Appendix - Table of Contents

METHODS

Terminology

A "Trust" is an organisational unit within the NHS that may provide hospital services, community services, or may act as a commissioner when sub-contracting care to other health services providers. In this study, the term "Trust" is synonymous with acute hospital.

A provider "spell" reflects a period of care in a single hospital, starting at the time of admission and ending at the time of discharge, death, or transfer out. Within a spell an "episode" reflects a defined period of care under a consultant. Thus, a spell comprises one or more episodes.(1) Episodes which share a patient identification number (HESID), admission date, provider code, and hospital provider spell number together compose a provider spell.(2)

Bed days reported by PHE(3) are collected by consultant main specialties and include either beds that are occupied at midnight on wards that are open overnight or beds occupied by at least one day-case on dayonly wards. The reporting excludes beds that have been commissioned from non-NHS providers, beds designated solely for use by well infants, critical care beds, and residential care beds.(4)

Identifying spells

Study authors Eric P. Budgell and A. Sarah Walker had access to the raw HES data extracts, which included 123,079,071 consultant episodes. Episodes were dropped if identified as perfect duplicates (n=14,672,193), duplicates across all fields except episode identifier number (epikey) (n=22,196), if admitted before 01 April 2009 (n=139,179), missing an admission date (n=5,804), missing a HESID $(n=70,619)$, or missing a hospital provider spell number $(n=340)$. Among the remaining 108,168,740 episodes, 88,718,419 unique spells were identified in 15,708,476 patients between 1 April 2009 and 31 March 2017. The remaining cleaning steps are outlined in Appendix Figure 1.

Imputing missing data

In the final analytic cohort (n=36,124,372 spells), the nearest preceding or subsequent spells were used to impute missing IMD score (n=83,642) (patients missing IMD score in all spells had already been dropped). Missing sex $(n=1,547)$ was imputed to the mode (female), as was missing admission method $(n=3,832)$ (mode: elective and other non-emergency), and missing admission source $(n=36,848)$ (mode: usual/other place of residence).

Preparation of antibiotic data

Only systemic antibiotics were included in DDD estimates; thus eye drops and other topical agents (gels, creams) were excluded, as were pessaries, beads, bone cement, and non-absorbed antibiotics such as neomycin (ATC code: J01GB05). A small number of other agents were excluded for miscellaneous reasons, including methenamine (this is metabolised to formalin to suppress urinary tract infections, but is not an antibiotic), demeclocycline (this is an antibiotic but is used primarily to treat hyponatremia), and colistin (ATC code: A07AA10) which was very rarely used. Rifabutin was excluded due to its probable use in tuberculosis treatment; other antibiotics used primarily in the treatment of tuberculosis were excluded in advance by PHE.

Broad-spectrum antibiotics were defined as co-amoxiclav, piperacillin/tazobactam, second (e.g. cefuroxime), third (e.g. ceftriaxone, ceftazidime), fourth (e.g. cefepime) or fifth (e.g. ceftobiprole) generation cephalosporins, cephalosporin-beta-lactamase inhibitor combinations (e.g. ceftazidimeavibactam, ceftolozane-tazobactam), carbapenems, quinolones, azithromycin, tigecycline, aztreonam, telithromycin, and polymyxins (Appendix Table 3). Antibiotics not meeting this definition were classified as narrow-spectrum.

Preparation of HES data and model selection

Following HES guidance,(5) the primary diagnosis code was used to derive 142 CCS diagnosis groups, while secondary diagnosis fields were used to derive a Charlson Comorbidity Index. Following an approach reported in previous analyses,(6) the CCS groups were then aggregated into 43 subgroups based on clinical feedback, and categories with low $(<0.5\%)$ observed mortality were collapsed into a single 'low risk' group. To improve model stability, the smallest remaining categories (cumulatively accounting for 6.11% of all spells) were then collapsed into two 'other' categories with 30-day mortality risk above and below the overall mortality risk, respectively. This process yielded a total of 29 CCS categories shown in Appendix Table 5. A binary indicator of immunosuppression was defined from secondary diagnosis code indicating metastatic cancer, severe liver disease, or HIV.

The IMD score was provided by NHS Digital based on the 2010 English index of deprivation associated with the Super Output Area Level of the patient's residence (postal codes were not collected); a more recent 2015 deprivation index is available, but HES continues to use the 2010 index.(5)

The categories used to define primary admission specialty were based on the descending prevalence of consultant specialties. First, spells with a main specialty code (HES: mainspef) or treatment specialty code (HES: tretspef) of 300 in the first or second consultant episode were defined as acute/general medicine. If this code was not present then the next most common admission specialty code was used. This logic was then repeated for each of the specialties included in the sample population (Figure 1, main text). To improve model stability, individual admission specialties comprising <1 million spells were grouped as "other", including endocrinology, diabetic medicine, acute internal medicine, respiratory medicine, infectious diseases, neurology, and rheumatology. In the primary analysis (n=36,124,372 spells), Royal Surrey NHS Foundation Trust had very few admissions (n=182) in the reference group of admission specialty (general medicine) and yielded the highest adjusted probability estimate for 30-day mortality (2.96%, 95% CI: 0.84,5.07). In meta-regression models this point estimate was therefore truncated down to the 99th percentile.

Ethnic categories (HES: ethnos) were defined as white (ethnos="British (White)", "Irish (White)", "Any other White background"), Asian (ethnos="Indian (Asian or Asian British)", "Pakistani (Asian or Asian British)", "Bangladeshi (Asian or Asian British)", "Any other Asian background", "Chinese (other ethnic group)"), black (ethnos="Caribbean (Black or Black British)", "African (Black or Black British)", "Any other Black background"), other (ethnos="White and Black Caribbean (Mixed)", "White and Black African (Mixed)", "White and Asian (Mixed)", "Any other Mixed background", "Any other ethnic group"), and unknown (ethnos="Not stated", "Not known"). Admission method categories (HES: admimeth) were defined as $A&E$ (admimeth=" $2A$ ", " 21 "), elective and other non-emergency (admimeth="11", "12", "13", "25", "31", "32", "81", "82", "83"), or emergency via GP or other source ("22", "23", "24", "28", "2B", "2D"). Admission source categories (HES admisorc) were defined as usual/other place of residence (admisorc="19", "29", "39", "66", "79") or NHS general ward/other care provider (admisorc="49", "51", "52", "53", "54", "65", "85", "87", "88"). The definition of each admission method and source category can be found in the HES Admitted Patient Care data dictionary.(7) Intended management was collected from HES but not included in regression models as it was unknown in 53.8% (n=19,417,403) of spells.

Patient classification categories were collapsed, with "Regular night attender" $(n=1,113)$ being combined with "Regular day attender", "Mothers and babies using only delivery facilities" (n=168) being combined with "Ordinary admission", and "Day-case admission" being combined with "Ordinary admission" if length of stay was >1 day (n=92).

If no outcomes were observed in one of the categories of a model factor (for example, no deaths among "regular day attenders"), the category was combined with the mode for that hospital Trust only. One exception was the "low risk" category of CCS group which was combined with the "Other" CCS group where mortality risk was below the overall risk (Appendix Table 5). In the primary analysis of 30-day mortality, such changes were made to 13/135 hospital Trusts and affected only one or two variables where categories concerned had very few admissions (often <0.1% of total admissions within the hospital).

Substantial collinearity was observed between alternate measures of previous hospital exposure, including the number of overnight admissions in the past year versus ever (Spearman's ρ: 0.73, p<0.0001) and the number of complex overnight admissions in the past year versus ever (Spearman's ρ: 0.75, p<0.0001). Each measure was assessed in a single model adjusting for all hospital Trusts, as was a binary indicator for ever having had a complex admission in the past year. Factors were selected for inclusion in multivariate models based on minimizing the BIC in these single models, and included overnight admissions in the past year and any complex admission in the past year.

To improve model stability, age at admission (years), Charlson Comorbidity Index, and overnight admissions in the past year were truncated at the 99th percentile. To account for non-linear effects of continuous variables, age, Charlson score, IMD score, and overnight admissions in the past year were fit as natural cubic splines (Stata command: mkspline, cubic), as they lowered BIC of models adjusting for all hospital Trusts. The number of knots (up to six) was chosen based on BIC, with five Harrell knots(8) chosen for age, six Harrell knots chosen for IMD score, six knots chosen for Charlson score (with knots placed at the at the 10th, 35th, 50th, 65th, 80th, and 90th percentile of the non-zero distribution), and three knots chosen for overnight admissions in the past year (with knots placed at the 10th, 70th, and 90th percentile of the non-zero distribution) (Appendix Figure 2). In the latter two covariates, placement of knots at fixed intervals was not possible due to limited variation in the truncated distributions. Adjusted models retained the same number of knots, which remained significant (p<0.001) and were visualized to assess over-fitting.

Nine interaction terms included in previous analyses(6) were added individually to a single multivariate model including all hospital Trusts and found to significantly improve the model fit (lowered the BIC); these interactions were between age and Charlson score, admission method and calendar year, admission method and overnight admissions in the past year, admission specialty and overnight admissions in the past year, admission specialty and Charlson score, overnight admissions in the past year and Charlson score, ever had a complex overnight admission in the past year and Charlson score, overnight admissions in the past year and ever had a complex overnight admission in the last year, age and overnight admissions in the last year.

The 16 admission factors in Table 1 were adjusted for irrespective of statistical significance to ensure the best possible adjustment for case-mix when deriving marginal effects, rather than to assess the relative effect of factors within models, and these admission factors have been identified as predictors of mortality in previous large studies.(6,9) For the same reason, each model was fitted to data from each hospital Trust individually. To account for patients with repeated hospital admissions (observations which are not independent), a clustered version of the sandwich estimator was used, which affects the variance– covariance matrix of the estimators (including standard errors) but has no impact on estimated coefficients, thus yielding wider confidence bounds around point estimates than would be observed if all observations were assumed to be independent.

All marginal effects were derived in Stata using "nlcom" for nonlinear combination of estimators; results were compared with output from Stata's "margins" command, which was computationally slow but

provided a reference check on all estimates and their standard errors. Random effects meta-regression was then performed in Stata using the "metareg" command.(10) This method is an extension of standard random effects meta-analysis(11–14) and is often performed on study-level summary data, allowing the underlying sources of statistical heterogeneity in effect estimates across hospital Trusts to be explored. The use of random-effects models accounts for the variance of hospital-level effects and residual (unexplained) sources of between-hospital heterogeneity, assuming hospital-level effects follow a normal distribution around the linear predictor.

For the primary mortality analysis, marginal effects were also derived using logistic regression.(15) These were very highly correlated (Spearman's ρ: 0.9992, p<0.0001) and similar in magnitude to estimates derived from Poisson models, so were not considered further.

APPENDIX TABLES

Table 1: Changes to HES Provider Codes Due to Splitting and Mergers Between NHS Hospital Trusts, April 2010 to March 2017

 \mathbf{a}^{\dagger} A mapping from old to new provider codes is available from NHS Digital(16). Note other Trust mergers and splitting occurred during the study timeframe (for example, RG3, RGZ, and RG2 merged to form RYQ in 2010/11), but are excluded from this table because these provider codes were either already updated in our raw data files or not included. Note some provider codes have changed since the close of the HES data in March 2017 (for example, between April 2017 and October 2019 all of the following provider codes had changed due to Trust mergers: RE9, RGQ, RJF, RLN, RNL, RQ6, RQQ, RR1, RW3).

ii RYQ split into RJZ, RJ2, and RPG. RJD split into RJE and RL4. Since RJZ and RJE had the most admissions per year(17), these provider codes were assumed.

Table 2: The RECORD Statement: A checklist of items that should be reported in observational studies using routinely collected health data, and the page on which each checklist item appears

 $N/A = Not$ applicable

Table 3: Classification of Antibiotics into Broad-spectrum, Access, Watch, and Reserve Categories, Adapted by Public Health England from the World Health Organization 2017 Essential Medicines List

 1 Included in meta-regression models as its own category, separate from Access and separate from Watch antibiotics

² Although second-generation cephalosporins were classified as broad-spectrum, Cefaclor was not as it is administered orally and is not well absorbed.

³Though not absorbed, A07AA12 and A07AA09 are important (alternate) treatment options for *Clostridium difficile* and were retained as they accounted for 0.8% of total DDDs.

Table 4: Number of Acute/General Medical Admissions in 135 Acute Care NHS Hospital Trusts, England, April 2010 to March 2017

Table 5: Clinical Classifications Software (CCS) Diagnosis Group Identified among Acute/General Medical Admissions in 135 Acute Care NHS Hospital Trusts, England, April 2010 to March 2017

Table 6: Random effects meta-regression of the association between the adjusted probability of death within 30 days of admission (in/out of hospital) and different metrics of hospital-level antibiotic use, among 36,124,372 acute/general medicine admissions to 135 NHS acute care hospital Trusts, April 2010 – March 2017.

¹ Probability of death derived from 135 separate (hospital-specific) Poisson models, with each model adjusting for all the factors in Table 1 (main text). Plots for each model are displayed in Appendix Figure 3.

Table 7: Random effects meta-regression of the association between the adjusted probability of death within 30 days of admission (in/out of hospital) and different metrics of hospital-level antibiotic use, among a more narrowly defined cohort of 19,023,144 acute/general medicine admissions, April 2010 – March 2017.

 1 ¹To improve model stability, five hospital Trusts with fewer than 50,000 admissions were excluded, including City Hospitals Sunderland NHS Foundation Trust, Harrogate and District NHS Foundation Trust, Royal Surrey County Hospital NHS Foundation Trust, The Hillingdon Hospitals NHS Foundation Trust, and The Whittington Hospital NHS Trust. Thus, the probability of death was derived from 130 separate (hospital-specific) Poisson models, with each model adjusting for all the factors in Table 1 (main text) excluding admission specialty. Two out of 130 models required the removal of an interaction term between admission method and admission year in order for the model to converge.

Table 8: Random effects meta-regression of the association between the adjusted probability of death within 30 days of admission (in/out of hospital) and different metrics of hospital-level antibiotic use, among 16,492,990 acute/general medicine admissions, 01/April/2014 – 31/March/2017.

 1 To improve model stability, four hospital Trusts with fewer than 50,000 admissions were excluded, including East Cheshire NHS Trust, George Eliot Hospital NHS Trust, Weston Area Health NHS Trust, and Wye Valley NHS Trust. Royal Surrey County Hospital was also excluded as it had just nine admissions in the reference group of admission specialty. City Hospitals Sunderland was excluded as it had no remaining admissions in the reference group of admission specialty. The probability of death was therefore derived from 129 separate (hospital-specific) Poisson models, with each model adjusting for all the factors in Table 1 (main text). Charlson comorbidity index and IMD score were modelled as linear predictors, rather than using natural cubic splines, as the use of splines led to poorly fitting models or the failure of models to converge.

Table 9: Random effects meta-regression of the association between the adjusted probability of death within 14 days of admission (in/out of hospital) and different metrics of hospital-level antibiotic use, among 36,124,372 acute/general medicine admissions, 01/April/2010 – 31/March/2017.

 1 Each model adjusting for all the factors in Table 1 (main text). However, 14 models would only converge with Charlson comorbidity index and IMD score modelled as linear predictors, rather than as natural cubic splines.

Table 10: Random effects meta-regression of the association between the adjusted probability of nonelective re-admission within 30 days of discharge and different metrics of hospital-level antibiotic use, among 34,427,698 acute/general medicine admissions discharged alive, 01/April/2010 – 28/February/2017.

 $\frac{1}{1}$ Each model adjusted for all the factors in Table 1 (main text). Charlson comorbidity score was modelled as a linear predictors, rather than using natural cubic splines, as the use of splines led to poorly fitting models or the failure of models to converge.

APPENDIX FIGURES

Figure 1: Order of Data Cleaning Steps Followed to Derive Analytic Cohort

 $\frac{1}{1}$. These spells are duplicated across adjacent HES data-years and thus the duplicate was dropped(1)

ⁱⁱ Raw data contained admissions between 2009-2017 (also see ^{vii} below). Patients with \geq 10 ages recorded were dropped, as age was judged to be invalid.

 $\frac{di}{dt}$ If a patient was recorded to have died in >1 spell, the earliest date of death was retained and subsequent spells were dropped

iv Acute/general medicine defined as in Figure 1 (main text)

^vAt this point, previous hospital exposure and re-admission after discharge were estimated, as all patients had at least 1 admission that met our inclusion criteria in Figure 1 (main text)

^{vi} Among patients ever admitted ≥ 16 years of age

^{vii} Admissions between 1 April 2009 to 31 March 2010 were used only to estimate previous hospital exposure for spells beginning from 01 April 2010

viii Includes one hospital Trust where most (84.9%) admissions were to clinical haematology (The Royal Marsden NHS Foundation Trust) and three hospital Trusts where all admissions were to either cardiology or respiratory medicine (Royal Brompton and Harefield NHS Foundation Trust, Papworth Hospital NHS Foundation Trust, and Liverpool Heart and Chest NHS Foundation Trust) – i.e. no admissions had a HES main specialty code (mainspef) or treatment specialty code (tretspef) equal to 300 in the first or second consultant episode

Figure 2: Non-linear Relative Risks of Death Within 30 Days of Admission (In/Out of Hospital) Among Model Covariates Fit As Natural Cubic Splines

Adjusted risk estimates derived from a single Poisson model containing NHS hospital Trust as a factor and all admission characteristics in Table 1 (main text).

Figure 3: Top 10 Most Commonly Used Antibiotics as a Fraction of Total DDDs Within Each of 135 NHS Acute Care Hospital Trusts (2016), Ranked by Amoxicillin/Clavulanic Acid

Figure 4: Random Effects Meta-Regression of the Association Between Adjusted Probability of Death Within 30 Days of Admission (In/Out of Hospital) and Hospital-level Antibiotic Use

4.1a) Total DDDs/1,000 bed-days*

4.1b) Total DDDs/1,000 admissions*

4.2a) Inpatient DDDs/1,000 bed-days*

4.2b) Inpatient DDDs/1,000 admissions*

4.3a) Outpatient DDDs/1,000 bed-days*

4.3b) Outpatient DDDs/1,000 admissions*

4.4b) Oral DDDs/1,000 admissions*

4.5b) Parenteral DDDs/1,000 admissions*

4.6a) Narrow-spectrum DDDs/1,000 bed-days*

4.6b) Narrow-spectrum DDDs/1,000 admissions*

4.7a) Broad-spectrum DDDs/1,000 bed-days*

4.7b) Broad-spectrum DDDs/1,000 admissions*

4.8b) Access DDDs/1,000 admissions*

4.9a) Watch DDDs/1,000 bed-days*

4.9b) Watch DDDs/1,000 admissions*

4.10a) Access or Watch (depending on indication) DDDs/1,000 bed-days*

4.10b) Access or Watch (depending on indication) DDDs/1,000 admissions*

4.11a) Reserve DDDs/1,000 bed-days*

4.11b) Reserve DDDs/1,000 admissions*

4.12a) Piperacillin/tazobactam and meropenem DDDs/1,000 bed-days*

4.12b) Piperacillin/tazobactam and meropenem DDDs/1,000 admissions*

* Marginal effects were derived from 135 separate (hospital-specific) models, represented here as circles sized according to the precision of the estimate (inverse of the within-hospital variance). Most (127/135) hospital-specific multivariate models (grey circles) included four spline terms (for age, Charlson score, IMD score, and overnight admissions in the past year), however models for eight hospital Trusts (orange

circles) would only converge with two spline terms (age and overnight admissions in the past year). Antibiotic use was truncated below the 2.5 percentile and above the $95th$ percentile. Probability estimates were truncated above the 99th percentile.

REFERENCES

1. Herbert A, Wijlaars L, Zylbersztejn A, Cromwell D, Hardelid P. Data Resource Profile: Hospital Episode Statistics Admitted Patient Care (HES APC). *Int J Epidemiol*. 2017;46(4):1093-1093i. Available from: http://www.ncbi.nlm.nih.gov/pubmed/28338941

2. Health and Social Care Information Centre. Methodology to create provider and CIP spells from HES APC data. 2014. Available from: https://digital.nhs.uk/binaries/content/documents/corporatewebsite/publication-system/ci-hub/compendium-indicators/compendiumindicators/publicationsystem:cilandingasset%5B3%5D/publicationsystem:Attachments%5B12%5D/p ublicationsystem:attachmentResource

3. Public Health England. MRSA bacteraemia: annual data. Annual counts and rates of meticillin resistant Staphylococcus aureus (MRSA) bacteraemia by acute trust and clinical commissioning group (CCG). 2018. Available from: https://www.gov.uk/government/statistics/mrsa-bacteraemia-annualdata

4. NHS England. KH03 Guidance: Data Definitions. 2010. Available from: https://www.england.nhs.uk/statistics/statistical-work-areas/bed-availability-and-occupancy/

5. NHS Digital. Summary Hospital-level Mortality Indicator (SHMI) Indicator specification. Version 1.32. 2020. Available from: http://digital.nhs.uk/SHMI

6. Walker AS, Mason A, Quan TP, et al. Mortality risks associated with emergency admissions during weekends and public holidays: an analysis of electronic health records. *Lancet*. 2017 Jul 1;390(10089):62–72. Available from: http://linkinghub.elsevier.com/retrieve/pii/S0140673617307821

7. NHS Digital. HES Data Dictionary Admitted Patient Care. 2018. Available from: https://digital.nhs.uk/data-and-information/data-tools-and-services/data-services/hospital-episodestatistics/hospital-episode-statistics-data-dictionary

8. Harrell F. Regression modeling strategies: with applications to linear models, logistic and ordinal regression, and survival analysis. Springer; 2015.

9. Freemantle N, Richardson M, Wood J, et al. Weekend hospitalization and additional risk of death: an analysis of inpatient data. *J R Soc Med*. 2012 Feb;105(2):74–84. Available from: http://www.ncbi.nlm.nih.gov/pubmed/22307037

10. Harbord R, Higgins J. Meta-regression in Stata. *The Stata Journal*. 2008;8(4):493–519.

11. Greenland S, O'Rourke K. Meta-Analysis. In: Rothman K, Greenland S, Lash T, editors. Modern Epidemiology. 3rd ed. Philadelphia, USA: Lippincott Williams & Wilkins; 2008. p. 652–82.

12. Thompson S, Higgins J. How should meta-regression analyses be undertaken and interpreted? *Stat Med*. 2002;21(11):1559–73.

13. Berkey CS, Hoaglin DC, Mosteller F, Colditz GA. A random-effects regression model for metaanalysis. *Stat Med*. 1995 Feb 28;14(4):395–411. Available from: http://www.ncbi.nlm.nih.gov/pubmed/7746979

14. Jackson D, White IR, Thompson SG. Extending DerSimonian and Laird's methodology to perform multivariate random effects meta-analyses. *Stat Med*. 2009 Apr 30;29(12):1282–97. Available from: http://www.ncbi.nlm.nih.gov/pubmed/19408255

15. Norton EC, Dowd BE, Maciejewski ML. Marginal Effects—Quantifying the Effect of Changes in Risk Factors in Logistic Regression Models. *JAMA*. 2019 Apr 2;321(13):1304. Available from: http://www.ncbi.nlm.nih.gov/pubmed/30848814

16. NHS Digital. Provider spells methodology: provider mapping files October 2019. 2019. Available from: https://digital.nhs.uk/data-and-information/publications/ci-hub/summary-hospital-levelmortality-indicator-shmi

17. NHS Digital. Hospital admitted patient care activity: provider level analysis. 2017. Available from: https://digital.nhs.uk/data-and-information/publications/statistical/hospital-admitted-patientcare-activity