



RESEARCH ARTICLE

REVISED **Feasibility of informing syndrome-level empiric antibiotic recommendations using publicly available antibiotic resistance datasets [version 2; peer review: 2 approved, 1 approved with reservations]**

Quentin J. Leclerc ^{1,2}, Nichola R. Naylor ¹, Alexander M. Aiken^{1,3}, Francesc Coll ^{4*}, Gwenan M. Knight ^{1,2*}

¹Department of Infectious Disease Epidemiology, London School of Hygiene & Tropical Medicine, London, WC1E 7HT, UK

²Centre for Mathematical Modelling of Infectious Diseases (CMMID), London School of Hygiene & Tropical Medicine, London, WC1E 7HT, UK

³Imperial College Healthcare NHS Trust, London, UK

⁴Department of Infection Biology, London School of Hygiene & Tropical Medicine, London, WC1E 7HT, UK

* Equal contributors

V2 **First published:** 24 Sep 2019, 4:140
<https://doi.org/10.12688/wellcomeopenres.15477.1>

Latest published: 24 Jun 2020, 4:140
<https://doi.org/10.12688/wellcomeopenres.15477.2>

Abstract

Background: Antibiotics are often prescribed empirically to treat infection syndromes before causative bacteria and their susceptibility to antibiotics are identified. Guidelines on empiric antibiotic prescribing are key to effective treatment of infection syndromes, and need to be informed by likely bacterial aetiology and antibiotic resistance patterns. We aimed to create a clinically-relevant composite index of antibiotic resistance for common infection syndromes to inform recommendations at the national level.

Methods: To create our index, we used open-access antimicrobial resistance (AMR) surveillance datasets, including the ECDC Surveillance Atlas, CDDEP ResistanceMap, WHO GLASS and the newly-available Pfizer ATLAS dataset. We integrated these with data on aetiology of common infection syndromes, existing empiric prescribing guidelines, and pricing and availability of antibiotics.

Results: The ATLAS dataset covered many more bacterial species (287) and antibiotics (52) than other datasets (ranges = 8-11 and 16-32 respectively), but had a similar number of samples per country per year. Using these data, we were able to make empiric prescribing recommendations for bloodstream infection, pneumonia and cellulitis/skin abscess in up to 44 countries. There was insufficient data to make national-level recommendations for the other six syndromes investigated. Results are presented in an interactive web app, where users can visualise underlying resistance proportions to first-line empiric antibiotics for infection syndromes and countries of interest.

Open Peer Review

Reviewer Status ✓ ✓ ?

	Invited Reviewers		
	1	2	3
version 2 (revision) 24 Jun 2020	✓ report	✓ report	
	↑	↑	
version 1 24 Sep 2019	? report	? report	? report

1. **Liam Shaw** , University of Oxford, Oxford, UK
2. **Marlieke de Kraker** , University of Geneva Hospitals, Geneva, Switzerland
3. **Ben S. Cooper** , University of Oxford, Oxford, UK

Any reports and responses or comments on the article can be found at the end of the article.

Conclusions: We found that whilst the creation of a composite resistance index for empiric antibiotic therapy was technically feasible, the ATLAS dataset in its current form can only inform on a limited number of infection syndromes. Other open-access AMR surveillance datasets are largely limited to bloodstream infection specimens and cannot directly inform treatment of other syndromes. With improving availability of international AMR data and better understanding of infection aetiology, this approach may prove useful for informing empiric prescribing decisions in settings with limited local AMR surveillance data

Keywords

Antimicrobial resistance, empiric therapy, guidelines, online tool, data linkage

Corresponding author: Quentin J. Leclerc (quentin.leclerc@lshtm.ac.uk)

Author roles: **Leclerc QJ:** Conceptualization, Data Curation, Methodology, Software, Visualization, Writing – Review & Editing; **Naylor NR:** Conceptualization, Data Curation, Investigation, Methodology, Software, Visualization, Writing – Review & Editing; **Aiken AM:** Conceptualization, Data Curation, Investigation, Methodology, Validation, Writing – Review & Editing; **Coll F:** Conceptualization, Data Curation, Investigation, Methodology, Validation, Visualization, Writing – Review & Editing; **Knight GM:** Conceptualization, Data Curation, Formal Analysis, Methodology, Software, Supervision, Visualization, Writing – Original Draft Preparation, Writing – Review & Editing

Competing interests: F.C. is consultant to Next Gen Diagnostics. All other authors have no competing interests.

Grant information: Q.L. is funded by a Medical Research Council London Intercollegiate Doctoral Training Program studentship (MR/N013638/1). N.N. is funded by is funded by CGIAR Research Program on Agriculture for Nutrition and Health (A4NH) – (CRP21-0B3-2017). F.C. is funded by a Sir Henry Wellcome Postdoctoral Fellowship (201344/Z/16/Z). G.K. is funded by the Medical Research Council (MR/P014658/1).

The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Copyright: © 2020 Leclerc QJ *et al.* This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

How to cite this article: Leclerc QJ, Naylor NR, Aiken AM *et al.* **Feasibility of informing syndrome-level empiric antibiotic recommendations using publicly available antibiotic resistance datasets [version 2; peer review: 2 approved, 1 approved with reservations]** Wellcome Open Research 2020, 4:140 <https://doi.org/10.12688/wellcomeopenres.15477.2>

First published: 24 Sep 2019, 4:140 <https://doi.org/10.12688/wellcomeopenres.15477.1>

REVISED Amendments from Version 1

This article has been updated in response to reviewer comments. Notably, the new [Figure 3](#) illustrates the calculation behind our resistance index, and [Figure 4](#) has been modified to better communicate our assumptions and decision-making process for deciding recommendations for treatment. Overall, our changes aimed to improve the clarity of our article, and expand the Discussion on our assumptions, the ATLAS dataset, and the implications of our results.

Any further responses from the reviewers can be found at the end of the article

Introduction

Worldwide, most bacterial infections are treated empirically, meaning that antibiotics are prescribed based on clinical judgement prior to the infectious agent and its susceptibilities to antibiotics determined by diagnostic tests¹. A point prevalence survey of antibiotic-prescribing in children showed that globally over 75% of antibiotics in neonatal treatment were given empirically². In low- and middle-income countries, the laboratory capacity that could inform appropriate empiric therapy choices is frequently lacking.

Empiric antibiotic prescribing guidelines contain recommendations on what antibiotics to use for specific infection syndromes. An infection syndrome is a clinical situation where the presence of a specific type of infection (e.g. pneumonia, urinary tract infection) is suspected. Such an infection syndrome might thus be caused by bacteria or a virus or other transmissible pathogens or even might be due to non-infectious inflammatory causes. Here we assume that a decision has already been made that antibiotics are required (i.e. there is a bacterial cause of the syndrome) and do not address the other key diagnostic issue of bacterial or viral agent. The use of prescribing guidelines has been associated with reductions in patient mortality³, particularly among the most critically ill patients⁴, though benefits vary by patient group and infection⁵. Guidelines have also been shown to reduce levels of inappropriate prescribing⁶, which leads to a reduction in the selective pressure for antimicrobial resistance (AMR). Empiric guidelines are important in low-income settings where microbiological confirmation rarely occurs due to infrastructural and resource constraints^{6,7}.

Guidelines for empiric antibiotic therapies are often set at the national level. For example, in England, Public Health England and the National Institute of Health and Care Excellence (NICE) produce national antimicrobial prescribing guidance⁸. Creation of such guidelines requires an understanding of both the aetiology (typical causative pathogens) and the prevalence of relevant antibiotic susceptibilities. Each anatomic site of infection (e.g. respiratory tract, urinary tract, skin and soft tissue, gastrointestinal tract) has typical infecting microorganisms. The aetiology of some infection syndromes and associated antibiotic susceptibilities varies by setting, age and even season (as is the case for pneumonia for example⁹), but some broad generalizations can be made, especially with the broad-spectrum nature of some antibiotic agents.

Though it is recommended that prescribing guidelines should be adapted by healthcare institutions to take into account local patterns of AMR, in practice, this is infrequently performed¹⁰. This may be due to a lack of resources to develop appropriate guidelines or a lack of appreciation of the need¹¹ – furthermore, the existence of guidelines is no guarantee that local prescribers will adhere to such recommendations. Providing readily available, easy-to-use, transparently created tools based on open-access international AMR surveillance data may help practitioners in resource-limited settings generate appropriately-tailored local prescribing guidelines.

Whilst antibiotic resistance levels and other clinical criteria form the basis for designing antibiotic prescribing guidelines, in practice, antibiotic use is also constrained by market factors, such as cost and access to antibiotics. This may be particularly true in the case of low- and middle-income countries, which can have limited healthcare budgets and access to medicines. Two antibiotic market factors which can be informed through open-access data are those of antibiotic supplier prices and antibiotic placement on the World Health Organization's (WHO's) Essential Medicines List¹².

Currently, antibiotic resistance surveillance tools typically present resistance data for individual bacteria-antibiotic (“bug-drug”) combinations¹³. We use a more clinically-oriented presentation of resistance proportions at the level of infection syndromes, which could be used to inform empiric antibiotic prescribing recommendations. Similar “indices” have been previously proposed¹⁴. The Drug Resistance Index was developed to quantify resistance to multiple antibiotics for individual bacterial species¹³ and communicate to policymakers and non-experts the combined impact resistance has on the antibiotics available for treatment, without directly supporting clinical care. A similar index, but used to assess the population-level appropriateness of empiric treatment regimens for complicated UTI in the Netherlands was also explored by Ciccolini *et al.*⁷ where the relative frequency of causative agents and frequency of resistance was combined with antibiotic usage data. A study in a Canadian intensive care unit explored the likely efficacy of empiric treatment for three device-associated infections by creating a composite syndrome level resistance¹⁵. A “basket” of bacterial agents causing each infection was used similarly to how economists measure the average price of a standard basket of consumer goods weighted by the relative importance of each good¹⁶. There are also the weighted incidence syndromic combination antibiograms, which aim to inform empiric prescribing by considering the local weighted incidence of causative pathogens for an infection syndrome¹⁷. Thus there are several examples of empiric therapy indices, however, they are either setting and/or syndrome specific, and do not present other potentially important information such as measures of drug access and/or cost information alongside clinical data.

We test the feasibility and robustness of creating a syndrome-level composite resistance index from open-access data sources, including international AMR surveillance datasets, and developed a user-friendly web-based application, the AR.IA App¹⁸, that

brings together all this information. The app does not aim to be a predictor for likelihood of viral versus bacterial infections, but rather to aid antibiotic prescribing choice where the infectious agent is presumed to be bacterial. This work was undertaken as part of the Wellcome Data Re-use Prize¹⁹, motivated by the release of a new open-access dataset (ATLAS) from Pfizer that contained 633,820 bacterial clinical isolates collected from 77 countries over a 14-year period¹⁹.

Methods

This work consists of three main objectives, where specific methods applied:

1. To compare antibiotic resistance proportions calculated using the ATLAS dataset with those estimated from other global AMR surveillance datasets.
2. To integrate data on antibiotic susceptibilities from the ATLAS dataset with the aetiology of infection syndromes to derive a syndrome-level composite resistance index; and combine such data with access to and cost of antibiotics.
3. To develop an interactive web app (AR.IA App) to access the above information and offer empiric therapy recommendations based on available data.

All of the above was conducted in R software²⁰, using the following packages: [shiny](#) 1.2.0²¹, [ggplot2](#) 3.1.0²², [dplyr](#) 0.7.8²³, [rworldmap](#) 1.3-1²⁴, [RColorBrewer](#) 1.0-5²⁵, [reshape2](#) 1.4.2²⁶, [DT](#) 0.5²⁷, [magrittr](#) 1.0.1²⁸, [fuzzyjoin](#) 0.1.4²⁹. These are available at: <https://cran.r-project.org/>.

Surveillance data comparison

The [ATLAS dataset](https://amr.theodi.org/programmes/atlas) (available for download at <https://amr.theodi.org/programmes/atlas>) is an open-access dataset on human AMR surveillance data generated by the commercial pharmaceutical company Pfizer that contains granular antibiotic susceptibility data, including ‘raw’ minimum inhibitory concentration (MIC) data, for 633,820 bacterial clinical isolates collected from 77 countries and spanning 14 years³⁰. This dataset also contains information on the gender and age group (age groups are: 0–2, 3–12, 13–18, 19–64, 65–84 and 85+) of the patients isolates were collected from. It also contains the specimen clinical source, indicating the anatomical site isolates were sampled from (e.g. skin, blood, nose, etc.). The ATLAS dataset was made publicly viewable in 2017, and downloadable in 2018 as part of the Wellcome Data Reuse Prize¹⁹. This was an initiative to encourage reuse of AMR data shared by industry and to facilitate the development of common methodological and metadata standards.

We additionally used the [European Centre for Disease Prevention and Control \(ECDC\) Surveillance Atlas](#)³¹, [ResistanceMap](#) by the Center for Disease Dynamics Economics and Policy (CDDEP)³² and the [Global Antimicrobial Resistance Surveillance System \(GLASS\) database](#) by the World Health Organization (WHO)³³. The first holds AMR data collected in European countries whilst the second and third hold global data from national AMR surveillance programmes. ResistanceMap and

GLASS both include all of the ECDC dataset, since they use it as a source for their European sepsis data. Data between 2004 and 2017 were considered to match the ATLAS time coverage, but only from 2017 for the GLASS dataset (the only year available for download at the time of our analysis). Missing susceptibility labels (i.e. “resistant”, “intermediate” or “susceptible”) in the ATLAS dataset (443,899/633,820) were assigned from available MIC data for other isolates within the ATLAS dataset. We did not use any external data on breakpoints to derive susceptibility labels for MICs which were not labelled elsewhere within the ATLAS dataset (see *Further Methods* in *Extended Data* for details)³⁴.

We estimated the “agreement” of the ATLAS dataset as the percentage of resistance proportions for all bug-drug combinations with point estimates falling within the corresponding 95% confidence intervals in the ECDC Surveillance Atlas and ResistanceMap databases (see *Further Methods* in *Extended Data* for details)³⁴. Sample sizes (i.e. number of samples per country per year in each dataset, where one sample is one combination of bacteria and antibiotic) were compared using boxplots. We matched susceptibility labels across datasets by assigning all isolates as “resistant” if they were non-susceptible (i.e. “intermediate” or “resistant”).

Data integration and Mapping

[Figure 1](#) shows the steps required to extract and integrate information from sources other than the ATLAS dataset to produce our composite resistance index. We focused on nine infection syndromes. Each infection syndrome was mapped to the corresponding causative bacteria (i.e. aetiology, informed by the scientific literature), antibiotics used to treat them empirically (informed by antibiotic prescribing guidelines and clinical consultation) and related specimen sources (informed by the ATLAS metadata and clinical consultation).

(a) Common infection syndromes. We first chose which infection syndromes to focus on. A comprehensive list of infection syndromes was extracted from NICE guideline⁸. We discarded syndromes predominantly caused by viral and fungal infections and kept the nine most common bacterial ones. Clinically, these syndromes are all identifiable with simple clinical examination and/or basic investigations, and occur worldwide.

(b) Mapping isolate source to infection syndrome. We linked isolates in the ATLAS dataset to the infection syndrome they most likely originated from informed by the clinical “source” description in the ATLAS metadata. Due to the diversity of sources, we kept sample types represented by at least 1,000 isolates. We discarded sample types not clearly linked to an infection syndrome (e.g. “wound”), as these samples might not necessarily represent infecting organisms, but rather colonizing bacterial flora.

(c) Antibiotics used to treat empirically infection syndromes. We extracted which antibiotics are used to treat empirically each of the nine infection syndromes from the NICE guidelines⁸ and then took a simplified set as advised by clinical consultation.

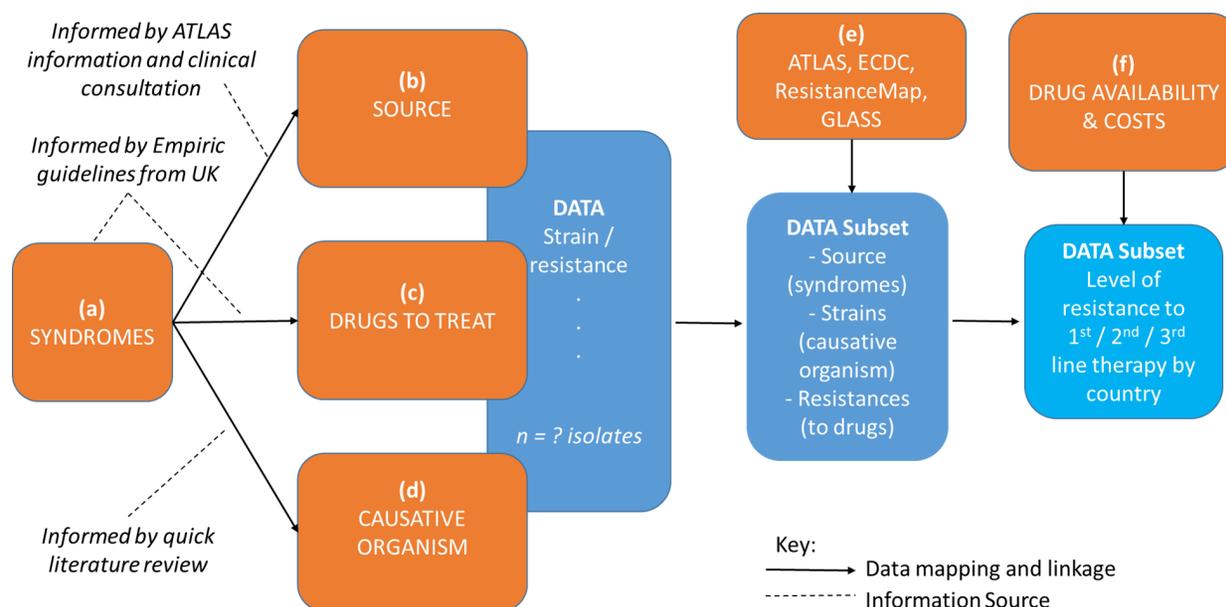


Figure 1. Flowchart of data analysis and inputs for this project. Orange boxes indicate different datasets from which information was extracted. Dark blue boxes represent the initial and subset ATLAS dataset used to inform levels of resistance (light blue box) to standard empiric therapy. Solid arrows joining boxes indicate mapping between data types. Dashed arrows are used to indicate the sources data were extracted from. The direction of the arrows indicate the order data types were extracted and integrated.

These empiric therapies represent typical current practice in the UK, though we attempted to make use of agents that were widely available at low cost internationally. For simplicity, we did not incorporate additional patient-level prescribing criteria (such as penicillin allergy status and pregnancy) when choosing these antibiotics, hence the need for clinical consultation alongside the complex NICE guidelines.

(d) Contributing pathogen distribution: syndrome aetiology.

To establish the distribution of causative pathogens for each infection syndrome, we identified reviews on the global aetiology for different syndromes and, if we could not find any, performed a rapid literature search for recent publications. Rapid, informal reviews were done by three of the authors (NRN, QJL, GMK). Individual syndromes were investigated by each author and searches of the syndrome plus terms like “aetiology” were performed in PubMed and Google in January 2019. Based on the literature found, the percentage of each infection syndrome caused by a bacterial agent were extracted from each study into a prespecified Excel data extraction sheet. One author (GMK) then integrated all of these results, per syndrome, creating a suggested pathogen distribution for each syndrome. Even where non-bacterial (i.e. viral or fungal) pathogens were found to be causative of the syndrome in the literature, only the proportion of the relevant bacterial pathogens were used in this distribution.

(e) Combining antibiotic susceptibility data from four AMR surveillance datasets. We extracted the antibiotic susceptibility data from the ATLAS dataset as well as from three more AMR surveillance datasets: ECDC, ResistanceMap and GLASS,

to allow the end-user of our ARIA App to select the underlying antibiotic susceptibility data.

(f) Combining with drug information datasets. We extracted data on supplied cost (and cost unit) for antibiotics from the Management Sciences for Health (MSH) International Medical Products Price Guide³⁵ which we inflated to the 2017 level using World Bank inflation data³⁶. We also included whether a recommended drug was on the WHO Essential Medicines List (EML) and on the AWaRE classification system, which builds on the EML to advise on what antibiotics to use for common infections (“access” category), for a small number of infections (“watch” category) and to be considered as last-resort options (“reserve” category)³⁷ (Figure 2).

(g) Mapping data to recommendations for therapy. We multiplied the frequency of each syndrome’s contributing bacteria by their resistance proportion to calculate a composite resistance index for empirically used antibiotics. An example of a made-up syndrome caused by two bacterial species can be seen in Figure 3. In this example, the composite resistance to the first-line antibiotic A is 7%, and since this is less than the default 15% cut-off, first-line treatment A would still be recommended.

Each infection syndrome was assumed to be caused entirely by bacterial species, based on the scope of our work. We noted that not all bacterial species were included in the ATLAS database, nor were all species tested for the antibiotics included in empiric therapies. We thus define the “causative pathogen availability” as the proportion of isolates from available species out of all

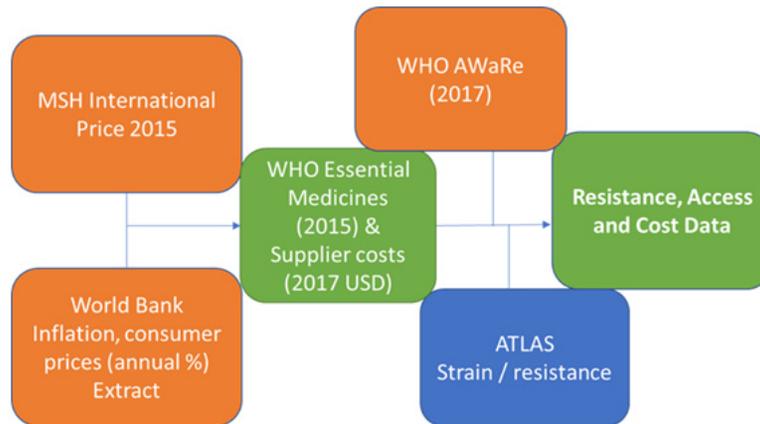


Figure 2. Flowchart of data linkage for further data. Orange boxes indicate additional datasets, blue boxes indicate ATLAS data and green boxes indicate resulting data utilised within this project.

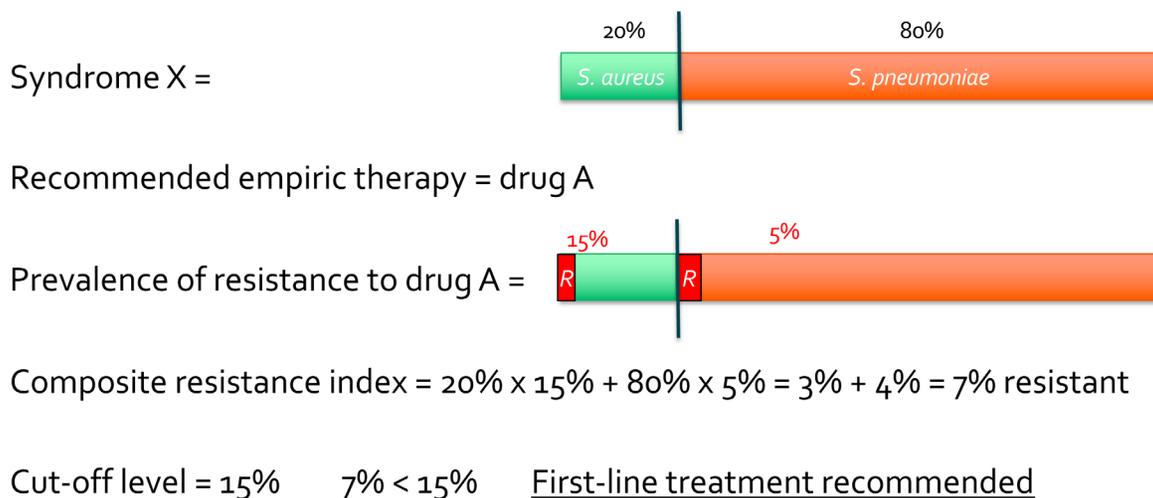


Figure 3. Example of calculation of the composite resistance index. We designed a simple hierarchical decision workflow (Figure 4) to inform on the appropriateness of using first-line empiric antibiotic therapy by comparing the syndrome-level composite resistance index calculated for each country against the chosen resistance cut-off, defined as the resistance proportion above which to escalate therapy, which is set to 15% by default in the AR.IA App.

syndrome-contributing species. For example, if 20% of the syndrome cases are due to bacteria X, 80% due to bacteria Y, but bacteria Y is not in the ATLAS database then the “causative pathogen availability” is 20%. We further looked at susceptibility coverage – not all species were tested for resistance to all drugs. Thus if only 50% of bacteria X had only been tested for resistance to all empiric therapies the coverage would only be 10% in this example. The composite resistance index is then calculated using the resistance proportions of available causative pathogens, assuming missing bacteria to be totally susceptible, which may bias towards using first line therapies. We report the causative pathogen availability, as well as whether fewer than 10 isolates

were available in the final recommendation table, but do not set a minimum threshold.

The AR.IA App creation

The Shiny R package²¹ was used to build an interactive web app (referred to as ‘the AR.IA App’) that hosts data from all four AMR surveillance datasets, integrated with external information, and the main output data used to recommend what antibiotics are appropriate to treat common infection syndromes in different regions of the world. Key parameters can be edited by the user including the underlying antibiotic susceptibility surveillance dataset used, the syndrome aetiology (proportion due

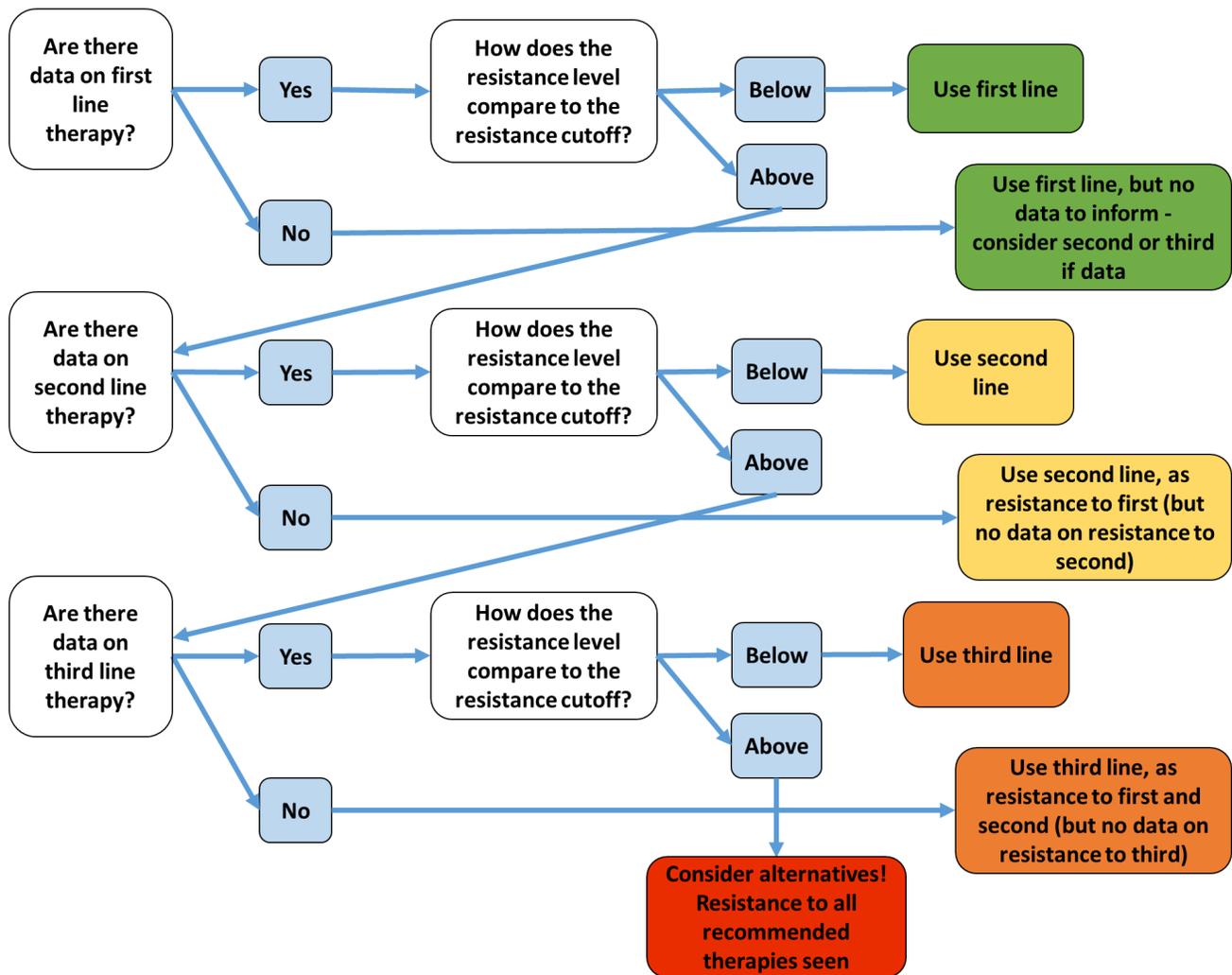


Figure 4. Description of decision-making process for deciding recommendations for treatment. If there is no data to inform therapy, we assume that the bacteria are susceptible and recommend the therapy, with a disclaimer that we have no data to inform. Note here that resistance is at the level of the syndrome. The “...” indicate where the decision making would continue onto third and higher-level therapy options.

to each included bacterial species) and the resistance cut-off for changing empiric therapy, which is set by default to 15% (see *Usage of the AR.IA App in Extended Data* for details)³⁸. The underlying data manipulation and Shiny R code can be found on [GitHub](#).

Results

Surveillance data comparison

We used four open-access datasets: ATLAS³⁰, ECDC³¹, ResistanceMap³² and GLASS³³ (Table 1).

Assigning susceptibility labels from MIC values in the ATLAS dataset reduced the number of isolates with incomplete records by 63% (to 164,918 isolates). The number of isolates per country and

year was similar across datasets (*Extended data*, Supplementary Figures 2 & 3 in *Further Results*)³⁹. The ATLAS dataset reports many more bacterial species and antibiotics tested than any other dataset (Table 1), resulting in smaller sample sizes for each country/year/species/antibiotic combination. The “agreement” of the ATLAS dataset by year as compared to the ECDC or ResistanceMap datasets ranged from 5–30% (*Extended data*, Supplementary Figure 1A in *Further Results*)³⁹. This agreement increases as the ATLAS sample size increases, suggesting that the low agreement is driven in part by small samples sizes, but stays around 25% after sample size exceeds 30,000 (*Extended data*, Supplementary Figure 1B in *Further Results*), pointing to differences in sampling as the main cause of the low agreement³⁹.

Table 1. Descriptors of surveillance datasets.

Database	ATLAS	ECDC	ResistanceMap	GLASS
Created by	Pfizer	ECDC	CCDEP	WHO
Year utilised in AR.IA App	2017	2017	2015	2017
Types of samples	All clinical samples	Blood (majority) and cerebrospinal fluid	Blood (majority) and cerebrospinal fluid	Blood, urine, genital and stool
Geographical area covered	Global	Europe	Global*	Global*
Number of countries covered	77	30	92	69
Bacterial species covered	287	8	11	8
Antibiotics tested	52	16	20	32
Sample size	1,689,267	1,299,941	1,288,325	1,683,765

*with ECDC data for European blood samples

Data integration and mapping

(a) Common infection syndromes. The nine chosen infection syndromes are shown in Table 2. These were chosen as they could be clearly linked to anatomic site or sample types, are common infections worldwide, and are caused by common bacterial species.

(b) Mapping isolate source to infection syndrome. We could map 366,001/633,820 isolates (58%) from the ATLAS dataset to the syndrome they likely originated from (Table 2). The major sources excluded were “INT: Wound” (n=96,306 isolates) and “Respiratory: Trachea” (n=19,278) as they could not be linked to a single syndrome and could represent colonizing flora. There was no accompanying clinical information available to help discriminate genuine infecting organisms from colonizers. The isolate “source” information was not sufficient to assign respiratory specimens (Table 2) to either community or hospital acquired pneumonia, thus we used the same pool of respiratory isolates for both syndromes, but with a different etiological make-up.

(c) Antibiotics used to treat infection syndromes empirically. The antibiotics used to treat the nine infection syndromes empirically are shown in Table 2 (Extended data, second line and third line presented in Supplementary Table 1 in Further Results)³⁹. Some of the antibiotics recommended for treatment were not tested against in the ATLAS dataset and thus we mapped them to their equivalent tested antibiotic where possible (Extended data, Supplementary Table 2 in Further Results)³⁹.

(d) Contributing pathogen distribution: syndrome aetiology. Except for bacterial meningitis, we could not find a consensus global aetiology for each infection syndrome. There are liable to be some regional differences in the aetiology of infections and also greater difficulty in obtaining reliable microbiological diagnosis in some parts of the world. We therefore relied on rapid literature reviews to find an approximate breakdown of the top bacterial species commonly isolated from each type of infection (see Further Results in Extended Data for details)³⁹.

However, the AR.IA App allows the bacterial aetiology to be changed by the user. Our syndrome aetiology from the literature included a total of 19 bacterial species. Of these, two were not present in the ATLAS dataset, including “*Streptococcus, viridans* group” and *Neisseria gonorrhoeae*, the latter responsible for the majority of purulent urethritis/cervicitis cases. We therefore excluded the latter syndrome from the AR.IA App.

(e) Combining antibiotic susceptibility data from four AMR surveillance datasets. We aggregated antibiotic susceptibility data across the four AMR datasets (ATLAS, ECDC, ResistanceMap and GLASS) by keeping isolates from the most recent year available (2017) for all except ResistanceMap, which had very few data points for 2016 and 2017 (n=46 and n=197, respectively) and so we used 2015 instead (n=684); by standardizing the spelling of antibiotics, species and countries; and by mapping single antibiotics in the ATLAS dataset to their corresponding antibiotic classes (reported in the rest of datasets). When combining the datasets, we average the resistance proportions reported in each dataset for each combination of country, bacteria and antibiotic class. We mapped each isolate source to their relevant infection syndrome as done for the ATLAS dataset. Datasets included isolates from different infection syndromes (Table 1).

(f) Combining with drug information datasets. Approximately 75% of the antibiotics tested in the ATLAS database and used to treat the chosen syndromes were found on the 2015 EML list and over 80% in the AWaRE classification system (see Further Results in Extended Data for details)³⁹. Only one of these antibiotics (fosfomycin) is classified on the AWaRE “reserve” group. As cost comparisons are difficult across different antibiotics that have different formulations, we allow the AR.IA App user to see exactly which formulation the available costs relate to by presenting the cost in “per specified unit”.

Summary of data integration and sub-setting. Following steps (a) - (f) above, we kept isolates in the ATLAS dataset that met

Table 2. Isolate, syndrome and antibiotic mapping results.

Infection Syndrome	Isolate source in ATLAS	Number of isolates in ATLAS	Total number of isolates in ATLAS	First Line	
				First Drug	Second Drug
Bloodstream infection	CVS: Blood	104,148	104,148	Amoxicillin	Gentamicin
Urinary tract infection	GU: Urine	82,086	84,689	Trimethoprim	-
	GU: Urinary Bladder	2,603			
Pneumonia	Respiratory: Sputum	81,669	122,291	Co-amoxiclav (Hospital-acquired)	-
	Respiratory: Bronchials	25,032			
	Respiratory: Bronchoalveolar lavage	8,335		Amoxicillin (Community-acquired)	
	Respiratory: Other	4,412			
	Respiratory: Lungs	2,843			
Cellulitis / skin abscess	INT: Abscess	15,904	33,374	Flucloxacillin	-
	INT: Skin	7,257			
	INT: Skin Ulcer	5,674			
	INT: Cellulitis/Erysipelas	2,501			
	INT: Burn	2,038			
Purulent urethritis / cervicitis	Genital/Urinary (GU)	1,416	2,613	Ceftriaxone	-
	GU: Urethra	1,197			
Upper respiratory tract infection	HEENT: Ears	5,127	15,785	Penicillin V	-
	HEENT: Throat	3,725			
	HEENT: Nose	3,197			
	Respiratory: Sinuses	2,415			
	HEENT: Other	1,321			
Bacterial meningitis	Bodily Fluids: CSF	1,811	1,811	Penicillin	Gentamicin
Septic arthritis	Bodily Fluids: Synovial	1,290	1,290	Oxacillin	-

the following criteria: were isolated from sources that could be mapped to infection syndromes (Table 2), belong to the list of bacterial species causing infection syndromes (Table 3), had assigned susceptibility status (susceptible/resistant) to at least one of the antibiotics used to treat infection syndromes, and were collected in 2017.

Applying these criteria resulted in a subset of 435,557 (69%) isolates from the ATLAS dataset that could be used to inform empiric guidelines. When grouped by country, species, syndrome and antibiotic, this resulted in 16,596 data points we could use in the ARIA App. These data points represent resistance levels to individual antibiotics in our empiric guidelines in species isolated from a syndrome source in a single country. This subset of ATLAS isolates came from 46 countries only (Figure 5A), out of an original 73, limiting the number of countries we could generate recommendations for.

(g) Resulting recommendations on the appropriateness of empiric antibiotic therapies. Recommendations existed for at least one of our nine syndromes for all 46 countries in the merged dataset. On average, each syndrome had recommendations for 32 countries ranging from 14 countries for bacterial meningitis to 44 countries for bloodstream infection (Table 3). Most of these countries were in Europe, the Americas and Asia, and only two in Africa (South Africa and Morocco) (Figure 5A). This reflects the underlying availability of isolates in the ATLAS dataset.

First-line antibiotics were recommended if the composite resistance index for the infection syndrome was lower than the default resistance cut-off (15%). This cut-off was chosen because this represented a resistance level recently used for a local change in empirical antibiotic use for treatment of suspected severe infection. It is expected that users might want to use different

Table 3. Summary table of the recommendations for therapy. Causative pathogen availability: the average proportion (over all countries) of syndrome causing bacterial species which had susceptibility information available from the dataset. We only include recommendations for syndromes with a causative pathogen availability of at least 0.5. If the composite resistance index is above 15% for an antibiotic, therapy is escalated to the next line.

Syndrome	Mean causative pathogen availability (S.D.)	Recommendation (as displayed in the app)	Number of countries	Therapy	Key driver
Bloodstream infection	0.70 (0.14)	Use second line, as resistance to first (but no data on resistance to second)	2	Cefuroxime and Gentamicin	<i>S. aureus</i> resistance
		Consider alternatives! Resistance to all recommended therapies seen	42		
Community acquired pneumonia	0.83 (0.18)	Use first line, but no data to inform – consider second or third if data	7	Amoxicillin and Clarithromycin	<i>S. pneumoniae</i> resistance
		Use second line, as resistance to first (but no data on resistance to second)	1	Co-Amoxiclav and Clarithromycin	
		Use second line	10	Co-Amoxiclav and Clarithromycin	
		Use third line (if exists)	24	Levofloxacin	
		Consider alternatives! Resistance to all recommended therapies seen	3		
Hospital acquired pneumonia	0.74 (0.16)	Use first line	2	Co-Amoxiclav	<i>S. aureus</i> and <i>P. aeruginosa</i> resistance
		Consider alternatives! Resistance to all recommended therapies seen	43		
Cellulitis / skin abscess	0.55 (0.20)	Use first line	30	Flucloxacillin	<i>S. aureus</i> resistance
		Use first line, but no data to inform – consider second or third if data	6	Flucloxacillin	
		Use second line	2	Clindamycin	
Bacterial meningitis	0.51 (0.32)	Use first line, but no data to inform – consider second or third if data	10	Penicillin and Gentamicin	*
		Use second line, as resistance to first (but no data on resistance to second)	4	Ceftriaxone and Gentamicin	
Septic arthritis	0.57 (0.09)	Use first line but no data to inform – consider second or third if data	15	Oxacillin	*

S.D., standard deviation.

*There were little data to establish key drivers.

thresholds for different syndromes, taking into account the trade-off between syndrome severity and other factors including cost and future promotion of antibiotic resistance. If susceptibility data to first line antibiotics was not available, we recommended first line therapy but clarified this as “Use first line, but no data to inform - consider second or third if data”.

Most syndromes had a mean causative pathogen availability, across all countries and antibiotics, of above 50% (Table 3),

except for complicated urinary tract infection (UTI) (2%) and upper respiratory tract infection (32.5%)

Empiric therapy recommendations by syndrome derived from the ATLAS dataset are given in Table 3, which lists the recommendations for syndromes with a mean causative pathogen availability of at least 50%. However, the recommendations for bacterial meningitis and septic arthritis rely on our assumption to treat missing susceptibility information as indicating full

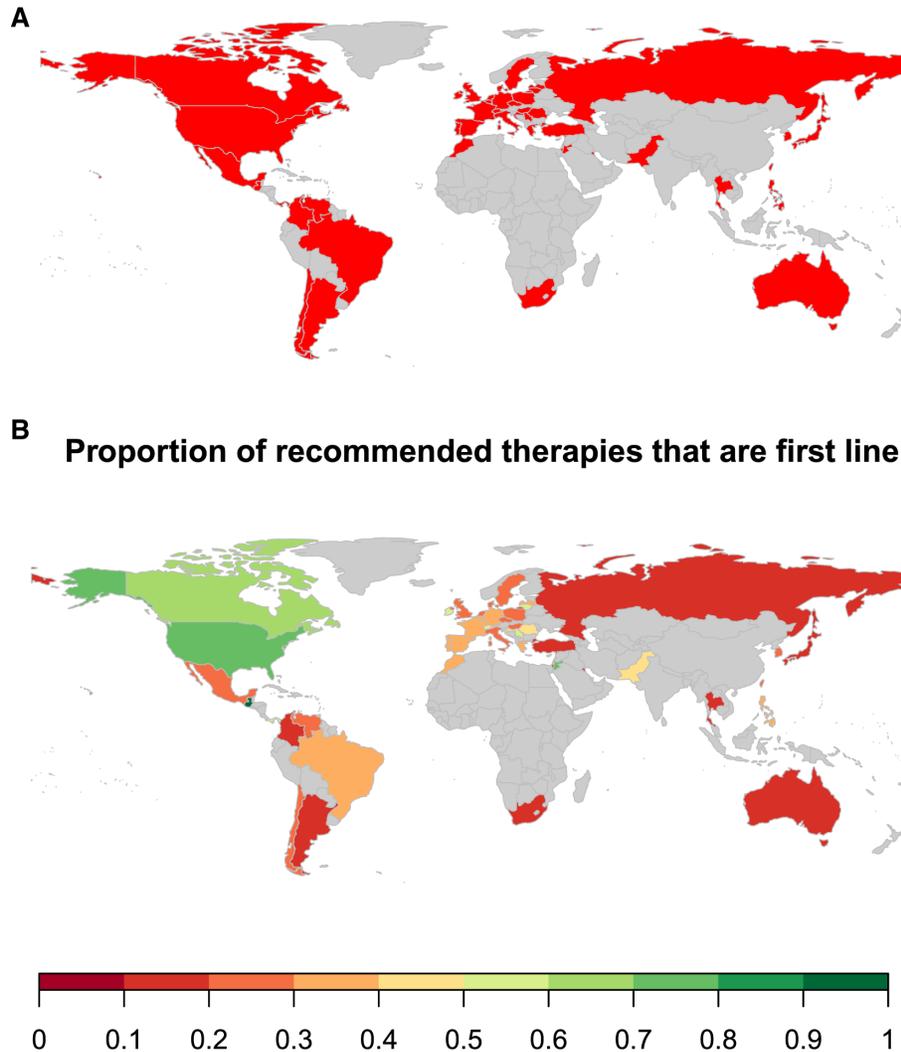


Figure 5. Map of ATLAS data. (A) Countries in red are represented in the subset of the ATLAS dataset used in our analysis. (B) Map of first line therapy proportion. Colours represent the proportion of syndromes for that country that we recommended to use first line therapy. Those in grey had no data for all syndromes.

susceptibility, as indicated by the “no data” mention. Therefore, we consider here that we cannot give robust recommendations for these two syndromes.

We found that 42/44 countries had resistance to all recommended therapies for the treatment of bloodstream infection. This was driven by high levels of amoxicillin (first-line therapy) resistance in *Staphylococcus aureus* (assumed to cause 25% of bloodstream infection cases). In total, 48% of *S. aureus* bloodstream infection isolates in the ATLAS database were resistant to oxacillin and hence phenotypically MRSA, which are also resistant to cefuroxime (the 2nd line agent). The high levels of resistance to first-line antibiotics seen in hospital-acquired pneumonia were driven by high proportions of beta-lactam resistance (cefuroxime in this therapy) in *S. aureus* and *P. aeruginosa* contributing 35% and 28% to this syndrome, respectively. A similar pattern was

seen in community-acquired pneumonia, in this case driven by *Streptococcus pneumoniae* resistance to amoxicillin. For cellulitis/skin abscess the opposite was true, with the majority (>75%) of countries being recommended to use the first line antibiotic (flucloxacillin).

E. coli causes around 70% of complicated UTI but this species was not commonly tested for trimethoprim susceptibility in the ATLAS dataset, which resulted in a mean causative pathogen availability of only 2% across all countries. Low causative pathogen availability for complicated UTI and respiratory tract infection meant that recommendations could not be strongly supported for these syndromes.

In summary, we could make recommendations for a limited number of infection syndromes using the ATLAS dataset: only

for bloodstream infection, pneumonia and cellulitis/skin abscess. Our algorithm frequently recommended use of last-resort therapies due to high levels of resistance in *S. aureus* and *Pseudomonas aeruginosa*. This contrasts with the lower antibiotic resistance rates derived from ResistanceMap, ECDC and GLASS datasets, which suggests that there may be systematic differences in the way these datasets were generated.

Most of our final recommendations included therapies that were in the WHO Essential Medicines List and none were on the AWaRE reserve list. This means that theoretically most of the recommended antibiotics should be available on the market, even in resource constrained settings.

The AR.IA App

The AR.IA App (available here: https://gwenknight.shinyapps.io/empiric_prescribing/)¹⁸ presents the underlying data and recommendations described above. User instructions on how to use the App are presented in the section *AR.IA App Documentation in Extended Data*³⁸. The AR.IA App allows the user to choose and combine multiple AMR surveillance datasets when calculating the syndrome-level composite resistance index. It allows the user to change multiple parameters, including syndrome type, resistance cut-off and aetiology. Recommending a change in antibiotic prescribing is a binary decision based on the resistance cut-off, users can change this cut-off to explore changes in antibiotic recommendations when the composite resistance level is close to the resistance cut-off. It can produce visual aids like the maps showed above (see [Figure 4B](#), showing a global map of the proportion of syndromes for each available country for which we still recommended to use first-line therapy).

Discussion

We aimed to determine whether open-access AMR surveillance datasets, such as the newly available ATLAS dataset, could be used to inform on the appropriateness of empiric antibiotic therapies to treat common infection syndromes. We integrated data on which antibiotics are commonly prescribed as empiric therapy, the bacterial aetiology of each syndrome and the antibiotic susceptibilities of syndrome-contributing bacteria to produce a syndrome-level composite resistance index. We presented our results on an interactive web app, the AR.IA App¹⁸, to allow users to explore the impact of resistance proportions on prescribing decisions. Our code is available in an open-access format and broken down into discrete sections that can be re-used and modified by any user. To our knowledge, this is the first time that antibiotic resistance estimates have been compared between multiple global AMR surveillance datasets and linked to the MSH International prices dataset to present a coalition of resistance, proxy cost and proxy access indicators.

Despite the variety of antibiotics tested, clinical sources, bacterial species and countries represented in the ATLAS dataset, we often found there were not enough isolates—from syndrome-causing bacteria, syndrome-relevant sources and tested for the antibiotic of interest—to calculate composite resistance indices

for most syndromes. As a result, we could only derive country-level recommendations for relatively few infections (bloodstream infection, pneumonia and cellulitis/skin abscess). We also noted what appeared to be an over-representation of antibiotic-resistant isolates in the ATLAS dataset, as compared to the ECDC and ResistanceMap datasets. While other surveillance datasets typically only include data on the first isolate per patient, mostly from blood and cerebrospinal fluid, the sampling methods for the ATLAS dataset are unclear to us. Our results suggest there may have been a sampling bias in the ATLAS dataset to test for non-susceptible isolates or particular types of infections with higher proportions of resistance. We are therefore more likely to observe resistance to first-line therapies in the ATLAS dataset, which leads to more frequent recommendation of last-line therapies. This is likely to be a common issue in AMR data generated from convenience sampling of clinical databases in settings with limited access to microbiology services. The ideal situation would be universal or representative sampling from all patients with suspected infection. Finally, the relatively low agreement values between ATLAS and the other AMR surveillance highlight the need for critical appraisal before using ATLAS to inform empiric prescribing in its current form. Nevertheless, to the best of our knowledge, the ATLAS dataset will remain freely accessible, and will be updated every 6 to 8 months, which could improve its usefulness to inform empiric prescribing in the future. An added improvement to the dataset would be the conversion of MIC values to sensitive/resistant classifications (e.g. using EUCAST guidelines) – currently there are measurements without classifications. We did not perform this decision making here, in part due to conflicting thresholds in different guidelines and due to the fact that many of the antibiotics with missing classifications are not used for empirical guidelines (e.g. colistin), but this would increase the data available for future analyses.

Our analysis has several limitations. We only included a limited set of infection syndromes and hence used only part of all available ATLAS entries. Future work should include other syndromes, such as purulent urethritis (typically caused by *Neisseria gonorrhoeae*), and sub-classify broad syndromes into narrower types of infections. The aim of this work is to inform prescribing after a decision, based on clinical examination, has been made as to the site and type of infection (i.e. the syndrome). Apart from bloodstream infections, which do often co-exist with other syndromes, our assumption that syndromes were independent is likely to hold but in practice few of the infection syndromes have entirely reliable identification. This simplification of syndromes is a limitation of our work but also of empiric prescribing in general.

Syndrome aetiology was informed only by basic literature reviews and will need to be supported by in-depth systematic reviews and account for regional, seasonal and host population differences. At this stage, we allow AR.IA App users to change the aetiological distributions. The choice of antibiotics used as empiric therapies could also be inputted by the user. The level

of recommendations (i.e. country-level) was dictated by the type of sampling available from global AMR surveillance datasets. Increased granularity will be needed to tailor antibiotic prescribing guidelines to local settings (e.g. hospital-level). Alternatively, resistance data could be pooled across multiple neighbouring countries with sparse availability of resistance information.

An important assumption we make is that bacteria with missing antibiotic susceptibility information are considered to be fully susceptible to that antibiotic. If we were to reverse this assumption, and consider that bacteria with missing information are fully resistant, this would push our recommendations towards second or third line therapy, or even to the conclusion that none of the therapies would be effective. As we provide the value of our resistance index for all levels of therapy in the AR.IA App regardless of our final recommendation, users can themselves choose to interpret missing values (indicated by “NA”) as being a synonym of resistance rather than susceptibility, and escalate therapy to the next line.

AMR surveillance datasets do not report the susceptibilities of antibiotics that bacteria are commonly sensitive or intrinsically resistant to. For example, *Pseudomonas aeruginosa* is intrinsically resistant to many beta-lactam antibiotics and thus it is never tested against these agents. AMR surveillance datasets will need to systematically incorporate these rules. Datasets other than ATLAS reported susceptibility to antibiotic classes (e.g. cephalosporins), instead of that to individual drugs (e.g. ceftriaxone), which was not helpful for empiric therapy design as resistance is not always common to all antibiotics belonging to the same class. Relatively low isolate numbers for some countries restricted the potential usefulness of the datasets, and this factor also forced us to rely on potentially outdated numbers from previous years (up to 2015 for ResistanceMap). Additionally, the data presented in these surveillance datasets focus on hospital-associated infections, which limits their relevance for community-associated infections. Given these limitations, these AMR surveillance datasets may not be the most reliable source of AMR proportions to base empirical treatment on. Identifying such an “optimum arrangement for recording and reporting of AMR data” is indeed one of the objectives of the UK Five-Year Antimicrobial Resistance Strategy⁴⁰. This includes using point prevalence surveys as the gold standard (as opposed to convenience sampling of isolates from clinical specimens that may be biased towards resistant strains) to estimate the total burden of AMR, determine the aetiology of common infection syndromes and ultimately inform empiric antibiotic guidelines.

This work focuses on improving guidelines for empiric treatment of infection, but there are limitations to the usefulness of this approach⁴¹. Firstly, individual patients identified by laboratory investigation to have antibiotic-susceptible infections can still be effectively treated with agents that show extensive resistance proportions at the population level. Secondly, empiric guidelines typically favour the use of broad-spectrum antibiotics to achieve the highest chance of treatment success, but this may not always be in the individual or population’s best

overall interest in terms of minimizing side-effects (such as *Clostridium difficile* infection) or conserving effectiveness of treatments. Thirdly, some drugs with low-levels of resistance but multiple other sub-optimal properties (such as vancomycin) may be recommended in later lines of empiric therapy at high levels of resistance. And fourthly, the assumption is made that infection syndromes are caused entirely by bacterial pathogens requiring antibiotic treatment, when in reality many infections are caused by viruses⁴² and would recover without needing antibiotics. Whilst future advances on rapid and cheap point-of-care diagnostics for AMR bacteria might remove the need for empiric therapy, these will continue to widely be used in many settings, especially in low-income countries.

Future iterations of this app should include user defined therapies as well as the option to upload local resistance data. We also had a binary resistance cut-off. This means that, for example, if the cut-off is set to 15% (and resistance to first line antibiotics is 16%), even if resistance to second-line antibiotics is 14% they will still be recommended. For now, users who suspect this may be an issue can alter this cut-off and see its effects on the recommendations (for example switch to 14% manually and see if there is a change in recommendation). Future iterations should include a range around the cut-off which would produce suitable warnings.

Conclusion

We have shown how independent sources of data can be combined with AMR surveillance information, such as the ATLAS dataset, to add clinical and policy-making value. Our results suggest that whilst the creation of a composite resistance index is technically feasible, the data needed to make robust prescribing recommendations for most infectious syndromes is currently lacking. In line with a move towards more evidence-based antibiotic prescribing, we believe this approach could be used to monitor the effectiveness of antibiotic empiric therapies, the cornerstone of current antibiotic prescribing practices. Such an approach can be applied to more robust data as these become available.

Data availability

Underlying data

Table 4 contains the underlying data used in this study; Table 5 contains these data compiled for use in the AR.IA App.

Extended data

Figshare: AR.IA paper Extended Data - Further Methods. <https://doi.org/10.6084/m9.figshare.9852029.v2>³⁴.

This project contains further information on the manipulation of the ATLAS dataset, the comparison of the ATLAS dataset with ECDC Surveillance Atlas and ResistanceMap, the creation of our drug-resistance index, and details on other drug information datasets used in our analysis.

Figshare: AR.IA paper Extended Data - Further Results. <https://doi.org/10.6084/m9.figshare.9852041.v2>³⁹.

Table 4. Source of the underlying data used in this study.

Title	Use	Reference
ATLAS	Source of laboratory data	https://atlas-surveillance.com/#/login and https://amr.theodi.org/programmes/atlas
ECDC Surveillance Atlas	Used in comparison of datasets	Dataset provided by ECDC based on data provided by WHO and Ministries of Health from the affected countries: https://ecdc.europa.eu/en/antimicrobial-resistance/surveillance-and-disease-data/data-ecdc
ResistanceMap	Used in comparison of datasets	The Center for Disease Dynamics Economics and Policy. ResistanceMap: Antibiotic Resistance. 2018. https://resistancemap.cddep.org/AntibioticResistance.php . Date accessed: February 13, 2019.
GLASS	Used in comparison of datasets	Dataset provided by the World Health Organization. https://www.who.int/glass/en/
MSH International Products Price Guide	Price of drug and placement on WHO Essential Medicines list	Management Sciences for Health. The International Medical Products Price Guide (2015 Edition) (2016).
World Bank CPI	Inflation of drug costs	The World Bank. Inflation, consumer prices (annual %). (2019) https://data.worldbank.org/indicator/FP.CPI.TOTL.ZG [Accessed 20/02/2019]
AWaRe	Placement on AWaRe list	Sharland, M., Pulcini, C., Harbarth, S., Zeng, M., Gandra, S., Mathur, S. and Magrini, N. Classifying antibiotics in the WHO Essential Medicines List for optimal use—be AWaRe. <i>The Lancet Infectious Diseases</i> , 18(1), pp.18–20. (2018). http://apps.who.int/medicinedocs/documents/s23413en/s23413en.pdf
NICE guidelines	Empiric antibiotic therapies	The National Institute for Health and Care Excellence (NICE) and Public Health England (PHE). Summary of antimicrobial prescribing guidance - managing common infections. 1–23 (2019): https://www.nice.org.uk/Media/Default/About/what-we-do/NICE-guidance/

Table 5. Data from Table 4 compiled for use in the final running of the AR.IA App.

Title	DOI	Use	Reference
AR.IA App all datasets	https://doi.org/10.6084/m9.figshare.9821996	Data from the ATLAS dataset, the ECDC Surveillance Atlas, ResistanceMap, and GLASS aggregated by species and drugs of interest in our analysis	43
AR.IA App drug breakdown	https://doi.org/10.6084/m9.figshare.9822077	Baseline empiric therapy recommendations	44
AR.IA App drug breakdown (groups)	https://doi.org/10.6084/m9.figshare.9822104	Baseline empiric therapy recommendations with drug groupings instead of individual drugs	45
AR.IA App species breakdown	https://doi.org/10.6084/m9.figshare.9822179	Baseline contributing bacteria distributions	46
AR.IA App all species	https://doi.org/10.6084/m9.figshare.9822146	List of species to include	47
AR.IA App economic data	https://doi.org/10.6084/m9.figshare.9822122	Cost, WHO EML and AWaRe	48

This project contains further information on the comparison of the ATLAS dataset with ECDC Surveillance Atlas and

ResistanceMap, the antibiotics used in our analysis, the summary of our review for the aetiology of the infection syndromes, the

process to combine the antimicrobial resistance datasets, and to combine the drug information datasets. Also contained are Supplementary Figures 1–3, and Supplementary Tables 1–6.

Figshare: AR.IA paper Extended Data - AR.IA App documentation. <https://doi.org/10.6084/m9.figshare.9852056.v2>³⁸.

This project contains instructions regarding the usage of the App developed as part of this project. Also contained is Supplementary Figure 4.

Extended data are available under the terms of the [Creative Commons Attribution 4.0 International license](#) (CC-BY 4.0).

Software availability

The AR.IA App is available at: https://gwenknight.shinyapps.io/empiric_prescribing/.

Source code available from: <https://github.com/gwenknight/empiricprescribing>.

Archived source code at time of publication: <https://doi.org/10.5281/zenodo.3418998>¹⁸.

Licence: [GNU General Public License v3.0](#).

Acknowledgments

The authors would like to acknowledge Dr Lisa Knight for consultation relating to clinical input. Analysis and results used in this publication were generated as part of the Wellcome Trust Data Re-use Prize: AMR Surveillance, obtained through Synapse ID syn17009517 and with support from the Open Data Institute, Pfizer, and Sage Bionetworks.

References

- Versporten A, Zarb P, Caniaux I, *et al.*: **Antimicrobial consumption and resistance in adult hospital inpatients in 53 countries: results of an internet-based global point prevalence survey.** *Lancet Glob Health.* 2018; **6**(6): e619–e629. [PubMed Abstract](#) | [Publisher Full Text](#)
- Versporten A, Bielicki J, Drapier N, *et al.*: **The Worldwide Antibiotic Resistance and Prescribing in European Children (ARPEC) point prevalence survey: developing hospital-quality indicators of antibiotic prescribing for children.** *J Antimicrob Chemother.* 2016; **71**(4): 1106–1117. [PubMed Abstract](#) | [Publisher Full Text](#)
- Schuts EC, Hulscher MEJL, Mouton JW, *et al.*: **Current evidence on hospital antimicrobial stewardship objectives: a systematic review and meta-analysis.** *Lancet Infect Dis.* 2016; **16**(7): 847–856. [PubMed Abstract](#) | [Publisher Full Text](#)
- Paul M, Shani V, Muchtar E, *et al.*: **Systematic review and meta-analysis of the efficacy of appropriate empiric antibiotic therapy for sepsis.** *Antimicrob Agents Chemother.* 2010; **54**(11): 4851–4863. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Piskin N, Aydemir H, Oztoprak N, *et al.*: **Inadequate treatment of ventilator-associated and hospital-acquired pneumonia: risk factors and impact on outcomes.** *BMC Infect Dis.* 2012; **12**: 268. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Fernández González F, Detrés J, Torrellas P, *et al.*: **Comparison of the appropriate use of antibiotics based on clinical guidelines between physicians in-training versus practicing physicians.** *Bol Asoc Med P R.* 2013; **105**(3): 21–4. [PubMed Abstract](#)
- Ciccolini M, Spoorenberg V, Geerlings SE, *et al.*: **Using an index-based approach to assess the population-level appropriateness of empirical antibiotic therapy.** *J Antimicrob Chemother.* 2015; **70**(1): 286–293. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- The National Institute for Health and Care Excellence (NICE) and Public Health England (PHE): **Summary of antimicrobial prescribing guidance - managing common infections.** 2019; 1–23. [Reference Source](#)
- Choe YJ, Smit MA, Mermel LA: **Seasonality of respiratory viruses and bacterial pathogens.** *Antimicrob Resist Infect Control.* 2019; **8**(1): 125. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Elias C, Moja L, Mertz D, *et al.*: **Guideline recommendations and antimicrobial resistance: the need for a change.** *BMJ Open.* 2017; **7**(7): e016264. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Schweizer ML, Perencevich EN, Eber ER, *et al.*: **Optimizing antimicrobial prescribing: Are clinicians following national trends in methicillin-resistant staphylococcus aureus (MRSA) infections rather than local data when treating MRSA wound infections.** *Antimicrob Resist Infect Control.* 2013; **2**(1): 28. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- The Selection and Use of Essential Medicines.** World Health Organisation (WHO), 2017. [Reference Source](#)
- Laxminarayan R, Klugman KP: **Communicating trends in resistance using a drug resistance index.** *BMJ Open.* 2011; **1**(2): e000135. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Spivak ES, Cosgrove SE, Srinivasan A: **Measuring Appropriate Antimicrobial Use: Attempts at Opening the Black Box.** *Clin Infect Dis.* 2016; **63**(12): 1639–1644. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Hughes JS, Hurford A, Finley RL, *et al.*: **How to measure the impacts of antibiotic resistance and antibiotic development on empiric therapy: new composite indices.** *BMJ Open.* 2016; **6**(12): e012040. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- International Monetary Fund: **Consumer Price Index Manual: Theory and Practice.** 2004. [Reference Source](#)
- Randhawa V, Sarwar S, Walker S, *et al.*: **Weighted-incidence syndromic combination antibiograms to guide empiric treatment of critical care infections: a retrospective cohort study.** *Crit Care.* 2014; **18**(3): R112. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Knight G: **gwenknight/empiricprescribing: First release (Version v1.0.0).** *Zenodo.* 2019. <http://www.doi.org/10.5281/zenodo.3418998>
- Wellcome: **New data re-use prizes help unlock the value of research | AMR surveillance.** 2019. [Reference Source](#)
- R Core Team: **R: A Language and Environment for Statistical Computing.** R Foundation for Statistical Computing. 2017. [Reference Source](#)
- Chang W, Cheng J, Allaire JJ, *et al.*: **shiny: Web Application Framework for R.** 2018. [Reference Source](#)
- Wickham H: **ggplot2: Elegant Graphics for Data Analysis.** Springer-Verlag New York. 2016. [Publisher Full Text](#)
- Wickham H, François R, Henry L, *et al.*: **dplyr: A Grammar of Data Manipulation.** 2019. [Reference Source](#)
- South A: **rworldmap: a new R package for mapping global data.** *The R Journal.* 2011; **3**(1): 35–43. [Publisher Full Text](#)
- Neuwirth E: **RColorBrewer: ColorBrewer Palettes.** 2014. [Reference Source](#)
- Wickham H: **Reshaping Data with the reshape Package.** *J Stat Softw.* 2007; **21**(12): 1–20. [Publisher Full Text](#)
- Xie Y, Cheng J, Tan X: **DT: A Wrapper of the JavaScript Library "DataTables".** 2018. [Reference Source](#)
- Bache SM, Wickham H: **magrittr: A Forward-Pipe Operator for R.** 2014. [Reference Source](#)

29. Robinson D: **fuzzyjoin: Join Tables Together on Inexact Matching**. 2018. [Reference Source](#)
30. Beardmore D, L'Henaff P, Shetty N, *et al.*: **Antimicrobial Resistance Surveillance: Sharing industry data - Post-pilot project report**. 2018. [Reference Source](#)
31. European Centre for Disease Prevention and Control: **Data from the ECDC Surveillance Atlas - Antimicrobial resistance**. 2017. [Reference Source](#)
32. E. & P. The Center For Disease Dynamics: **ResistanceMap**. 2017. [Reference Source](#)
33. World Health Organization (WHO): **Global Antimicrobial Resistance Surveillance System (GLASS)**. WHO, 2018. [Reference Source](#)
34. Leclerc Q, Naylor NR, Aiken AM, *et al.*: **AR.IA paper Extended Data - Further Methods**. *figshare*. Journal contribution. 2019. <http://www.doi.org/10.6084/m9.figshare.9852029.v2>
35. MSH (Management Sciences for Health): **International Medical Products Price Guide, 2015 Edition**. 2016. [Reference Source](#)
36. I. F. S. and data files International Monetary Fund, Inflation, consumer prices (annual %) | **Data**. 2019. [Reference Source](#)
37. Sharland M, Pulcini C, Harbarth S, *et al.*: **Classifying antibiotics in the WHO Essential Medicines List for optimal use-be AWaRe**. *Lancet Infect Dis*. 2018; **18**(1): 18–20. [PubMed Abstract](#) | [Publisher Full Text](#)
38. Leclerc Q, Naylor NR, Aiken AM, *et al.*: **AR.IA paper Extended Data - AR.IA App documentation**. *figshare*. Journal contribution. 2019. <http://www.doi.org/10.6084/m9.figshare.9852056.v2>
39. Leclerc Q, Naylor NR, Aiken AM, *et al.*: **AR.IA paper Extended Data - Further Results**. *figshare*. Journal contribution. 2019. <http://www.doi.org/10.6084/m9.figshare.9852041.v2>
40. UK Department of Health: **UK Five Year Antimicrobial Resistance Strategy 2013 to 2018**. 2013. [Reference Source](#)
41. Friedman ND, Temkin E, Carmeli Y: **The negative impact of antibiotic resistance**. *Clin Microbiol Infect*. 2016; **22**(5): 416–422. [PubMed Abstract](#) | [Publisher Full Text](#)
42. Pneumonia Etiology Research for Child Health (PERCH) Study Group: **Causes of severe pneumonia requiring hospital admission in children without HIV infection from Africa and Asia: the PERCH multi-country case-control study**. *Lancet*. 2019; **394**(10200): 757–779. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
43. Leclerc Q, Naylor NR, Aiken AM, *et al.*: **AR.IA App all datasets**. *figshare*. Dataset. 2019. <http://www.doi.org/10.6084/m9.figshare.9821996.v1>
44. Leclerc Q, Naylor NR, Aiken AM, *et al.*: **AR.IA App drug breakdown**. *figshare*. Dataset. 2019. <http://www.doi.org/10.6084/m9.figshare.9822077.v1>
45. Leclerc Q, Naylor NR, Aiken AM, *et al.*: **AR.IA App drug breakdown (groups)**. *figshare*. Dataset. 2019. <http://www.doi.org/10.6084/m9.figshare.9822104.v1>
46. Leclerc Q, Naylor NR, Aiken AM, *et al.*: **AR.IA App species breakdown**. *figshare*. Dataset. 2019. <http://www.doi.org/10.6084/m9.figshare.9822179.v1>
47. Leclerc Q, Naylor NR, Aiken AM, *et al.*: **AR.IA App all species**. *figshare*. Dataset. 2019. <http://www.doi.org/10.6084/m9.figshare.9822146.v1>
48. Leclerc Q, Naylor NR, Aiken AM, *et al.*: **AR.IA App economic data**. *figshare*. Dataset. 2019. <http://www.doi.org/10.6084/m9.figshare.9822122.v1>

Open Peer Review

Current Peer Review Status:   

Version 2

Reviewer Report 30 June 2020

<https://doi.org/10.21956/wellcomeopenres.17655.r39220>

© 2020 Shaw L. This is an open access peer review report distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.



Liam Shaw 

Nuffield Department of Clinical Medicine, University of Oxford, Oxford, UK

Thanks to the authors for their revisions and responses. I'm happy to approve the manuscript in its revised form.

Competing Interests: As stated in my first review: I entered the same competition as the authors (Wellcome AMR Data Reuse Prize) and received a runner-up award. I also did a short masters project supervised by Gwen Knight in 2014. I do not feel this has unduly biased my reviews.

Reviewer Expertise: My main research area is bacterial genetics, including antimicrobial resistance. I also have direct experience of analysing the ATLAS dataset used in this work.

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Reviewer Report 30 June 2020

<https://doi.org/10.21956/wellcomeopenres.17655.r39218>

© 2020 de Kraker M. This is an open access peer review report distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.



Marlieke de Kraker 

Infection Control Program, Faculty of Medicine, World Health Organization Collaborating Centre on Patient Safety (Infection Control and Improving Practices), University of Geneva Hospitals, Geneva, Switzerland

I thank the authors for the responses, and I approve the current version of the manuscript.

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: I am an epidemiologist working in the field of antimicrobial resistance, and I have ample experience with AMR surveillance data.

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Version 1

Reviewer Report 25 February 2020

<https://doi.org/10.21956/wellcomeopenres.16928.r36746>

© 2020 Cooper B. This is an open access peer review report distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.



Ben S. Cooper 

Nuffield Department of Medicine, University of Oxford, Oxford, UK

This feasibility study represents a valuable addition to the surprisingly small literature on how to make empiric antibiotic recommendations. It is clearly written and well-organised, and the framework presented describes a sensible approach to using existing data-sets. Limitations of the approach, most of which are related to current limitations in publicly available data-sets, are clearly acknowledged. Suggestions for improvements below are mostly relatively minor, though points 5 and 6 below seem important to address.

1. Page 4, paragraph 2. "Concordance" would be a better word than "accuracy here" as there is no gold standard surveillance system. "Accuracy" is also used elsewhere in the manuscript.
2. Figure 1. Can the caption explain what the direction of the arrows indicates?
3. Page 5 "...which is set to 15% by default". What is the reason for choosing 15%? Perhaps this point could be considered in the discussion. Surely it would make sense to choose different cut-offs for different syndromes as the trade-off between impact on patient outcomes and future selection for resistance will be different.
4. Page 5 "missing isolates". I don't think I understand why the term "missing isolates" is used. Were the isolates missing or were the records just incomplete?
5. Figure 3: I don't completely understand the logic of this. If there are no data on the first line therapy wouldn't it make sense to make no recommendation but to collect more data, rather than immediately going on and recommend the second line therapy.
6. Table 3: Some of the recommendations here seem incompatible with Figure 3. e.g the first

row for bacteraemia recommends using second line, but says no data on resistance to second line, while in the flow chart (Fig 3) the second line can only be recommended if resistance level is below the cut-off. Similar issues for septic arthritis. It would also be useful to have some indication of the strength of the evidence for the recommendation in the table.

7. Page 9, paragraph 2. Should make it clear that the cut-off of 15% can be changed in the app.

Is the work clearly and accurately presented and does it cite the current literature?

Yes

Is the study design appropriate and is the work technically sound?

Yes

Are sufficient details of methods and analysis provided to allow replication by others?

Yes

If applicable, is the statistical analysis and its interpretation appropriate?

Yes

Are all the source data underlying the results available to ensure full reproducibility?

Yes

Are the conclusions drawn adequately supported by the results?

Yes

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Communicable disease epidemiology & antimicrobial resistance.

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Author Response 19 Jun 2020

Quentin Leclerc, London School of Hygiene & Tropical Medicine, London, UK

This feasibility study represents a valuable addition to the surprisingly small literature on how to make empiric antibiotic recommendations. It is clearly written and well-organised, and the framework presented describes a sensible approach to using existing data-sets. Limitations of the approach, most of which are related to current limitations in publicly available data-sets, are clearly acknowledged. Suggestions for improvements below are mostly relatively minor, though points 5 and 6 below seem important to address.

Thank you for your comments - we are glad you found it a valuable addition. We have left

your comments in bold and italics below, while our answers are written in plain text.

- 1. Page 4, paragraph 2. "Concordance" would be a better word than "accuracy here" as there is no gold standard surveillance system. "Accuracy" is also used elsewhere in the manuscript.**

Thank you for this suggestion. In light of other reviewer comments, we have decided to replace "accuracy" by "agreement" throughout the manuscript. We hope that this is a satisfactory change.

- 1. Figure 1. Can the caption explain what the direction of the arrows indicates?**

We have now clarified in the caption that the arrows linking the boxes indicate the process of integrating the different open-access datasets together to generate our index to inform empiric therapy. In addition, following suggestions from other reviewers, we have modified the figure to include two types of arrows to improve readability.

- 1. Page 5 "...which is set to 15% by default". What is the reason for choosing 15%? Perhaps this point could be considered in the discussion. Surely it would make sense to choose different cut-offs for different syndromes as the trade-off between impact on patient outcomes and future selection for resistance will be different.**

Such a cut-off is always going to be a somewhat arbitrary choice – we used 15% as this value had recently been used in the clinical practice for a similar question for one of us. In other situations, we have found that other cut-off values have been used – for example 25% for treatment of pneumonia in one recommendation from USA (see reference in text). We agree that different cut-offs would probably be selected in different circumstances by different users for different syndromes – hence the value of having a slider for this parameter in the app. However, in the main text, for simplicity we chose to use just one value (15%).

- 1. Page 5 "missing isolates". I don't think I understand why the term "missing isolates" is used. Were the isolates missing or were the records just incomplete?**

Thank you for spotting this; we have corrected this phrase to indicate that the isolates were indeed not missing, but had incomplete records (MIC values were missing the corresponding susceptibility labels).

- 1. Figure 3: I don't completely understand the logic of this. If there are no data on the first line therapy wouldn't it make sense to make no recommendation but to collect more data, rather than immediately going on and recommend the second line therapy.**

Thank you for highlighting this – there was a mistake in this figure in that we do currently recommend first line therapy if there is no data. In the app recommendations, where relevant data is missing, we have:

- "Use second line, as resistance to first (but no data on resistance to second)"
- "Use third line, as resistance to both first and second (but no data on resistance to third)"
- "Use first line, but no data to inform - consider second or third if data"

Where there is resistant to the lower rank therapy, then we recommend use of the next level – "Use second line" in this app is purely because there is resistance to first line. Figure 3 has now been updated to better reflect the logic we use in our recommendations, and highlight that we assume bacteria are susceptible if there are no data to inform on potential resistance.

The aim of our work was to see what recommendations could be made at the empiric

prescribing level by deriving a syndrome-level index. Therefore, we are interested in the immediate decision a prescriber would make solely based on this index. You are correct when saying that ideally, more data would have to be collected before making any decision. However, we assume that we are in a situation where a patient requires immediate antibiotic treatment, so we report what the app recommends given this set of data. In line with comments from another reviewer we have included more discussion to highlight our focus on this secondary part of the empiric prescribing process.

- 1. Table 3: Some of the recommendations here seem incompatible with Figure 3. e.g the first row for bacteraemia recommends using second line, but says no data on resistance to second line, while in the flow chart (Fig 3) the second line can only be recommended if resistance level is below the cut-off. Similar issues for septic arthritis. It would also be useful to have some indication of the strength of the evidence for the recommendation in the table.**

Thank you for spotting this; as discussed in response to your comment 5), there was an error in Figure 3, which has now been corrected. As for an indication of the strength of the evidence, we have not mentioned this in Table 3 as it would vary between countries, but in the AR:IA App we now indicate the number of isolates used to generate the recommendation as a measure of strength of evidence.

- 1. Page 9, paragraph 2. Should make it clear that the cut-off of 15% can be changed in the app.**

Thank you for highlighting this. We did already mention that the cut-off could be changed in the app in the "Methods, The AR:IA App creation" section, however we have added another phrase in the paragraph you mentioned to hopefully make it clearer that we were referring to the 15% cut-off.

Competing Interests: No competing interests were disclosed.

Reviewer Report 05 February 2020

<https://doi.org/10.21956/wellcomeopenres.16928.r37548>

© 2020 de Kraker M. This is an open access peer review report distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.



Marlieke de Kraker

Infection Control Program, Faculty of Medicine, World Health Organization Collaborating Centre on Patient Safety (Infection Control and Improving Practices), University of Geneva Hospitals, Geneva, Switzerland

General comments

I would like to congratulate the authors with this original, comprehensive article. It describes the development of a freely available app to provide an evidence base for syndrome- and country-specific empirical prescribing. The resistance data is derived from four different datasets, including

one commercial dataset, ATLAS, and three public datasets, EARS-Net, ResistanceMap and GLASS. This research was carried out as part of the Wellcome Data Re-use Prize, promoting the use of pharmaceutical AMR data.

While it is an interesting tool, part of the objective of this study was to look at the applicability of ATLAS data. The authors mention that the resistance proportions reported by ATLAS are higher than those reported by the other datasets. I would have liked to know more about the ATLAS database in order to understand these differences and avoid misuse in the future. For example, the type of samples collected (colonization, infection, community, hospital), the sampling strategy (first isolate per species, per patient, all isolates or only those with susceptibility results for last-resort antibiotics, i.e. with confirmed resistance to first-line antibiotics), method of antimicrobial susceptibility testing (central, local, automated, applied breakpoints etc.) etc.

Based on the discrepancies they found between ATLAS and the public datasets, I think the app should have a warning indicating that the reported data are not necessarily representative, and mainly based on samples from the hospital setting. This should also be emphasized in the conclusion of the abstract and article.

The app could also have been further improved by indicating the number of isolates that the advice was based on per country so users can appreciate the validity of the data for their setting.

Detailed comments

Abstract

The introduction lacks a clear objective, which seems to be included in the Methods section "to derive .. syndromes". I suggest to move this to the introduction and make the objective PICO. The Methods section could explicitly mention the other datasets, and provide a bit more detail of what was done.

In the Results section, I would like to see some explicit results, like the number of isolates and countries that were available to inform the system. The specific syndromes that could be selected, etc.

In the conclusions, "rates of AMR" is used I think "AMR proportions" is more correct.

Introduction

In paragraph 4, it is described that "Providing .. guidelines based on open access international AMR surveillance data may help practitioners". I think what is meant is providing a tool, not guidelines to help inform empirical prescribing.

In the last paragraph it is mentioned that "We propose a more clinically-oriented presentation of resistance rates" and "we designed a syndrome-level composite resistance index". This suggests that this is a novel idea of the authors, but as described in the discussion such a kind of index has already been proposed by multiple others. If I understood correctly, the Weighted Incidence Syndromic Combination is exactly the same.¹ I suggest to move the paragraph about these indices from the discussion to the introduction, and make it clear that this is not a novel concept, but that this concept is now taken further by developing an easy-to-use tool informed by publicly available resistance data.

Methods

Surveillance data comparison

It is stated that "The ATLAS dataset .. contains high-quality AST data". However, it is not explained

why it is considered “high-quality”. Please elaborate.

Could you add an address where the ATLAS data can be found for interested researchers?

Please report about the sample selection procedure for all datasets, as this heavily influences resistance proportions. For the surveillance systems it is commonly the first isolate per sample type per patient per year, and as they focus on blood samples all are clinical samples, mainly from the hospital setting. It would be important to have the same information for the ATLAS dataset.

The type of variables reported to ATLAS would be informative as well. In the app you can select age groups, for example, but availability of this data has never been discussed in the paper.

Please add the type and year of publication of the breakpoints that were applied to MIC data with missing S/I/R interpretation.

In the fifth paragraph, “resistance rates” are often referred to, this should be “resistance proportions”.

Here it is also stated that “ ECDC and ResistanceMap .. derive from the same source data”. I think it is more correct to state that ResistanceMap includes all ECDC data, and uses other sources for AMR data from other continents. Moreover, the same can be explained for GLASS, all European sepsis data is derived from ECDC data.

“Accuracy” might not be the best term, perhaps “agreement” is more clear.

Step c) Why was additional clinical consultation needed, and how were the NICE guidelines and the clinical consultation integrated?

Step g) I think “syndrome coverage” can be explained a bit more clearly. Perhaps another term should be used, like “causative pathogen availability”.

Step g) If data is missing full susceptibility is assumed. This is a strong assumption, which would have warranted a sensitivity analysis for a best-case worst-case scenario, which could even have been incorporated in the app.

Results

In table 1, and supplementary figures 2 and 3, please report the number of countries included in each dataset.

In table 1, I would also like to see the total number of isolates per dataset, and the overlap in countries between datasets clearly displayed.

Figure 1. I think the file names can be removed. “Other datasets” could be replaced by the real names, including ATLAS.

Figure 3. The arrow from “Are there data on first-line therapy - No”, results in second line therapy being suggested (if data is available). This does not seem to be in line with what is discussed in step g) where missing data is assumed “fully susceptible” please align.

Figure 4. I would not give the same (grey) colour to countries that had no data, or had no syndrome for which first-line therapy was recommended. These are two, very distinctive categories.

Surveillance data comparison

The “accuracy” here is shockingly low, it would have been interesting to see whether certain specific bug-drug combinations were responsible for the difference in AMR proportions, and how large the differences in AMR proportions were. The proportion of overlap by itself is not very informative.

Step e)

How does it work if the datasets are combined? Are all isolates just aggregated without any weighting? And if you switch between only showing ATLAS data, or showing data from 2 or more datasets including ATLAS are ATLAS data then transformed to class resistance versus single antibiotic resistance when only ATLAS is selected? Please clarify.

How much overlap was there between the datasets in country coverage? It would have been nice to have maps indicating country participation for the different datasets.

While you mention that you could only provide recommendations for bacteremia, pneumonia and cellulitis, in the app and table 3, meningitis and arthritis are still included?

App

I think a line should be included explaining that the included AMR proportions are from hospital-associated infections. And a warning should be included that the data may not be representative of the national prevalence of AMR due to selective sampling (mainly tertiary care hospital data, selection of isolates with full AST data, i.e. the more resistant ones, etc.).

The tool could have been even more useful if uploading of local resistance data was made easy. The possibility to export country-specific data could be helpful too. If this is easy to add, I would like to ask the authors to add this functionality, if it is a lot of work this would not be required.

The cut-off for selecting first or second line antibiotics is binary which means that second line therapeutics could be advised even if resistance proportions only differ 1% point. Could it somehow be incorporated that only above a clinically relevant difference in resistance the advice would switch from first to second line antibiotics? As above, if this is easy to add, I would like to ask the authors to add this functionality, if it is a lot of work this would not be required. Then, I would like to see this remark added to the paper.

Discussion

In the limitations section, it should be discussed that these data may not be the most reliable source of AMR proportions to base empirical treatment on, especially since for certain countries the number of isolates included are very low and not very up-to-date (2017).

It should also be mentioned that the provided data focus on hospital-associated infections (and tertiary care centres) and do not apply to community-associated infections.

Finally, there should be a warning against the future use of ATLAS data without critical appraisal; there is very low “accuracy”, and too little information about the sampling scheme and its

representativeness. On certain websites (<https://www.amrindustryalliance.org/case-study/antimicrobial-testing-leadership-and-surveillance-atlas/>) they claim that "ATLAS is the only database of its kind capable of providing such a broad scope of reliable, readily available information in an easy to use platform" which is "crucial for clinicians and public health officials", which could have important implications.

Please also report about the future prospects of ATLAS data, and whether this database will remain available and/or will be updated with new data?

References

1. Randhawa V, Sarwar S, Walker S, Elligsen M, et al.: Weighted-incidence syndromic combination antibiograms to guide empiric treatment of critical care infections: a retrospective cohort study. *Crit Care*. 2014; **18** (3): R112 [PubMed Abstract](#) | [Publisher Full Text](#)

Is the work clearly and accurately presented and does it cite the current literature?

Partly

Is the study design appropriate and is the work technically sound?

Yes

Are sufficient details of methods and analysis provided to allow replication by others?

Yes

If applicable, is the statistical analysis and its interpretation appropriate?

Partly

Are all the source data underlying the results available to ensure full reproducibility?

Yes

Are the conclusions drawn adequately supported by the results?

Partly

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: I am an epidemiologist working in the field of antimicrobial resistance, and I have ample experience with AMR surveillance data.

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Author Response 19 Jun 2020

Quentin Leclerc, London School of Hygiene & Tropical Medicine, London, UK

I would like to congratulate the authors with this original, comprehensive article. It describes the development of a freely available app to provide an evidence base for syndrome- and country-specific empirical prescribing. The resistance data is derived from

four different datasets, including one commercial dataset, ATLAS, and three public datasets, EARS-Net, ResistanceMap and GLASS. This research was carried out as part of the Wellcome Data Re-use Prize, promoting the use of pharmaceutical AMR data.

Many thanks for your comments. We have left your comments in bold and italics below, while our answers are written in plain text.

While it is an interesting tool, part of the objective of this study was to look at the applicability of ATLAS data. The authors mention that the resistance proportions reported by ATLAS are higher than those reported by the other datasets. I would have liked to know more about the ATLAS database in order to understand these differences and avoid misuse in the future. For example, the type of samples collected (colonization, infection, community, hospital), the sampling strategy (first isolate per species, per patient, all isolates or only those with susceptibility results for last-resort antibiotics, i.e. with confirmed resistance to first-line antibiotics), method of antimicrobial susceptibility testing (central, local, automated, applied breakpoints etc.) etc. Based on the discrepancies they found between ATLAS and the public datasets, I think the app should have a warning indicating that the reported data are not necessarily representative, and mainly based on samples from the hospital setting. This should also be emphasized in the conclusion of the abstract and article.

We agree that we are relying on a database where the collection methods are unclear but believe it to be a convenience sample of isolates taken from hospital surveillance systems with a bias towards sampling from longer and more complex infections, that are proving harder to treat. We believe this to be the case for most AMR databases with isolate samples including the WHO GLASS surveillance. We need better surveillance with pro-active sampling including of those infections that are not complex and do respond to treatment to understand syndrome aetiology and to get a handle on the overall burden of AMR. We found it difficult to add more to our abstract and believe that highlighting the use of ATLAS there then allows readers to move to understanding further what ATLAS is within the paper. We have added further discussion in the discussion section to what we believe the samples in ATLAS and the wider AMR datasets represent.

The app could also have been further improved by indicating the number of isolates that the advice was based on per country so users can appreciate the validity of the data for their setting.

We thank the reviewer for this suggestion – we now report the number of isolates in the Therapy Recommendations tab in the app.

Detailed comments

Abstract

The introduction lacks a clear objective, which seems to be included in the Methods section “to derive .. syndromes”. I suggest to move this to the introduction and make the objective PICO.

The Methods section could explicitly mention the other datasets, and provide a bit more detail of what was done.

In the Results section, I would like to see some explicit results, like the number of isolates and countries that were available to inform the system. The specific syndromes that could be selected, etc.

Thank you for these suggestions to improve the Abstract. We have tried to incorporate as much of the requested information as possible, while respecting the 300 words limit.

In the conclusions, "rates of AMR" is used I think "AMR proportions" is more correct.
We have now made the recommended change from "rates" to "proportions".

Introduction

In paragraph 4, it is described that "Providing .. guidelines based on open access international AMR surveillance data may help practitioners". I think what is meant is providing a tool, not guidelines to help inform empirical prescribing.

Thank you for spotting this mistake, we have now corrected it since we indeed meant "tool" here.

In the last paragraph it is mentioned that "We propose a more clinically-oriented presentation of resistance rates" and "we designed a syndrome-level composite resistance index". This suggests that this is a novel idea of the authors, but as described in the discussion such a kind of index has already been proposed by multiple others. If I understood correctly, the Weighted Incidence Syndromic Combination is exactly the same.¹ I suggest to move the paragraph about these indices from the discussion to the introduction, and make it clear that this is not a novel concept, but that this concept is now taken further by developing an easy-to-use tool informed by publicly available resistance data.

Thank you for this suggestion. We have moved this discussion paragraph to the penultimate paragraph of the introduction. We have clarified in the last paragraph of the introduction that the novelty of our framework comes from using open-access datasets and presenting the results in an online tool.

Methods

Surveillance data comparison

It is stated that "The ATLAS dataset .. contains high-quality AST data". However, it is not explained why it is considered "high-quality". Please elaborate.

Many thanks for this comment, we have clarified this by changing the wording of this sentence.

Could you add an address where the ATLAS data can be found for interested researchers?

The link to the official ATLAS website (<https://atlas-surveillance.com>) is already provided as a hyperlink when clicking on "ATLAS dataset" at the beginning of the Methods section.

However, we have now added a second link (<https://amr.theodi.org/programmes/atlas>) to facilitate download of the ATLAS dataset. These links are also present in Table 4 in the "Data availability" section.

Please report about the sample selection procedure for all datasets, as this heavily influences resistance proportions. For the surveillance systems it is commonly the first isolate per sample type per patient per year, and as they focus on blood samples all are

clinical samples, mainly from the hospital setting. It would be important to have the same information for the ATLAS dataset.

The types of samples for each dataset are presented in Table 1. We have also added a sentence discussing this difference in sampling between AMR surveillance datasets in the second paragraph of the Discussion section.

The type of variables reported to ATLAS would be informative as well. In the app you can select age groups, for example, but availability of this data has never been discussed in the paper.

Thank you for pointing this out; we have now added two sentences highlighting that the ATLAS dataset also contains information on gender, age and specimen source in the second paragraph of the Methods section.

Please add the type and year of publication of the breakpoints that were applied to MIC data with missing S/I/R interpretation.

We did not apply external data on breakpoints to the MIC data with missing S/I/R interpretation. Rather, we chose to look for other isolates within the ATLAS dataset of the same bacterial species and MIC which had a susceptibility label, and applied the same label to the isolates which only had the raw MIC values (for details, please refer to our Further Methods in Extended Data).

To make this clearer to the reader, we have modified the last sentence of the 2nd paragraph of the Methods – Surveillance data comparison section to:

“Missing susceptibility labels (i.e. “resistant”, “intermediate” or “susceptible”) in the ATLAS dataset (443,899/633,820) were assigned from available MIC data for other isolates within the ATLAS dataset. We did not use any external data on breakpoints to derive susceptibility labels for MICs which were not labelled elsewhere within the ATLAS dataset (see Further Methods in Extended Data for details).”

In addition, we have added a sentence after the third paragraph of our Further Methods to elaborate on our decision to not use external breakpoints data.

In the fifth paragraph, “resistance rates” are often referred to, this should be “resistance proportions”.

Thank you for this suggestion, we have now replaced “resistance rates” by “resistance proportions” throughout the manuscript.

Here it is also stated that “ ECDC and ResistanceMap .. derive from the same source data”. I think it is more correct to state that ResistanceMap includes all ECDC data, and uses other sources for AMR data from other continents. Moreover, the same can be explained for GLASS, all European sepsis data is derived from ECDC data.

We have now added a sentence in paragraph 4 to state that ECDC data is included in ResistanceMap and GLASS.

“Accuracy” might not be the best term, perhaps “agreement” is more clear.

Thank you for this suggestion, we have now replaced “accuracy” by “agreement” throughout the manuscript.

Step c) Why was additional clinical consultation needed, and how were the NICE guidelines

and the clinical consultation integrated?

In this first iteration of the app, we needed to decide upon a simple set of empiric antibiotic guidelines with one choice for each infection syndrome, ignoring for now the complexities of pregnancy, age and allergies. For this we took the NICE guidelines as a starting point and then used clinical consultation to give the most commonly used antibiotic treatment. For this we relied on our clinical author's (AA) experience and overview of what happens in several UK Trusts. We have modified the paragraph to reflect this and hope in future iterations to allow for user input to avoid this reliance on one possible therapy.

Step g) I think "syndrome coverage" can be explained a bit more clearly. Perhaps another term should be used, like "causative pathogen availability".

Thank you for this suggestion, we have replaced "syndrome coverage" by "causative pathogen availability" throughout the manuscript.

Step g) If data is missing full susceptibility is assumed. This is a strong assumption, which would have warranted a sensitivity analysis for a best-case worst-case scenario, which could even have been incorporated in the app.

We agree that this is a strong assumption. The consequence of this reverse assumption would be to push therapies towards second or third line, or even to the recommendation "Consider alternatives! Resistance to all recommended therapies seen". To better highlight this, and make sure that the reader understands this assumption and the consequence of reversing it, we have added a paragraph on this topic in the section about limitations in the Discussion, and reworked Figure 4 to clearly show this assumption in the decision-making process.

Results***In table 1, and supplementary figures 2 and 3, please report the number of countries included in each dataset.***

We have now added this information in Table 1. However, we have not added this to Supplementary Figures 2 and 3 as we believe this would make the Figures harder to read (as the number of countries in each dataset each year can change, and we would therefore need to add a number for every year).

In table 1, I would also like to see the total number of isolates per dataset, and the overlap in countries between datasets clearly displayed.

We have now added a footnote in Table 1 to indicate that ResistanceMap and GLASS use ECDC data for their European blood samples, as well as the total sample size per dataset. Please note that this sample size corresponds to the total number of unique bacteria and drug measurements made, rather than number of isolates (i.e. number of lines) in the ATLAS dataset. This is the definition we also use when comparing sample size in our Further Results, to align with the measures provided in other surveillance datasets. We have clarified this in the Surveillance data comparison section in Methods.

Figure 1. I think the file names can be removed. "Other datasets" could be replaced by the real names, including ATLAS.

This has now been adapted in the Figure to match the suggestions given.

Figure 3. The arrow from “Are there data on first-line therapy - No”, results in second line therapy being suggested (if data is available). This does not seem to be in line with what is discussed in step g) where missing data is assumed “fully susceptible” please align.

Thank you for spotting this mistake. Figure 3 has now been updated to better reflect the logic used the app, and the fact that we assume missing data is “fully susceptible”.

Figure 4. I would not give the same (grey) colour to countries that had no data, or had no syndrome for which first-line therapy was recommended. These are two, very distinctive categories.

Apologies, this was an error in the figure and accompanying legend. Grey now only represents countries that had no data. The colour scheme has been updated to better align with other figures, countries with no syndrome for which first-line therapy was recommended would now be in dark red.

Surveillance data comparison

The “accuracy” here is shockingly low, it would have been interesting to see whether certain specific bug-drug combinations were responsible for the difference in AMR proportions, and how large the differences in AMR proportions were. The proportion of overlap by itself is not very informative.

We have now also changed the phrasing of accuracy to agreement in line with previous reviewer suggestions. We agree that such a question would be worth investigating, however we feel that such an in-depth comparison would stand as an entirely different piece of work from our own analysis here, beyond the scope of this paper.

Step e)

How does it work if the datasets are combined? Are all isolates just aggregated without any weighting? And if you switch between only showing ATLAS data, or showing data from 2 or more datasets including ATLAS are ATLAS data then transformed to class resistance versus single antibiotic resistance when only ATLAS is selected? Please clarify.

We have indicated in step e) that we aggregated data across multiple datasets “by mapping antibiotics in the ATLAS dataset to their corresponding antibiotic classes (reported in the rest of datasets)”. We have clarified that we are mapping “single” antibiotics. When combining the datasets, we average the resistance rates reported in each dataset for each combination of country, bacteria and antibiotic class; we have now clarified this in the main text to incorporate this comment’s suggestion.

How much overlap was there between the datasets in country coverage? It would have been nice to have maps indicating country participation for the different datasets.

Thank you for this suggestion. We have decided not to include a new Figure, as we feel that this would lengthen the paper without directly relating to our research question. We now clearly state that there is geographical overlap in Table 1 by saying that ECDC data underlies the European data in ResistanceMap and GLASS, and we already stated the general geographical coverage of the datasets in that Table (Europe or Global).

While you mention that you could only provide recommendations for bacteremia, pneumonia and cellulitis, in the app and table 3, meningitis and arthritis are still included?

The only recommendations we can provide for meningitis and arthritis come with the

caveat that there is actually no data to inform bacterial susceptibility, and therefore rely on our assumption to consider lack of data equivalent to full susceptibility. This is stated in Table 3. As a consequence, we consider that we cannot give robust recommendations for these two syndromes. We have added a couple of sentences in the “Resulting recommendations on the appropriateness of empiric antibiotic therapies” section of the Results to highlight this.

App

I think a line should be included explaining that the included AMR proportions are from hospital-associated infections. And a warning should be included that the data may not be representative of the national prevalence of AMR due to selective sampling (mainly tertiary care hospital data, selection of isolates with full AST data, i.e. the more resistant ones, etc.).

We have now added the following disclaimer line in the app: “Resistance prevalence reported here may not be representative of the national prevalence. This tool is for research use only. Do not use for clinical decision making.”

The tool could have been even more useful if uploading of local resistance data was made easy. The possibility to export country-specific data could be helpful too. If this is easy to add, I would like to ask the authors to add this functionality, if it is a lot of work this would not be required.

We agree that it would be a nice feature to add to the app. Unfortunately, this would be complicated to add in practice, therefore we have not considered it at this stage, but something that we are looking into for future work.

The cut-off for selecting first or second line antibiotics is binary which means that second line therapeutics could be advised even if resistance proportions only differ 1% point. Could it somehow be incorporated that only above a clinically relevant difference in resistance the advice would switch from first to second line antibiotics? As above, if this is easy to add, I would like to ask the authors to add this functionality, if it is a lot of work this would not be required. Then, I would like to see this remark added to the paper.

We agree that including subtleties around the levels of resistance and how this affects treatment choice would be great additions to this app, but we have decided not to include them in this iteration due to the level of work involved and the small likelihood that this happens frequently. As the user can alter the 15% cut-off this can be used to explore sensitivity and resistance levels. We have added a new section into the discussion about future iterations of the app and how this could be included.

Discussion

In the limitations section, it should be discussed that these data may not be the most reliable source of AMR proportions to base empirical treatment on, especially since for certain countries the number of isolates included are very low and not very up-to-date (2017).

We now mention this in the fifth paragraph of the discussion.

It should also be mentioned that the provided data focus on hospital-associated infections (and tertiary care centres) and do not apply to community-associated infections.

We now mention this in the fifth paragraph of the discussion.

Finally, there should be a warning against the future use of ATLAS data without critical appraisal; there is very low “accuracy”, and too little information about the sampling scheme and its representativeness. On certain websites (<https://www.amrindustryalliance.org/case-study/antimicrobial-testing-leadership-and-surveillance-atlas/>) they claim that “ATLAS is the only database of its kind capable of providing such a broad scope of reliable, readily available information in an easy to use platform” which is “crucial for clinicians and public health officials”, which could have important implications.

We now mention some cautionary advice regarding the ATLAS database in the second paragraph of the discussion and in the conclusion. Our primary purpose in this study wasn't to appraise the ATLAS database, but to see how it could be used.

Please also report about the future prospects of ATLAS data, and whether this database will remain available and/or will be updated with new data?

We have now added this information in the second paragraph of the discussion. To the best of our knowledge, the ATLAS database will remain available and be updated every 6 to 8 months.

Competing Interests: No competing interests were disclosed.

Reviewer Report 23 October 2019

<https://doi.org/10.21956/wellcomeopenres.16928.r36747>

© 2019 Shaw L. This is an open access peer review report distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

? **Liam Shaw** 

Nuffield Department of Clinical Medicine, University of Oxford, Oxford, UK

This is an important proof-of-concept study. The aim is to provide a "clinically-oriented presentation of resistance rates" from available AMR surveillance datasets. The main conclusion is that while this is technically possible for some syndromes, the datasets are not yet adequate.

I enjoyed reading the manuscript and anticipate that it will be read with interest by other researchers interested in this general approach. Thanks to the authors for carrying out this work!

Positive points:

- An impressive integration with existing datasets beyond the central dataset (ATLAS).
- A range of sources of data have been synthesised – not just resistance rates but also information on syndromes, guidelines, etc.

- Clear presentation of methodology and results.
- Online interactive app so readers can experiment themselves and get a feel for the approach.
- Reproducibility: I could successfully clone the github repository and run the app myself in under a minute.

I have selected "Approved with reservations" as I do have some suggestions for improvement, which I hope the authors will consider.

Major points

1. Infection syndromes:

I think a strict definition of what is meant by "infection syndrome" should be given, as readers may not be that familiar with the term (I admit I wasn't totally clear what was meant without looking at the examples). With the caveat that I am not a clinician, I think it is underplayed that the infection syndromes looked at are often not known to be due to bacteria. For example, upper respiratory tract infections – the majority are viral (the Further Methods state 90%). And for bacteraemia, patients can present only with fever, at which point a decision has to be made about treatment. It seems to me it would be less common to have a scenario where a patient has a known bacterial infection syndrome but nothing is known beyond that. This may be my own ignorance; I'd welcome any supporting evidence about how often this is the case in practice.

I know the authors are aware of this problem, but I think it should be stated explicitly that the problem considered here is: "an ideal world where one was confident that an infection was bacterial but did not know the species or the resistance profile" (I think that's a reasonable formulation?). The effect of including the proportion expected to be non-bacterial and therefore unaffected by antibiotics would be to reduce the likelihood of prescribing antibiotics. The high proportion of recommendations to treat with antibiotics beyond first-line antibiotics is driven by this exclusively bacterial assumption (and the resistance datasets used). This should be emphasised.

To be clear, I'm not suggesting redoing all this work taking into account the proportion of patients presenting with symptoms X/Y/Z who have a bacterial vs. viral infection. The work here stands in its own right. However, the major problem in antibiotic prescribing is the decision to prescribe antibiotics *at all* rather than deciding whether to prescribe first/second/third-line antibiotics. I think this needs to be mentioned up front; at present it comes very late as the fourth limitation in the discussion. In practice this is what most readers will think of when thinking about antibiotic stewardship and how prescribing apps might help.

2. Accuracy

I appreciate this is always put in quotes to emphasise the restricted sense of "accurate" that is meant, but I still think this is the wrong term. I suggest changing all uses of it and avoiding the word entirely. "Accurate" suggests that the ground truth or right answer can be derived from the ResistanceMap or ECDC dataset, and that because the ATLAS point

estimate lies outside the 95% confidence interval from these datasets it is “inaccurate”, which I think is misleading. The right answer for the ‘true’ prevalence of resistance isn’t known in this circumstance, and in fact resistance prevalence is context-dependent on the sampling (community vs. hospital etc.). I’d suspect variation in sampling is the driver, as suggested, and also that the ATLAS dataset for some countries might include only a few sites. What is being defined is not accuracy, but agreement between datasets. So, “concordance” or “agreement” would be a better term.

I would also be interested to see a plot using data presented in Supplementary Figure 1, but adding the agreement against the number of samples that went into producing the ATLAS point estimate. If the problem driving low agreement is not enough samples in the ATLAS dataset, then more samples would produce more agreement with ECDC. If the problem is differences in sampling (e.g. bias towards resistant organisms in ATLAS due to specific clinical setting) then more samples in ATLAS would produce *less* agreement with ECDC. It might not be as clear-cut as I’m assuming, but at the moment this isn’t clear. It seems like an important distinction to shed light on.

3. Comparison with other datasets

I agree that ECDC data is merged into ResistanceMap and the resistance rates for those countries are therefore the same. Therefore, I’m a bit confused why a comparison is made of ATLAS vs. ECDC *and* ATLAS vs. ResistanceMap without apparently removing ECDC as far as I could tell? Fundamentally the ECDC data is inside ResistanceMap. It seems to me that the important comparisons would be ATLAS vs. ECDC and ATLAS vs. (ResistanceMap – ECDC). I think for this reason there’s currently double-counting going on in Supplementary Figure 1, 2, 3 which could be rectified. If I misunderstood and this has already been removed, apologies!

Minor points

I have ordered these by section.

Introduction

- Para 1: First sentence, suggest replacing “being known” by “determined by tests” (as highlights that this can only be really known by some sort of test/procedure outside normal clinical assessment, and also that sometimes tests are inconclusive).
- Para 2: “infection syndrome” --> “infection syndromes”.
- Para 3: suggest providing example(s) of a bacterial infection syndrome where the aetiology varies by season.
- Para 4: think the usual phrase is “resource-limited” rather than “resource constrained” (I could be wrong)
- Para 5: “which have limited healthcare budgets and difficulties to access medicines” --> “which can have limited healthcare budgets and access to medicines”
- Para 6: “infection syndromes” need to be defined earlier. Suggest in Para 2.

Methods

- Para 2: “This dataset was made public in 2017” – I accept this was when the ATLAS website went live, but the dataset was made browsable, not downloadable. Worth explaining that the whole dataset as csv was actually released in 2018 to participants in Data Reuse Prize.
- Para 5, Surveillance data comparison: see major point about changing “accuracy” (and change elsewhere as well).
- (a) “Common infection syndromes”: I’m not sure if all these syndromes are identifiable with simple clinical examination and/or radiology, so just want to check this point i.e. what’s meant by “clinical examination”. As I mentioned, as far as I’m aware patients can present only with fever for bacteraemia, so I’m not sure how it can be easily identified. Also, the syndromes are assumed to be independent of each other i.e. patients present with only one. This is a limitation as in practice they can be linked e.g. bacterial meningitis develops from initial bacteraemia. Or e.g. NICE guidelines highlight link between hospital-acquired pneumonia and bacteraemia.
- “NICE guideline” --> “NICE guidelines”. Also, the link to the NICE guidelines (reference 8) is wrong. It takes me to PHE’s 119-page document “Summary of antimicrobial prescribing guidance – managing common infections”. I think the correct link should be: <https://www.nice.org.uk/Media/Default/About/what-we-do/NICE-guidance/antimicrobial%20guidance/summary-antimicrobial-prescribing-guidance.pdf> Is this correct? Assuming these are the correct guidelines, I became slightly confused about how the syndromes were identified. For example, upper respiratory tract infection as defined in Table 2 doesn’t include “acute otitis media”, which is included in this syndrome in the NICE guidelines. The first-choice treatment for acute otitis media is amoxicillin, rather than penicillin V. I have not investigated all the other syndromes but I just want to check why this was excluded? Apologies if I missed this somewhere.
- On this point, I couldn’t find the guidelines for bacteraemia in that NICE link. “Bacteraemia” is the presence of bacteria in the blood; “sepsis” is probably a better term for the infection syndrome. I noticed that the app uses “sepsis”, but the main paper uses “bacteraemia”, suggest to make them consistent.
- (d) “Contributing pathogen distribution: syndrome aetiology”: on “rapid literature search”, it would be good to know how this review worked in practice e.g. searching PubMed / Google Scholar / the wider web? One author assigned to a syndrome? One author doing all syndromes? How were results integrated? There’s no problem at all with it being rapid and informal, but a little more info on the process would be great.

Results

- Table 3: the summary table of the recommendations for therapy is based on the arbitrary 15% cut-off. This should be stated in the caption of the table. I think it would be good to present the actual distributions in a supplementary figure (i.e. the percentages for the index), but this isn’t essential. I would also state that only those syndromes with an average above 50% were included (which I infer from the associated text). Column stating “Average syndrome coverage” – presume “Mean syndrome coverage” is meant instead.

- (g) "Resulting recommendations on the appropriateness of empiric antibiotic therapies"
Para 3: bacterial meningitis has an "average syndrome coverage" of 0.51 in Table 3, so why does a sentence here state it has lower than 50%? Also, suggest stating the syndrome coverages for these syndromes that are not listed in Table 3.
- Figures: in general, please check the resolution of the figures. For example, Figure 3 seems quite low-res to me. For Figure 1: the arrows on the left-hand side aren't clear, I got a bit confused with the arrows pointing to other arrows. Suggest using a colour key for these arrows instead.
- Table 2: "Infectious syndrome" --> "Infection syndrome" (first column)

Discussion

- Para 3: "over-representation of antibiotic-resistant isolates in the ATLAS datasets" – see previous comment about a plot of ATLAS sample size against agreement with ECDC which might provide some more support for the hypothesis causing this.
- Para 5: "antimicrobial resistance" --> AMR

Conclusion

- "the WHO calls for" --> "the WHO's call for". Also provide a reference for this call for "evidence-based prescribing". I wasn't sure what was being referred to.

Further methods

- Section 1: no new MIC thresholds were included, which is acknowledged as a limitation. These would not in principle be difficult to include from EUCAST and could bump up the dataset. For example, *E. coli* and ceftaroline-avibactam: there are 13,203 isolates with a reported MIC. The EUCAST breakpoint is 0.5 mg/L. So using my version of the ATLAS dataset, 139/13,202 isolates (1.05%) are resistant. I accept compiling bug/drug combinations for breakpoints (and considering both EUCAST and CLSI) would be tedious, and so much data integration has already been done! I'd be happy to approve the paper without this extra work. But it would be good to have some idea of how much doing this would increase sample sizes, which could be done without external data integration.
- Section 3: This paragraph is quite dense and a brief overview in the main paper would be useful, because the index is a key concept. In a talk based on this work by Francesc Coll at a SEDRIC meeting on October 9th 2019, I saw a great diagram, I think of the example given here. I would strongly suggest including this as a figure in the main paper, as it quickly conveys the idea of the index visually.

Further results

- Section 3: Upper respiratory tract infection has a reference (49) which is missing from the references at the end of this document.

Is the work clearly and accurately presented and does it cite the current literature?

Partly

Is the study design appropriate and is the work technically sound?

Yes

Are sufficient details of methods and analysis provided to allow replication by others?

Yes

If applicable, is the statistical analysis and its interpretation appropriate?

Yes

Are all the source data underlying the results available to ensure full reproducibility?

Yes

Are the conclusions drawn adequately supported by the results?

Yes

Competing Interests: I entered the same competition as the authors (Wellcome AMR Data Reuse Prize) and received a runner-up award. I also did a short masters project supervised by Gwen Knight in 2014. I do not feel this has unduly biased my review.

Reviewer Expertise: My main research area is bacterial genetics, including antimicrobial resistance. I also have direct experience of analysing the ATLAS dataset used in this work.

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Author Response 19 Jun 2020

Quentin Leclerc, London School of Hygiene & Tropical Medicine, London, UK

This is an important proof-of-concept study. The aim is to provide a "clinically-oriented presentation of resistance rates" from available AMR surveillance datasets. The main conclusion is that while this is technically possible for some syndromes, the datasets are not yet adequate.

I enjoyed reading the manuscript and anticipate that it will be read with interest by other researchers interested in this general approach. Thanks to the authors for carrying out this work!

Thank you for your comments and review, much appreciated. We have left your comments in bold and italics below, while our answers are written in plain text.

Positive points:

- ***An impressive integration with existing datasets beyond the central dataset (ATLAS).***

- ***A range of sources of data have been synthesised – not just resistance rates but also information on syndromes, guidelines, etc.***
- ***Clear presentation of methodology and results.***
- ***Online interactive app so readers can experiment themselves and get a feel for the approach.***
- ***Reproducibility: I could successfully clone the github repository and run the app myself in under a minute.***

I have selected "Approved with reservations" as I do have some suggestions for improvement, which I hope the authors will consider.

Major points

1. ***Infection syndromes: I think a strict definition of what is meant by "infection syndrome" should be given, as readers may not be that familiar with the term (I admit I wasn't totally clear what was meant without looking at the examples). With the caveat that I am not a clinician, I think it is underplayed that the infection syndromes looked at are often not known to be due to bacteria. For example, upper respiratory tract infections – the majority are viral (the Further Methods state 90%). And for bacteraemia, patients can present only with fever, at which point a decision has to be made about treatment. It seems to me it would be less common to have a scenario where a patient has a known bacterial infection syndrome but nothing is known beyond that. This may be my own ignorance; I'd welcome any supporting evidence about how often this is the case in practice. I know the authors are aware of this problem, but I think it should be stated explicitly that the problem considered here is: "an ideal world where one was confident that an infection was bacterial but did not know the species or the resistance profile" (I think that's a reasonable formulation?). The effect of including the proportion expected to be non-bacterial and therefore unaffected by antibiotics would be to reduce the likelihood of prescribing antibiotics. The high proportion of recommendations to treat with antibiotics beyond first-line antibiotics is driven by this exclusively bacterial assumption (and the resistance datasets used). This should be emphasised. To be clear, I'm not suggesting redoing all this work taking into account the proportion of patients presenting with symptoms X/Y/Z who have a bacterial vs. viral infection. The work here stands in its own right. However, the major problem in antibiotic prescribing is the decision to prescribe antibiotics *at all* rather than deciding whether to prescribe first/second/third-line antibiotics. I think this needs to be mentioned up front; at present it comes very late as the fourth limitation in the discussion. In practice this is what most readers will think of when thinking about antibiotic stewardship and how prescribing apps might help.***

Thank you for this comment. We have now added a clearer definition of "infection syndrome" in the second paragraph of the Introduction. In addition, we have added a statement in the last paragraph of the Introduction to clearly mention that we are assuming that we are in a scenario where an antibiotic has to be prescribed, and we therefore aim to answer the question of which antibiotic should be prescribed.

1. ***Accuracy I appreciate this is always put in quotes to emphasise the restricted sense of "accurate" that is meant, but I still think this is the wrong term. I suggest***

changing all uses of it and avoiding the word entirely. "Accurate" suggests that the ground truth or right answer can be derived from the ResistanceMap or ECDC dataset, and that because the ATLAS point estimate lies outside the 95% confidence interval from these datasets it is "inaccurate", which I think is misleading. The right answer for the 'true' prevalence of resistance isn't known in this circumstance, and in fact resistance prevalence is context-dependent on the sampling (community vs. hospital etc.). I'd suspect variation in sampling is the driver, as suggested, and also that the ATLAS dataset for some countries might include only a few sites. What is being defined is not accuracy, but agreement between datasets. So, "concordance" or "agreement" would be a better term.

Thank you for this suggestion to improve clarity; we have now replaced "accuracy" by "agreement" throughout the manuscript.

I would also be interested to see a plot using data presented in Supplementary Figure 1, but adding the agreement against the number of samples that went into producing the ATLAS point estimate. If the problem driving low agreement is not enough samples in the ATLAS dataset, then more samples would produce more agreement with ECDC. If the problem is differences in sampling (e.g. bias towards resistant organisms in ATLAS due to specific clinical setting) then more samples in ATLAS would produce less agreement with ECDC. It might not be as clear-cut as I'm assuming, but at the moment this isn't clear. It seems like an important distinction to shed light on.

Thank you for this suggestion. We have now divided Supplementary Figure 1 into two sections; A) is the old Supplementary Figure 1, and B) is a new plot showing agreement against number of ATLAS samples used to calculate that agreement. One sample is one isolate, for one bacteria, one antibiotic, and one country. This new plot suggests that more ATLAS samples produce more agreement with ECDC and ResistanceMap, however this does not seem clear-cut as the increase in agreement is arguably not significant past 30,000 samples. Therefore, this suggests that while increasing sample size can increase agreement, there is still a difference between the datasets, which could be due to differences in sampling. We now discuss this in the Extended Data – Further Results section, and have added a sentence on the difference in sampling between the datasets in the discussion section in the main text.

- 1. Comparison with other datasets I agree that ECDC data is merged into ResistanceMap and the resistance rates for those countries are therefore the same. Therefore, I'm a bit confused why a comparison is made of ATLAS vs. ECDC and ATLAS vs. ResistanceMap without apparently removing ECDC as far as I could tell? Fundamentally the ECDC data is inside ResistanceMap. It seems to me that the important comparisons would be ATLAS vs. ECDC and ATLAS vs. (ResistanceMap – ECDC). I think for this reason there's currently double-counting going on in Supplementary Figure 1, 2, 3 which could be rectified. If I misunderstood and this has already been removed, apologies!**

Thank you for this suggestion. Our aim behind this comparison was to see how well the ATLAS dataset agreed with other AMR surveillance datasets, therefore we made all the comparisons against the full datasets (i.e. without removing the ECDC data from ResistanceMap). We believe this is the valid comparison to make as this will be the one that

others are doing too: how do results from ATLAS analysis compare to results from ResistanceMap analysis rather than subsetting the latter database and comparing. We agree that a more in-depth comparison between the AMR surveillance datasets would be interesting, however this would require a much larger analysis to be correctly examined. We therefore think that this is outside the scope of this paper as it would represent a different research question.

Minor points

I have ordered these by section.

Introduction

- ***Para 1: First sentence, suggest replacing “being known” by “determined by tests” (as highlights that this can only be really known by some sort of test/procedure outside normal clinical assessment, and also that sometimes tests are inconclusive).***

Thank you for this suggestion, we have made the recommended change to “determined by diagnostic tests”.

- ***Para 2: “infection syndrome” --> “infection syndromes”.***

Thank you for spotting this mistake, we have now corrected it.

- ***Para 3: suggest providing example(s) of a bacterial infection syndrome where the aetiology varies by season.***

We have now added the example of pneumonia as a bacterial infection syndrome where aetiology varies by season.

- ***Para 4: think the usual phrase is “resource-limited” rather than “resource constrained” (I could be wrong)***

We have made the recommended change to “resource-limited”.

- ***Para 5: “which have limited healthcare budgets and difficulties to access medicines” --> “which can have limited healthcare budgets and access to medicines”***

We have rephrased this section as recommended.

- ***Para 6: “infection syndromes” need to be defined earlier. Suggest in Para 2.***

We have now added a definition of “infection syndromes” in paragraph 2 of the Introduction.

Methods

- ***Para 2: “This dataset was made public in 2017” – I accept this was when the ATLAS website went live, but the dataset was made browsable, not downloadable. Worth explaining that the whole dataset as csv was actually released in 2018 to participants in Data Reuse Prize.***

Thank you for this clarification, we have added a sentence in paragraph 2 of the Methods to explain this.

- ***Para 5, Surveillance data comparison: see major point about changing “accuracy” (and change elsewhere as well).***

“accuracy” has now been replaced by “agreement” throughout the manuscript.

- ***(a) “Common infection syndromes”: I’m not sure if all these syndromes are identifiable with simple clinical examination and/or radiology, so just want to check***

this point i.e. what's meant by "clinical examination". As I mentioned, as far as I'm aware patients can present only with fever for bacteraemia, so I'm not sure how it can be easily identified. Also, the syndromes are assumed to be independent of each other i.e. patients present with only one. This is a limitation as in practice they can be linked e.g. bacterial meningitis develops from initial bacteraemia. Or e.g. NICE guidelines highlight link between hospital-acquired pneumonia and bacteraemia.

Infection syndromes cover a wide range of possibilities and you are right that bacteraemia is notoriously difficult to pin down characteristic features for – fever is really the only reliably present element. Clinical examination essentially encompasses all the elements of medical examination. This would start with taking temperature and performing basic observations through to examining different parts of the body to look for localized evidence of infection. Few of the infection syndromes have entirely reliable identification in practice. You are right that bacteraemia often co-exists with some of the other syndromes, but apart from that, overlap is relatively rare.

We have included further discussion of this in the discussion (paragraph 3).

- ***"NICE guideline" --> "NICE guidelines". Also, the link to the NICE guidelines (reference 8) is wrong. It takes me to PHE's 119-page document "Summary of antimicrobial prescribing guidance – managing common infections". I think the correct link should be: <https://www.nice.org.uk/Media/Default/About/what-we-do/NICE-guidance/antimicrobial%20guidance/summary-antimicrobial-prescribing-guidance.pdf> Is this correct? Assuming these are the correct guidelines, I became slightly confused about how the syndromes were identified. For example, upper respiratory tract infection as defined in Table 2 doesn't include "acute otitis media", which is included in this syndrome in the NICE guidelines. The first-choice treatment for acute otitis media is amoxicillin, rather than penicillin V. I have not investigated all the other syndromes but I just want to check why this was excluded? Apologies if I missed this somewhere.***

Thank you for the corrected link. For "acute otitis media" (AOM), even though this infection is frequently caused by bacteria, there are typically no bacterial samples sent for patients. This is because AOM occurs on the inside of eardrum, hence it is hard to access unless the eardrum is perforated. We confined our work to infection syndromes where there were easily corresponding bacterial isolates. To make recommendations about AOM, one would normally draw inferences from bacteria isolates from other body sites, which we have avoided doing here.

Broadly, we agree with the reviewer that this is a limitation of our analysis. We aimed to provide a tool that could be used at the point of prescribing antibiotics to treat an "infection syndrome". We therefore make no clear statements about whether these are independent or upon what level of evidence the decision is made. We have added this as a limitation into our discussion section (paragraph 3).

- ***On this point, I couldn't find the guidelines for bacteraemia in that NICE link. "Bacteraemia" is the presence of bacteria in the blood; "sepsis" is probably a better term for the infection syndrome. I noticed that the app uses "sepsis", but the main paper uses "bacteraemia", suggest to make them consistent.***

You are correct that there is no mention of bacteraemia in the NICE guidelines – we felt that we should add this as a recommendation category. The terms sepsis and bacteraemia are often used interchangeably, though they refer to different phenomena. For consistency, we will use the term "bloodstream infection" as the standard description for both the app and

the paper.

- ***(d) “Contributing pathogen distribution: syndrome aetiology”: on “rapid literature search”, it would be good to know how this review worked in practice e.g. searching PubMed / Google Scholar / the wider web? One author assigned to a syndrome? One author doing all syndromes? How were results integrated? There’s no problem at all with it being rapid and informal, but a little more info on the process would be great.***

Thank you for this suggestion. We have now added this information in the main text.

Results

- ***Table 3: the summary table of the recommendations for therapy is based on the arbitrary 15% cut-off. This should be stated in the caption of the table. I think it would be good to present the actual distributions in a supplementary figure (i.e. the percentages for the index), but this isn’t essential. I would also state that only those syndromes with an average above 50% were included (which I infer from the associated text). Column stating “Average syndrome coverage” – presume “Mean syndrome coverage” is meant instead.***

We now mention the 15% cut-off in the caption, and have corrected the column to indicate “Mean syndrome coverage”. In the Extended Data (Further Methods) section we acknowledge that the 15% cut-off is rather conservative compared to the 25% resistance rate recommended by European guidelines for managing hospital-acquired and ventilator-acquired pneumonia (1[QL1]). Hence why we allow the user to vary this in the app.

- ***(g) “Resulting recommendations on the appropriateness of empiric antibiotic therapies” Para 3: bacterial meningitis has an “average syndrome coverage” of 0.51 in Table 3, so why does a sentence here state it has lower than 50%? Also, suggest stating the syndrome coverages for these syndromes that are not listed in Table 3.***

Thank you for spotting this mistake, bacterial meningitis should not have been in this sentence. We have now corrected that, and give the syndrome coverage values in the text for the two syndromes with a coverage of less than 50%.

- ***Figures: in general, please check the resolution of the figures. For example, Figure 3 seems quite low-res to me. For Figure 1: the arrows on the left-hand side aren’t clear, I got a bit confused with the arrows pointing to other arrows. Suggest using a colour key for these arrows instead.***

Thank you for pointing this out, we have now edited all Figures to be of higher resolution, and have added two types of arrows in Figure 1 to hopefully make this figure clearer.

- ***Table 2: “Infectious syndrome” --> “Infection syndrome” (first column)***

We have now corrected this.

Discussion

- ***Para 3: “over-representation of antibiotic-resistant isolates in the ATLAS datasets” – see previous comment about a plot of ATLAS sample size against agreement with ECDC which might provide some more support for the hypothesis causing this.***

Thanks again for this comment – we have now added an extra sentence here to highlight the differences in sampling between the datasets.

- ***Para 5: “antimicrobial resistance” --> AMR***

We have corrected this to the recommended abbreviation.

Conclusion

- ***“the WHO calls for” --> “the WHO’s call for”. Also provide a reference for this call for “evidence-based prescribing”. I wasn’t sure what was being referred to.***

Thank you for highlighting this, we have corrected the sentence and added a reference for this call.

Further methods

- ***Section 1: no new MIC thresholds were included, which is acknowledged as a limitation. These would not in principle be difficult to include from EUCAST and could bump up the dataset. For example, E. coli and ceftaroline-avibactam: there are 13,203 isolates with a reported MIC. The EUCAST breakpoint is 0.5 mg/L. So using my version of the ATLAS dataset, 139/13,202 isolates (1.05%) are resistant. I accept compiling bug/drug combinations for breakpoints (and considering both EUCAST and CLSI) would be tedious, and so much data integration has already been done! I’d be happy to approve the paper without this extra work. But it would be good to have some idea of how much doing this would increase sample sizes, which could be done without external data integration.***

Thank you for raising this interesting point. We did not consider this further data integration when completing this work, however you are undeniably correct when saying that it would increase the number of useful data points in the ATLAS dataset. The issue would come with deciding which guideline to follow and to split our analysis further by EUCAST and CLSI for example would indeed be a large amount of extra work. We did do some of this work by using existing breakpoints in the data for isolates without a classification but an MIC value.

We have now added a few sentences at the end of this section to clarify that we did not do this, but that it could potentially be applied to increase the value of the dataset. We also would note that many of the antibiotics where the additional data would be added (eg ceftaroline-avibactam, colistin etc) are very “high-level” antibiotics that would not feature in empirical guidelines.

- ***Section 3: This paragraph is quite dense and a brief overview in the main paper would be useful, because the index is a key concept. In a talk based on this work by Francesc Coll at a SEDRIC meeting on October 9th 2019, I saw a great diagram, I think of the example given here. I would strongly suggest including this as a figure in the main paper, as it quickly conveys the idea of the index visually.***

We thank you for this suggestion and have now included a further figure (Figure 3) based on the slide presented.

Further results

- ***Section 3: Upper respiratory tract infection has a reference (49) which is missing from the references at the end of this document.***

Thank you for spotting this mistake. We have corrected this to indicate the correct reference number (17), and have added the reference in the bibliography.

[QL1]A. Torres *et al.*, International ERS/ESICM/ESCMID/ALAT guidelines for the management of hospital-acquired pneumonia and ventilator-associated pneumonia. *European Respiratory Journal*. **50**, 1700582 (2017).

Competing Interests: No competing interests were disclosed.

