Filling the gaps in the global prevalence map of clinical antimicrobial resistance

Rik Oldenkamp[a,b,1] , Constance Schultsz[a,b] , Emiliano Mancini[a,b,c,d] , and Antonio Cappuccio[a,b,d,e,1]

[a]Amsterdam Institute for Global Health and Development, 1105 BP Amsterdam, The Netherlands; [b]Department of Global Health, Amsterdam University Medical Centres, Location AMC, 1105 AZ Amsterdam, The Netherlands; [c]Institute of Life Sciences and Technologies, Daugavpils University, LV-5401 Daugavpils, Latvia; [d]Institute for Advanced Study, University of Amsterdam, 1012 GC Amsterdam, The Netherlands; and [e]Department of Neurology, Icahn School of Medicine at Mount Sinai, New York, NY 10029

Surveillance is critical in containing globally increasing antimicrobial resistance (AMR). Affordable methodologies to prioritize AMR surveillance efforts are urgently needed, especially in low- and middle-income countries (LMICs), where resources are limited. While socioeconomic characteristics correlate with clinical AMR prevalence, this correlation has not yet been used to estimate AMR prevalence in countries lacking surveillance. We captured the statistical relationship between AMR prevalence and socioeconomic characteristics in a suite of beta-binomial principal component regression models for nine pathogens resistant to 19 (classes of) antibiotics. Prevalence data from ResistanceMap were combined with socioeconomic profiles constructed from 5,595 World Bank indicators. Cross-validated models were used to estimate clinical AMR prevalence and temporal trends for countries lacking data. Our approach provides robust estimates of clinical AMR prevalence in LMICs for most priority pathogens (cross-validated $q^2 > 0.78$ for six out of nine pathogens). By supplementing surveillance data, 87% of all countries worldwide, which represent 99% of the global population, are now informed. Depending on priority pathogen, our estimates benefit 2.1 to 4.9 billion people living in countries with currently insufficient diagnostic capacity. By estimating AMR prevalence worldwide, our approach allows for a data-driven prioritization of surveillance efforts. For carbapenem-resistant *Acinetobacter baumannii* and third-generation cephalosporin-resistant *Escherichia coli*, specific countries of interest are located in the Middle East, based on the magnitude of estimates; sub-Saharan Africa, based on the relative prevalence increase over 1998 to 2017; and the Pacific Islands, based on improving overall model coverage and performance.

antimicrobial resistance | surveillance | global health | carbapenem-resistant *Acinetobacter baumannii* | third-generation cephalosporin-resistant *Escherichia coli*

Antimicrobial resistance (AMR) increases morbidity, mortality, and health-care expenditures worldwide. Surveillance has a key role in supporting policies and stewardship to contain AMR. Surveillance strategies are implemented globally through surveillance networks such as the World Health Organization's (WHO) Global Antimicrobial Surveillance System (GLASS) (1), regional and national surveillance programs, and small-scale efforts within hospitals and laboratories. These fragmented data have been aggregated to describe national AMR trends in Europe (2) and across the globe through initiatives such as ATLAS (Antimicrobial Testing Leadership and Surveillance) (3), ResistanceOpen (4), and ResistanceMap (5). To date, ResistanceMap is the most comprehensive resource, with data from 14.3 million tested isolates covering 74 countries during the period 1998 to 2017.

As setting up and maintaining surveillance are expensive in terms of direct costs, human resources, and infrastructure, the status of clinical AMR in low- and middle-income countries (LMICs) remains largely unknown. This lack of information limits the support for empirical and targeted treatment guidelines, risking further exacerbation of AMR levels due to inappropriate antibiotic usage.

Affordable alternative ways to systematically estimate AMR prevalence in LMICs are thus urgently needed.

While many countries lack structural data on clinical AMR, their socioeconomic status is more extensively characterized. The World Bank collates information on 10,000+ standardized metrics across all world countries over time (6). Various previous works showed significant correlations between AMR prevalence and socioeconomic factors for a large number of countries, including LMICs (19 to 103 countries per study) (7–15). Examples of socioeconomic determinants of AMR prevalence are the income status of individual countries (8), the extent of their out-of-pocket health expenditure (7), and poor governance and corruption (11). Recently, these correlations were utilized to complete the historical record for 52 countries and forecast their national trends toward 2030 for eight priority pathogens (13). However, no previous study has yet leveraged these correlations to systematically estimate AMR prevalence in regions where no historical data are available.

To fill current gaps in the global prevalence map of clinical AMR, we developed a statistical model able to predict AMR prevalence based on exhaustive, unbiased socioeconomic World Bank profiles. Here we show that the model, trained and evaluated through robust cross validation, has high predictive accuracy for

Significance

While antimicrobial resistance is an urgent global problem, substantial clinical surveillance gaps exist in low- and middle-income countries (LMICs). We fill the gaps in the global prevalence map of nine pathogens, resistant to 19 (classes of) antibiotics (representing 75 unique combinations), based on the robust correlation between countries' socioeconomic profiles and extensive surveillance data. Our estimates for carbapenem-resistant *Acinetobacter baumannii* and third-generation cephalosporin-resistant *Escherichia coli* benefit over 2.2 billion people in countries with currently insufficient diagnostic capacity. We show how structural surveillance investments can be prioritized based on the magnitude of prevalence estimated (Middle Eastern countries), the relative prevalence increase over 1998 to 2017 (sub-Saharan African countries), and the improvement of model performance achievable with new surveillance data (Pacific Islands).

some of the most critical pathogens (*Acinetobacter baumannii, Enterobacter aerogenes/cloacae, Enterococcus faecalis, Enterococcus faecium, Escherichia coli, Klebsiella pneumoniae, Pseudomonas aeruginosa, Staphylococcus aureus,* and *Streptococcus pneumoniae*) (16).

Methods

Data Description. We obtained the entire World Bank database of 10,489 standardized metrics on 9 April 2019. We used a subset of 5,595 indicators, excluding those with limited coverage or variation (*SI Appendix*, section S1), to construct 3,948 World Bank profiles, each reflecting the socioeconomic status of one unique country–year combination. For 710 of these profiles, resistance rates for at least one pathogen were available from ResistanceMap (5). Resistance rates are reported as the fraction of all invasive isolates, cultured from blood or cerebrospinal fluid, that exceeds susceptibility breakpoints set by the Clinical and Laboratory Standards Institute (CLSI) (∼40% of all isolates in ResistanceMap) or by the European Committee on Antimicrobial Susceptibility Testing (EUCAST) (∼60% of all isolates in ResistanceMap). We obtained data for 75 unique combinations of pathogens and (classes of) antibiotics on 9 October 2019, representing AMR prevalence in 74 countries during the period 1998 to 2017 (*SI Appendix*, Fig. S1).

Model Development. We performed a principal component regression to develop a suite of candidate models per pathogen in an iterative manner. Principal components of the 5,595 indicators were computed with the nonlinear iterative partial least squares algorithm, enabling imputation of incomplete datasets such as the World Bank database. We used the entire dataset of 3,948 World Bank profiles to ensure our models' applicability to countries without AMR prevalence data and selected the first 30 principal components for model development, as these retained the underlying variability structure of the full set of indicators well (Mantel test with a *P* value of 0.001). Using all country–year combinations for which resistance data were available, beta-binomial vector generalized linear models were developed per pathogen. These models handle overdispersion in the data while accounting for variation in the number of isolates tested. In addition to the principal components, we incorporated year of isolation as a continuous predictor, including its quadratic component, and antibiotic (class) as a categorical predictor. The most parsimonious models were then identified using a backward Akaike Information Criterion-based stepwise reduction of the full models. To identify countries with a relatively large influence on the model, we performed a grouped Cook's distance analysis. As a cutoff for disproportionate influence, we used the median point of the *F* distribution, $F_{0.5}(p, n − p)$, with *P* being the number of predictors and *n* being the number of countries with data. If the highest Cook's distance exceeded the cutoff, we repeated the procedure without the respective country. Otherwise, the suite of candidate models was finalized for the specific pathogen. *SI Appendix*, section S1 extensively describes the computational procedure.

Model Selection and Performance. Per pathogen, the candidate model with the highest Bhattacharyya coefficient was selected for application (17). This coefficient reflects the overlap between the socioeconomic profiles used in model development and those for which prevalence is to be estimated. Consequently, the model with the highest Bhattacharyya coefficient should result in the most accurate estimates. Moreover, we analyzed the extent of geographical autocorrelation present in the surveillance data and in our prevalence estimates to ensure that no artificial (lack of) autocorrelation was introduced by our models (*SI Appendix*, section S7).

The predictive performance of the selected models was quantified by their predictive correlation coefficient q^2, obtained via a five times repeated fivefold groupwise cross validation. Antibiotics were balanced between the folds, and all records from the same country were always included in the same fold. This ensures that individual countries are never represented in multiple folds, effectively forcing the testing of more distant records and thus reducing optimism in error estimates. Cross-validation results were further used in a logistic recalibration to assess their level of over- or underfitting and potential bias and, in the computation of a calibration belt, to identify specific ranges of over- or underestimation (18).

Model Application. We estimated the prevalence of clinical AMR for all country–year combinations not represented in ResistanceMap, but with sufficient World Bank data. National estimates were combined with measured data and aggregated into human population-weighted mean values, both at the global scale and per World Bank–assigned income class. This allowed the assessment of supernational time trends of clinical AMR over the period 1998 to 2017. We also calculated the mean yearly increase factor for individual countries.

To quantify how overall predictions might improve by adding new surveillance data, we generated a set of 30 fictitious isolates per country, the threshold number for CLSI antibiogram reporting and the minimum number of isolates required for inclusion in ResistanceMap (5). Each set of isolates reflected that country's most recent estimate for the specific combination of pathogen and antibiotic (class). By adding one of these sets to the existing resistance data and retraining and reapplying the model for the respective pathogen, we quantified the relative reduction in the total (summed) prediction error. We distinguished between direct effects, i.e., error reductions for the country with new surveillance data, and indirect effects, i.e., error reductions for other countries due to overall model improvement.

Literature Review. To corroborate our estimates of carbapenem resistance (CR) prevalence in *Acinetobacter baumannii* (AB) and third-generation cephalosporin resistance (3GCR) prevalence in *Escherichia coli* (EC), we reviewed the available literature for reviews and meta-analyses of resistant invasive isolates in the 10 countries where our estimates were highest. *SI Appendix*, section S2 contains a detailed description of the procedure followed. We identified 11 studies on CR in AB that fulfilled our criteria and 3 on 3GCR in EC. References to underlying studies were assessed if they explicitly included blood or cerebrospinal fluid isolates. Through this cross-referencing, 17 additional local studies were identified for CR in AB, but none were identified for 3GCR in EC.

Results

Robustness and Performance. We modeled the statistical association between antibiograms from the ResistanceMap repository and exhaustive socioeconomic profiles from the World Bank database. We checked model accuracy through a robust cross-validation procedure. Overall, the models showed high predictive accuracy, with q^2 values between 0.78 and 0.86 for six out of nine pathogens. The predictive accuracy was lower for *Pseudomonas auruginosa* ($q^2 = 0.58$) and *Staphylococcus aureus* ($q^2 = 0.56$) (Fig. 1*A*). The performance for the (predominantly) community-associated pathogen *Streptococcus pneumoniae* was substantially less ($q^2 = 0.27$).

Prediction errors were largest for *Enterobacter aerogenes/cloaecae* and *Enterococcus faecium* and for *E. coli* estimates in Iran and Indonesia specifically (Fig. 1*B*). *E. aerogenes/cloaecae* and *E. faecium* are represented in ResistanceMap by a relatively small number of 13 and 46 countries, respectively (*SI Appendix*, Fig. S1*C*). Consequently, while accurately predicting clinical AMR prevalence in similar countries, performance is lower for target countries with divergent characteristics. Indeed, prediction errors increase when the socioeconomic profiles of target countries lie further from the countries used in model development (*SI Appendix*, Fig. S3*A*).

To assess the stability of our estimates under the influence of model parameterization uncertainty, we constructed a multivariate normal distribution per model with their fitted coefficients as the mean and their variance–covariance matrix as the variance (13). From these, we drew 1,000 sets of possible coefficients, with each computing a new estimate per target country. Uncertainty intervals for each target country's most recent estimates are shown in *SI Appendix*, Table S7. Stable rankings of all countries indicate little impact of parameter uncertainty on the overall prioritization of surveillance when based on our estimates (Spearman's ρ > 0.93 for all pathogens; *SI Appendix*, Table S8). However, estimates for Indonesia, Iran, Russia, and Brazil were more volatile, i.e., disproportionally influenced by parameter uncertainty. This volatility can, at least partly, be explained by the major contributions of population- and gross domestic product-related indicators to the first principal components (*SI Appendix*, Fig. S5).

To further examine the validity of our models, we conducted a calibration analysis per pathogen. These analyses indicated limited over- or underfitting for all pathogens except *S. pneumoniae* and no systematic bias for all pathogens except *S. pneumoniae*
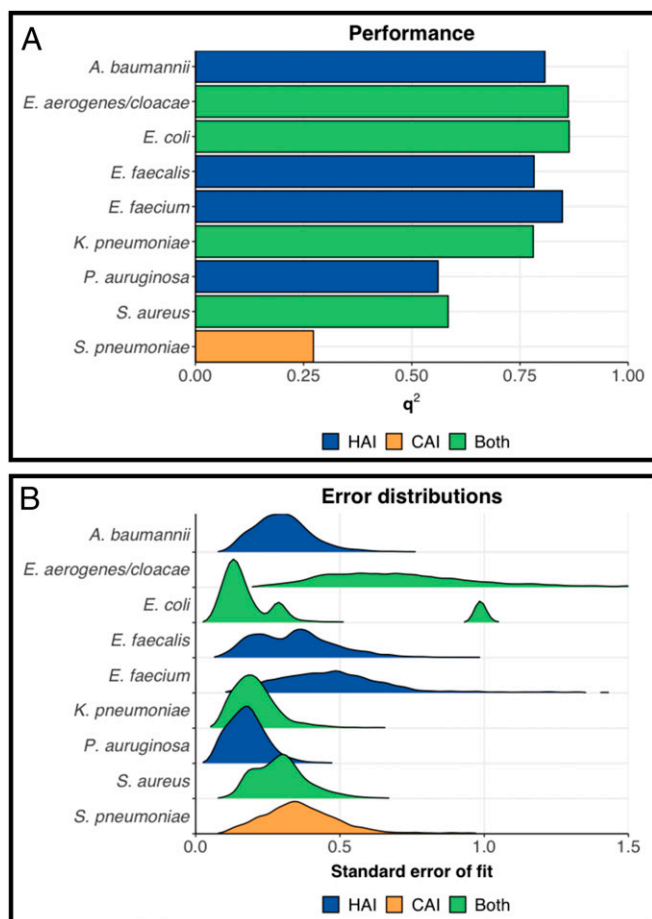
Fig. 1. Performance and robustness of the models constructed: (*A*) square correlation coefficient q^2, derived via a five times repeated fivefold group-wise cross validation and (*B*) distribution of prediction errors at the scale of the linear predictors. HAI: (predominantly) hospital-associated pathogens; CAI: (predominantly) community-associated pathogens; Both: pathogens associated with both hospital and community. Classification is based on WHO's description of priority pathogens list.

and *K. pneumoniae* (*SI Appendix*, Fig. S3). The prevalence of resistance in *K. pneumoniae* was slightly overestimated over the full range of prevalence values. Our models further tend to overestimate the higher prevalence of resistance for most pathogens and underestimate the lower prevalence of resistance for some pathogens (*SI Appendix*, Fig. S4). Despite these deviations at the limits of the resistance prevalence spectrum, calibration analyses confirmed that for the majority of pathogens we can provide actionable insights for surveillance prioritization.

Here we highlight the implications for two priority pathogens: 1) carbapenem-resistant *Acinetobacter baumannii* (CRAB), the leading cause of hospital-associated infections, and 2) *Escherichia coli* resistant to third-generation cephalosporins (3GCREC), widely responsible for both community- and hospital-acquired infections. Both are classified as critical priority pathogens by WHO (16). An overview of the final models of all pathogens is included in *SI Appendix*, Table S1.

Filling Geographical Gaps. We improved the coverage of the global prevalence map of clinical AMR from 24% (CRAB; Fig. 2*A*) and 29% (3GCREC; Fig. 2*D*) to 87% of all 235 countries and areas specified by the United Nations (19). The remaining 13% are overseas territories or disputed areas with limited World Bank

information. Our estimates represent populations of ~2.5 (CRAB) and 2.2 (3GCREC) billion people, increasing coverage to 99% of the global population. Most of the added population lives in low-income (0.7 billion) and lower middle-income (1.2 billion for CRAB; 0.9 billion for 3GCREC) countries.

Per pathogen, *SI Appendix*, Table S7 provides our most recent national estimates for countries not included in ResistanceMap. Estimates of CR in AB and 3GCR in EC were generally highest in countries located in the (Greater) Middle East, most notably in Iran, with 100% and 73% prevalence in 2017, respectively. Iran has been enrolled in GLASS since 2016 but submitted few isolates to the GLASS call for 2017 data: 10 *E. coli* blood isolates, indicating 90% prevalence of 3GCR (95% confidence interval [CI] 60 to 100%), and 8 *A. baumannii* blood isolates indicating 100% prevalence of CR (95% CI 68 to 100%), both in line with our estimates (1). Our literature review yielded four meta-analyses assessing CRAB in Iran, reporting aggregate resistance ranging from 55 to 91% (*SI Appendix*, Table S2). These percentages are lower than our 100% estimate but reflect long-term pooled prevalence rates of isolates with a mixed origin sampled between 1995 and 2018. In fact, one local study from Iran that included only invasive isolates ($n = 59$; collected in 2012 to 2013) reported 100% resistance to both imipenem and meropenem (20). Further, we identified a recent meta-analysis of *E. coli* resistance against the fourth-generation cephalosporin cefepime, reporting 62% resistance (95% CI 57–67%), based on 347 isolates from survey years 2012 to 2014 (*SI Appendix*, Table S3) (21). These estimates are highly concordant with our estimates for the same period, i.e., 62% in 2012 to 75% in 2014.

Our 2017 estimates of the prevalence of CR in AB also fill important gaps for other Middle Eastern countries, including Lebanon (92% prevalence estimated), Jordan (86%), Egypt (85%), Syria (83%), and Sudan (77%); for 3GCR in EC they are Yemen (61%), Afghanistan (60%), Syria (57%), and Sudan (56%). A comprehensive review of countries of the Arab League reported values similar to, albeit somewhat lower than, our predictions for CR in AB: 77 to 91% in Lebanon (three local studies, 2010 to 2015), 64% in Jordan (one local study, 2009 to 2010), 70 to 96% in Egypt (two local studies, 2012 to 2016), 71% in Syria (one local study, 2008 to 2011), and 89% in Sudan (one local study, 2011 to 2014) (22). For 3GCR in EC, 53 to 62% resistance derived from local Syrian data (four studies) was in line with our prediction of 57%, while 35% resistance for Sudan (one study) was lower than our estimate of 56% (22).

Altogether, our highest estimates of carbapenem resistance in invasive *A. baumannii* and third-generation cephalosporin resistance in invasive *E. coli* are consistent with the sparse local data available from the open literature (*SI Appendix*, Tables S2 and S3).

Filling Temporal Gaps. To obtain insights into temporal trends of resistance, i.e., globally and per World Bank–assigned income class, we aggregated national estimates and measurements into (population-weighted) mean yearly prevalence values. The global prevalence of CR in AB increased from below 40% in 1998 to over 60% in 2008, after which it stabilized, fluctuating between 60% and 70% over the last decade (Fig. 3*A*). Trends across income classes were similar, although the relative increase in low-income countries was larger, as the steepest increases were estimated in low- and lower middle-income countries (Fig. 3*B*). This particularly relates to countries in sub-Saharan Africa, e.g., Angola (yearly increase with a factor of 1.58), Cameroon (factor of 1.36), Burkina Faso, and Equatorial Guinea (factor of 1.35).

In contrast with CR in AB, the prevalence of 3GCR in EC has steadily increased since surveillance started (Fig. 3*C*). With an average 1.0% yearly increase, global resistance rose from 34 to 54% over the period 1998 to 2017. Moreover, trends were indiscriminate of income class or geography (Fig. 3*D*).

Oldenkamp et al.
Filling the gaps in the global prevalence map of clinical antimicrobial resistance

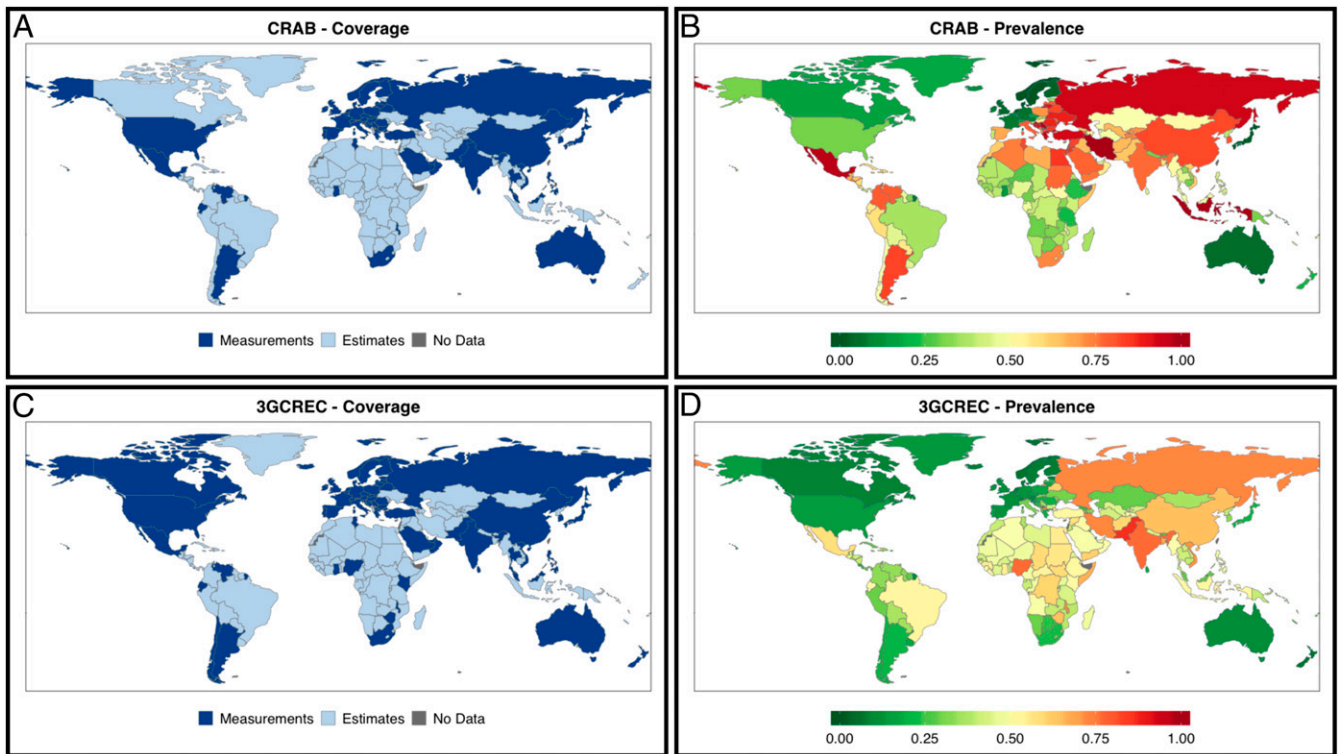PNAS | 3 of 7
https://doi.org/10.1073/pnas.2013515118

Fig. 2. Global coverage and prevalence of carbapenem resistance in *A. baumannii* and third-generation cephalosporin resistance in *E. coli*. (*A*) Coverage of CRAB measurements (24% of countries and areas; 66% of global population) increases by supplementing with estimates (87% of countries; 99% of global population). (*B*) Measured and estimated CRAB for the most recent year available or possible. (*C*) Coverage of 3GCREC measurements (29% of countries and areas; 70% of global population) increases by supplementing with estimates (87% of countries; 99% of global population). (*D*) Measured and estimated 3GCREC for the most recent year available or possible.

The Benefits of New Surveillance Data. The introduction of structural surveillance in a given country may have both direct and indirect benefits mediated by the model, as new local data might improve estimates for other similar countries. This notion enables us to identify countries where surveillance could best improve overall estimates of CR in AB (Fig. 4*A*) and 3GCR in EC (Fig. 4*B*). Based on the principal components included in the respective models, an agglomerative hierarchical clustering analysis identified clusters of similar countries (Roman numerals in Fig. 4). Their composition is shown in *SI Appendix*, Tables S4 and S5. Estimates for Indonesia (singleton v in Fig. 4*A*; singleton xix in Fig. 4*B*) and Iran (singleton vi in Fig. 4*A*; singleton xx in Fig. 4*B*) were least accurate, as indicated by the direct error reduction achievable with new surveillance data (orange bars in Fig. 4). Since these countries do not cluster, new data generated in Iran or Indonesia will marginally improve estimates for other countries. In contrast, cluster xii in Fig. 4*A* and, to a lesser extent, cluster xv in Fig. 4*B* represent a group of data-poor countries where new surveillance data could substantially improve the parameter space covered by the models (purple bars in Fig. 4). These clusters consist of various smaller (predominantly) island states.

Discussion

Our global maps of national resistance prevalence provide data-driven criteria for the prioritization of surveillance efforts in LMICs. First, surveillance might be prioritized in countries where high resistance rates are expected. Our analyses of carbapenem-resistant *A. baumannii* and third-generation cephalosporin-resistant *E. coli* indicate such countries are often located in the (Greater) Middle East (Fig. 2). While acknowledging that the calibration results indicate that our models have a slight tendency to overestimate resistance prevalence at the upper limit of the

spectrum (*SI Appendix*, Fig. S4), estimates for CRAB and 3GCREC were consistent with available literature (*SI Appendix*, section S2). As such, our results support concerns raised on the alarming potential for AMR to develop into a "perfect storm" in parts of the Middle East that are burdened by conflict, such as Syria, Yemen, and Iraq (23).

Model performance for the community-associated pathogen *S. pneumoniae* was much lower compared to the other pathogens. This might indicate that resistance prevalence in pathogens predominantly acquired within the community correlate less well with countries' socioeconomic characteristics than resistance prevalence in hospital-associated pathogens and that other factors might play a role. A more likely explanation is the fact that vaccination against pneumococcal disease reduces resistance prevalence in *S. pneumoniae*, both through direct action and indirectly by lowering antibiotic consumption (24, 25), and global immunization initiatives such as GAVI (Global Alliance for Vaccines and Immunization), providing vaccines to LMICs, have successfully weakened plausible correlations between pneumococcal conjugate vaccine coverage and countries' socioeconomic status (26). In contrast, vaccines for the other pathogens studied are not available.

Second, in addition to countries with high estimated AMR, surveillance efforts might be prioritized in countries where AMR is estimated to increase rapidly. Our prevalence trends highlight various countries in sub-Saharan Africa as being important (Fig. 3), especially for carbapenem-resistant *A. baumannii*, and support previous calls for systematic surveillance to inform empirical treatment guidelines in this region of the world (27). In contrast to CR in AB, the prevalence of 3GCR in EC has increased steadily and indiscriminate of geography (Fig. 3*D*). Globally, we estimate an average yearly prevalence increase of
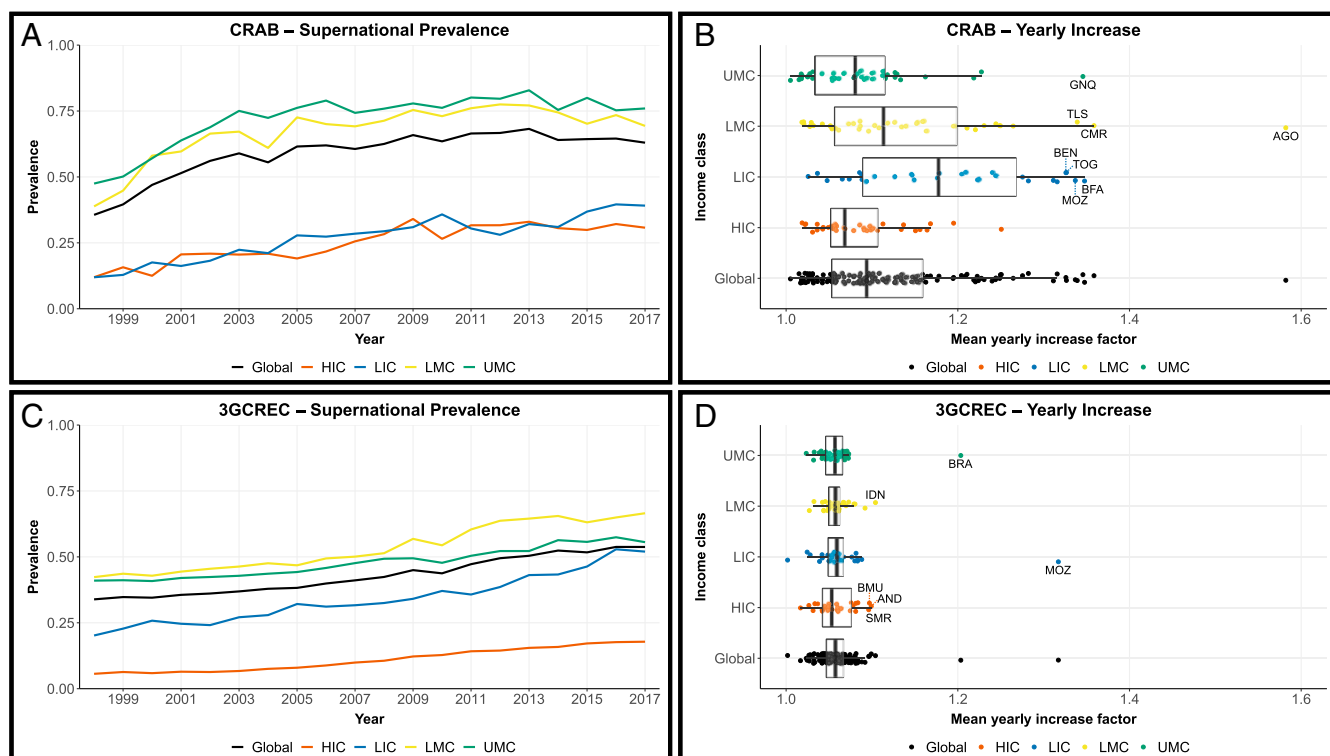
Fig. 3. Temporal trends of resistance worldwide (black), in low-income countries (LIC; blue), lower middle-income countries (LMC; yellow), upper middle-income countries (UMC; green), and high-income countries (HIC; orange) over the period 1998 to 2017. (*A* and *C*) Human population-weighted mean resistance prevalence of carbapenem resistance in *A. baumannii* and third-generation cephalosporin-resistance in *E. coli*, respectively. (*B* and *D*) Tukey box and whisker plots of mean yearly increase factors in all countries without systematic national surveillance for CRAB and 3GCREC, respectively. Positive outliers, based on the global box whisker plots, are labeled with their ISO3 code (International Organization for Standardization alpha-3 code). AGO: Angola; AND: Andorra; BEN: Benin; BFA: Burkina Faso; BMU: Bermuda; BRA: Brazil; CMR: Cameroon; GNQ: Equatorial Guinea; IDN: Indonesia; MOZ: Mozambique; SMR: San Marino; TLS: Timor-Leste; TOG: Togo.

1.0%, as its global prevalence rose from 34 to 54% over the period 1998 to 2017 (Fig. 3*C*). Recently, Alvarez-Uria and colleagues forecasted a global trend with a similar slope of 0.83%, but with a 13% higher intercept (28). This is likely due to the high (~80%) resistance prevalence in India, which, with its low per capita income and large population size, had a major influence on their global estimate. In our model, however, India's unique profile makes it less informative for other countries, reducing its impact on the global mean estimate.

Third, prioritization of surveillance efforts might focus on improving model predictions. The Pacific Islands form an interesting cluster of countries where surveillance of AMR has been limited so far (29) but where new surveillance data would most improve overall model performance (Fig. 4). In addition to these regions, Latin America and the Caribbean (LAC) form another extensive region largely absent from ResistanceMap. While they did not explicitly emerge as priority countries, LAC countries of interest for CRAB are Colombia (79% resistance estimated), El Salvador (74%), and Guatemala (67%). Countries of interest for 3GCREC are Haiti (55%), Brazil (52%), and Guatemala (45%). Although most LAC countries participate in the long-running Latin American Antimicrobial Resistance Surveillance Network (ReLAVRA), the sample type and origin of isolates were not reported to ReLAVRA before 2018. Since our research specifically focuses on AMR in invasive clinical isolates, these data could not be included. Nevertheless, following extensive curation, aggregation, and harmonization, additional data from other (supra)national surveillance networks (30), as well as from alternative sources like academic institutes, pharmaceutical companies, and private laboratories (31), should

be harnessed to further fill the global map of clinical AMR and increase the reliability of our estimates.

It is widely acknowledged that laboratory-based surveillance is prone to selection bias, especially in LMICs, where access to quality microbiology diagnostics is poor and routine diagnostics are often limited to treatment failures and specialized tertiary care facilities (32). These limitations propagate into our prevalence estimates, and they should thus be interpreted as an approximation of the true prevalence in a country. Future unbiased surveillance will likely require novel approaches, e.g., population-based surveillance (33, 34), for which the necessary resources and infrastructure are not yet in place. Our predictive framework provides a means to support the prioritization of investments required for this.

Comparisons across countries are complicated by discrepancies between CLSI and EUCAST susceptibility break points, with the latter generally more stringent and thus resulting in higher resistance rates (35). Any potential bias introduced into our models by these discrepancies could be reduced by further international harmonization of methods and breakpoints (36).

Our approach has several limitations. While our estimates of AMR prevalence can support national surveillance strategies, subnational heterogeneity in resistance prevalence reduces their local usefulness, especially in large and highly populated countries such as India (37). Due to its generic nature, however, our approach inherently allows for assessments at a more local scale, provided that both socioeconomic indicators and clinical AMR data are available at such a local level. The benefit of expanding our approach subnationally is further emphasized by the fact that parameter uncertainty specifically affects prevalence estimates in
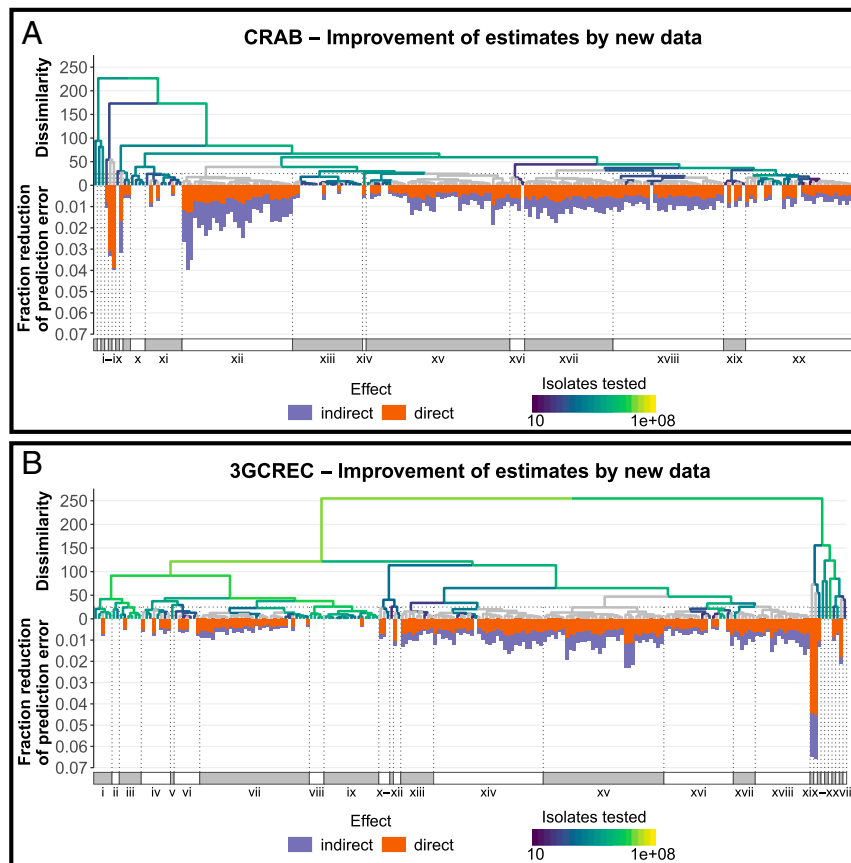
Oldenkamp et al.
Filling the gaps in the global prevalence map of clinical antimicrobial resistance

PNAS | 5 of 7
https://doi.org/10.1073/pnas.2013515118

MICROBIOLOGY

SOCIAL SCIENCES

Fig. 4. Potential reduction of the total (summed) prediction error in all estimated resistance rates for 2017, should resistance data for 2017 become available for the specific country. Hierarchical clustering of all countries based on the set of principal components used in the respective models, with gradient indicating number isolates available from ResistanceMap tested for carbapenem resistance in *A. baumannii* (*A*) or third-generation cephalosporin resistance in *E. coli* (*B*). Gray: no data. Roman numerals indicate clusters of countries based on a dissimilarity threshold of 25 (total within-cluster sum of squares), with their composition listed in *SI Appendix*, Tables S4 and S5, respectively. Bars represent indirect (purple) and direct (orange) reductions of the summed prediction error over all estimated resistance rates for 2017, should surveillance data from 2017 become available for the specific country.

emerging economies with large population sizes (*SI Appendix, Table S7*).

We leveraged the full extent of the World Bank database in a principal component regression, reducing its high dimensionality while retaining the underlying variability structure. This comes with limitations of interpretability. Because principal components are obtained from only the World Bank indicators, the constraints imposed on their coefficients have no direct relationship with how they might affect prevalence rates, obscuring univariate relationships between underlying indicators and AMR prevalence (38). Additionally, by applying our models predictively we inherently assumed that the correlations found for predominantly upper middle- and high-income countries are universally valid. This also limits the applicability of our approach to pathogens that are globally relevant and therefore broadly included in surveillance programs. Consequently, priority pathogens that specifically burden LMICs, e.g., fluoroquinolone-resistant *Salmonella* species (16), would require a more regional or local approach since truly global data are not, and will not, become available.

As AMR forms the quintessential One Health issue, indicators from other domains might improve prevalence estimates (39). The flexibility of our approach allows for integration of such indicators in future analyses when sufficient data become available. Interesting candidates are metagenomics data from human sewage (40, 41) and climatic parameters such as temperature, which correlates with clinical AMR (42). Furthermore, antibiotic

consumption was no predictor in our models. Although it is commonly considered one of the key drivers of AMR (43, 44), data scarcity would either require extensive imputation (13) or substantially reduce the number of complete records available for model development. Instead, we considered antibiotic consumption a midpoint indicator, thus assuming international variability in antibiotic consumption to be intrinsically captured by our socioeconomic profiles. This assumption is supported by a sensitivity analysis using antibiotic consumption data available from ResistanceMap (45) (*SI Appendix*, section S6) and by previous studies showing antibiotic consumption parallel socioeconomic indicators such as the Human Development Index and aging of the population (46), gross domestic product per capita (43), out-of-pocket health expenditure, population by age group, population density, and GDP (13).

The prevalence of AMR in clinical priority pathogens is unknown in many regions of the world, and more than 2 billion people live in countries without data on even the most surveyed pathogens. The majority of these countries are LMICs. Our improved global maps of clinically relevant AMR provide a means to close this knowledge gap in a most efficient way.

Data Availability. All study data are included in the article and *SI Appendix*.

1. World Health Organization, *Global Antimicrobial Resistance Surveillance System (GLASS) Report—Early Implementation 2017-2018* (World Health Organization, Geneva, Switzerland, 2018).
2. A. Cassini *et al*.; Burden of AMR Collaborative Group, Attributable deaths and disability-adjusted life-years caused by infections with antibiotic-resistant bacteria in the EU and the European Economic Area in 2015: A population-level modelling analysis. *Lancet Infect. Dis.* **19**, 56–66 (2019).
3. D. Beardmore *et al*., *Antimicrobial Resistance Surveillance: Sharing Industry Data—Post-Pilot Project Report* (Wellcome Trust, London, UK, 2018).
4. D. R. MacFadden *et al*., A platform for monitoring regional antimicrobial resistance, using online data sources: ResistanceOpen. *J. Infect. Dis.* **214**, S393–S398 (2016).
5. Center for Disease Dynamics Economics & Policy, ResistanceMap. https://resistancemap.cddep.org/AntibioticResistance.php/. Accessed 9 October 2019.
6. World Bank, World DataBank: World Development Indicators. https://databank.worldbank.org/reports.aspx?source=world-development-indicators. Accessed 20 September 2019.
7. M. Alsan *et al*., Out-of-pocket health expenditures and antimicrobial resistance in low-income and middle-income countries: An economic analysis. *Lancet Infect. Dis.* **15**, 1203–1210 (2015).
8. G. Alvarez-Uria, S. Gandra, R. Laxminarayan, Poverty and prevalence of antimicrobial resistance in invasive isolates. *Int. J. Infect. Dis.* **52**, 59–61 (2016).
9. G. Alvarez-Uria, M. Midde, Trends and factors associated with antimicrobial resistance of Acinetobacter spp. invasive isolates in Europe: A country-level analysis. *J. Glob. Antimicrob. Resist.* **14**, 29–32 (2018).
10. A. Blommaert *et al*., Determinants of between-country differences in ambulatory antibiotic use and antibiotic resistance in Europe: A longitudinal observational study. *J. Antimicrob. Chemother.* **69**, 535–547 (2014).
11. P. Collignon, P.-C. Athukorala, S. Senanayake, F. Khan, Antimicrobial resistance: The major contribution of poor governance and corruption to this growing problem. *PLoS One* **10**, e0116746 (2015).
12. P. Collignon, J. J. Beggs, T. R. Walsh, S. Gandra, R. Laxminarayan, Anthropological and socioeconomic factors contributing to global antimicrobial resistance: A univariate and multivariable analysis. *Lancet Planet. Health* **2**, e398–e405 (2018).
13. T. Cravo Oliveira Hashiguchi, D. Ait Ouakrim, M. Padget, A. Cassini, M. Cecchini, Resistance proportions for eight priority antibiotic-bacterium combinations in OECD, EU/EEA and G20 countries 2000 to 2030: A modelling study. *Euro Surveill.* **24**, 1800445 (2019).
14. C. García-Rey, A. Fenoll, L. Aguilar, J. Casal, Effect of social and climatological factors on antimicrobial use and Streptococcus pneumoniae resistance in different provinces in Spain. *J. Antimicrob. Chemother.* **54**, 465–471 (2004).
15. W. W. Lim *et al*., Determinants of methicillin-resistant Staphylococcus aureus (MRSA) prevalence in the Asia-Pacific region: A systematic review and meta-analysis. *J. Glob. Antimicrob. Resist.* **16**, 17–27 (2019).
16. World Health Organization, *Global Priority List of Antibiotic-Resistant Bacteria to Guide Research, Discovery, and Development of New Antibiotics* (World Health Organization, Geneva, Switzerland, 2017).
17. A. Bhattacharyya, On a measure of divergence between two multinomial populations. *Sankhyā: Indian J. Stat.* **7**, 401–406 (1946).
18. S. Finazzi, D. Poole, D. Luciani, P. E. Cogo, G. Bertolini, Calibration belt for quality-of-care assessment based on dichotomous outcomes. *PLoS One* **6**, e16110 (2011).
19. United Nations, Department of Economic and Social Affairs, Population Division, "World population prospects 2019: Volume I: Comprehensive Tables" (Rep. ST/ESA/SER.A/426, United Nations, New York, NY, 2019).
20. B. Pourabbas, R. Firouzi, G. Pouladfar, Characterization of carbapenem-resistant Acinetobacter calcoaceticus-baumannii complex isolates from nosocomial bloodstream infections in southern Iran. *J. Med. Microbiol.* **65**, 235–239 (2016).
21. S. M. Bechashk, G. Moradi, B. Mohsenpour, R. Ramazanzadeh, Prevalence of cefepime-resistant Escherichia coli in Iran: A meta-analysis (2007-2016). *Iran. J. Public Health* **48**, 603–611 (2019).
22. R. A. Moghnieh *et al*., Epidemiology of common resistant bacterial pathogens in the countries of the Arab League. *Lancet Infect. Dis.* **18**, e379–e394 (2018).
23. S. Devi, AMR in the Middle East: "A perfect storm." *Lancet* **394**, 1311–1312 (2019).
24. K. P. Klugman, S. Black, Impact of existing vaccines in reducing antibiotic resistance: Primary and secondary effects. *Proc. Natl. Acad. Sci. U.S.A.* **115**, 12896–12901 (2018).
25. J. A. Lewnard, N. C. Lo, N. Arinaminpathy, I. Frost, R. Laxminarayan, Childhood vaccines and antibiotic use in low- and middle-income countries. *Nature* **581**, 94–99 (2020).
26. G. Ikilezi, O. J. Augusto, J. L. Dieleman, K. Sherr, S. S. Lim, Effect of donor funding for immunization from Gavi and other development assistance channels on vaccine coverage: Evidence from 120 low and middle income recipient countries. *Vaccine* **38**, 588–596 (2020).
27. S. J. Leopold, F. van Leth, H. Tarekegn, C. Schultsz, Antimicrobial drug resistance among clinically relevant bacterial isolates in sub-Saharan Africa: A systematic review. *J. Antimicrob. Chemother.* **69**, 2337–2353 (2014).
28. G. Alvarez-Uria, S. Gandra, S. Mandal, R. Laxminarayan, Global forecast of antimicrobial resistance in invasive isolates of Escherichia coli and Klebsiella pneumoniae. *Int. J. Infect. Dis.* **68**, 50–53 (2018).
29. N. D. Foxlee, N. Townell, L. McIver, C. L. Lau, Antibiotic resistance in Pacific Island countries and territories: A systematic scoping review. *Antibiotics (Basel)* **8**, 29 (2019).
30. E. A. Ashley *et al*., An inventory of supranational antimicrobial resistance surveillance networks involving low- and middle-income countries since 2000. *J. Antimicrob. Chemother.* **73**, 1737–1749 (2018).
31. E. A. Ashley *et al*., Harnessing alternative sources of antimicrobial resistance data to support surveillance in low-resource settings. *J. Antimicrob. Chemother.* **74**, 541–546 (2019).
32. S. Ombelet *et al*.; Bacteriology in Low Resource Settings Working Group, Clinical bacteriology in low-resource settings: Today's solutions. *Lancet Infect. Dis.* **18**, e248–e258 (2018).
33. F. Ginting *et al*., Rethinking antimicrobial resistance surveillance: A role for lot quality assurance sampling. *Am. J. Epidemiol.* **188**, 734–742 (2019).
34. A. K. Sugianli *et al*., Laboratory-based versus population-based surveillance of antimicrobial resistance to inform empirical treatment for suspected urinary tract infection in Indonesia. *PLoS One* **15**, e0230489 (2020).
35. T. P. Cusack *et al*., Impact of CLSI and EUCAST breakpoint discrepancies on reporting of antimicrobial susceptibility and AMR surveillance. *Clin. Microbiol. Infect.* **25**, 910–911 (2019).
36. G. Kahlmeter, Defining antibiotic resistance-towards international harmonization. *Ups. J. Med. Sci.* **119**, 78–86 (2014).
37. R. Laxminarayan, R. R. Chaudhury, Antibiotic resistance in India: Drivers and opportunities for action. *PLoS Med.* **13**, e1001974 (2016).
38. H. Artigue, G. Smith, The principal problem with principal components regression. *Cogent Math. Stat.* **6**, 1622190 (2019).
39. A. Chatterjee *et al*., Quantifying drivers of antibiotic resistance in humans: A systematic review. *Lancet Infect. Dis.* **18**, e368–e378 (2018).
40. R. S. Hendriksen *et al*.; Global Sewage Surveillance Project Consortium, Global monitoring of antimicrobial resistance based on metagenomics analyses of urban sewage. *Nat. Commun.* **10**, 1124 (2019).
41. F. M. Aarestrup, M. E. J. Woolhouse, Using sewage for surveillance of antimicrobial resistance. *Science* **367**, 630–632 (2020).
42. D. R. MacFadden, S. F. McGough, D. Fisman, M. Santillana, J. S. Brownstein, Antibiotic resistance increases with local temperature. *Nat. Clim. Chang.* **8**, 510–514 (2018).
43. E. Y. Klein *et al*., Global increase and geographic convergence in antibiotic consumption between 2000 and 2015. *Proc. Natl. Acad. Sci. U.S.A.* **115**, E3463–E3470 (2018).
44. G. Allwell-Brown *et al*., Trends in reported antibiotic use among children under 5 years of age with fever, diarrhoea, or cough with fast or difficult breathing across low-income and middle-income countries in 2005-17: A systematic analysis of 132 national surveys from 73 countries. *Lancet Glob. Health* **8**, e799–e807 (2020).
45. The Center for Disease Dynamics Economics & Policy, ResistanceMap: Antibiotic use. https://resistancemap.cddep.org/AntibioticUse.php. Accessed 26 August 2020.
46. R. Oldenkamp, A. H. W. Beusen, M. A. J. Huijbregts, Aquatic risks from human pharmaceuticals–Modelling temporal trends of carbamazepine and ciprofloxacin at the global scale. *Environ. Res. Lett.* **14**, 034003 (2019).

MICROBIOLOGY

SOCIAL SCIENCES

Oldenkamp et al.
Filling the gaps in the global prevalence map of clinical antimicrobial resistance

PNAS | 7 of 7
https://doi.org/10.1073/pnas.2013515118