 8

4.1.3 Primary diagnostic position of each episode in a continuous inpatient spell . . . 9

4.2 Estimating the costs of hospital admissions that are attributed to tobacco and alcohol consumption 9

4.2.1 Cost the entire spell . . . 9

4.2.2 Cost only the highest costing episode . . . 10

4.2.3 Cost only the episodes with a relevant primary diagnosis . . . 10

5 Our choice of methods **10**

5.1 Step-by-step protocol . . . 10

5.2 Variable costs . . . 11

5.2.1 Excess bed days . . . 11

5.2.2 Unbundled costs . . . 11

5.3 Calculation of unit costs . . . 11

5.4 Calculation of the rates secondary care use . . . 11

5.4.1 Partitioning people among single-morbidity health states . . . 12

5.4.2 Calculating the expected number of hospital visits in a year (the multiplier) . . . 12

References **12**

1 Introduction

The NHS is under increasing pressure from a growing demand for secondary care, and inpatient care is the main contributor to the total cost of healthcare provision. A large proportion of inpatient care is preventable because it stems from past patient exposure to risk factors for ill-health, of which tobacco and alcohol rank highly. Being able to link past tobacco and alcohol consumption to current inpatient care is therefore important to inform the potential health and economic consequences of public policy decisions that could affect tobacco and/or alcohol consumption. From a health sector perspective, public health policies aim to reduce the demand for and consequent public financial costs of healthcare. In doing so there is also the additional aim of narrowing the societal inequalities in health, i.e. it would be better for the greatest savings in the costs of care to come from those sectors of society who have the highest demand for care.

2 Background

For England, NHS Digital statistics estimated that in England in 2017/18 there were 489,300 hospital admissions attributable to smoking (4% of all hospital admissions) (NHS Digital 2019b), and 337,870 hospital admissions attributable to alcohol (narrow measure) (NHS Digital 2019a). However, these publications stopped short of estimating the financial cost of these admissions. A report by Public Health England on the costs of tobacco to the NHS in England, estimated that in 2015/16, there were 520,000 smoking attributable hospital admissions, amounting to a cost of £851.6 million (Public Health England 2017).

2.1 The HES Admitted Patient Care data

We use data from the Admitted Patient Care (APC) component of the English Hospital Episode Statistics (HES) for patients resident in England, which provide a huge resource of data for the estimation of the rates and costs of tobacco and alcohol related hospital care (Herbert et al. 2017). Each HES record contains a wide range of information about a patient including clinical information about diagnoses and operations, patient information such as age, sex, and socio-economic conditions (deprivation quintile of residence), and administrative information such as dates of admission and discharge.

2.1.1 Episodes and spells

HES APC information is recorded at the finished consultant episode (FCE) level, with an episode being a period of care within a particular consultant specialty at a single hospital provider. Episodes are frequently joined together to create provider spells, which refer to the time spent within a single hospital, or continuous inpatient spells (CIPS), which refer to the entire period of continuous inpatient care across multiple hospitals. An episode can be finished because of discharge to another consultant, transfer to another hospital or to place of residence or if the patient dies. As a result, a patient may have more than one FCE during an admission to hospital, most usually because they are transferred from one speciality to another (one estimate was that 86.7% of all spells only contained a single episode (Geue et al. 2012)).

2.1.2 Diagnosis and procedure codes

Each FCE has a series of codes to indicate which procedures and diagnoses have been carried out. HES uses the ICD-10 (International Classification of Disease and Related Health Problems, 10th revision) classification to record diagnosis information. An episode can contain up to twenty diagnostic codes, with the first diagnostic code being the primary diagnosis. The OPCS-4 (Office of Population, Censuses and Surveys: Classification of Interventions and Procedures, 4th revision) classification is used to record details of any procedures or interventions performed.

2.1.3 Costs

Healthcare Resource Group (HRG) grouper software is used to assign an HRG code to every FCE. The software assigns an HRG code based on diagnosis, procedure performed, gender, age, length of stay, and the discharge method. Additional variable costs might then be added to account for long stays in hospital and some procedures that are not considered within the HRG algorithm. Unit costs for FCEs can be derived from national-level NHS reference costs (Department of Health 2016), which are linked to the HRG code and the type of episode (elective, non-elective, day case, regular attender). A HRG code can also be assigned to the entire inpatient spell, and this is done using a further algorithm. It should be expected that the spell-level HRG is associated with a cheaper reference cost than the sum of the episode-level HRGs, e.g. because the episode-level costs might double-count some of the costs associated with a spell.

2.2 Programme of work

Investigating the past trends in hospital admissions and costs associated with tobacco and alcohol consumption is fundamental to evaluating the effect of past policy and appraising the potential effects of new policy. We use HES APC data from 2002/3 for an on-going programme of work at The University of Sheffield to develop public health economic models for policymaker decision support in the fields of alcohol and tobacco control. This work builds on the methods used in the Sheffield Alcohol Policy Model (Brennan et al. 2009, 2015) and is being conducted at a range of geographic scales, including national and local authority level.

2.2.1 Objectives

The programme of work has two objectives:

1. To investigate past trends in alcohol and/or smoking related hospital admissions and costs. The data will also be used to investigate past policy effects on hospital admissions and associated costs.

2. To investigate the potential future effects on hospital admissions and costs of proposed changes to alcohol and/or tobacco policy. We typically use the most recent years of data to inform the baseline rates of hospital admissions. We then update these rates according to estimated policy effects and project the outcomes over around 30+ years.

2.2.2 Diseases attributable to tobacco and/or alcohol

We use the HES APC data to investigate the rates and costs of hospital admissions associated with a set of conditions that have been classed as wholly- or partially-attributable to the consumption of tobacco or alcohol. Our current list of conditions is large (84 separate ICD-10 categories in total) and includes several cancers, cardiovascular diseases and respiratory diseases. We consider 52 diseases related to tobacco (Webster et al. 2018) and 45 diseases related to alcohol (Angus et al. 2018) (we expand the alcohol list to 48 conditions when merging our tobacco and alcohol disease lists by subdividing some of the ICD-10 categories). Within our merged disease list, there are 15 conditions that are related to both tobacco and alcohol.

3 The aim of this report

There are three main problems to be solved in linking past patient tobacco and alcohol consumption to trends in the rates and costs of inpatient care:

1. How to attribute a case of a disease in an individual to their current or past tobacco and/or alcohol consumption. For this, the NHS Digital statistics use survey data on tobacco or alcohol consumption

and estimates of the relative risks of disease associated with tobacco or alcohol consumption to calculate Population Attributable Fractions (PAFs), which give an estimate of the percentage of disease cases that are likely to be due to tobacco or alcohol consumption.

2. How to attribute inpatient care to those specific diseases.

3. How to estimate the costs of hospital admissions that are attributed to tobacco and alcohol consumption.

We address (1) in a separate report on tobacco and alcohol attributable fractions [*link to that report*]. This report is focused on explaining our choice of methodology for (2) Attributing inpatient care to specific diseases, and (3) Estimating the costs of hospital admissions that are attributed to tobacco and alcohol consumption.

4 Review of existing methods

We reviewed statistical reports and peer-reviewed publications that measure patient admissions using the HES APC data for England. We time-limited our search to publications after 1998, when patients' NHS numbers became a mandated return from hospitals, allowing episodes to be linked to the same patient within the HES data. Whilst we primarily searched for publications related to tobacco or alcohol, methods used by other analyses will still be useful and so we did not limit our search to exclusively tobacco or alcohol related publications. We focus in particular on methods already used in work by The University of Sheffield (see the Sheffield Alcohol Research Group webpage), and in the English national statistics on tobacco and alcohol published by NHS Digital (NHS Digital 2019b, 2019a).

4.1 Attributing inpatient care to specific diseases

Following an admission to hospital, an individual might see multiple consultants resulting in multiple episodes of care. Within each episode of care, there can be up to twenty diagnostic codes assigned. Assigning a single diagnosis as the cause of an admission can therefore be a task filled with uncertainty. Table 2 shows the main methods that we consider.

Table 2: Four alternative methods to attribute inpatient care to specific diseases.

Name	Description
Primary diagnostic position of the admission (first) episode	Uses admission episode only (we take this to mean the first episode of a CIPS). Most common to use the primary diagnosis of the admission episode. Might also scan for relevant diagnosis codes in the first three or all diagnostic positions of the admission episode.
Narrow method for alcohol	Uses admission episode only. Looks for alcohol-related diagnosis codes in the primary diagnoistic position, and looks for alcohol-related external causes in all diagnostic positions.
Broad method for alcohol	Uses all episodes in a spell (we take this to mean a CIPS). Chooses the alcohol-related diagnosis with the highest alcohol PAF. If two diagnoses have the same PAF, then the earlier recorded diagnosis is chosen.
Primary diagnostic position of each episode in a continuous inpatient spell	Looks at the first diagnostic position of all episodes within a CIPS and retains episodes that have a relevant diagnosis in the primary diagnostic position. External causes can be included in the attribution of a diagnosis to each episode following the narrow method for alcohol. The admission is then attributed to the diagnosis in the earliest retained episode.

4.1.1 Primary diagnostic position of the admission (first) episode

The NHS Digital statistics on the number of hospital admissions attributable to smoking (NHS Digital 2019b) refer to the number of finished admission episodes that have a primary diagnosis related to smoking, i.e. the

ICD-10 disease code that appears in the first position in the list of disease diagnoses associated with the admission episode. This method was also used in an analysis of the costs of smoking by Public Health England (Public Health England 2017). However, using this method to select risk-related admissions is likely to underestimate the rates of hospital admissions associated with a specific diagnosis.

Another approach used in the literature to attribute a disease diagnosis as the cause of an admission to hospital is to look at whether the diagnosis code for that disease occurs within the *first three diagnostic positions* of the admission episode (Keeping et al. 2015; Hobbelen et al. 2016). This was designed to balance the risk of incorrectly excluding admissions as being attributable to a diagnosis by using too strict a definition with the risk of incorrectly including admissions not due to the condition by using too broad a definition. Hobbelen et al (2016) then conducted sensitivity analyses to investigate how using the first three diagnostic positions compared to:

1. Using the primary diagnosis of the admission episode (which reduced the yearly number of attributed hospital admissions by 25%);
2. Using all diagnostic fields of the admission episode (which increased the yearly number of attributed hospital admissions by 3.4%).

4.1.2 Narrow and broad methods for alcohol

The NHS Digital statistics on alcohol (NHS Digital 2019a) use two methods to associate alcohol-related diagnoses with inpatient care:

1. The 'narrow' measure looks only at whether the primary diagnosis of the admission episode is alcohol-related or whether an alcohol-related external cause has been recorded within the list of diagnoses related to that episode, e.g. assault (external conditions, which are environmental causes of injury occurring outside the body, do not feature as the primary diagnosis). This is in an effort to select the condition that is the cause of the episode. If there is an alcohol-related external cause, then this is assigned the cause of the episode; if there is no alcohol-related external cause, then the diagnosis in the primary diagnostic position is used. If this results in more than one alcohol-related diagnosis, then the diagnosis with the largest alcohol PAF is used. Green (2017) analysed trends in alcohol-related admissions to hospital by age, sex and socio-economic deprivation in England using the narrow measure.
2. The 'broad' measure looks across all the diagnoses associated with a spell in hospital (which we take to mean a CIPS) and attributes the admission to hospital to the diagnosis most strongly associated with alcohol, i.e. the diagnosis with the largest alcohol PAF. If two or more diagnoses have the same PAF, then the earliest recorded diagnosis is used. This method has been used to inform estimates of alcohol attributable hospital admissions (Jones et al. 2008); the Sheffield Alcohol Policy Model also uses it to estimate the baseline rates of hospitalisations used in modelling the potential impact of policies on alcohol attributable hospital admissions (Brennan et al. 2009).

NHS Digital (NHS Digital 2019a) explain that "*The 'broad' measure is a better indicator of the total burden that alcohol has on health services as it takes more account of secondary diagnoses than the 'narrow' measure. However, since secondary diagnosis fields have become better populated over time*[1]*, this impacts upon time series comparisons for the 'broad' measure as increases can be partly due to an improvement in data quality rather than a real effect. Consequently, the 'narrow' measure is a better indicator of changes over time.*" For example in 2017/18, it was estimated that alcohol was the main reason for 2.1% of all admissions to hospital using the narrow measure, but using the broad measure, it was estimated that 7.2% admissions were linked to alcohol.

[1]The number of possible diagnosis codes has increased from 14 to 20, and the clinical practice has changed so that more secondary diagnosis codes are being recorded.

4.1.3 Primary diagnostic position of each episode in a continuous inpatient spell

This method looks at the first diagnostic position of all episodes within a CIPS and retains episodes that have a relevant diagnosis in the primary diagnostic position (i.e. we look at all episodes, not just the admission episode). External causes can be included in the attribution of a diagnosis to each episode following the narrow method for alcohol. This is the method that we applied to estimate the cost of tobacco to secondary care for Chapter 3 of the Royal College of Physician's report "Hiding in plain sight: Treating tobacco dependency in the NHS" (Tobacco Advisory Group of the Royal College of Physicians 2018).

The admission is then attributed to the diagnosis in the earliest retained episode, e.g. if the second and third episodes within a CIPS both have relevant diagnoses in the primary diagnostic position, then the admission is attributed to the diagnosis in the primary diagnostic position of the second episode.

The method is a merger of the narrow and broad methods, but with the removal of the use of the PAF in attributing a diagnosis as the cause of the admission. It balances the risk of incorrectly excluding admissions as being attributable to a diagnosis by looking across all episodes within an admission with the risk of incorrectly including admissions not due to the condition by using only the primary diagnostic positions of each episode. This method should also give a better indication of the total burden that tobacco and alcohol have on health services by taking account of all episodes of care. Further, it should be a robust indicator of changes over time by only making use of the primary diagnostic position within each episode.

4.2 Estimating the costs of hospital admissions that are attributed to tobacco and alcohol consumption

There are many variations of costing HES APC data and we draw on four sources to understand these and to support choosing our own method (Geue et al. 2012; Leal, Manetti, and Buchanan 2018; Asaria, Grasic, and Walker 2016; Gaughan et al. 2012). Our aim in choosing a method was to balance the risk of incorrectly excluding costs of admissions that are due to the diagnosis assigned to the admission with the risk of incorrectly including costs that are *not* due to the diagnosis assigned to the admission. Table 3 shows the three methods that we consider. Below, we critically review these methods.

Table 3: Three alternative methods to estimate the costs of hospital admissions that are attributed to tobacco and alcohol consumption.

Name	Description
Cost the entire spell	Cost all episodes associated with an admission.
Cost only the highest costing episode	For patients who had a single episode, take the cost of that episode. For patients who had multiple episodes, take the cost of the spell to be the cost of the highest costing episode within that spell plus any variable costs from other episodes.
Cost only the episodes with a relevant primary diagnosis	For patients who had a single episode with a relevant primary diagnosis, take the cost of that episode. For patients who had multiple episodes with the same primary diagnosis, take the cost of the spell to be the cost of the highest costing episode with the diagnosis within that spell plus any variable costs from other episodes with the same primary diagnosis.

4.2.1 Cost the entire spell

An analysis of the costs of smoking to the NHS due to hospital admissions by Public Health England (Public Health England 2017) assigned admissions to a smoking-related diagnosis based on the primary diagnostic position of the admission episode, and then estimated the cost of the full provider spell rather than just the cost of the admission episode. The Healthcare Resource Group (HRG) was identified using the HES field "SUS generated core Spell HRG", i.e. the spell-level HRG. Where this field was unavailable, the HRG associated with the episode was used to cost the admission.

However, there is evidence to suggest that for admissions comprising more than one episode, where the cost estimated from the spell-level would be expected to be less that the sum of the episode-level costs (e.g. due to potential double-counting of costs in the episode-level HRGs), the estimated cost can be higher when using the spell-level HRG (Leal, Manetti, and Buchanan 2018). Leal et al. (2018) interpreted this as raising questions about the spell-based HRG allocation algorithms and the accuracy of the validation checks of spell- and episode-level costs submitted by each hospital provider.

4.2.2 Cost only the highest costing episode

For patients who had multiple episodes, a simple way to estimate the cost of the spell is to use the cost of the highest-costing episode (Geue et al. 2012; Asaria, Grasic, and Walker 2016). Geue et al. (2012) then recommended adding the variable costs associated with other episodes to this cost (e.g. the costs of excess bed days and certain procedures that are not accounted for by the standard episode-level cost).

4.2.3 Cost only the episodes with a relevant primary diagnosis

The method that we applied to estimate the cost of tobacco to secondary care for Chapter 3 of the Royal College of Physician's report (Tobacco Advisory Group of the Royal College of Physicians 2018) retained all episodes within a spell that had a relevant diagnosis in the primary diagnostic position. For the RCP report, we then summed the costs of these episodes by diagnosis, before attributing the cost to smoking using the diagnosis specific smoking PAF. Whilst this method was conservative by only counting the cost of episodes with a smoking-related primary diagnosis, two further rules should be applied to arrive at a conservative unit cost for a smoking-related admission to hospital.

1. Within a CIPS, only the costs of episodes with a primary diagnosis that matches the diagnosis assigned to the admission should be counted. This would mean excluding all costs associated with episodes of care within the CIPS that do not have this diagnosis code, even if they are associated with a different smoking-related diagnosis.

2. If there are multiple episodes within the CIPS that share a primary diagnosis with the diagnosis assigned to the admission, then the cost of the admission should be the cost of the highest-costing of these episodes plus any variable costs from other episodes with the same primary diagnosis.

5 Our choice of methods

The method that we choose is based on attributing inpatient care to specific diseases using the primary diagnostic position of each episode in a continuous inpatient spell (see Section 4.1.3), and estimating the costs of hospital admissions that are attributed to tobacco and alcohol consumption by costing only the episodes with a relevant primary diagnosis (see Section 4.2.3).

5.1 Step-by-step protocol

1. Look at the first diagnostic position of all episodes within a CIPS and retain episodes that have a relevant diagnosis in the primary diagnostic position. Also look for whether a relevant external cause has been recorded within the list of diagnoses related to that episode.

2. Attribute the admission to the diagnosis in the earliest retained episode.

3. Estimate the cost of the retained episodes using the episode-level HRG codes.

4. Within a CIPS, select the cost of the highest-cost episode.

5. If there are multiple episodes within the CIPS that share a primary diagnosis with the diagnosis assigned to the admission, then add any variable costs from these episodes to the cost of the highest-cost episode.

5.2 Variable costs

We include two types of variable cost: excess bed days; unbundled costs.

5.2.1 Excess bed days

A per diem costing was applied to excess bed days that occurred during the HES data year beyond the standard number of days anticipated for a given HRG (known as the "trim point"). Trim points are used to define a threshold: bed days occurring within the trim point (truncated bed days); and bed days occurring above the trim point (excess bed days). Any excess bed days are accounted for by calculating how many days over the trim point an episode lasted, and this is then multiplied by the unit cost of an excess bed day and summed to the total cost an of episode.

5.2.2 Unbundled costs

We account for additional costs of treatment given to regular attenders. These costs are generally high, related to a limited number of patients and reported separately by providers. For chemotherapy, radiotherapy and renal dialysis, many episodes have a zero-cost code for same day treatment (there is no national tariff due to a wide regional variation in costs and practice (Keeping et al. 2017)). For episodes with a primary diagnosis of cancer, we added chemotherapy and/or radiotherapy costs; these are costed separately as high-cost treatments. For episodes with a primary diagnosis of chronic kidney disease or end-stage renal failure, we included additional renal dialysis costs. Our episode costs for episodes with a primary diagnosis of cancer or chronic kidney disease or end stage renal disease therefore include a cost based on the SUS generated HRG code and one or more procedures.

5.3 Calculation of unit costs

To inform our modelling, we calculate the unit costs of a hospital admission for each tobacco and/or alcohol related diagnosis category, and stratify these costs by age category, sex and Index of Multiple Deprivation (IMD) quintile. We calculate these costs using a single year of the HES APC data (currently our unit costs are calculated based on the 2016/17 data). When calculated, the unit costs therefore corresponded to the monetary value in that year. If we need the unit costs to correspond to a different year, then we deflate or inflate them as necessary using the Hospital and Community Health Services (HCHS) pay and price inflation index (Curtis and Burns 2016).

5.4 Calculation of the rates secondary care use

Our modelling makes the conservative assumption that an individual can only have one disease related to tobacco or alcohol, i.e. there is no multi-morbidity in our model, only single morbidity. This assumption is made following the methodology of the Sheffield Alcohol Policy Model (SAPM) in defining health states in terms of *person-specific single morbidity* (PSSM), which means that *we assign each individual to either one of our tobacco or alcohol related diagnosis categories, or to "no tobacco or alcohol related diseases"* (2009). Since we consider such a large set of tobacco and alcohol related diseases (Section 2.2.2), there would be a vast array of multi-morbid health states through which we would need to track individual transitions, and this was not necessary or computationally efficient for our modelling.

Below, we describe how we use our HES APC analysis to:

1. Partition the people in each year of the HES APC data among single-morbidity health states;

2. Calculate the average number of times in a year that people in each single-morbidity health state would be expected to have an admission to hospital associated with the disease assigned to that health state. We refer to this as the *multiplier* because multiplying the number of individuals in each single-morbidity health state by these values gives the total number of admissions expected to be associated with that disease in a year.

5.4.1 Partitioning people among single-morbidity health states

To partition people among single-morbidity health states, we start with the admission-level dataset that shows the diagnosis assigned to each admission. We then look across all the admissions that each individual has within each year of the HES APC data, and assign each individual the diagnosis associated with their first tobacco- or alcohol-related admission of the year. This creates an individual-level "*person-specific single-morbidity dataset*". We then filter the admission-level dataset to retain only the admissions of an individual that share the diagnosis assigned to that individual. This creates an admission-level "*admission-specific single-morbidity dataset*".

Thus, if the same person was admitted on two separate occasions for oesophageal cancer during the year, then this person would be assigned "oesophageal cancer". If an individual is admitted on two different occasions for two different reasons, first for oesophageal cancer and subsequently for asthma, then they would also be assigned "oesophageal cancer".

The corresponding number of people in the "no tobacco or alcohol related diseases" is calculated by subtracting the number of people assigned to a tobacco- or alcohol-related health state from the mid-year population size. The proportion of people in each health state is calculated by dividing the annual number of people in each health state by the mid-year population size. We stratify the partitioning of people among health states by age category, sex and IMD quintile.

5.4.2 Calculating the expected number of hospital visits in a year (the multiplier)

For individuals assigned to a tobacco- or alcohol-related health state, we use the HES APC data to calculate the average number of times in a year that they would be expected to have an admission to hospital associated with the disease assigned to that health state. We refer to this value as the *multiplier* because multiplying it by the number of individuals in each health state gives the total number of admissions associated with that health state in a year. We calculate multipliers as the disease-specific ratio between the number of admissions in the *admission-specific single-morbidity dataset* and the number of people in the *person-specific single-morbidity dataset*. We also stratify our calculation of the multipliers by age category, sex and IMD quintile.

References

Angus, Colin, M Henney, L Webster, and Duncan Gillespie. 2018. *Alcohol-Attributable Diseases and Dose-Response Curves for the Sheffield Alcohol Policy Model Version 4.0*. https://doi.org/10.15131/shef.data.6819689.v1.

Asaria, Miqdad, Katja Grasic, and Simon Walker. 2016. "Using Linked Electronic Health Records to Estimate Healthcare Costs: Key Challenges and Opportunities." *Pharmacoeconomics* 34 (2): 155–60. https://link.springer.com/content/pdf/10.1007/s40273-015-0358-8.pdf.

Brennan, Alan, Petra Meier, Robin Purshouse, Rachid Rafia, Yang Meng, Daniel Hill-Macmanus, Colin Angus, and John Holmes. 2015. "The Sheffield Alcohol Policy Model – a Mathematical Description." Journal Article. *Health Economics* 24 (10): 1368–88. https://doi.org/10.1002/hec.3105.

Brennan, A, R Purshouse, K Taylor, R Rafia, A Booth, D O'Reilly, T Stockwell, et al. 2009. "Modelling the Potential Impact of Pricing and Promotion Policies for Alcohol in England: Results from the Sheffield Alcohol Policy Model Version 2008 (1-1). Independent Review of the Effect of Alcohol Pricing and Promotion: Part B." Report. University of Sheffield. https://www.sheffield.ac.uk/polopoly_fs/1.95621!/file/PartB.pdf.

Curtis, L, and A Burns. 2016. "Unit Costs of Health and Social Care 2015. Canterbury: PSSRU; 2015." https://www.pssru.ac.uk/project-pages/unit-costs/unit-costs-2015/.

Department of Health. 2016. "NHS Reference Costs 2015 to 2016." https://www.gov.uk/government/publications/nhs-reference-costs-2015-to-2016.

Gaughan, James Michael, Anne Mason, Andrew David Street, and Padraic Ward. 2012. "English Hospitals Can Improve Their Use of Resources: An Analysis of Costs and Length of Stay for Ten Treatments." http://eprints.whiterose.ac.uk/136538/.

Geue, Claudia, James Lewsey, Paula Lorgelly, Lindsay Govan, Carole Hart, and Andrew Briggs. 2012. "Spoilt for Choice: Implications of Using Alternative Methods of Costing Hospital Episode Statistics." *Health Economics* 21 (10): 1201–16. https://onlinelibrary.wiley.com/doi/pdf/10.1002/hec.1785.

Green, Mark A, Mark Strong, Lucy Conway, and Ravi Maheswaran. 2017. "Trends in Alcohol-Related Admissions to Hospital by Age, Sex and Socioeconomic Deprivation in England, 2002/03 to 2013/14." *BMC Public Health* 17 (1). https://doi.org/10.1186/s12889-017-4265-0.

Herbert, Annie, Linda Wijlaars, Ania Zylbersztejn, David Cromwell, and Pia Hardelid. 2017. "Data Resource Profile: Hospital Episode Statistics Admitted Patient Care (Hes Apc)." *International Journal of Epidemiology* 46 (4): 1093–1093i. https://doi.org/10.1093/ije/dyx015.

Hobbelen, Peter HF, Julia Stowe, Gayatri Amirthalingam, Liz Miller, and Albert-Jan van Hoek. 2016. "The Burden of Hospitalisation for Varicella and Herpes Zoster in England from 2004 to 2013." *Journal of Infection* 73 (3): 241–53. https://doi.org/10.1016/j.jinf.2016.05.008.

Jones, Lisa, Mark A Bellis, Dan Dedman, Harry Sumnall, and Karen Tocque. 2008. "Alcohol-Attributable Fractions for England: Alcohol-Attributable Mortality and Hospital Admissions." *Liverpool: Centre for Public Health, Liverpool John Moores University.* https://fingertips.phe.org.uk/documents/alcoholattributablefractions.pdf.

Keeping, Sam T., Michael J. Tempest, Stephanie J. Stephens, Stuart M. Carroll, and Vijay K Sangar. 2015. "Penile Cancer Treatment Costs in England." *BMC Public Health* 15 (1). https://doi.org/10.1186/s12889-015-2669-2.

Keeping, S. T., M. J. Tempest, S. J. Stephens, S. M. Carroll, R. Simcock, T. M. Jones, and R. Shaw. 2017. "The Cost of Oropharyngeal Cancer in England: A Retrospective Hospital Data Analysis." *Clinical Otolaryngology* 43 (1): 223–29. https://doi.org/10.1111/coa.12944.

Leal, José, Stefania Manetti, and James Buchanan. 2018. "The Impact of Hospital Costing Methods on Cost-Effectiveness Analysis: A Case Study." *PharmacoEconomics* 36 (10): 1263–72. https://doi.org/10.1007/s40273-018-0673-y.

NHS Digital. 2019a. *Statistics on Alcohol, England.* https://digital.nhs.uk/data-and-information/publications/statistical/statistics-on-alcohol/2019.

———. 2019b. *Statistics on Smoking, England.* https://digital.nhs.uk/data-and-information/publications/statistical/statistics-on-smoking/statistics-on-smoking-england-2019.

Public Health England. 2017. *Cost of Smoking to the Nhs in England: 2015.* https://www.gov.uk/government/publications/cost-of-smoking-to-the-nhs-in-england-2015/cost-of-smoking-to-the-nhs-in-england-2015.

Tobacco Advisory Group of the Royal College of Physicians. 2018. *Hiding in plain sight: Treating tobacco dependency in the NHS*. https://www.rcplondon.ac.uk/projects/outputs/hiding-plain-sight-treating-tobacco-dependency-nhs.

Webster, L, C Angus, A Brennan, and D. Gillespie. 2018. *Smoking and the Risks of Adult Diseases.* The University of Sheffield. http://eprints.whiterose.ac.uk/140480/1/Smokingandtherisksofdisease.pdf.