Quantifying the Primary and Secondary Effects of Antimicrobial Resistance on Surgery Patients: Methods and Data Sources for Empirical Estimation in England

Nichola R Naylor*†[1,2,3], Stephanie Evans*[3], Koen B Pouwels[4,5], Rachael Troughton[1], Theresa Lamagni[3], Berit Muller-Pebody[3], Gwenan M Knight[2], Rifat Atun[6,7], Julie V Robotham[3,1]

*Joint first author

† Corresponding author: nichola.naylor@phe.gov.uk

1. The National Institute for Health Protection Research Unit in Healthcare Associated Infection and Antimicrobial Resistance at Imperial College London, London, England
2. AMR Centre, Department of Infectious Disease Epidemiology, Faulty of Epidemiology and Population Health, London School of Hygiene & Tropical Medicine, London, England
3. Healthcare Associated Infection and Antimicrobial Resistance Division, United Kingdom Health Security Agency, London, England.
4. Health Economics Research Centre, Nuffield Department of Population Health, University of Oxford, Oxford, England.
5. The National Institute for Health Research Health Protection Research Unit in Healthcare Associated Infections and Antimicrobial Resistance at the University of Oxford, Oxford, England.
6. Department of Global Health and Population, Harvard TH Chan School of Public Health, Harvard University, Boston, USA
7. Department of Global Health and Social Medicine, Harvard Medical School, Harvard University, Boston, USA

Abstract
Antimicrobial resistance (AMR) may negatively impact surgery patients through reducing the efficacy of treatment of surgical site infections (the 'primary effects' of AMR). Previous estimates of the burden of AMR have largely ignored the potential 'secondary effects', such as changes in surgical care pathways due to AMR (for example different infection prevention procedures or reduced access to surgical procedures altogether), with literature providing limited quantifications of this potential burden. Former conceptual models and approaches for quantifying such impacts are available, though they are often high-level and difficult to utilise in practice. We therefore expand on this earlier work to incorporate heterogeneity in antimicrobial usage, AMR, and causative organisms, providing a detailed decision-tree-Markov-hybrid conceptual model to estimate the burden of AMR on surgery patients. We collate available data sources in England and describe how routinely collected data could be used to parameterise such a model, providing a useful repository of data systems for future health economic evaluations. The wealth of national-level data available for England provides a case study in describing how current surveillance and administrative data capture systems could be used in the estimation of transition probability and cost parameters. However, it is recommended that such data are utilised in combination with expert opinion (for scope and scenario definitions) to robustly estimate both the primary and secondary effects of AMR over time. Though we focus on England, this discussion is useful in other settings with established and/or developing infectious diseases surveillance systems that feed into AMR National Action Plans.

Key Words
Antimicrobial resistance, secondary effects, surgical site infection, burden.

Introduction

Surgical site infections (SSIs) place a substantial burden on healthcare systems [1]. SSIs, known to be "infections of superficial or deeper tissue occurring within 30 days of non-implant surgery, or within one-year for implant-related procedures" [2], are costly to the National Health Service (NHS) in England, creating prolonged hospital stays for patients and increased costs for hospitals [3]. When considering the additional costs of lost productivity and reduced workforce, the economic burden could be substantially more.

Antimicrobials have allowed us to develop safe patient-care pathways that would previously have put patients at a high risk of SSIs, by the integration of protocols of antibiotic prophylaxis into these pathways. Consequently, the threat of antimicrobial resistance (AMR) goes beyond a reduction of treatment effectiveness for acute infections. Increased AMR reduces the effectiveness, and therefore benefits, of the antimicrobial prophylaxis currently protecting patients from potential infections (such as those undergoing surgery at risk of SSIs) [4, 5]. AMR then has the potential to disrupt the care of these patients through multiple processes, often referred to as the 'secondary effects' of AMR [4, 6]. Incorporating all such potential future costs for different evaluations of associated interventions is key for efficient policy making. To quantify the burden of these secondary effects of AMR from healthcare system and societal perspectives, potential patient health outcomes, payer and provider cost, and socio-economic data are needed [4].

The primary and secondary use of health and economic data (with primary defined as 'use for intended clinical, public health, societal and/or research purposes stated *a priori'*, and secondary for purposes other than those stated as primary [7]) has been highlighted as a key way to tackle AMR through increased knowledge of epidemiology and economic burden, subsequently helping shape the allocation of finite resources. For example, the Global Antimicrobial Resistance Surveillance System was launched in 2015 to promote the collecting of not just microbiological but also outcome data associated with AMR [8], whilst the Wellcome Data Re-use Prize promoted the secondary use of similar AMR data in generating policy recommendations [9].

Though there have been discussions on the importance and use of SSI surveillance data in terms of preventing and improving quality of care in relation to SSIs and surgery patients correspondingly [10], the use of such data in informing the quantification of the secondary effects of AMR has not yet been discussed in detail. There have been reviews that include the quantification of AMR burden in relation to SSIs [11, 12]. However, these focus on narrow definitions of AMR burden in relation to SSIs, namely the direct effectiveness of AMR in preventing and treating infections in past or current cohorts of surgical patients. As we have seen from the recent COVID-19 pandemic, there are potentially major costs borne through a capacity or risk threshold being met (e.g. an implicit or explicit 'occupancy of hospital beds' or 'risk of death' threshold being met in unmitigated AMR scenarios). As such, broader treatment behaviours for many patients, and the general population, may need to be changed (for instance cancelling elective procedures in hospitals), which in turn can lead to other health and economic burdens to society [13].

We aim to provide a practical discussion of approaches and data sources that have been and/or could be used to estimate the total (primary and secondary effects) of AMR in relation to surgery, based on literature and available data in England (where there is a wealth of national-level data on SSIs, AMR, hospital admissions, population demographics and economic measures [2, 14-16]). The objectives of this study are to; (i) discuss previous methods used to estimate the primary and secondary effects of AMR in relation to surgery, (ii) discuss potential health and economic data sources available in England, and (iii) based on the methods and data described, propose a conceptual model for quantifying the potential total burden of AMR on surgery patients in England.

Previous Estimates of the Secondary Effects of Antimicrobial Resistance for Surgery Patients

AMR can impact surgical patients in three main ways; (i) increasing the risk that antimicrobials will fail to treat an acquired SSI leading to additional and/or different therapeutic recommendations needed to treat the SSI , including (but not limited to) increased length of hospital stay, the need for revisional surgeries and chronic infection care, (ii) by increasing the risk of acquiring SSIs through failure of prophylaxis and (iii) by changing patient pathways for surgical treatment altogether (such as reducing the provision of certain procedures and/or introducing new screening protocols) [6]. In the case of (i) and (ii), the increase in SSIs and AMR may not only affect patient health and wellbeing, but also the economic burden on health system and society through an increase in the number of patients experiencing SSIs that need therapeutic treatment (which in turn could increase future cases through onward transmission of healthcare associated infections) [1]. In (iii), the risk-to-benefit balance for a procedure may tip in favour of not recommending patients undergo surgery due to an increased risk of SSI or associated illness and mortality, producing a scenario where surgery is only recommended for certain patient groups or surgical categories where the risk of SSIs and/or serious consequences as a result of SSIs is 'low' [6], thereby reducing access to surgical procedures that may otherwise increase patients' quality of life and reduce future health system costs [17]. For example, in the case of joint arthroplasty, patients with a damaged joint considered ineligible for surgical intervention would be reliant on a pathway of physiotherapy and pain management as an alternative [18, 19]. Such considerations are in place in England, where around 80 out of 135 Clinical Commissioning Groups (purchasers of care) already apply a body mass index threshold to funding orthopaedic surgery because of associated risks of surgical complications [20].

Two previously published literature reviews investigate AMR impacts on SSI [11, 12]. The first, a more general review of multi-drug resistant impact on healthcare associated infections, highlighted two studies with contrasting impact estimates in relation to AMR impact on post-SSI mortality and length of stay [11]. The second review recommends a standardised definition of SSI, a standardised surveillance system, and priority setting of antibiotic resistant pathogens in relation to AMR [12], based on the current gaps in the literature. Matched cohort studies are proposed as a method to account for other covariates and potential confounders in quantifying the direct cost burden of AMR in SSIs, where data and resources are available [12]. There is a call for more studies addressing the wider societal impact of AMR through SSIs, though currently there is more of focus on SSI and AMR surveillance data collection protocols for primary effect estimation (such as prospective, matched cohort studies) than on methods to estimate total future, potential burden.

The effects of AMR on the loss of antimicrobials as a tool for infection prevention and treatment are rarely incorporated in estimates of the potential future burden of AMR [21]. To our knowledge, studies that have estimated the health burden of AMR due to secondary effects do so through fairly simplistic linear combinations of AMR and infection incidence and associated outcomes found through the literature [4]. Such studies have estimated that a reduction in antimicrobial effectiveness of 30% could lead to an additional 120,000 SSIs and 6,300 infection-related deaths per year in the USA, whilst a study involving selected countries in Europe estimated that a 100% reduction in antimicrobial effectiveness could lead to 439,000 additional postoperative infections each year, equivalent to a 50% increase compared to current SSI levels [4, 22]. Though there is little inclusion of parameter uncertainty surrounding these estimates, beyond discussing the impact of changing the level of reduction in antimicrobial effectiveness. As a proxy for pathway (iii), the O'Neill Antimicrobial Resistance Review estimated that the economic cost of AMR on critical surgeries (for example Caesarean sections and organ transplants) and cancer treatments could be around 120 trillion USD over a 35-year period [5]. However, this used un-referenced statements of "caesarean sections contribute about 2% to world Gross Domestic Product" and approximated world Gross Domestic Product equivalents [5]. These numbers suggest that the secondary effects of AMR on foregone surgery could be substantial. Hence there is a need for further investigation using more robust and transparent modelling techniques.

One method of more formally estimating these effects was proposed by Smith and Coast, whereby a decision-tree model which included a “no-antimicrobial prophylaxis” surgery pathway as a proxy for a scenario of high AMR was suggested [6]. Turning to past literature that has investigated the cost-effectiveness of antimicrobial prophylaxis in surgical procedures, we can find decision-tree models for hip and knee replacements (common elective surgeries) showing a dominant effect in using tested ‘antimicrobial prophylaxis’ compared to ‘no antimicrobial prophylaxis’ (i.e. lower SSIs and lower costs in the intervention scenario compared to a baseline of no antimicrobial prophylaxis) [23, 24]. However, these models have only considered specific pathogens (e.g. *Staphylococcus aureus* [23, 24]) and focused just on direct costs to the healthcare systems. Nevertheless, these studies could still be useful in the parameterisation of a “no-antimicrobial prophylaxis” arm of such models in the event of no other data being available. Additionally, studies comparing alternatives to surgery for such procedures, such as a Markov model comparing delayed and no-surgery strategies to total hip replacement [25], can be used to provide insight to the longer-term costs and effects of elective surgery not being accessible to certain patients. The 2013 study, set in Italy, estimated delayed surgery for men and women to be cost-effective compared to medical treatment across tested age groups. For example, there was an incremental cost per quality-adjusted life year gained of €463 and €82 for males and females respectively. Comparing this against a willingness-to-pay threshold of €36,500-€60,000 per QALY gained (as €463 is well below €36,500), this is highly cost-effective. This suggests potentially large efficiency losses in a “no surgery due to AMR” world [25].

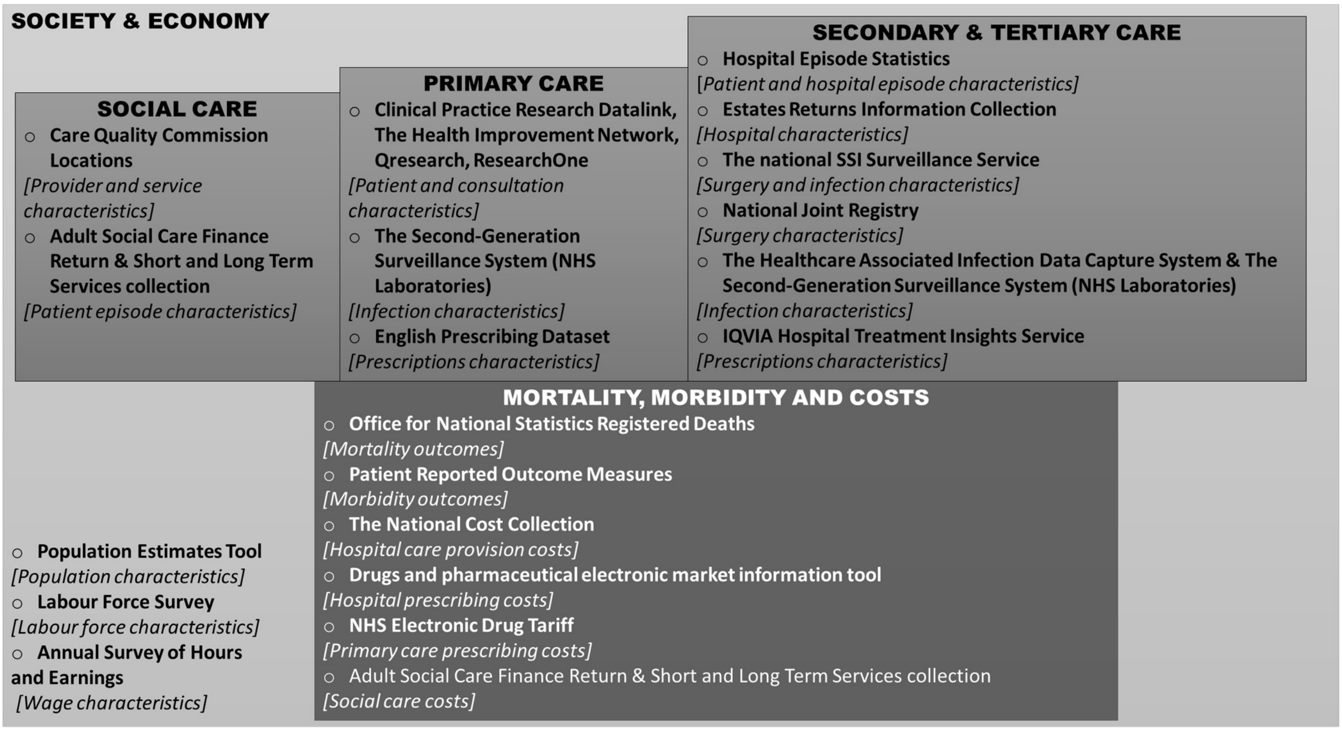
Hence, decision trees, Markov models and model parameters used in previously conducted health economic evaluations could be combined to estimate the total potential primary and secondary effects of AMR on SSIs and fill a gap in the current literature. This approach also allows for the explicit inclusion of parameter uncertainty within the modelling process (unlike some of the more scenario-based, step-by-step calculations performed previously [4, 26]). Aforementioned studies also give indications of the type of data needed (infection epidemiology, procedures performed, associated health outcome, healthcare-cost data and macroeconomic data). We next explore what health and economic data are available for such future burden analyses within the English setting through already established data capture systems.

Potential Health and Economic Data in England for Quantifying Total Secondary Effects

Figure 1 (expanded in Supplementary Material Table A1) summarises some of the key datasets available for epidemiology and health economics research for AMR and SSIs in England, highlighting the large breadth of data sources across the healthcare system and wider economy.

Figure 1. Overview of Datasets for the Estimation of Antimicrobial Resistance Impact on Surgery Patients in England

Boxes represent the setting of the data, bullet points represent the name of the dataset, followed by a high-level description the use of the dataset in square brackets, for the purposes of Surgical Site Infection and Antimicrobial Resistance research. For more information on data sources see Supplementary Material Table A1. Abbreviations: NHS – National Health Service, SSI - surgical site infection.



In regards to health data, England has a centralised administrative data capture and processing system for hospital admissions and care, known as Hospital Episode Statistics (HES) [16]. Secondary use of HES data has, in the past, included the linkage of HES data to other datasets (such as to those listed in Figure 1 under "Primary Care" and "Mortality, morbidity and costs") to have a more complete picture of patients and their care. For example, as HES collects only data based on what happens within hospitals, further information might be needed on post-discharge mortality. Therefore, HES has been linked with Office for National Statistics (ONS) mortality data to incorporate post-discharge mortality [27]. Additionally, HES data can be linked to Second Generation Surveillance System (SGSS) data to get more information about infection characteristics (such as microbe type and susceptibility to antibiotics), whilst antibiotic prescribing data and hospital characteristics may be available through linkage with related datasets (see Supplementary Material Table A1). This is possible due to granularity of these data capture systems, namely the inclusion of patient-level identifiers (such as unique NHS numbers, names, and date of birth) and hospital identifiers (unique provider codes). A patient-level data set, linked across these sources, therefore could be used to estimate the transition probabilities of patients acquiring types of infections, undergoing revision surgery, and the subsequent impact of these different treatment pathways on mortality, as done previously [17].

A previous review suggests the use of prospective, matched cohort studies to estimate the burden of SSIs by infection type [12]. However, such studies are resource intensive and can have low external validity unless conducted on a national/international scale. Secondary use of these national microbiology and HES data sources paired with appropriate statistical methods have been utilised in the past to estimate associated mortality from healthcare-associated infections [28-30], and this could be further expanded to capture post-discharge mortality rates [27].

Linkage of hospital patient data to primary care data (see Figure 1) could allow for exploration on the need for additional patient pathways, such as increased primary care identification of infections, treatment and/or visits following certain infections or procedures. However, although post-discharge Surveillance is encouraged through the national SSI Surveillance Service (SSISS) [2], there can be a delay from the initial procedure to the time that the associated infection is detected, making it harder to define case exposures. For example, there could be up to 1 year from surgery to infection for surgeries requiring the placement of an implant e.g. artificial joints. When such patients do require hospitalisation there is no guarantee that patients will return to the same healthcare facility in which they underwent the related surgery, so any records of infections and surgical procedures need to be cross-referenceable within and between healthcare facilities. This means long-term surveillance of SSIs is required. If basing such estimates on established primary care administrative systems, all three of the primary-care-based administrative systems rely on voluntary inclusion from GP practices and patients, with varying degrees of sample sizes and representativeness across the three systems [31]. However, these have still been used for previous analyses of primary care healthcare utilisation and population health outcomes within England [32].

Given the median age of elective-surgery patients covered within the national surveillance reports ranges from approximately 50 to 85 years old [2], long-term care facility data may also be useful, with a Care Quality Commission directory highlighting post-codes of such facilities that can be matched to patient postcodes [33]. Additionally, other social care data sources listed in Figure 2 (and described further in Supplementary Material Table A1) provide information on long and short-term forms of social care that could be useful for costing purposes, if/when applicable to the patient groups of interest.

The cost-of-illness impact, from the NHS (payer and provider) perspective, of surgery, SSIs and drug-resistant infections can be estimated using the aforementioned linked surveillance-administrative datasets to estimate length of stay and/or 'Health Resource Group' (HRG) impact [1, 28]. If working directly with patient data, the National Costing Grouper, which confers a core Health Resource Group (HRG) to patients' hospital stays, can be used alongside the National Costing Collection workbook (which provides monetary unit costs per HRG) to calculate patient level costs. If working with excess length of stay estimates, the acute patient level activity and costing for 2019-2020 (unbundled activity) gives, by speciality within critical care, the total number of days and total cost (£) submitted to the PLICs Acute collection [34], which could hypothetically be used to estimate applicable proxy costs of an excess bed day (e.g. using "Surgical adult patients (unspecified specialty) had a total of 31,807 days and £47,911,153.33 costs across all data submitted"). For any additional costs of antibiotic prophylaxis or treatments in the hospital or community, unit costs are readily available across the 'drugs and pharmaceutical electronic market information tool' (eMIT), the English Prescribing Dataset and the NHS Electronic Drug Tariff for NHS Trust-hospitals, General Practices and community providers respectively [35].

For AMR and SSI burden estimation from a societal perspective, utilising standard methods (e.g. human capital methods where each year of life lost is costed to be equivalent to average annual earnings), the use of labour activity and earnings to estimate lost wages through illness and death is needed, and is available (see Supplementary Material Table A1)[36]. However, with these data, those over 60+ and/or 65+ are grouped making it hard to disentangle contributions across cohorts of interest concentrated above 60 years of age, though of course assumptions can be made on the distribution of wage values across ages 60 and 100. Other methods for calculating a value for a statistical life year, which are broader in scope and/or more nuanced in calculation, are available but require more primary data collection in the English context [36].

As well as monetary costs, policy/AMR-scenario impact on population utility, is an important outcome, necessary for cost-utility analyses utilised by national policy makers [37]. Such outcomes are generally a function of mortality impacts (discussed above) and morbidity impacts. Patient-reported outcome measure data ,collected by NHS Digital, have been previously linked to other patient data (such as SSISS) to estimate QALY impacts of different infection prevention strategies for primary hip prothesis in England between 2009-2012, with more updated data now available [17, 38].

In terms of data access, although all of the patient-based data sources outlined in Figure 1 are not fully open-access due to patient identifiable data and subsequent data safeguarding, summary statistics are published openly (such as through annual reports or summary Excel files downloadable from government websites), which still could be used to estimate incidence rates at a national-level. Access to patient-level data may be permitted subject to asset owner approval processes, allowing for linkage across systems if appropriate data protection protocols are in place.

In the future there may be an increased ease of secondary use of the data sources outlined in Figure 1 for burden estimation purposes through systems similar to OpenSAFELY [39]. OpenSAFELY was created to allow for urgent research involving electronic health records in primary care (using TPP SystmOne software and EMIS), SGSS, ONS data and A&E attendance data for COVID-19 [39]. Researchers write and test code on dummy versions of health data locally, then the code is submitted to be run on secure servers which hold the real versions of the health data, checked in terms of "disclosivity" before then being released for publication purposes [39].

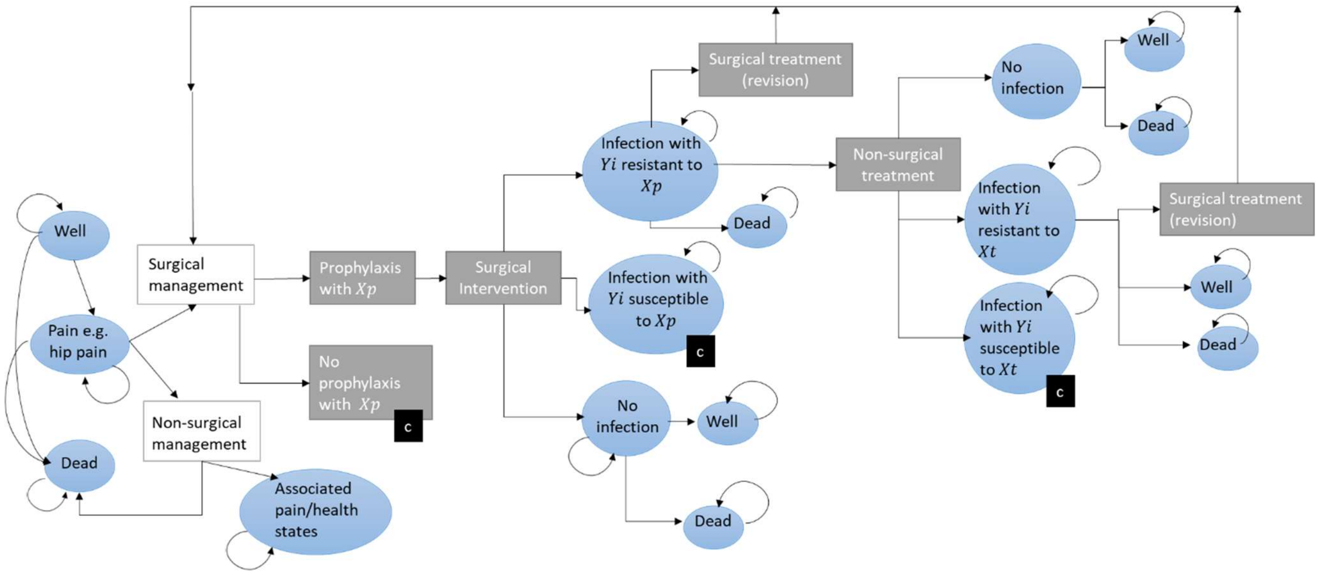
A Conceptual Model for Estimating the Secondary Effects for Surgery Patients

Based on both i) the existing AMR and SSI literature, and ii) the data available described above, we propose a state-transition model estimating the primary and secondary effects of AMR, incorporating decision-trees outlining treatment strategies and Markov models outlining potential health states (Figure 2).

For the primary effects (i.e. the direct burden of increased AMR in patients who get surgery and develop SSIs), these pathways can be parameterised with the secondary use of surveillance, administrative and economic data available in England. If outcome data are available at the patient level across hospital stays, statistical models can be utilised to estimate transitions along the patient pathways, adjusting for patient, provider, and socioeconomic characteristics. There are methods available to account for potential sources of bias when dealing with healthcare-associated infections, such as time varying confounding, and other complexities that may occur regardless of study design, such as competing events [28, 30, 40, 41]. This allows for more robust patient outcome inclusion, such as cost-of-SSI and mortality impacts (see Supplementary Material Figure A1 for more detail).

Figure 2. A Conceptual Model for Estimating the Impact of Antimicrobial Resistance on Surgery Patients.

The progress of patients through a surgical management pathway with the potential for infection is depicted following from left to right. Patients initially start (far left) in one of three health states (blue circles) and progress through treatment decisions (rectangles). Circles represent health states, rectangles represent pathway treatment decisions, [c] represents a collapsed branch that mirrors that of another branch within that level (for example the pathway following "No Prophylaxis" includes the same transitions and states as "Prophylaxis". Yi represents microbes where $i=1,…,m$ different microbes of clinical importance; Xp,t represents antibiotics given for prophylaxis ($p=1,…,n$) at time t, Tx represents antibiotics given for treatment of SSI ($t=1,….,q$). It may be that $p=t$. The length of time patients spend in each state (indicated by a curved returning arrow) is variable.



Across the prophylaxis pathway depicted in Figure 2, we highlight that there are the potentially numerous patient pathways with different drug regimens and potential for infection with different organisms. Figure 2 highlights the complexity of this issue, with transition rates and outcomes potentially heterogenous for different drug and bacteria combinations. In practice there is even variation in the antibiotics used across hospitals, with gentamicin and flucloxacillin used for prophylaxis by 57 out of 147 NHS Hospital Trusts and other (differing) antibiotics used across others [42]. In a previous cost-effectiveness analysis of strategies to reduce risk of infection following hip replacement therapy, where AMR effects were not incorporated into the equations, a weighted average of the cost of different prophylaxis guidelines between Trusts was used to account for this [17]. At an individual level, the specific microbe that caused the infection and type of antimicrobial used to treat it could be important in terms of pathways (and subsequently outcomes and costs), thus the scope of different combinations of prophylactic and therapeutic drug-pathogen exposure definitions (of *Yi* and *Xp,t* in Figure 2 representing different microbes, prophylaxis and treatment antibiotics respectively) warrant consideration by experts in SSIs depending on the scale of the research question (e.g. local, regional or national).

Inclusion of primary care pathways, in terms of direct effects of AMR on post-surgery patients could be parameterised through linkage of patient data across primary and secondary care settings. Health states presented can have utility values, hospital costs and societal costs attached from Patient Reported Outcome Measure, National Cost Collection and Office for National Statistics datasets respectively, as discussed in the above section [38, 43, 44]. However, expert opinion would be needed to define appropriate case definitions (e.g. time of GP consultation post-surgery/discharge and associated diagnoses “Read” codes). However, for hip surgeries (and other similar, short-stay procedures) it was estimated half of SSIs were captured through readmission surveillance, therefore post-discharge surveillance and linked-HES data to follow (and cost) patients within and across hospitals across the shown post-surgery pathways is key. Average adjusted-wage losses per excess death and days unable-to-work could be estimated through ONS data on employment and wages, and combined with excess deaths and days (e.g. in hospital) counted through Figure 2 [44].

In a situation of the datasets described in the above section not being accessible at the patient-level, given their coverage they can still be useful in quantifying probabilities of acquiring SSIs. As an example, for hip replacement surgery specifically (which is under the mandatory surveillance for SSIs under which 60% undertook continuous surveillance for 2019/20); SSISS reports for 2019/20 stated that hip surgery had a risk of SSIs of 0.5% , of which 30.2% and 4.3% were caused by methicillin-resistant and methicillin-susceptible *Staphylococcus aureus* respectively [14]. These figures can therefore be used to estimate transition probabilities across Figure 2 for contracting a methicillin-resistant Staphylococcus SSIs following hip replacement surgery in England, alongside assumptions of prophylaxis impact, in the absence of more granular data.

As England has access to longitudinal microbiology data, trends can be determined using these data and simple techniques (such as simple linear regression or random-walk models) through to more complex transmission and forecasting modelling methods (such including seasonality and non-linearity) [45]. These trends can be included as potential AMR and infection risk scenarios run through the conceptual model [22]. However, to parameterise the case of a pan- or extensive-drug-resistant world where current patients would stay in the "no-surgery" pathway (i.e. the secondary effects pathway), previous cost-effectiveness analyses of the actual surgeries would need to be utilized (for example previous economic evaluations of hip replacement surgeries [46-49]). There is also the challenge of the increased cost of requiring long-term assisted living for individuals with non-operative management of end-stage osteoarthritis in the elderly population, these costs could be incorporated into Figure 2 by separating out non-surgical management of hip pain and non-surgical treatment of SSIs by settings of care, potentially parameterised by the social care and society demographic data highlighted by Figure 1.

Though previous cost-effectiveness analyses can tell us the general transition probabilities, cost, and utility estimates for a non-surgery scenario, they can't tell us under what AMR and patient-characteristic situations they would occur. Moreover, current SGSS and SSISS data report data on AMR and microbes currently circulating within the healthcare system: a wider scope of scenarios is needed to include microbes and associated drug resistances that may be important in the future, but that aren't currently seen in the data due to low or no numbers (e.g. colistin resistance in Gram-negative infections or a multi-drug resistant *Candida auris*) [6]. For this we need expert elicitation of resistance cut-off levels (i.e. when the "no-surgery scenario" would take effect) and epidemiological forecasting of AMR and infectious disease trends incorporating expert elicitation of predictions for future microbe and AMR importance.

Discussion

We first highlight that there are three potential ways for AMR to impact surgery patients including increasing SSI risk, treatment failure risk and risk of operations being unavailable altogether. Highlighted literature indicates that secondary effects could play a substantial role in the burden of AMR in the future, with estimates for lessening antimicrobial effectiveness including an additional 6,300 deaths per year in the USA and a loss of 2% of world Gross Domestic Product, across different scenarios [4, 26]. However, many of the discussed estimates of burden did not sufficiently incorporate uncertainty and/or did not use an explicit mathematical modelling framework that can be practically used and adapted, according to need. A conceptual model utilising decision trees and Markov models could be used in estimating the potential impacts of AMR on surgery patients, if scoped appropriately and parameterised robustly. The conceptual model constructed within this study highlights the nuance of AMR for SSIs across all pathways, through the acknowledgement of the different potential antimicrobial usage exposures, microbe exposures and treatment options.

The scope for the secondary use of health data in establishing SSI and AMR burden for parameterising transition probabilities, costs and mortality for associated infections occurring in a health system is large. The retrospective use of such data may allow for reduced research burden in parameterising our proposed conceptual model. Even national-level, aggregated data can be used in estimating transition probabilities if data are externally valid. This, in turn, highlights another benefit to SSI surveillance, which already has been shown to reduce SSI rates themselves through benchmarking and outlier identification functions [10]. The secondary use of such data, as described here, may be a consideration in the cost-benefit case of public health surveillance itself. Additionally, many of the data sources collated and described within this study are of use in estimating impacts of AMR on other syndromes and clinical specialities, such as respiratory or bloodstream infections treated in primary and secondary care settings.

Earlier published studies have highlighted the need for standardised SSI surveillance protocols such as defined follow up length and data entry methods [12]. Our review highlights the benefits of established surveillance systems being able to be readily linked across microbe-, susceptibility- and mortality-surveillance and administrative datasets at the patient-level. Such linkage allows for a greater understanding of the impact of AMR and SSIs on patient outcomes and health system costs, with this being feasible in the NHS through the capture of consistent patient identifiers (unique NHS number, date of birth, sex) across systems [50, 51]. However, with patient identifiable and confidential data comes a responsibility to have robust information governance and data protection protocols in place [7, 51]. For example, for the surveillance system within England, there is strict adherence to handling patient data in accordance with the Data Protection Act 2018, General Data Protection Regulations (GDPR) and the Caldicott Guidelines [52]. Examples of processes that aid this include establishing policy on how long data are held for, who can access the data and how data can be shared. It has been suggested that specifying ethical and privacy principals, and linking these to governance and data access can help with public trust in data capture systems [53], a key factor in having robust data for primary and/or secondary use. With new data access frameworks being explored, such as OpenSAFELY [39], there is increased scope for a reduction in the transfer of patient-level data across parties for research purposes in the future.

However, even with access to current data, data completeness needs to be reviewed. Taking completeness to be in terms of documentation (i.e. are all the available fields filled in and available for use)[54], about 80% of patients in SSISS had a NHS number in a previous analysis (even after doing additional patient tracing to retrieve some missing numbers) and therefore patient-level linkage across numerous datasets may bias subsequent estimates of transitions and outcomes if there are systematic reasons for data non-completeness [17]. The completeness of patient data in the SSISS is high for mandatory surveillance in terms of case identification [2], though it is only mandatory to carry out surveillance "for a minimum of 3 consecutive months per financial year in at least one of 4 orthopaedic categories: hip replacement, knee replacement, repair of neck of femur or reduction of long bone fracture" [2]. This means data for other surgeries may not be fully representative of English surgical patients and SSIs (such as for caesarean or lower bowel surgery patients who represent a large proportion of the overall burden to the NHS [1, 55]), however the number of operations submitted for 2019/20 for voluntary SSI surveillance showed a 9% increase in comparison to 2018/19, with 27,877 procedures submitted voluntarily in 2019/20 [2]. Furthermore, one could use weighting or post-stratification techniques to obtain representative estimates if variables determining selection into the sample are available [56].

Only a few infection types (mainly bacteraemia and notifiable infections) are listed as mandatory surveillance within relevant data capture systems, and as such the epidemiological data for other pathogens could be biased. However, such data has been routinely used to present AMR data at the national level in England [14] and a 2020 report comparing mandatory and voluntary submissions found a high ascertainment rate comparing across the systems (for bacteria present in both systems) [57]. In the absence of surveillance systems, routinely collected HES data may be useful for infection rates and patient outcomes, with routinely collected data to estimate rates of SSIs being found to have sensitivities ranging from 60% to 98% [10, 58], though this would likely not provide information on microbe or AMR.

Though PROMs data are theoretically available for certain patients, it is only available for certain surgeries (hip and knee replacement) [59]. Moreover, even when these data are available, they may not be useful for our intended purpose. A previous analysis had to revert back to using literature as PROMs data weren't available for their SSI case definitions (e.g. within 14 days of the date of infection) [17]. For international comparisons where Disability-Adjusted Life Years may be wanted (instead of Quality-adjusted life years), a large European study is available, where disability weights for SSI states are based on previous observational studies which have elicited utility values [60]. From a patient perspective, administrative datasets in NHS England described in this review do not account for patient-level costs, though these data may be available in insurance-based healthcare systems [7].

Regarding the HRG and excess length of stay unit costing, England is currently undergoing the NHS England and NHS Improvement's Costing Transformation Programme. This was piloted in 2016 and had annual stages of increased implementation subsequently. Therefore, currently such data come with potential data quality issues given these are the first few years of the new Patient-Level Information and Costing system, though this bias should decrease over time if systems and processes remain unchanged [61].

Even with complete and secure data systems in place, this review deduces that the secondary use of health data cannot be used solely to parameterise a secondary effects model for AMR and SSIs. Prospective trials of SSI prevention measures, patient and public elicitation for utility values, alongside expert elicitation studies for "post-antibiotic" scenario understanding are needed. A 2019 literature review calls for more evidence from primary studies on the intervention effectiveness of different SSI prevention techniques [12]. Such data normally come from randomized-control trials rather than secondary use of health data, although even trials are now making use of routinely collected data to inform primary trials and/or for longer follow-up period [62].

Additionally, the model currently depicts a simplified picture of 'prophylaxis' vs 'no prophylaxis comparison of surgical management pathways, as in addition to antimicrobial prophylaxis, there are several interventions to prevent SSIs, such as using sterile gowns or changing surgical instruments prior to wound closure [12], many of which are recommended for surgery undertaken in England [63]. These pathways can be added to the core framework outlined here as and when necessary, with scope of the model pathways best extended based on expert opinion for specific surgeries or settings. As has been done in more general AMR burden estimation models [22]. The proposed conceptual model recommends AMR scenario dynamics be included through external trend analyses and expert elicitation to then feed epidemiological parameters directly into the state transition model, which has been done in more general AMR burden estimation models [22]. The conceptual model could be expanded into incorporate transmission dynamics but would require more health states (representing other reservoirs of antibiotic usage and resistance) and therefore potentially more data.

While here we do not make specific recommendations related to the general health economic approach of quantifying our conceptual model, general guidelines are available elsewhere for health economic modelling [37, 64], the reporting of which is currently lacking from some studies that have attempted to quantify secondary effects of AMR [4, 26]. Based on standard guidelines, the England case study should take an NHS perspective (as recommended by NICE for the base case [37]), cover the lifetime of a hypothetical cohort to capture the potential long-term impacts, and use a 3.5% discount rate for future costs and a 1.5% discount rate for quality-adjusted life years declining over 30 years as recommended by the Treasury [65]. Using this approach, parameter and methodological uncertainty can be tested by varying parameter values, discount rates and time horizon through one-way and probabilistic sensitivity analyses. Structural uncertainty and heterogeneity could also be explored in further iterations of Figure 1, by adapting pathways and including specific subgroups if sufficient data are available. With such modelling approaches, a broader perspective is enabled through the inclusion of labour productivity cost proxied by national wage and employment data

[66]. However, there may be a need in the future to explore the effects of presenteeism impacts (i.e. incorporate not only loss of work productivity through hospital stay or death, but also general loss of work productivity for patients along different pathways) [67], and also explore informal market production impacts [36, 66].

Though we have currently focused on the secondary use of health data in the NHS, the findings are applicable to settings where similar datasets are available. A 2018 review found 56 healthcare-associated infection and AMR surveillance systems from 20 countries within Europe, with 32 SSI surveillance systems included [68]. This indicates that there is already a large potential resource for understanding the secondary effects of AMR considering the proposal outlined in our review. We have also focused on SSIs, but secondary use of cancer patient data may be explored in a similar manner; through national cancer registration data linked to other surveillance and administrative data [51].

In conclusion, AMR is a complex phenomenon which has the potential to alter health outcomes for patients who contract drug resistant SSIs, and change surgery patient pathways due to secondary effects. Though the secondary use of health data, in the English setting, has the potential to parameterise models quantifying the former, it falls short of being able to quantify the latter in isolation. However, such data can be combined with expert elicitation to parameterise a health state transition model that incorporates primary and secondary impacts of AMR on surgery patients over time. With growing SSI and AMR surveillance systems globally, alongside expert elicitation and investigations into potential future epidemiological scenarios, we can begin to understand the potential secondary effects of AMR through the application of the proposed conceptual model in other settings, and therefore understand how to more efficiently deal with this phenomenon.

Abbreviations:

AMR – antimicrobial resistance

HES – Hospital Episode Statistics

NHS – National Health Service

ONS – Office for National Statistics

PROMs – Patient Reported Outcome Measures

SGSS - Second Generation Surveillance System

SSI – surgical site infection

SSISS – SSI Surveillance Service

Funding:

The research was funded by the National Institute for Health Research Health Protection Research Units (NIHR HPRU) in Healthcare Associated Infections and Antimicrobial Resistance at Imperial College London [grant number HPRU-2012–10047 funding NN’s time and with NN subsequently holding an honorary contract through HPRU-200876] and the University of Oxford [grant number HPRU-2012-10041], all in partnership with Public Health England (PHE) [now United Kingdom Health Security Agency (UKHSA)]. GMK was supported by a fellowship from the UK Medical Research Council (MR/P014658/1). The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR, the Department of Health, or the United Kingdom Health Security Agency (formerly Public Health England).

Conflicts of Interest:

None declared by all co-authors

Ethical Approval

This review cites only data from publicly available reports and/or published material and does not conduct any primary research nor utilise any patient data, therefore was not subject to ethical approval processes.

References

1. Troughton, R., et al., *Mapping national surveillance of surgical site infections in England: needs and priorities.* J Hosp Infect, 2018. **100**(4): p. 378-385.
2. Public Health England, *Surveillance of surgical site infections in NHS hospitals in England: April 2019 to March 2020*. 2020.
3. Leaper, D.J., C.E.J. Edmiston, and C.E. Holy, *Meta-analysis of the potential economic impact following introduction of absorbable antimicrobial sutures.* Br J Surg, 2017. **104**(2): p. e134-e144.
4. Teillant, A., et al., *Potential burden of antibiotic resistance on surgery and cancer chemotherapy antibiotic prophylaxis in the USA: a literature review and modelling study.* The Lancet Infectious Diseases, 2015. **15**(12): p. 1429-1437.
5. The Review on Antimicrobial Resistance, *Antimicrobial Resistance: Tackling a crisis for the health and wealth of nations*. 2014.
6. Smith, R. and J. Coast, *The economic burden of antimicrobial resistance: Why it is more serious than current studies suggest*, L.S.o.H.T. Medicine, Editor. 2012: London.
7. Safran, C., et al., *Toward a national framework for the secondary use of health data: an American Medical Informatics Association White Paper.* Journal of the American Medical Informatics Association, 2007. **14**(1): p. 1-9.
8. World Health Organization, *Global antimicrobial resistance surveillance system (GLASS) report: early implementation 2020.* 2020.
9. Leclerc, Q.J., et al., *Feasibility of informing syndrome-level empiric antibiotic recommendations using publicly available antibiotic resistance datasets.* Wellcome Open Research, 2019. **4**.
10. Troughton, R., *Investigating opportunities to improve surgical site infection prevention through social and technological innovation*. 2018, Imperial College London.
11. Serra-Burriel, M., et al., *Impact of multi-drug resistant bacteria on economic and clinical outcomes of healthcare-associated infections in adults: Systematic review and meta-analysis.* PLoS One, 2020. **15**(1): p. e0227139.
12. Iskandar, K., et al., *Highlighting the gaps in quantifying the economic burden of surgical site infections associated with antimicrobial-resistant bacteria.* World J Emerg Surg, 2019. **14**: p. 50.
13. Maringe, C., et al., *The impact of the COVID-19 pandemic on cancer deaths due to delays in diagnosis in England, UK: a national, population-based, modelling study.* The lancet oncology, 2020. **21**(8): p. 1023-1034.
14. Public Health England, *English Surveillance Programme for Antimicrobial Utilisation and Resistance (ESPAUR) Report 2019 to 2020*. 2020.
15. Public Health Transition Team, *Public Health Surveillance - Towards a Strategy for Public Health England*, D.o. Health, Editor. 2012.
16. NHS Digital, *Hospital Episode Statistics (HES) Analysis Guide*. 2019: December 2019.
17. Graves, N., et al., *A cost-effectiveness modelling study of strategies to reduce risk of infection following primary hip replacement based on a systematic review.* Health Technol Assess, 2016. **20**(54): p. 1-144.
18. Bennell, K., *Physiotherapy management of hip osteoarthritis.* Journal of Physiotherapy, 2013. **59**(3): p. 145-157.

19. Bennell, K.L., R. Buchbinder, and R.S. Hinman, *Physical therapies in the management of osteoarthritis: current state of the evidence.* Current opinion in rheumatology, 2015. **27**(3): p. 304-311.
20. Wilton, T. *Hip and knee replacements: BMI and smoking thresholds are not the answer*. 2015 26/05/2021]; Available from: https://www.rcseng.ac.uk/news-and-events/blog/hip-and-knee-replacements/.
21. Smith, R. and J. Coast, *The economic burden of antimicrobial resistance: Why it is more serious than current studies suggest.* LSHTM Research Online, 2012.
22. OECD, *Stemming the Superbug Tide*. 2018.
23. Courville, X.F., et al., *Cost-effectiveness of preoperative nasal mupirocin treatment in preventing surgical site infection in patients undergoing total hip and knee arthroplasty: a cost-effectiveness analysis.* Infect Control Hosp Epidemiol, 2012. **33**(2): p. 152-9.
24. Rennert-May, E., et al., *A cost-effectiveness analysis of mupirocin and chlorhexidine gluconate for Staphylococcus aureus decolonization prior to hip and knee arthroplasty in Alberta, Canada compared to standard of care.* Antimicrob Resist Infect Control, 2019. **8**: p. 113.
25. Mota, R.E., *Cost-effectiveness analysis of early versus late total hip replacement in Italy.* Value Health, 2013. **16**(2): p. 267-79.
26. Resistance, A., *Tackling a Crisis for the Health and Wealth of Nations.* Review on Antimicrobial Resistance, 2014.
27. Digital, N. *Linked HES-ONS mortality data*. 2021 [cited 2021 17/02/2021]; Available from: https://digital.nhs.uk/data-and-information/data-tools-and-services/data-services/linked-hes-ons-mortality-data.
28. Naylor, N.R., et al., *The health and cost burden of antibiotic resistant and susceptible Escherichia coli bacteraemia in the English hospital setting: A national retrospective cohort study.* PloS one, 2019. **14**(9): p. e0221944.
29. Green, N., et al., *Quantifying the burden of hospital-acquired bloodstream infection in children in England by estimating excess length of hospital stay and mortality using a multistate analysis of linked, routinely collected data.* Journal of the Pediatric Infectious Diseases Society, 2015. **4**(4): p. 305-312.
30. Hernán, M.A., et al., *Specifying a target trial prevents immortal time bias and other self-inflicted injuries in observational analyses.* Journal of Clinical Epidemiology, 2016. **79**: p. 70-75.
31. Kontopantelis, E., et al., *Spatial distribution of clinical computer systems in primary care in England in 2016 and implications for primary care electronic medical record databases: a cross-sectional population study.* BMJ Open, 2018. **8**(2): p. e020738.
32. Kneale, D., M. Khatwa, and J. Thomas, *Identifying and appraising promising sources of UK clinical, health and social care data for use by NICE*. 2016, EPPI-Centre, Social Science Research Unit, UCL Institute of Education, University College London: London.
33. Roselló Gilchrist, A., *Antibiotic resistant Gram-negative bacteria in long-term care facilities, an epidemiological and dynamic modelling study*. 2018, UCL (University College London).
34. Digital, N. *Provisional Monthly Hospital Episode Statistics for Admitted Patient Care, Outpatient and Accident and Emergency data April 2020 - December 2020*. Hospital Episode Statistics for Admitted Patient Care and Outpatient Data 2021 [cited 2021 17/02/2021]; Available from: https://digital.nhs.uk/data-and-information/publications/statistical/hospital-episode-statistics-for-admitted-patient-care-outpatient-and-accident-and-emergency-data/april-2020---december-2020.
35. NHS Business Services Authority. *NHS Electronic Drug Tariff*. 2021 [Accessed: 01/09/2021]; Available from: https://www.drugtariff.nhsbsa.nhs.uk/#/00805984-DC/DC00805981/Home.
36. Basu, A., *Understanding Productivity Benefits and Related Future Research Needs in Cost-Effectiveness Analysis.* Value & Outcomes Spotlight, 2018. **July/August**.

37. National Institute for Health and Care Excellence, *The guidelines manual*. 2012.
38. NHS Digital. *Provisional Patient Reported Outcome Measures (PROMs) in England - or Hip and Knee Replacement Procedures (April 2020 to March 2021)*. 2021 [Accessed: 22/09/2021]; Available from: https://digital.nhs.uk/data-and-information/publications/statistical/patient-reported-outcome-measures-proms/hip-and-knee-replacement-procedures-april-2020-to-march-2021#.
39. University of Oxford for The DataLab 2021. *OpenSAFELY documentation*. 2020 [Accessed: 20/09/2021].
40. Pouwels, K.B., et al., *Estimating the Effect of Healthcare-Associated Infections on Excess Length of Hospital Stay Using Inverse Probability-Weighted Survival Curves.* Clin Infect Dis, 2020. **71**(9): p. e415-e420.
41. Coeurjolly, J.F., et al., *Attributable risk estimation for adjusted disability multistate models: application to nosocomial infections.* Biom J, 2012. **54**(5): p. 600-16.
42. Hickson, C.J., et al., *Prophylactic antibiotics in elective hip and knee arthroplasty: an analysis of organisms reported to cause infections and National survey of clinical practice.* Bone Joint Res, 2015. **4**(11): p. 181-9.
43. NHS. *National Cost Collection for the NHS*. 2020 [cited 2021 24/02/2021].
44. Office for National Statistics. *Employment and labour market*. 2021 [cited 2021 24/02/2021]; Available from: https://www.ons.gov.uk/employmentandlabourmarket.
45. Hyndman, R.J., & Athanasopoulos, G., *Forecasting: principles and practice*. 2021, OTexts: Melbourne, Australia.
46. Kunkel, S.T., et al., *The Cost-Effectiveness of Total Hip Arthroplasty in Patients 80 Years of Age and Older.* J Arthroplasty, 2018. **33**(5): p. 1359-1367.
47. Kamaruzaman, H., P. Kinghorn, and R. Oppong, *Cost-effectiveness of surgical interventions for the management of osteoarthritis: a systematic review of the literature.* BMC Musculoskelet Disord, 2017. **18**(1): p. 183.
48. Nwachukwu, B.U., et al., *Current status of cost utility analyses in total joint arthroplasty: a systematic review.* Clin Orthop Relat Res, 2015. **473**(5): p. 1815-27.
49. Gordon, M., et al., *Age- and health-related quality of life after total hip replacement: decreasing gains in patients above 70 years of age.* Acta Orthop, 2014. **85**(3): p. 244-9.
50. NHS Business Services Authority & NHS Digital, *Medication Safety -Indicators Specification: August 2019*. 2019.
51. Padmanabhan, S., et al., *Approach to record linkage of primary care data from Clinical Practice Research Datalink to other health-related patient data: overview and implications.* Eur J Epidemiol, 2019. **34**(1): p. 91-99.
52. Public Health England, *Laboratory reporting to Public Health England: A guide for diagnostic laboratories*. 2020.
53. De Lusignan, S., et al., *Using routinely collected health data for surveillance, quality improvement and research: Framework and key questions to assess ethics, privacy and data access.* J Innov Health Inform, 2016. **22**(4): p. 426-32.
54. Weiskopf, N.G., et al., *Defining and measuring completeness of electronic health records for secondary use.* Journal of biomedical informatics, 2013. **46**(5): p. 830-836.
55. Wloch, C., et al., *Cost-benefit analysis of surveillance for surgical site infection following caesarean section.* BMJ Open, 2020. **10**(7): p. e036919.
56. Pouwels, K.B., et al., *Community prevalence of SARS-CoV-2 in England from April to November, 2020: results from the ONS Coronavirus Infection Survey.* The Lancet Public Health, 2021. **6**(1): p. e30-e38.
57. Public Health England, *Mandatory Healthcare Associated Infection Surveillance: data quality statement: October 2020*. 2020.
58. Freeman, R., et al., *Advances in electronic surveillance for healthcare-associated infections in the 21st Century: a systematic review.* J Hosp Infect, 2013. **84**(2): p. 106-19.

59. Digital, N. *Finalised Patient Reported Outcome Measures (PROMs) in England - April 2015 to March 2016*. Patient Reported Outcome Measures (PROMs) 2017 [cited 2021 22/02/2021]; Available from: https://digital.nhs.uk/data-and-information/publications/statistical/patient-reported-outcome-measures-proms/finalised-patient-reported-outcome-measures-proms-in-england-april-2015-to-march-2016.
60. Cassini, A., et al., *Attributable deaths and disability-adjusted life-years caused by infections with antibiotic-resistant bacteria in the EU and the European Economic Area in 2015: a population-level modelling analysis.* The Lancet infectious diseases, 2019. **19**(1): p. 56-66.
61. NHS Digital. *Patient Level Information and Costing System (PLICS) Data Collections - NHS Digital* 2021 [Accessed: 20/09/2021]; Available from: https://digital.nhs.uk/data-and-information/data-tools-and-services/data-services/patient-level-information-and-costing-system-plics-data-collections.
62. Lensen, S., et al., *Access to routinely collected health data for clinical trials - review of successful data requests to UK registries.* Trials, 2020. **21**(1): p. 398.
63. National Institute for Health and Care Excellence, *Surgical site infections: prevention and treatment: NICE guideline*. 2019.
64. Husereau, D., et al., *Consolidated Health Economic Evaluation Reporting Standards (CHEERS) statement.* Eur J Health Econ, 2013. **14**(3): p. 367-72.
65. HM Treasury, *The Green Book: Central Government Guidance on Appraisal and Evaluation*. 2020.
66. Sanders, G.D., et al., *Recommendations for Conduct, Methodological Practices, and Reporting of Cost-effectiveness Analyses: Second Panel on Cost-Effectiveness in Health and Medicine.* JAMA, 2016. **316**(10): p. 1093-1103.
67. Tang, K., *Estimating productivity costs in health economic evaluations: a review of instruments and psychometric evidence.* Pharmacoeconomics, 2015. **33**(1): p. 31-48.
68. Nunez-Nunez, M., et al., *The methodology of surveillance for antimicrobial resistance and healthcare-associated infections in Europe (SUSPIRE): a systematic review of publicly available information.* Clin Microbiol Infect, 2018. **24**(2): p. 105-109.

Supplementary Material for 'Quantifying the Primary and Secondary Effects of Antimicrobial Resistance on Surgery Patients: Methods and Data Sources for Empirical Estimation in England'

Table A1. Summary of Health and Economic Data Sources Discussed for Antimicrobial Resistance Burden Estimation for Surgery Patients in England

Datasets are listed in alphabetical order where possible. *The list of variables provided in this column may not be exclusive but describes the types of data held relevant to this study. † Datasets listed as 'United Kingdom Health Security Agency' were managed by Public Health England at the time of reference sourcing but are transferred to the United Kingdom Health Security Agency following October 1st, 2021.

Dataset *[Information asset manager]*	Primary Purpose	Data Available*	References
Adult Social Care Finance Return (ASC-FR) & Short and Long Term Services (SALT) collection [*NHS Digital*]	The Adult Social Care Activity and Finance data files are produced from a combination of the ASC-FR and SALT collections, collected from Councils with Adult Social Services Responsibilities (CASSRs) in England to provide insight into adult social care activity and expenditure on an annual basis.	This dataset contains by region, age bands and care type (long term versus short term, or specific support type such as "physical support" provided in a residential home) data on; expenditure (gross and net), income (total and by source), unit costs and number of patient/client completed episodes, number of clients accessing care and new requests for care.	[1]
Analysis of population estimates tool *[Office for National Statistics (ONS)]*	The aim of the ONS is to 'collect, analyse and disseminate statistics about the UK's economy, society and population.' All the cited datasets from ONS are stated to 'have a wide range of uses' by 'Central government, local government and the health sector use them for planning, resource allocation and managing the economy.'	This dataset provides an 'interactive analysis of estimated annual, mid-year population changes for England and Wales, by geography, age and sex.'	[2, 3]
Annual Survey of Hours and Earnings (ASHE) *[Office for National Statistics (ONS)]*	The aim of the ONS is to 'collect, analyse and disseminate statistics about the UK's economy, society and population.' All the cited datasets from ONS are stated to 'have a wide range of uses' by 'Central government, local government and the health sector use them for planning, resource allocation and managing the economy.'	The Annual Survey of Hours and Earnings contains estimates of earnings for employees, including (but not limited to): • gross weekly pay • weekly pay excluding overtime • overtime pay • gross hourly pay • gross annual pay These data are available by sex and full-time or part-time status, by region, occupation, industry, age group and public or private sector.	[4]
Clinical Practice Research Datalink (CPRD) (~700 GP Practices)	Each of these data sources are based on information from GP practice administrative, which use hierarchical clinical coding systems. The primary	Data variables are those based on the result of GP consultations. This includes patient characteristics (NHS number, birth month and year, marital status etc.), practice characteristics (unique	[5-9] [10]

[*Department of Health and Social Care*] AND The Health Improvement Network (THIN) (~600 GP Practices) [*The Health Improvement Network Ltd*] AND QResearch (~1500 GP Practices) [*University of Oxford and EMIS*] AND ResearchOne (~400 GP Practices) [*The Phoenix Partnership (TPP) and the University of Leeds*]	purposes of the administrative electronic health record systems have been previously stated to mainly include clinical practice (aiding in the recall of previous treatment or diagnoses to clinicians). However other functions are financial (billing and budgeting) and statistical (such as research use).	identifier, region), consultation characteristics (staff, dates, type, medical codes), test and treatment characteristics (product/test types, amounts, results (for tests), dosages (for treatments). *Please note that much information here is taken from the cited Kontopantelis 2017 paper and as such exact numbers and descriptors may have changed subsequently.*	
CQC Locations [*Care Quality Commission (CQC)*]	A directory of providers that offer care that are registered with the CQC, whereby the CQC monitor, inspect and regulate services offered (including for care homes, dentists, hospitals, General Practitioner Practices, community and mental health services).	The weekly-updated dataset, provides the following variables for each registered facility; Name, Also known as, Address, Postcode, Phone number, Service's website (if available), Service types, Date of latest check, Specialisms/service, Provider name, Local Authority, Region, Location URL, CQC Location, CQC Provider ID.	[11, 12]
Deaths registered weekly/monthly in England and Wales *[Office for National Statistics]*	Surveillance of mortality events based on all death registrations in England.	Total counts of death by sex, age and week or month are available through this dataset. Non-open-access data extracts also include patient level data such as NHS number, date of birth, date of death and listed cause of death.	[13-15]
Drugs and pharmaceutical electronic market information tool (eMIT) [*Department of Health and Social Care*]	To provide information about the prices and usage for generic drugs/pharmaceutical products used by NHS hospitals.	The data is based on "Pharmex" data (data on 13 million line-order-entries for over 10,000 products), held by the Commercial Medicines Unit. The data eMIT provides (for the specified 12month period) includes: (i) Drug name and pack-size. (ii) Quantity: an estimate of NHS hospital-sector usage from English trusts for each of the products.	[16]

		(iii) Weighted average price: the average price paid for that product over the last 4 months of the period. (iv) A measure of how much that average changed.	
English Prescribing Dataset (EPD) [*NHS Business Services Authority*]	This dataset aims to collect and provide community-level prescribing information to NHS stakeholders primarily to aid performance management, financial planning and improve clinical practice.	It includes for each General Practitioner (GP) Practice (at the GP-level)/cost-centre: GP identifiers, prescribed and dispensed medicines and dressings (by British National Formulatory Chapter, Chemical Substance and Presentation); prescribed and dispensed dressings and appliances; total number of items, the quantity for each individual item, the total quantity prescribed and dispensed, the 'Net Ingredient Cost') (based on published prices), the 'Actual Cost' (accounts for the national average discount and some payments to dispensers), average Daily Quantity (the typical daily dose of a medication, prescribed to adult patients.	[17, 18]
Estates Returns Information Collection (ERIC) *[NHS Digital]*	To serve as a mandatory collection for all NHS organisations providing NHS funded secondary care, in England, to monitor 'efficiencies and funding of the NHS estate'.	Trust code, name and type (e.g. acute teaching versus community) data are available. Additionally, it houses a breakdown of costs of providing, maintaining and/or consuming the NHS Estate including buildings, equipment, utilities and services such as food and laundry.	[19]
Hospital Episode Statistics (HES) *[NHS Digital]*	Its primary purpose is to facilitate the payment for services from NHS England to NHS hospitals and independent sector health care providers (providing NHS-commissioned care).	Data available include; Trust code, patient characteristics (such as NHS number, age and sex, postcode), episode characteristics (such as admission and discharge dates) and procedures undertaken (such as type of surgery) during hospital admission, recorded outpatient service appointments and attendance at Accident & Emergency units.	[20-22]
IQVIA Hospital Treatment Insights Service [*IQVIA*]	This dataset is used to describe and understand 'how drugs are being used and to conduct studies to monitor the use, effectiveness and safety of drugs when treating diseases' within the hospital setting.	This dataset contains HES fields and antibiotic prescription information (in non-patient-identifiable form the NHS number has been removed). Information collected by NHS Trusts includes "a patient's Name, Address, Post Code, NHS Number and Date of Birth, only the hospital Trust and NHS Digital hold these details. NHS Digital links the data to its Hospital Episode Statistics database, then removes all personal details before the data is sent to IQVIA."	[23]
Labour Force Survey (LFS) *[ONS]*	The aim of the ONS is to 'collect, analyse and disseminate statistics about the UK's economy, society	There are different individual datasets available based on the LFS available through the ONS website, such as "A05 SA:	[24]

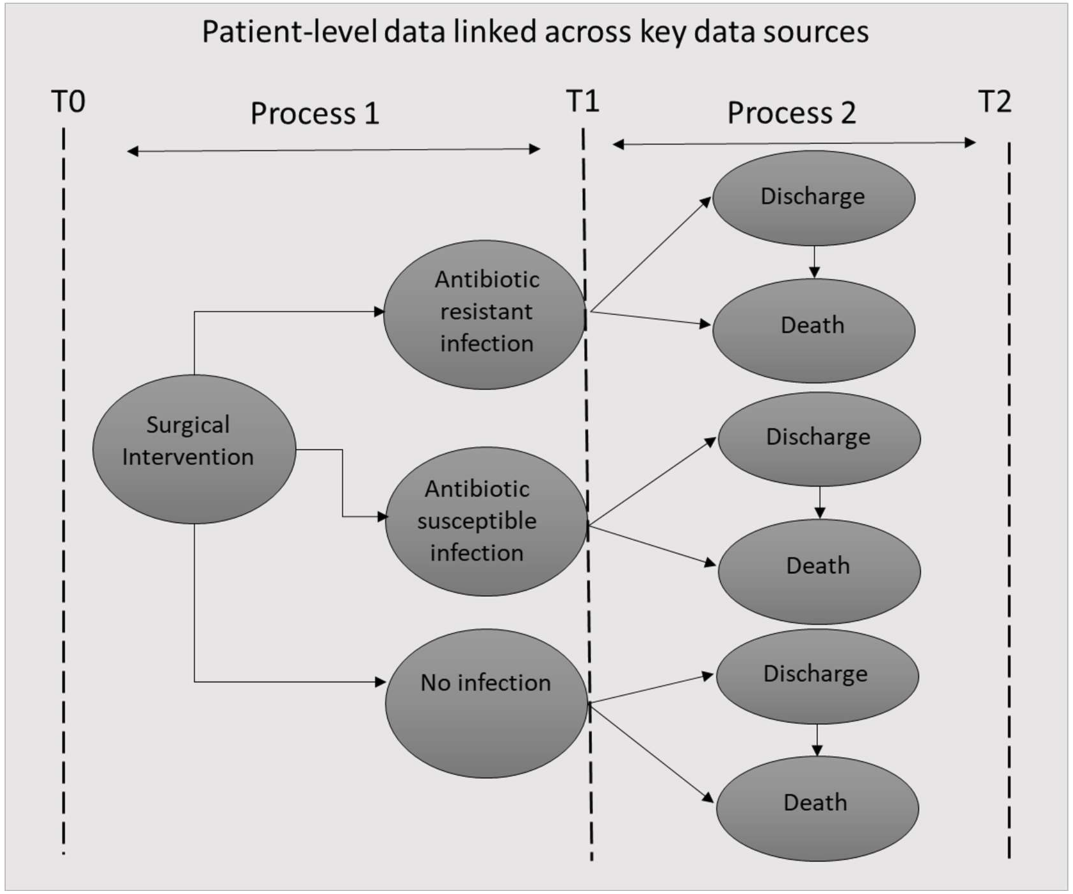
	and population.' All the cited datasets from ONS are stated to 'have a wide range of uses' by 'Central government, local government and the health sector use them for planning, resource allocation and managing the economy.'	Employment, unemployment and economic inactivity by age group (seasonally adjusted)", which provides employment, unemployment, activity and inactivity rates by sex and age groups.	
National Cost Collection *[NHS Digital, NHS England & NHS Improvement]*	To provide insights on the cost of care and how money is spent across the NHS to then help support delivery of 'high-quality care' and 'better value', nationally mandated across all secondary and tertiary care providers. The National Cost Collection (NCC) is based on the Patient Level Information Costing System data set, which is used to; inform new methods and approaches of pricing NHS services, inform the association of costs of care with provider characteristics and patient characteristics, and inform benchmarking for regulatory purposes.	Previously, NHS Reference cost reports were published within annual reports based on aggregated cost data submitted by NHS Trusts (such as unit cost per excess bed day). However, as part of the Costing Transformation Programme, National Cost Collection patient stay average costs are now estimated. Internal Trust-based Patient Level Information and Costing Systems will have patient characteristics, care characteristics and cost and income per patient per hospital episode estimates. The National Costing Grouper confers a core Health Resource Group (HRG) to every finished consultant episode of care for a patient. With access to the patient-level data one could use the NCC workbook, published HRG unit costs and the HRG grouper to calculate patient level costs (see referenced NHS Digital websites for more detail on this) as done at the Trust-level. The National Tariff annex provides HRG Name, "Outpatient procedure tariff (£)", "Combined day case /ordinary elective spell tariff (£)","Day case spell tariff (£)","Ordinary elective spell tariff (£)","Ordinary elective long stay trim point (days)","Non-elective spell tariff (£)","Non-elective long stay trim point (days)","Per day long stay payment (for days exceeding trim point) (£)" and information on Reduced short stay emergency tariffs applicable to that HRG. Additionally, total number and cost of crucial care days are submitted to the Patient Level Information and Costing System Acute collection, by critical care unit function (such as "non-specific general, adult critical care patients predominate" and "Surgical adult patients (unspecified speciality)". The acute patient level activity and costing for 2019-2020 (unbundled activity) has been published by total and trust provider level. This	[25-27]

		gives, by speciality within critical care, the total number of days and total cost (£) submitted to the PLICs Acute collection.	
NHS Electronic Drug Tariff [*NHS Business Services Authority*]	The Drug Tariff, which is sent to Pharmacies and GP Practices, outlines what will be paid to pharmacy contractors for NHS services provided either for reimbursement or for remuneration and rules to be observed when providing such services.	The monthly produced online and offline data can be used to estimate the basic cost price of drugs used and supply cost in community settings through: ➢ The basic prices of the drugs (and appropriate deductions as stated in the tariff). This includes drug, quantity of drugs, basic price and availability of generic versions of the drug. ➢ Professional fees. ➢ Cost of consumables and containers as set out in the tariff.	[28]
National life tables: UK *[ONS]*	The aim of the ONS is to 'collect, analyse and disseminate statistics about the UK's economy, society and population.' All the cited datasets from ONS are stated to 'have a wide range of uses' by 'Central government, local government and the health sector use them for planning, resource allocation and managing the economy.'	This data can be used to estimate background mortality rates, as this provides (i) the central rate of mortality, (ii) the mortality rate between age x and (x +1), that is the probability that a person aged x exact will die before reaching age (x +1), (iii) the number of survivors to exact age x of 100,000 live births of the same sex who are assumed to be subject throughout their lives to the mortality rates experienced in the three year period to which the National Life Table relates, (iv) the number dying between exact age x and (x +1) and (v) the average number of years that those aged x exact will live thereafter.	[29]
National Joint Registry (NJR) *[Healthcare Quality Improvement Partnership (HQIP)]*	The purpose of the NJR data collection is to provide an early warning system for patient safety issues. Furthermore, 'to continue to improve quality and cost-effectiveness of care, data collected about joint replacement surgery in the UK are used to report on, and monitor, patient outcomes and to support research.'	From this registry, data on the types of procedures performed (including surgical and incision approach) and if/when a revision was performed on the patient for hip, knee, ankle, elbow and shoulder joint replacements. The full non-open-access dataset includes patient identifiers (including NHS number and date of birth).	[30]
The national SSI Surveillance Service (SSISS) *[United Kingdom Health Security Agency †]*	Active surveillance programme collecting data on surgery and related SSIs submitted by hospitals through prospective follow-up of patients. Data are used to enhance the quality of patient care by providing national benchmarks which local providers can compare themselves to. Seventeen categories of major surgery are included within the SSISS, with	Patient-level and surgery-related characteristics by surgical category, SSI characteristics (such as incidence and time to infection, inpatient and/or readmission, causative organism), patient and surgical risk factors. The full non-open-access dataset includes patient identifiers (including NHS number and date of birth).	[13, 31, 32]

	surveillance targeting open surgical procedures, with orthopaedic surgery SSI surveillance mandatory.		
Patient Reported Outcome Measures (PROMs) *[NHS Digital]*	PROMs are a means of collecting information on the effectiveness of care delivered to NHS patients as perceived by the patients themselves.	The patient- and episode-level dataset provides patient answers to (i) condition specific measures for self-reported health status (such as the Oxford Hip Score (OHS) questions), (ii) generic measures for self-reported health status (such as the EQ-5D) and additional questions about patients' general health. Additionally, the dataset has patient-identifiable information for linkage (not available for general, wider analyses). This is available for hip and knee replacement surgery patients (those that complete it), it was historically also available for groin hernia and varicose vein surgeries.	[13, 33]
The Healthcare Associated Infection Data Capture System (HCAIDCS) and The Second-Generation Surveillance System (SGSS) *[United Kingdom Health Security Agency †]*	The system captures results from NHS laboratories to provide early detection and longer-term trends in changes in epidemiology, enabling the protection of public health. The system aims to determine burden of infectious disease (including epidemiology and mortality), enable timely action if necessary, monitor intervention impact, provide information to support the development of guidance on the clinical management of patients, inform material provided to the general public about infectious disease risks and ensure that the UK makes its full contribution to international efforts to protect health. The Second Generation Surveillance System (SGSS) is a voluntary surveillance data, whilst the mandatory surveillance data reported to PHE via a real-time web-based surveillance system (Healthcare Associated Infection Data Capture System (HCAIDCS)).	Data available include patient characteristics (such as NHS number, age and sex), organism characteristics (type of microbe and antibiotic susceptibility - in some cases), specimen characteristics (date and sample-type taken, source and reporting laboratories). Reporting and publication of data openly is on specific microbes that are determined to be of particular interest by the government advisory group, requiring mandatory reporting from NHS trusts (this includes *Methicillin-resistant Staphylococcus aureus* bacteraemia, *Methicillin-sensitive Staphylococcus aureus* bacteraemia, *Escherichia coli* bacteraemia, *Klebsiella spp.* bacteraemia, *Pseudomonas aeruginosa* bacteraemia and *Clostridium difficile* infections).	[34-36]

Figure A1. Potential Statistical Modelling Processes for Surgery Treatment Pathways within Hospitals

T0 is the individual-level start time (when surgery occurs) whereby outcomes of interest are observed before T2 (study end date). A time-dependent data set (with daily time intervals) can be defined including variables to highlight; T0, time of infection (T1), time of discharge and/or death. Multi-state model processes may be split into inverse probability weighting and/or cox-proportional hazards models to model the risk of experiencing infection events, discharge events and death events across either and/or both processes 1 and 2. Potential confounding impacts of different prophylaxis, treatment protocols, patient and hospital characteristics can be integrated into these models as seen in previous literature [37-39].



References for Supplementary Material

1. NHS Digital. *Adult Social Care Activity and Finance Report, England - 2019-20.* 2020 [Accessed: 30/09/2021]; Available from: https://digital.nhs.uk/data-and-information/publications/statistical/adult-social-care-activity-and-finance-report/2019-20#resources.
2. Office for National Statistics. *Population and migration*. [Accessed: 01/09/2021]; Available from: https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration.
3. Office for National Statistics. *Analysis of population estimates tool*. [Accessed: 01/09/2021]; Available from: https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/populationestimates/datasets/analysisofpopulationestimatestool.
4. Office for National Statistics. *Annual Survey of Hours and Earnings (ASHE)*. [Accessed: 01/09/2021].
5. Chaudhry, Z., et al., *Outputs and growth of primary care databases in the United Kingdom: bibliometric analysis.* Journal of innovation in health informatics, 2017. **24**(3).
6. Medicines & Healthcare Products Regulatory Agency & National Insitute for Health Research. *Clinical Practice Research Datalink.* 2021 30/09/2021]; Available from: https://www.cprd.com/.
7. The Health Improvement Network Ltd. *About us*. 2021 30/09/2021]; Available from: https://www.the-health-improvement-network.com/about.
8. QResearch. *About QResearch.* 2021 30/09/2021]; Available from: https://www.qresearch.org/about/.
9. Kneale, D., et al., *Identifying and appraising promising sources of UK clinical, health and social care data for use by NICE*. 2016, UCL Institute of Education, University College London, London.: EPPI-Centre, Social Science Research Unit.
10. McMillan, B., et al., *Primary Care Patient Records in the United Kingdom: Past, Present, and Future Research Priorities.* J Med Internet Res, 2018. **20**(12): p. e11293.
11. Care Quality Commission. *What we do.* 2021 [Accessed: 30/09/2021]; Available from: https://www.cqc.org.uk/what-we-do.
12. Care Quality Commission. *CQC care directory - zip*. 29/09/2021 [Accessed: 30/09/2021]; Available from: https://www.cqc.org.uk/files/cqc-care-directory-zip.
13. Graves, N., et al., *A cost-effectiveness modelling study of strategies to reduce risk of infection following primary hip replacement based on a systematic review.* Health Technol Assess, 2016. **20**(54): p. 1-144.
14. Health and Social Care Information Centre, *A Guide to Linked Mortality Data from Hospital Episode Statistics and the Office for National Statistics*. 2015.
15. Office for National Statistics. *Deaths registered weekly in England and Wales, provisional*. 2021 [Accessed: 27/05/2021]; Available from: https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/deaths/datasets/weeklyprovisionalfiguresondeathsregisteredinenglandandwales.
16. UK Department of Health and Social Care. *Drugs and pharmaceutical electronic market information tool (eMIT)*. 2011 [Accessed: 26/08/2021]; Available from: https://www.gov.uk/government/publications/drugs-and-pharmaceutical-electronic-market-information-emit.
17. NHS Business Services Authority. *Prescription data*. 2021 [cited 2021 18/02/2021]; Available from: https://www.nhsbsa.nhs.uk/prescription-data.
18. NHS Business Services Authority. *English Prescribing Dataset (EPD)*. 2021 [Accessed: 20/09/2021].

19. NHS Digital, *Estates Return Information Collection (ERIC) 2018/19: Data Quality Report 2019*. 2019.
20. NHS Digital, *Hospital Episode Statistics (HES) Analysis Guide*. 2019: December 2019.
21. HES Data Quality Team, *The HES Processing Cycle and HES Data Quality*, N. Digital, Editor. 2016.
22. Boyd, A., et al., *Understanding Hospital Episode Statistics (HES).* London, UK: CLOSER, 2017.
23. IQVIA. *Hospital Treatment Insights Service*. 2021 [cited 2021 22/02/2021]; Available from: https://www.iqvia.com/locations/united-kingdom/information-for-members-of-the-public/hospital-treatment-insights-service.
24. Office for National Statistics. *A05 SA: Employment, unemployment and economic inactivity by age group (seasonally adjusted)*. [Accessed: 01/09/2021]; Available from: https://www.ons.gov.uk/employmentandlabourmarket/peopleinwork/employmentandemployeetypes/datasets/employmentunemploymentandeconomicinactivitybyagegroupseasonallyadjusteda05sa.
25. NHS England and NHS Improvement, *2020/21 National Tariff Payment System*. 2020.
26. NHS England and NHS Improvement, *Volume 4: National Cost Collection – community sector*. 2020.
27. NHS England and NHS Improvement, *Volume 6: National Cost Collection – community sector*. 2020.
28. NHS Business Services Authority. *NHS Electronic Drug Tariff*. 2021 01/09/2021]; Available from: https://www.drugtariff.nhsbsa.nhs.uk/#/00805984-DC/DC00805981/Home.
29. Office for National Statistics. *National life tables – life expectancy in the UK: 2017 to 2019*. 2020 01/09/2021]; Available from: https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/lifeexpectancies/bulletins/nationallifetablesunitedkingdom/2017to2019.
30. National Joint Registry. *NJR Reports*. [cited 2021 06/09]; Available from: https://reports.njrcentre.org.uk/.
31. Troughton, R., et al., *Mapping national surveillance of surgical site infections in England: needs and priorities.* J Hosp Infect, 2018. **100**(4): p. 378-385.
32. Public Health England, *Surveillance of surgical site infections in NHS hospitals in England: April 2019 to March 2020*. 2020.
33. Secondary Care Analysis (PROMs) NHS Digital, *Patient Reported Outcome Measures (PROMs) in England*. 2017.
34. Public Health England. *AMR local indicators* Fingertips 2021 18/02/2021]; Available from: https://fingertips.phe.org.uk/profile/amr-local-indicators/data.
35. Public Health England, *English Surveillance Programme for Antimicrobial Utilisation and Resistance (ESPAUR) Report 2019 to 2020*. 2020.
36. Public Health England, *Mandatory Healthcare Associated Infection Surveillance: data quality statement: October 2020*. 2020.
37. Pouwels, K.B., et al., *Estimating the Effect of Healthcare-Associated Infections on Excess Length of Hospital Stay Using Inverse Probability-Weighted Survival Curves.* Clin Infect Dis, 2020. **71**(9): p. e415-e420.
38. Pouwels, K.B., et al., *Does appropriate empiric antibiotic therapy modify intensive care unit-acquired Enterobacteriaceae bacteraemia mortality and discharge?* J Hosp Infect, 2017. **96**(1): p. 23-28.
39. Coeurjolly, J.F., et al., *Attributable risk estimation for adjusted disability multistate models: application to nosocomial infections.* Biom J, 2012. **54**(5): p. 600-16.