



# Analysis of research intensity on infectious disease by disease burden reveals which infectious diseases are neglected by researchers

Yuki Furuse<sup>a,b,1</sup>

<sup>a</sup>Institute for Frontier Life and Medical Sciences, Kyoto University, 606-8507 Kyoto, Japan; and <sup>b</sup>Hakubi Center for Advanced Research, Kyoto University, 606-8501 Kyoto, Japan

Edited by Barry R. Bloom, Harvard T. H. Chan School of Public Health, Boston, MA, and approved November 26, 2018 (received for review August 22, 2018)

**Infectious diseases are associated with considerable morbidity and mortality worldwide. Although human, financial, substantial, and time resources are limited, it is unknown whether such resources are used effectively in research to manage diseases. The correlation between the disability-adjusted life years to represent disease burden and number of publications as a surrogate for research activity was investigated to measure burden-adjusted research intensity for 52 infectious diseases at global and country levels. There was significantly low research intensity for paratyphoid fever and high intensity for influenza, HIV/acquired immunodeficiency syndrome, hepatitis C, and tuberculosis considering their disease burden. We identified the infectious diseases that have received the most attention from researchers and those that have been relatively disregarded. Interestingly, not all so-called neglected tropical diseases were subject to low burden-adjusted research intensity. Analysis of the intensity of infectious disease research at a country level revealed characteristic patterns. These findings provided a basis for further discussion of the more appropriate allocation of resources for research into infectious diseases.**

infectious diseases | neglected tropical diseases | public health | epidemiology | philology

Infectious diseases are caused by microorganisms, such as bacteria, viruses, fungi, and parasites. Approximately 10 million people died of infectious diseases in 2016, accounting for one-fifth of all deaths worldwide (1). The highest mortality among infectious diseases comes from lower respiratory tract infections, followed by enteric infections causing diarrhea, tuberculosis, AIDS caused by the HIV, and malaria (1). Disability caused by infectious diseases also has a substantial impact on public health. For example, although trachoma and onchocerciasis are not fatal, they can result in loss of vision, a significant disease burden that affects quality of life and leads to economic loss (2). One indicator to quantify disease burden is the disability-adjusted life years (DALYs), calculated as the sum of years of life lost and years lived with disability (3).

The first vaccine against smallpox was developed in 1796, and the first antibiotic medication, penicillin, was discovered in 1928 (4). Since then, increasing types of vaccines and antimicrobials have been developed, dramatically reducing the infectious disease burden (5). This success, ironically, resulted in a reduced focus on infectious diseases during the 1960s and 1970s, assuming that the battle with infectious diseases had been won (5, 6).

Neglected tropical diseases (NTDs), such as Chagas disease and schistosomiasis, are thought to attract especially low attention. Although the disease burden of some NTDs can be reduced dramatically by implementing simple strategies, such as mass drug administration and vector control (7, 8), the limited endemicity of these diseases in low-income countries and low disease burden in high-income countries have meant that these strategies have not been followed completely.

In addition, threats from emerging and reemerging infectious diseases have increased globally; this was seen in the devastating outbreak of Severe Acute Respiratory Symptoms coronavirus in 2003 and Ebola virus in 2014 (9). Another concern is the emergence and increase in antimicrobial-resistant bacteria, such as methicillin-resistant *Staphylococcus aureus* and carbapenem-resistant *Enterobacteriaceae* in many areas (10, 11). The issue causes significant clinical and economic burden with important consequences for individual patients and public health (12, 13).

The threat of infectious disease will never be zero. Because human, financial, substantial, and time resources are limited, they should be used effectively to manage the diseases. This also applies to research on infectious diseases. We investigated which infectious diseases have been neglected by researchers and which have received attention through an analysis of disease burden (measured in DALYs) and research activity (measured by the number of publications).

## Results

**Infectious Diseases That Have Received the Most Attention from or Been Relatively Disregarded by Researchers.** In total, 52 infectious diseases were included in this study (Table 1 and *SI Appendix, Table S1*). The disease burden (in DALYs) and number of publications were well correlated, and the regression line between the two variables in 2010–2017 was drawn excluding diseases considered outliers (see *Materials and Methods* for detail;

## Significance

**Threats from emerging and reemerging infectious diseases have increased globally. However, neglected tropical diseases, such as Chagas disease and schistosomiasis, are believed to attract low attention. We investigated which infectious diseases have been neglected by researchers and which have received attention through an analysis of disease burden and research activity. We found, for example, that influenza and HIV/AIDS have attracted high research attention in relation to their disease burden, while paratyphoid fever has attracted low attention considering the disease burden. Interestingly, not all so-called neglected tropical diseases were subject to low research intensity. Further discussion must occur with regard to the appropriate allocation of resources for research into infectious diseases.**

Author contributions: Y.F. designed research, performed research, analyzed data, and wrote the paper.

The author declares no conflict of interest.

This article is a PNAS Direct Submission.

Published under the PNAS license.

<sup>1</sup>Email: furusey.kyoto@gmail.com.

This article contains supporting information online at [www.pnas.org/lookup/suppl/doi:10.1073/pnas.1814484116/-DCSupplemental](http://www.pnas.org/lookup/suppl/doi:10.1073/pnas.1814484116/-DCSupplemental).

Published online December 31, 2018.

**Table 1. Infectious diseases and countries analyzed**

Category	Item
Diseases [label used for tables and figures]	Adenovirus infection [Adenovirus]; <i>Aeromonas</i> gastroenteritis [ <i>Aeromonas</i> ]; African trypanosomiasis [Af trypanosomiasis]; Amoebiasis [Amoebiasis]; Ascariasis [Ascariasis]; Pseudomembranous enterocolitis by <i>Clostridium difficile</i> [ <i>C. difficile</i> ]; <i>Campylobacter</i> enteritis [ <i>Campylobacter</i> ]; Chagas disease [Chagas]; Chlamydial infection including sexual diseases and trachoma [Chlamydia]; Cholera [Cholera]; Cryptosporidiosis [Cryptosporidiosis]; Cysticercosis [Cysticercosis]; Dengue fever [Dengue]; Diphtheria [Diphtheria]; Dracunculiasis (Guinea worm disease) [Dracunculiasis]; Cystic echinococcosis [Echinococcosis]; Genital herpes [Genital herpes]; Gonococcal infection [Gonorrhoea]; Acute hepatitis A [HepA]; Hepatitis B (acute and chronic hepatitis/cirrhosis/liver cancer) [HepB]; Hepatitis C (acute and chronic hepatitis/cirrhosis/liver cancer) [HepC]; Acute hepatitis E [HepE]; <i>Haemophilus influenzae</i> type B meningitis and pneumonia [Hib]; HIV/AIDS [HIV/AIDS]; Hookworm disease [Hookworm]; Influenza [Influenza]; Leishmaniasis [Leishmaniasis]; Leprosy [Leprosy]; Lymphatic filariasis [Lymphatic filariasis]; Malaria [Malaria]; Measles [Measles]; Meningococcal meningitis [Meningococcal meningitis]; Salmonella infections other than typhoid/paratyphoid fevers [Nontyphoidal salmonellosis]; Norovirus infection [Norovirus]; Onchocerciasis [Onchocerciasis]; Paratyphoid fever [Paratyphoid fever]; Pertussis (whooping cough) [Pertussis]; Rabies [Rabies]; Rotaviral enteritis [Rotavirus]; Respiratory syncytial virus infection [RSV]; Pneumococcal meningitis and pneumonia [ <i>S. pneumoniae</i> ]; Schistosomiasis [Schistosomiasis]; Shigellosis (bacillary dysentery) [Shigellosis]; Syphilis [Syphilis]; Tetanus [Tetanus]; Food-borne trematodiasis [Trematodiasis]; Trichomoniasis [Trichomoniasis]; Trichuriasis [Trichuriasis]; Tuberculosis [Tuberculosis]; Typhoid fever [Typhoid fever]; Varicella and herpes zoster [Varicella/Zoster]; Yellow fever [Yellow fever]
Countries/territory	Argentina; Australia; Bangladesh; Brazil; Canada; Chile; China; Colombia; Croatia; Czech Republic; Egypt; France; Germany; Hungary; India; Indonesia; Iran; Israel; Italy; Japan; Kenya; Malaysia; Mexico; Mongolia; Morocco; Nepal; New Zealand; Nigeria; Pakistan; Papua New Guinea; Poland; Russia; Saudi Arabia; Singapore; South Africa; South Korea; Spain; Sri Lanka; Taiwan; Thailand; Tunisia; Turkey; United Kingdom; USA; Vietnam

$r = 0.40$ ,  $P = 0.0058$ ; Fig. 1). Elasticity of the correlation under the double logarithmic linear relationship was 0.2, indicating that the number of publications increased by 0.2% in response to a 1% change in DALYs. For a given disease we defined the “burden-adjusted research intensity” (BARI) index as the residual from regression line, in order to quantify deviation from expected trends. The BARI index was significantly high (>95% prediction interval) for influenza, HIV/AIDS, hepatitis C, and tuberculosis and significantly low for paratyphoid fever.

We also investigated the correlation at the linear scale without logarithmic conversion to determine the robustness of the result. Models by linear correlation showed similar results for BARI; for example, influenza, HIV/AIDS, and hepatitis C showed significantly high BARI in both models (Fig. 1 and *SI Appendix, Fig. S1*). Hepatitis B showed significantly high BARI only in the model by linear correlation. Although the intensity for hepatitis B did not reach statistical significance in a model assuming logarithmic correlation, the index remained high in the model as well. The BARI of pneumococcal meningitis and pneumonia [*S. pneumoniae*] was low only when the correlation was tested at linear scale, although the result was not significant. Deviation of the BARI index increased statistically along with increment of DALYs when the correlation was tested at a linear scale ( $P < 0.001$ ). In contrast, the trend was not observed in the model at a logarithmic scale ( $P = 0.20$ ; *SI Appendix, Fig. S1C*). This suggested that, in the linear scale, the BARI index tended to show statistical significance, whether high or low, when their DALYs were high. Because our purpose was partly to evaluate the research intensity for diseases with low burden, we decided to use the logarithmic correlation during further analysis.

For most infectious diseases, the index did not change dramatically between the 1990s and 2010s (Fig. 2). Exceptions included significant increases in the BARI for influenza, dengue fever, dracunculiasis, and norovirus and the significant decrease for HIV/AIDS.

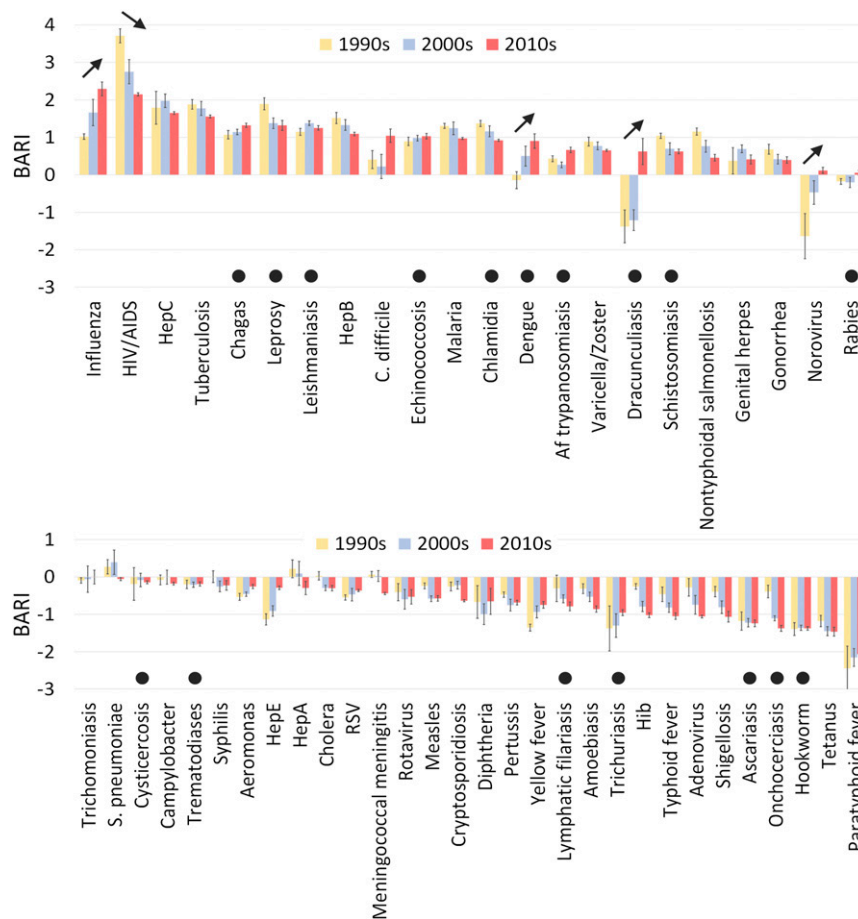
An analysis of whether the BARI differed by category of infectious disease showed no obvious tendencies by category (all  $P > 0.05$ ), although the median indices were slightly lower for diseases with middle and high compared with low disease burden, diseases due to bacteria and parasites compared with viral diseases, those transmitted from the environment compared with those with other transmission modes, diseases for which a vaccine is available compared with those without a vaccine, and non-NTDs compared with NTDs (Fig. 3A).

**Burden-Adjusted Research Intensity Patterns of Infectious Diseases by Country.** Taking the first author’s country of affiliation in a given study as a proxy for the country of that study, we analyzed the BARI at country level for 45 countries (Table 1). We identified six characteristic patterns in the indices (Fig. 3B and Table 2), including diseases for which the country level BARI: (i) was high in most countries, such as HIV/AIDS; (ii) was middle in most countries, such as chlamydial infection; (iii) was low in most countries, such as tetanus; (iv) depended on region or a country’s economic level (for example, there was high research intensity for *Campylobacter* enteritis in high-income countries and a low intensity in low and lower-middle income countries, and there was high research intensity for cholera in Asia, Western Europe, and North America, but low in Central/South America and Eastern Europe); (v) was high among affected countries and with a considerable number of publications also from nonaffected countries, such as malaria; and (vi) was low or middle from affected countries, although many nonaffected countries reported articles about the disease, such as ascariasis. *SI Appendix, Fig. S2* shows the results of the country-level analysis for all 52 infectious diseases.

## Discussion

We identified infectious diseases that have attracted high as well as low research attention considering their disease burden. No clear universal feature determined the BARI for a disease (Fig.





**Fig. 2.** Indices of BARI for the 52 infectious diseases in the 1990s, 2000s, and 2010s, calculated for each year and averaged by decade. The diseases are ordered according to the index in the 2010s. Error bars show SD of the indices of each year. Black dots indicate neglected tropical diseases, and slanted arrows indicate diseases with significant change in the BARI over the three decades.

and *SI Appendix*, Fig. S2). Paratyphoid fever, which is a non-NTD with significantly low BARI, attracted insufficient attention from researchers in most countries. Raising awareness at a global level to facilitate research may be required not only for those NTDs, but for non-NTDs with low research intensity.

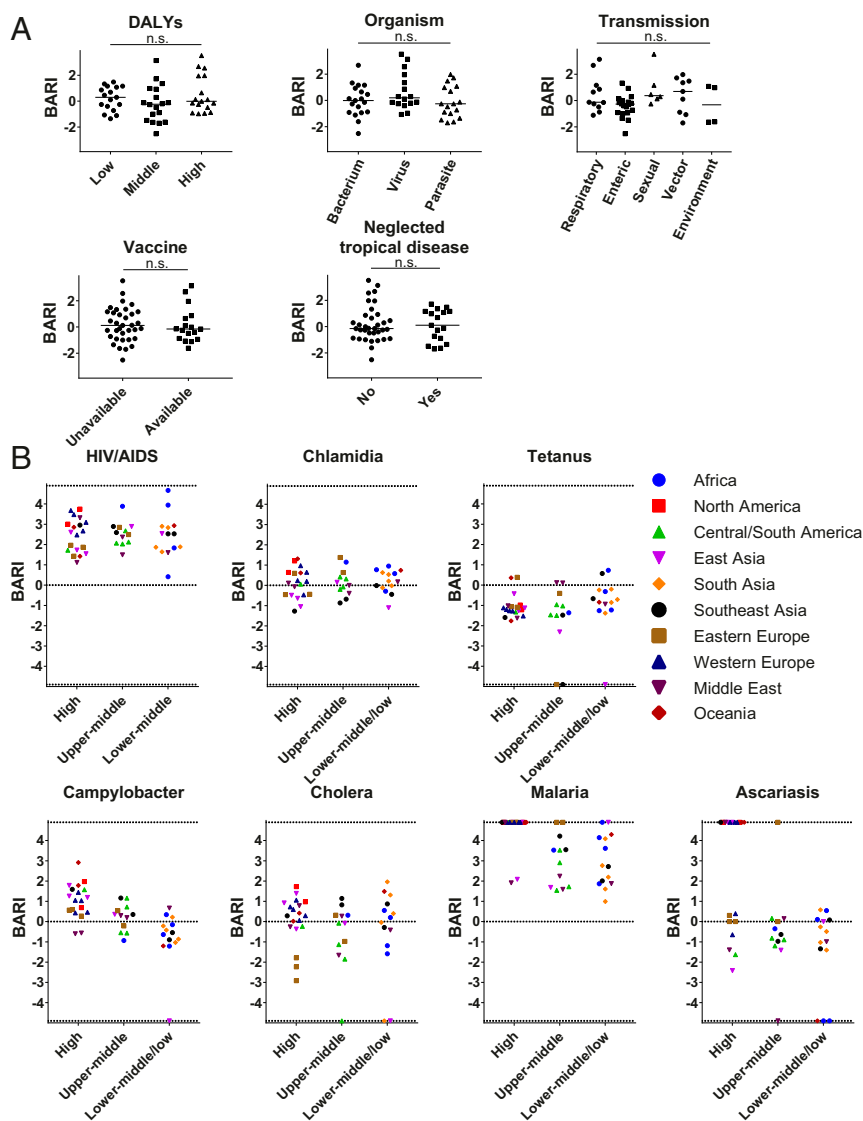
Infectious diseases with an exceptionally high disease burden, including HIV/AIDS and tuberculosis, showed high research intensity even taking their high disease burden into consideration. The large number of patients with these high-burden diseases may have encouraged a high level of investment in research resources and motivated researchers to conduct studies about the diseases. The expense for HIV/AIDS reportedly is exceptionally high for its disease burden, but attention for the disease has been decreasing (14, 15). Our results also showed high, but decreasing, research intensity for the disease (Fig. 2).

Many potential reasons exist for the highest research intensity for influenza. Its continuous genetic evolution and antigenic drift encourage researchers to study its molecular epidemiology and vaccine effectiveness every season (16). The implementation of molecular techniques, such as PCR and DNA sequencing, even in lower-middle and low income countries may have facilitated research about the disease worldwide (17, 18). The influenza pandemic in 2009 had a great impact on the number of publications about the disease (19), and the emergence and spread of the highly pathogenic avian influenza and the importance to prepare for future influenza pandemics could be further reasons for the increased research intensity (20).

Emerging infections, such as Zika fever and Ebola virus disease, do not yet cause significant DALYs worldwide, but nonetheless receive substantial research effort because of the public health risk posed by a future pandemic threat (21). Unfortunately, we did not analyze those diseases in this study because of the limited data on their disease burden.

We evaluated the relationship between BARI for infectious diseases using the number of publications and disease burden measured in DALYs, assuming a double logarithmic linear relationship between the two indicators. We also tested the correlation in linear scale and confirmed the robustness of our results (*SI Appendix*, Fig. S1). Still, we remain uncertain how research activity should be increased for diseases with a higher disease burden. The simple correlation approach in the present study will not specify how much research intensity diseases should be receiving.

Another concern is that the DALYs might underestimate the burden of some diseases. For example, deaths from renal disease caused by schistosomiasis may have been placed in the renal disease category rather than the schistosomiasis category in DALYs, measured by the global burden of disease (22). This can be especially relevant when interpreting the high research intensity for diseases with a low disease burden, such as Chagas disease and leishmaniasis. Underestimation of the disease burden might cause overestimation of research intensity for such diseases. Lastly, we referred to patterns of research activity by country, depending on the affiliation of the first author of the report. However, this also should be interpreted carefully. For example,



**Fig. 3.** Analysis of the BARI index of infectious diseases at global and country levels. (A) BARI index by disease classification at a global level. The median for each group is shown by a horizontal line. The disease burden in DALYs was grouped by tertile. Difference in indices between and among categories was tested by the Kruskal–Wallis Test. n.s., not significant. (B) The index for representative diseases at a country level by region and economic level. When there is one or more publications about a disease that has no DALYs there, dots were plotted on the top horizontal broken line. When there is no publication about a disease that has no DALYs there, dots were plotted on the middle horizontal broken line. When there is no publication about a disease that has DALYs there, dots were plotted on the bottom horizontal broken line. Results for all 52 diseases can be found in *SI Appendix, Fig. S2*.

in our study, a study on African trypanosomiasis conducted in Kenya by an investigator affiliated with an American institute was counted as a study in the United States.

Despite the limitations, our study identified infectious diseases that have received research attention or have been neglected by researchers from one viewpoint. We hope these findings provide a basis for further discussion about the more appropriate allocation of research resources to infectious diseases.

## Materials and Methods

**Data.** Infectious diseases included in the analysis and their disease burden in DALYs were extracted from the database of the Global Burden of Disease Study 2016 (1). Table 1 lists the infectious diseases and countries included in our study. The diseases were classified according to the tertile of their disease burden (low, middle, or high), organism (bacterium, virus, or parasite), transmission mode (respiratory, enteric, sexual, vector, or environmental), availability of vaccine, and designation as NTDs (*SI Appendix, Table S1*) (23, 24).

The number of studies on each infectious disease (research intensity) was obtained from PubMed using Medical Subject Headings (MeSH), a vocabulary for indexing publications in the life sciences (25). The name of a disease and causing agent were used as search terms. For example, the number of publications about malaria was obtained using the search term ["Malaria" OR "Plasmodium"]. The MeSH search terms used for each infectious disease in this study are listed in *SI Appendix, Table S1*. All types of publications, including original research, meta-analyses, reviews, and case reports, were included in the publication count.

Disease burden (measured in DALYs) and the number of publications were acquired at global and country levels. The number of publications by country was obtained using data on affiliation of the first author. The country-level analysis included five countries/territories that published the highest number of publications in the field of medicine in each region (Africa, North America, Central and South America, East Asia, South Asia, Southeast Asia, Eastern Europe, Western Europe, Middle East, and Oceania); these were selected using data from the SCImago Journal & Country Rank (<https://www.scimagojr.com/>). Countries that published <1,000 medical articles between 1996 and 2017 were excluded. Table 1 lists the countries analyzed in this study. The countries also were classified by economic level (high, upper-middle,

**Table 2. Patterns of BARI of infectious diseases at country level**

BARI pattern	Diseases
1: High in most countries	HepB; HepC; HIV/AIDS; Influenza; Nontyphoidal salmonellosis; Rabies; Tuberculosis
2: Middle in most countries	<i>Aeromonas</i> ; <i>Chlamydia</i> ; Cryptosporidiosis; Gonorrhoea; HepE; Measles; Rotavirus; <i>S. pneumoniae</i> ; Syphilis; Typhoid fever; Varicella/Zoster
3: Low in most countries	Adenovirus; Diphtheria; Genital herpes; Hib; Meningococcal meningitis; Paratyphoid fever; RSV; Tetanus
4: Depended on region or economic level	Amoebiasis; <i>C. difficile</i> ; <i>Campylobacter</i> ; Cholera; Cysticercosis; Echinococcosis; HepA; Norovirus; Pertussis; Shigellosis; Trichomoniasis
5: High in affected countries and with a considerable number of publications from nonaffected countries	Chagas; Dengue; Leishmaniasis; Leprosy; Lymphatic filariasis; Malaria; Onchocerciasis; Schistosomiasis; Trematodiasis
6: Low in affected countries although many nonaffected countries published articles about the disease	Af trypanosomiasis; Ascariasis; Dracunculiasis; Hookworm; Trichuriasis; Yellow fever

lower-middle, or low income) according to a report by the World Bank in 2017 (*SI Appendix, Table S2*) (26). All data used in the study were accessed and acquired between May 28, 2018, and October 12, 2018.

**BARI Index on Infectious Diseases.** The correlation between disease burden (in DALYs) and the number of publications was tested under a double logarithmic linear relationship with the method of least squares. For each infectious disease, its residual from the regression line was calculated. Then, we identified diseases with outlier values (outside of the two SDs) in DALYs, number of publications, and/or residual from the regression line. The correlation between DALYs and the number of publications was tested again excluding diseases with outlier values. The residual from the new regression line was calculated for each infectious disease. After standardization by dividing by SD, the values were used as the BARI index. The index was calculated using DALYs in 2010 and the number of publications in 2010–2017. Elasticity, the ratio of the percentage change in the number of publications to the

percentage change in DALYs, was calculated for this model. The correlation also was tested in the linear scale to check the robustness of the results.

Diseases with the number of publications higher and lower than 95% predication intervals of the regression line indicated significantly high and low BARI, respectively. The indices between and among categories of the diseases were compared using the Kruskal–Wallis Test. To determine the temporal trend of BARI, the index also was calculated by year and averaged by decade: 1990s, 2000s, and 2010s (Fig. 2). A change in the index of >1.0 (=single SD) over the three decades with  $P < 0.05$  by the Jonckheere test was considered significant.

**ACKNOWLEDGMENTS.** We thank Prof. Yoshio Koyanagi for support. This work was supported by the Leading Initiative for Excellent Young Researchers from the Ministry of Education, Culture, Sport, Science & Technology of Japan and Japan Society for the Promotion of Science Grant 16809810. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

- Abajobir AA, et al.; GBD 2016 DALYs and HALE Collaborators (2017) Global, regional, and national disability-adjusted life-years (DALYs) for 333 diseases and injuries and healthy life expectancy (HALE) for 195 countries and territories, 1990–2016: A systematic analysis for the Global Burden of Disease Study 2016. *Lancet* 390:1260–1344.
- Whitcher JP, Srinivasan M, Upadhyay MP (2001) Corneal blindness: A global perspective. *Bull World Health Organ* 79:214–221.
- Murray CJ (1994) Quantifying the burden of disease: The technical basis for disability-adjusted life years. *Bull World Health Organ* 72:429–445.
- Hajar R (2015) History of medicine timeline. *Heart Views* 16:43–45.
- Lederberg J (2000) Infectious history. *Science* 288:287–293.
- Washer P (2010) The conquest of infectious disease. *Emerging Infectious Diseases and Society* (Palgrave Macmillan, London), pp 20–44.
- Hotez PJ, et al. (2007) Control of neglected tropical diseases. *N Engl J Med* 357:1018–1027.
- Keenan JD, et al.; MORDOR Study Group (2018) Azithromycin to reduce childhood mortality in sub-Saharan Africa. *N Engl J Med* 378:1583–1592.
- Reperant LA, Osterhaus ADME (2017) AIDS, Avian flu, SARS, MERS, Ebola, Zika... what next? *Vaccine* 35:4470–4474.
- Okeke IN, et al. (2005) Antimicrobial resistance in developing countries. Part I: Recent trends and current status. *Lancet Infect Dis* 5:481–493.
- World Health Organization (2014) *Antimicrobial Resistance Global Report on Surveillance*. (World Health Organization, Geneva).
- Cosgrove SE, Carmeli Y (2003) The impact of antimicrobial resistance on health and economic outcomes. *Clin Infect Dis* 36:1433–1437.
- Maragakis LL, Perencevich EN, Cosgrove SE (2008) Clinical and economic burden of antimicrobial resistance. *Expert Rev Anti Infect Ther* 6:751–763.
- Global Burden of Disease Health Financing Collaborator Network (2018) Spending on health and HIV/AIDS: Domestic health spending and development assistance in 188 countries, 1995–2015. *Lancet* 391:1799–1829.
- Abdool Karim Q, Abdool Karim SS (2018) HIV—No time for complacency. *Science* 360:1153–1153.
- Carrat F, Flahault A (2007) Influenza vaccine: The challenge of antigenic drift. *Vaccine* 25:6852–6862.
- Daar AS, et al. (2002) Top ten biotechnologies for improving health in developing countries. *Nat Genet* 32:229–232.
- Bresee J, et al.; WHO Working Group on the Burden of Influenza Disease (2018) Progress and remaining gaps in estimating the global disease burden of influenza. *Emerg Infect Dis* 24:1173–1177.
- Lister SA, Redhead CS (2009) The 2009 influenza pandemic: An overview. *CRS Report for Congress* (Congressional Research Service, Washington, DC).
- Poovorawan Y, Pyunporn S, Prachayangprecha S, Makkoch J (2013) Global alert to avian influenza virus infection: From H5N1 to H7N9. *Pathog Glob Health* 107:217–223.
- Mehand MS, Al-Shorbaji F, Millett P, Murgue B (2018) The WHO R&D Blueprint: 2018 review of emerging infectious diseases requiring urgent research and development efforts. *Antiviral Res* 159:63–67.
- Herricks JR, et al. (2017) The global burden of disease study 2013: What does it mean for the NTDs? *PLoS Negl Trop Dis* 11:e0005424.
- World Health Organization (2017) Neglected tropical diseases. Available at [www.who.int/neglected\\_diseases/en/](http://www.who.int/neglected_diseases/en/). Accessed May 28, 2018.
- Centers for Disease Control and Prevention (2018) Which diseases are considered neglected tropical diseases? Available at <https://www.cdc.gov/globalhealth/ntd/diseases/index.html>. Accessed May 28, 2018.
- U.S. National Library of Medicine (2018) MeSH. Available at <https://www.nlm.nih.gov/mesh/meshhome.html>. Accessed June 17, 2018.
- World Bank (2018) World Bank country and lending groups. Available at <https://datahelpdesk.worldbank.org/knowledgebase/articles/906519-world-bank-country-and-lending-groups>. Accessed June 10, 2018.