Potential for reducing inappropriate antibiotic prescribing in English primary care

Timo Smieszek1,2*, Koen B. Pouwels1–3, F. Christiaan K. Dolk1,3, David R. M. Smith1, Susan Hopkins4,5, Mike Sharland6, Alastair D. Hay7, Michael V. Moore8 and Julie V. Robotham1

1Modelling and Economics Unit, National Infection Service, Public Health England, London, UK; 2MRC Centre for Outbreak Analysis and Modelling, Department of Infectious Disease Epidemiology, Imperial College School of Public Health, London, UK; 3PharmacoTherapy, -Epidemiology & -Economics, Department of Pharmacy, University of Groningen, Groningen, The Netherlands; 4Antimicrobial Resistance and Stewardship and Healthcare-Associated Infections Programme, Public Health England, London, UK; 5Directorate of Infection, Royal Free London NHS Foundation Trust, London, UK; 6Paediatric Infectious Diseases Research Group, St George’s University of London, London, UK; 7Centre for Academic Primary Care, NIHR School for Primary Care Research, Population Health Sciences, Bristol Medical School, University of Bristol, UK; 8Academic Unit for Primary Care and Population Sciences, University of Southampton, Southampton, UK

*Corresponding author. Modelling and Economics Unit, National Infection Service, Public Health England, London NW9 5EQ, UK. Tel: +44 20 8327 6707; E-mail: timo.smieszek@phe.gov.uk

Objectives: To identify and quantify inappropriate systemic antibiotic prescribing in primary care in England, and ultimately to determine the potential for reduction in prescribing of antibiotics.

Methods: Primary care data from 2013–15 recorded in The Health Improvement Network (THIN) database were used. Potentially inappropriate prescribing events in the database were identified by: (i) comparing prescribing events against treatment guidelines; (ii) comparing actual proportions of consultations resulting in prescription for a set of conditions with the ideal proportions derived from expert opinion; and (iii) identifying high prescribers and their number of prescriptions above an age- and body-system-specific benchmark.

Results: Applying the most conservative assumptions, 8.8% of all systemic antibiotic prescriptions in English primary care were identified as inappropriate, and in the least conservative scenario 23.1% of prescriptions were inappropriate. All practices had non-zero reduction potentials, ranging from 6.4% to 43.5% in the middle scenario. The four conditions that contributed most to inappropriate prescribing were sore throat (23.0% of identified inappropriate prescriptions), cough (22.2%), sinusitis (7.6%) and acute otitis media (5.7%). One-third of all antibiotic prescriptions lacked an informative diagnostic code.

Conclusions: This work demonstrates (i) the existence of substantial inappropriate antibiotic prescribing and (ii) poor diagnostic coding in English primary care. All practices (not just the high prescribers) should engage in efforts to improve antimicrobial stewardship. Better diagnostic coding, more precise prescribing guidelines and a deeper understanding of appropriate long-term uses of antibiotics would allow identification of further potential for reductions.

Introduction

The use of antimicrobial drugs puts evolutionary pressure on both pathogens and commensals, which inevitably results in adaptation through selection of antimicrobial resistance (AMR)1 and represents a fundamental challenge to public health. There is a direct link between the magnitude of antimicrobial use and the burden of AMR,2 suggesting that more prudent antimicrobial use may decelerate the emergence and subsequent spread of AMR. Nevertheless, global human consumption of antibiotics has increased by over a third since 2000,3 despite a mounting consensus that a substantial share of antibiotic use is inappropriate.4–7

Although much of the global rise in human antibiotic consumption is attributed to developing and transitional countries (where significant antibiotic under-prescribing and lack of access to healthcare can also be found), substantial variation within and between developed countries8–11 is indicative of antibiotic overuse and, hence, indicative of the potential for safe reductions in antibiotic prescribing in those countries and regions with comparatively high levels of use.

In the UK, the Review on Antimicrobial Resistance12 highlighted the scale and urgency of the AMR threat, prompting the UK Prime Minister’s pledge of ‘halving the inappropriate prescription of antibiotics in humans by 2020’.13 In order to meet this ambition,
England needed to quantify the extent of inappropriate prescribing. Although previous work has identified inappropriate prescribing for selected indications and syndromes in a variety of different countries, very few studies have attempted to quantify the totality of inappropriate antibiotic prescribing (see, for example, the work by Fleming-Dutra and colleagues on inappropriate prescribing in US ambulatory care). Furthermore, it is necessary to quantify the relative contributions of different clinical syndromes and conditions to the overall level of inappropriate prescribing in order to allow decision makers and clinicians to prioritize reduction efforts.

The goal of this paper is to quantify inappropriate prescribing in English primary care to inform policy in the context of this government ambition. Inappropriate prescribing can involve different types of failings, e.g. prescribing when antibiotic treatment is not or only marginally beneficial, not prescribing an antibiotic when it is necessary, or selecting a suboptimal type of antibiotic. This study concentrates exclusively on ‘overprescribing’ and defines inappropriate prescribing as any antibiotic prescribing that is likely to have marginal or zero patient benefit and be outweighed by the potential risks of prescribing.

This paper is part of a series of papers on antibiotic use and inappropriate prescribing in primary care in England. It synthesizes the findings of the preceding papers, which: (i) illustrated the current use of antibiotics in primary care in England; (ii) presented approaches to define inappropriate prescribing; (iii) analysed levels of antibiotic prescribing by condition and variation between practices, and (iv) modelled the influence of comorbidities, steroid prescriptions and consultation rates on practices’ variation in antibiotic prescribing. First, guidelines and expert opinion were used to classify antibiotic prescribing as appropriate or inappropriate; second, the proportion of inappropriate antibiotic prescriptions in the period 2013–15 was quantified; and third, variation in inappropriate prescribing between practices was analysed.

**Materials and methods**

**Ethics**

The Health Improvement Network (THIN) data were used for this work. The data collection scheme for THIN is approved by the UK Multicentre Research Ethics Committee (reference number 07H1102103). In accordance with this approval, the study protocol was reviewed and approved by an independent Scientific Review Committee (reference numbers 16THIN071 and 16THIN071-A1).

**Data**

Data were obtained from THIN, a primary care electronic database that contains anonymized patient, prescribing practice, and consultation data (for details see Dolk et al.). Systemic antibiotic prescriptions, except antimycobacterial and antileprosy drugs, recorded during the years 2013–15 were extracted for all participating English practices that met minimal data quality criteria (see below). These prescriptions were the totality against which reduction potentials were determined.

Practices were excluded from analyses if less than half of their antibiotic prescriptions could be linked with one or more informative diagnostic code(s) (prescription–diagnosis linkage is described in detail in Dolk et al.), i.e. those practices that generally fail to document the clinical indications underlying their prescriptions. Further, practices were excluded if they were in the lowest decile of annual consultation rates for more than one (out of 11) conditions included in the expert elicitation, i.e. those with atypically low consultation rates for common conditions (suggesting poor coding habits). Finally, practices were excluded if they had not been contributing to THIN for at least one full year during the study period. Comparisons between included and excluded practices were performed. Of note, inclusion and exclusion criteria for practices and prescriptions differed between the papers in this Supplement and so some results, such as the proportion of prescriptions without a linked diagnostic code, will also differ.

**Approach for classifying appropriateness**

The appropriateness of antibiotic prescriptions was determined in a three-step approach. First, prescriptions with a diagnostic code were compared against treatment guidelines. Then, estimates from an expert elicitation were used to determine appropriateness of additional prescriptions that could not be classified as appropriate or inappropriate using guidelines alone. Finally, for prescriptions covered neither by guidelines nor by the expert elicitation, distributions of prescribing proportions and rates were used to identify practices with very high prescribing and to flag their ‘excess prescribing’ as potentially inappropriate.

If a prescription was linked with multiple diagnostic codes that (potentially) justify antibiotic prescribing, the diagnostic code most likely to warrant antibiotic prescribing was assumed to underlie the prescription.

**Guideline-based classification**

Diagnostic codes linked with antibiotic prescriptions were compared with infection treatment guidance/guidelines issued by PHE, NICE and professional societies. For some codes, guidelines unambiguously indicate whether the use of antibiotics is clearly appropriate (prescribing is always warranted or necessary) or clearly inappropriate (prescribing is never indicated). For many codes, however, guidelines indicate that antibiotics should only be prescribed in certain circumstances (e.g. for upper respiratory tract infection (URTI), antibiotic prescribing is only warranted in the event of specific combinations of symptoms and/or a severe clinical presentation). Since severity markers and combinations of symptoms were mostly unavailable in the THIN dataset, using guidelines to reliably classify these prescriptions was not possible. Finally, in many cases diagnostic codes were missing, or the codes described conditions for which no English guidelines exist. Here, any judgement on appropriateness based on guidelines was not possible, with the exception of nitrofurantoin being prescribed to men in the absence of a diagnostic code, because nitrofurantoin is only used to treat urinary tract infection (UTI) and antibiotics should always be prescribed to treat male UTI.

**Expert-based classification**

For a defined list of common conditions that may or may not require antibiotic treatment depending on other factors, an expert elicitation was conducted to generate probability distributions of ‘ideal’ antibiotic prescribing proportions, i.e. the proportions of patients that should be prescribed systemic antibiotics when presenting to primary care with a given condition (described in detail elsewhere). For each of these conditions, all consultations were extracted for which patient characteristics matched the case description of the elicitation question. Some questions excluded comorbidities or focused on a specific age group or gender, e.g. UTI is more frequent but typically less complicated in women; treatment guidelines for acute otitis media (AOM) are different for young children (described in detail elsewhere). The main rationale for exclusions was to focus on seemingly uncomplicated presentations of conditions, with the exception of acute exacerbations of COPD and to separate demographic groups for which treatment guidelines differ.

The proportion of consultations that resulted in a systemic antibiotic prescription was calculated for these conditions in each practice included in the analysis. Practice- and condition-specific prescribing proportions were
then compared with ideal prescribing proportions derived from the distributions of the pooled expert opinions. Different quantiles of these distributions were used as benchmarks of appropriate prescribing in different scenario analyses (see below). If the measured prescribing proportion of a practice was below the estimate of ideal prescribing for a specific condition, all prescriptions for that condition were deemed appropriate. Conversely, if a practice’s prescribing proportion was higher than the expert estimate, then all excess prescribing, i.e. the proportion of all prescriptions above the ideal prescribing benchmark, was deemed inappropriate.

**Distribution-based classification**

A large proportion of prescriptions were not classifiable based on either guidelines or expert opinion, partly due to poor coding. Therefore, differences in prescribing behaviour between practices were used as an additional means by which to identify more cases of suspected overprescribing. All prescriptions that were not covered by either guidelines or expert opinion were stratified by age group (<18, 18–65, >65 years) and body system according to diagnostic code [respiratory tract infection (RTI) including ear, nose and throat (ENT); urogenital; skin and wounds; gastrointestinal; miscellaneous]. For each of these groups, all respective consultations were extracted from THIN to establish age group-specific prescribing proportions for these body systems. Age group-specific prescribing rates per practice were also established for prescriptions without a diagnostic code.

Then, different quantiles were determined for all groups as benchmarks, based on practice variation in prescribing (see below). If a practice was prescribing more antibiotics than given by these distribution-based benchmarks, its excess prescribing was marked as inappropriate. This approach allowed quantification of the contribution of high prescribers to inappropriate prescribing.

**Scenario construction**

Both the expert- and distribution-based classifications require assumptions about the ‘correct’ threshold that separates appropriate from inappropriate prescribing. Experts’ uncertainty regarding their estimates of ideal prescribing was measured and the distributions of each expert were pooled to generate average estimates, but no value-free and objectively correct benchmark can be derived from these distributions. The same is true for the benchmarks derived from between-practice variation used in the distribution-based classification, for which no data exist to objectively inform a decision on where overprescribing begins and appropriateness ends.

Three different scenarios will be presented to capture a reasonable range of possible assumptions for these classification approaches. The most conservative scenario uses the third quartile of pooled expert estimates (i.e. the ‘most generous’ estimates of antibiotic appropriateness from the experts) as benchmarks for the expert-based classification and no distribution-based classification is used; the middle scenario uses the median of the expert-based classification and the third quartile of distribution-based classification (i.e. only the highest 25% of prescribers would be considered to be overprescribing in the distribution-based classification); and the least conservative scenario uses the first quartile of expert-based classification (i.e. the ‘strictest’ end of the expert opinion distribution) and the median of the distribution-based classification such that those practices prescribing above the median were considered to be overprescribing.

**Condition- and practice-level analyses**

The main analyses present proportions of inappropriate and appropriate prescriptions at an aggregate level (i.e. proportions of all prescriptions of all English practices included in the analyses). In addition to these aggregate analyses, contributions of individual conditions to the overall proportion of inappropriate prescriptions are highlighted. Finally, scatter plots relate each included practice’s antibiotic prescribing rate to their identified (i) proportion and (ii) rate of inappropriate prescriptions.

**Results**

**Included versus excluded practices**

The inclusion criteria were met by 260 out of 349 (74.5%) English practices that contributed to THIN in the study period. The included practices issued 3740186 prescriptions (81.8%; all practices: 4574373).

The following comparisons between included and excluded practices are based on the year 2013.

**Excluded practices**

Excluded practices tended to be smaller than included practices, with a median of 694.3 (IQR 4498–1698) registered patients (mid-year) versus 10853 (IQR 7476–15221) registered patients.

**Included and excluded practices**

Included and excluded practices had similar prescribing rate distributions with medians of 48.5 per 100 practice population (excluded; IQR 39.6–57.3) and 52.8 (included; IQR 45.7–62.3) respectively. Also the trimethoprim to trimethoprim and nitrofurantoin ratio and the co-amoxiclav to co-amoxiclav and amoxicillin ratio—both indicators for good prescribing practice as both the ratios should be low—were similar between excluded and included practices. Medians for trimethoprim were 0.70 (IQR 0.60–0.74) in excluded and 0.67 (IQR 0.59–0.73) in included practices. For co-amoxiclav, medians were 0.15 (IQR 0.10–0.21) in excluded and 0.15 (IQR 0.10–0.21) in included practices.

**Coverage (guideline- and expert-based)**

Only 22.6% of all prescriptions could be classified as appropriate or inappropriate: 6.4% were classified based on treatment guidelines and 16.2% were prescriptions for conditions included in the expert elicitation. An additional 33.2% of prescriptions were covered by guidelines, but appropriateness could not be determined because (i) appropriateness depended on often unavailable patient-specific information (e.g. symptom severity) and (ii) they were not included in the expert elicitation. Conditions not covered by guidelines and not included in the expert elicitation accounted for 11.0% of prescriptions and 33.2% of all prescriptions were not documented well enough to determine appropriateness, i.e. they had either no or only non-specific diagnostic codes. These figures differ slightly from those reported previously, because here some practices were excluded based on the quality of their coding.

**Proportions of appropriate and inappropriate prescriptions**

Using the medians of pooled expert opinions as benchmarks for ideal prescribing proportions (as in the middle scenario) in combination with guidelines, 11.3% of all prescriptions were found to be inappropriate and 11.3% were found to be appropriate (Figure 1). The remaining 77.4% could not be directly classified based on guidelines or expert opinion. For the prescriptions that could not be directly classified as either appropriate or inappropriate, the distribution-based approach was used to identify excess prescribing.

Combining all three classification approaches (i.e. guideline-, expert- and distribution-based), in the least conservative scenario 23.1% of all prescriptions were classified as inappropriate, in the
middle scenario 15.4%, and in the most conservative scenario 8.8% (Figure 2).

**Contributions by condition**

In the middle scenario, the conditions that contributed most to inappropriate prescribing were sore throat (23.0% of all identified inappropriate prescriptions), cough (22.2%), sinusitis (7.6%), AOM in patients older than 6 months and younger than 18 years (5.7%), urinary tract infections (3.4%), acne (2.1%), impetigo (1.8%), and bronchitis (1.6%). In total, in the middle scenario, more than 60% of all identified inappropriate prescriptions were related to RTI and ENT conditions. In the other scenarios, the relative importance of some conditions shifted, but the ranking of the four most influential conditions remained the same.

**Variation in inappropriate prescribing between practices**

Practices varied in the proportions of their prescriptions that were determined to be inappropriate. All practices had a non-zero proportion of inappropriate prescriptions. In the least conservative scenario, inappropriate proportions of individual practices’ prescribing ranged from 9.5% to 52.9%, in the middle scenario from 6.4% to 43.5%, and in the most conservative scenario from 3.6% to 16.9%.

Figure 3 shows the proportion (Figure 3a) and the rate (Figure 3b) of inappropriate prescriptions by practice, based on guidelines and expert opinion (median values) only. No relationship between practices’ antibiotic prescribing rates and their proportions of inappropriate prescriptions was found (Figure 3a), when only considering inappropriate prescribing based on guidelines and expert opinion (excess prescribing identified with the distribution-based approach is, by definition, dependent on the quantity of prescribing, since we defined excess prescribing as the prescribing of practices above a pre-defined threshold; compare with the Distribution-based classification section in Materials and methods).

**Discussion**

In all scenarios, inappropriate prescribing in English primary care could be identified, ranging from 8.8% to 23.0% of all prescriptions, while 33.2% of all prescriptions were without an informative diagnostic code and, hence, could not be assessed regarding their appropriateness. Furthermore, inappropriate prescribing was identified in all included practices, ranging from 3.6% of a practice’s prescriptions (minimum of most conservative scenario) to 52.9% (maximum of least conservative scenario). The conditions that contributed the most to inappropriate prescribing were sore throat, cough, sinusitis and otitis media.

**Strengths and limitations**

To our knowledge, this is the first attempt to quantify inappropriate prescribing in English primary care, and, globally, is one of only a few attempts to quantify inappropriate prescribing at the national level.

Using guidelines and expert opinion, only a relatively small percentage (22.6%) of all prescriptions could be classified as either appropriate or inappropriate, although this value includes the conditions that are believed to contribute most to inappropriate prescribing. In particular, RTI consultations—many of which are covered in our work via guidelines and expert opinion—have repeatedly been proposed as main drivers of overprescribing since RTIs constitute a dominant part of all prescriptions,19 and, particularly in the case of URTI, only rarely warrant antibiotic use.20,22 Nonetheless, a substantial proportion of inappropriate prescribing could not be identified by our approach, in particular because of poor coding and because the expert elicitation was limited to a subset of common conditions and excluded complicated presentations. While potential inappropriate prescribing among unidentified prescriptions is covered by the distribution-based approach (in the middle and least conservative scenarios), this approach is not mechanistic and, hence, thresholds have been set relatively conservatively. Therefore, it is likely that the amount of inappropriate prescribing has been substantially underestimated for conditions...
Figure 2. Proportions of inappropriate and appropriate antibiotic prescription in three scenarios: most conservative, middle and least conservative. Appropriateness was determined using treatment guidelines, expert opinion of ideal prescribing proportions for defined conditions, and by identifying ‘excess’ prescribing based on unusually high prescribing proportions and rates.

Figure 3. Prescribing rates of English practices versus their inappropriate prescribing [based on guidelines and expert opinion (medians) only]. (a) Inappropriate prescribing as a percentage of the totality of a practice’s prescriptions. (b) Inappropriate prescriptions per practice population (mid-year). Dot colours: purple, practices in the highest inappropriate prescribing quartile; cyan, practices between median and third quartile; green, practices between first quartile and median; red, lowest quartile.
not included in the expert elicitation and not unambiguously covered by guidelines.

One quarter of practices had to be excluded from the analysis for data quality reasons and the data quality of the included practices is not ideal. While excluding practices could, in principle, introduce bias, we found that excluded and included practices were reasonably similar regarding three different indicators for antibiotic prescribing quantity and quality (antibiotic prescribing rate, trimethoprim to trimethoprim and nitrofurantoin ratio and co-amoxiclav to co-amoxiclav and amoxicillin ratio). So we believe the analysed sample is representative. Deficient coding in the included practices could also have resulted in an underestimation of the denominators of the prescribing proportions (e.g. if prescribers were more likely to enter a diagnostic code when prescribing). As a result, the fraction of inappropriate prescriptions would have been overestimated. However, as discussed elsewhere, the very low prescribing proportions found for some conditions indicate that such a bias, if it exists, is unlikely to have been substantial.

Antibiotic prescribing is not equal to antibiotic consumption because patients do not always fill a prescription and dispensed antibiotics are not always used for the current illness (but may be kept and used for self-medication during another period of illness). One reason for the difference between antibiotic prescribing and consumption is delayed (or 'back-up') prescribing, a strategy whereby prescribers issue an antibiotic prescription but advise the patient only to fill the prescription and use the antibiotic if symptoms worsen or fail to resolve within a certain period of time. Delayed prescribing has been proven to reduce antibiotic consumption, but its capacity to reduce consumption varies depending on the specific condition, how it is implemented and other factors. There are codes in THIN for delayed prescribing, there will be further potential for reductions in antibiotic prescribing quantity and quality (antibiotic prescribing rate, trimethoprim to trimethoprim and nitrofurantoin ratio and co-amoxiclav to co-amoxiclav and amoxicillin ratio), but prescribers use them very rarely (a slight positive correlation between prescribing proportions found for some conditions indicate that such a bias, if it exists, is unlikely to have been substantial).

This work concentrated on the identification of inappropriate prescribing events among acute and seemingly uncomplicated presentations of a selection of RTIs, UTIs and skin conditions. The UK is in the lower half of European countries with respect to antibiotic prescribing rates and it has been suggested that gener-
certain infections and should, ideally, prescribe to a lower proportion of consultations than low-prescribing practices. This work has not been able to prove that high prescribers have relatively more potential for safe reduction of antibiotic prescribing, even though this seems likely (compare with work presented in this Supplement that found that comorbidities do not explain much of the variation in practice-level prescribing10).

Finally, we were able to identify conditions and syndromes for which a large volume of inappropriate prescriptions was issued. This allows practitioners and policy makers to prioritize their efforts in trying to reduce unnecessary prescribing. A particular focus should be on prescriptions for sore throat, cough and sinusitis.

Conclusions
This work has shown there is potential for reduction in inappropriate antibiotic prescribing of between 8.8% and 23.0% of all prescriptions (relative to 2013–15) in English primary care, depending on which assumptions were made. The real reduction potential is probably higher and it is important to note that the totality of inappropriate prescribing is a moving target that may change with better data, new scientific insights and novel diagnostic tools becoming available. All practices included in the analyses had some potential to reduce their antibiotic prescribing, which suggests that currently all English practices can be expected to reduce antibiotic prescribing (without withholding appropriate and necessary prescriptions). For improving future assessment of antibiotic prescribing in primary care, substantially improved documentation of diagnoses and severity is vital. For the period 2013–15, one-third of all prescriptions completely lacked diagnostic information. Immediate versus delayed prescribing should also be clearly distinguished in data.

Acknowledgements
The authors are grateful for the support of a group of subject-matter experts who formed a Modelling Oversight Group (MOG) that convened bimonthly to discuss and review this work: Charles Alessi, Eleanor Anderson, Diane Ashiru-Oredope, Chris Butler, Elizabeth Beech, André Charlett, Tim Chadborn, Mary Dixon-Woods, Heather Edmonds, Martin Gulliford, Alastair Hay, Susan Hopkins, Yingfen Hsia, Alan Johnson, Paul Little, Simon de Lusignan, Marion Lyons, Clio McNulty, Graham Mitchell, Michael Moore, Puja Myles, Rebecca Owens, Calum Semple, Mike Sharland and Karen Tan. Please note that MOG membership does not automatically indicate endorsement of the findings presented here. S. H. and J. V. R. are affiliated with the National Institute for Health Research Health Protection Research Units (NIHR HPRU) in Healthcare Associated Infection and Antimicrobial Resistance at both Imperial College London and University of Oxford in partnership with PHE. Further, we thank Sandra Bösch for improving the figures presented in this paper.

Funding
This paper is part of a Supplement supported and resourced by Public Health England (PHE). Only internal resources were used for this paper.

Transparency declarations
M. S., A. D. H. and M. V. M. are members of the UK Government Advisory Committee on Antimicrobial Prescribing, Resistance and Healthcare Associated Infection (APRHAI). J. V. R. is a co-opted member of APRHAI. All other authors: none to declare.

References
Quantifying inappropriate antibiotic prescribing


