Visualization in clinical decision support system for antibiotic treatment

Humberto Garcia-Caballero[1], Manuel Campos*[1], Jose M. Juarez[1], and Francisco Palacios[2]

[1] Computer Science Faculty. University of Murcia
[2] University Hospital of Getafe

Abstract. The inappropriate use of antibiotics causes a decrease in their efficacy to defeat infections due to the resistance developed by microorganisms. In a hospital context, the antibiograms play a fundamental role to select the right antibiotic for each patient. The objective of this work is to use and extend data visualization techniques to boost the understanding of the antibiograms in a hospital as part of clinical decision support system for infections management. In this work we present and analyse three models of visualization that cover the information needs in the identified clinic tasks and the assessment of the users' performance and the users' experience developed by physicians.

1 Introduction

Community-acquired and nosocomial[3] infections are considered nowadays as one of the main causes of the global morbidity [1]. Since the creation of antibiotics, they have been recognized as the principal weapon for infection treatment. However, the passing time showed that a continuous use of antibiotics produced that the organism would generate resistance to them. This fact poses a threat for the effective treatment of bacterium infectious diseases. A huge amount of factors influence in the increasing of the antibiotics resistance [2]. Some factors depend on the patient's health status, such as immunocompromised (e.g. VIH, neoplasia, etc.), or long hospitalizations with prolonged treatments and permanent devices (e.g. peripheral lines or catheter) which may stimulate the growth of resistant bacterial strains.

Empirical treatment is the treatment which is done firstly before laboratory results are available, that is, before having information about the microorganism that caused the infection. For this reason, this treatment is based on heuristics and expert rules. Information about previous patients that can guide the prescription of antibiotics is given in the yearly cumulative antibiogram. The Clinical and Laboratory Standards Institute (CLSI) [10] defines an antibiogram as a profile of all microorganisms susceptibilities respect to the antibiotics.

* Corresponding author:manuelcampos@um.es

[3] An infection is nosocomial or hospital-acquired if it appears 48 hours after patient's admission.

A very relevant aspect of clinical decision support systems is the visualization [4]. The aim of techniques for visual analysis of data is to explode human perception to help in the understanding, analysis and communication of data, models and concepts. A number of visualization tools has been previously used in clinical contexts. For example, the goal of IPBC [3] was to improve the therapeutic adjustment in haemodialysis using a three-dimensional visualization. KHOSPAD [8] was designed to identify temporal relationships between primary care doctors, the patient and hospital stays. KNAVE-II [9] aims at visualizing oncological treatment information. VISITORS system [5] combines intelligent temporal analysis and information visualization techniques to integrate, process and visualize information belonging to multiple patients and information sources. Nevertheless, little attention has been paid explicitly to clinical decision support systems for empiric treatment.

The aim of this work is to develop visualization techniques of the cumulative antibiograms to help the clinicians in the prescription of empiric treatment. In these models, the most relevant clinic concepts must be represented for the antibiotics selection, such as the prevalence and efficacy. Apart from analysing the models properties, clinic users have answered to a survey for evaluation of users' performance (UP) and the users' experience (UE) [6]. In the survey we asked the to solve the clinical tasks needed for the empiric treatment we have previously identified.

The structure of this paper is as follows. In Section 2 we present the concepts and visualization models adapted to the specific tasks. In Section 3 we evaluate the characteristics of the models and show the survey submitted by clinicians. Finally, in Section 4 we provide the conclusions and future works of this research.

2 Concepts and visualization models

The underlying clinical concepts behind the visual models for empiric treatment are prevalence, sensitivity, resistance and efficacy. In the clinic practice, the sensitivity of a microorganism towards an antibiotic is proved through a laboratory test concluding that the microorganism is sensitive for the antibiotic in case it can defeat it, or resistant otherwise. The prevalence shows the probability of a microorganism to be cause of an infection. We define these concepts as follows.

Definition 1. *Let* $M = (M_1, M_2, M_3, \cdots, M_n)$ *be a vector of positive integer values such as for each* $i = 1, 2, 3, \cdots, n$, M_i *represents the number of samples for i-th microorganism. Let* $|M|$ *be the sum of all microbiology cultures,* $|M| = \sum_{i=0}^{n} M_i$. *The i-th microorganism's prevalence is defined as*

$$Prevalence(M_i) = \frac{M_i}{|M|} \cdot 100.$$

Definition 2. *Let* $S = (S_{1,1}, S_{1,2}, \cdots, S_{1,m}, S_{2,1}, S_{2,2}, \cdots, S_{m,n})$ *and let* $R = (R_{1,1}, R_{1,2}, \cdots, R_{1,m}, R_{2,1}, R_{2,2}, \cdots, R_{m,n})$ *with* $0 < i \leq n, 0 < j \leq m$ *where* $S_{i,j}$ *represents the number of sensitive samples for the i microorganism and the*

j antibiotic, and $R_{i,j}$ represents the number of resistant samples for the i-th microorganism and j-th antibiotic. Therefore, i-th microorganism's sensitivity and resistance for the j-th antibiotic is defined as

$$Sensitivity(M_i, A_j) = \frac{S_{i,j}}{S_{i,j} + R_{i,j}} \cdot 100,\ Resistance(M_i, A_j) = \frac{R_{i,j}}{S_{i,j} + R_{i,j}} \cdot 100.$$

In epidemiology, in order to communicate in an effective way the result of a procedure, to be more specific of a treatment, it is normally used a measurement that is expressed in a number of patients instead of in a percentage [7]. We have defined efficacy in terms of Number Needed to Fail (NNF) and Number Needed to Succeed (NNS) as shown in Definition 3.

Definition 3. *Efficacy or NNF and NNS are defined as*

$$NNF(M_i, A_j) = \begin{cases} \lceil \frac{100}{Resistance(M_i, A_j)} \rceil & \text{if } R_{i,j} \geq 1 \\ 100 & \text{in other case.} \end{cases}$$

$$NNS(M_i, A_j) = \begin{cases} \lceil \frac{100}{Sensitivity(M_i, A_j)} \rceil & \text{if } S_{i,j} \geq 1 \\ 100 & \text{in other case.} \end{cases}$$

Regarding Definition 3, the principal difference between NNF and NNS is the divisor used. In the first case, the percentage of the resistance is used, while in the second case the sensitivity percentage is considered. The range of these two measures is [1, 100]. High values of NNF indicate more effective antibiotics. However the NNS must be interpreted in a reverse way. That is, high values would indicate that the antibiotic is very ineffective. In the Definition 4 we introduce the accumulated efficacy as the weighted efficacy according to its prevalence.

Definition 4. *The accumulated efficacy or NNFA and NNSA are defined as*

$$NNF_{accumulated}(A_j) = \frac{\sum_{i=1}^{m} NNF(M_i, A_j) \cdot Prevalence(M_i)}{\sum_{i=1}^{m} Prevalence(M_i)}.$$

$$NNS_{accumulated}(A_j) = \frac{\sum_{i=1}^{m} NNS(M_i, A_j) \cdot Prevalence(M_i)}{\sum_{i=1}^{m} Prevalence(M_i)}.$$

From a technical point of view, in this work we propose these models to represent and compare several concepts simultaneously: intensity of one to many and many to many relationships, and proportions. From the clinical needs point of view, we have identified the following visual models needs that will be used later in the evaluation:

Item 1: To detect all the microorganisms found in a specific type of sample.
Item 2: To detect all the microorganisms that an antibiotic is active against.
Item 3: To detect all the microorganisms that an antibiotic is not active against.

Item 4: To detect all the antibiotics that are effective for a microorganism.
Item 5: To detect all the relationships between microorganisms and antibiotics.
Item 6: To use the same model for the empiric treatment and the directed one.
Item 7: To rank microorganisms and antibiotics according to prevalence and efficacy.
Item 8: To represent the prevalences and to compare proportions.

Next, we introduce the Sunburst, Bipartite and Tree models for the representation of the previous concepts and the resolution of clinical tasks.

2.1 Sunburst model extended for empiric treatment

The Sunburst model presents a hierarchical circular multilevel representation that is commonly used in bioinformatics [11]. Each one of the levels is divided into parts in different sections to show the relationships between them. The Sunburst model zoomed in Figure 1 is focused on one microorganism. The next level deals with the active antibiotics against the selected microorganism. Finally, the third level shows the susceptibility (sensitivity and resistance) of the microorganism for each antibiotic. The user's interaction as in the original model allows to display information of a level filtered by its prior level.

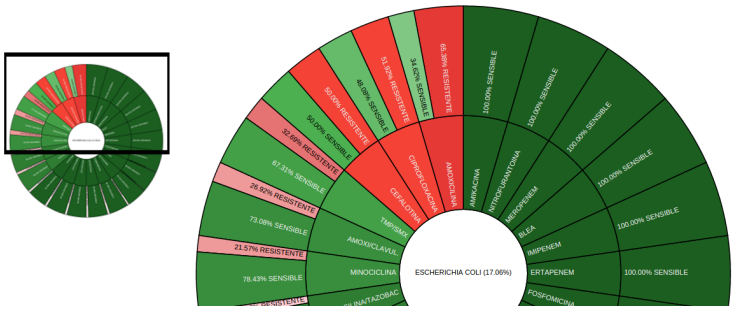

Fig. 1. Sunburst model for the *escherichia coli* microorganism.

The microorganisms are sorted clockwise, starting at 0 hours, from lower to higher sensitivity of to each one of the antibiotics. The length of the arch shows the proportion of the relationships.

We have defined a color codification in the visualization models elements that is common to the three models used. The more sensitive a microorganism is to an antibiotic, the more intense and dark green color. Conversely, the more resistant it is, it will turn into a more intense and dark red color.

This model has also several limitations. Readability is only good for the concepts in the "first hours" of the clock, and difficult for the last ones. In this model it is not easy to compare the sensitivity of two microorganisms that are far in the model without textual information; it is clear that arch length are not so easy to compare are the areas or length of bars. Besides, antibiotics in the second

level are repeated for different microorganisms and relations between several microorganisms with several antibiotics are difficult to identify and compare.

2.2 Bipartite model extended for empiric treatment

The Bipartite model showed in Figure 2 is a type of graphic that allows to represent through a different width channel the amount of relationships between two set of separate elements, the microorganisms and antibiotics in this case.

As in the previous model, we have included textual information to indicate with enough precision (apart from the visual elements) the sensitivity, the resistance and the efficacy. The user's interaction allows to highlight the relations between antibiotics and microorganisms.

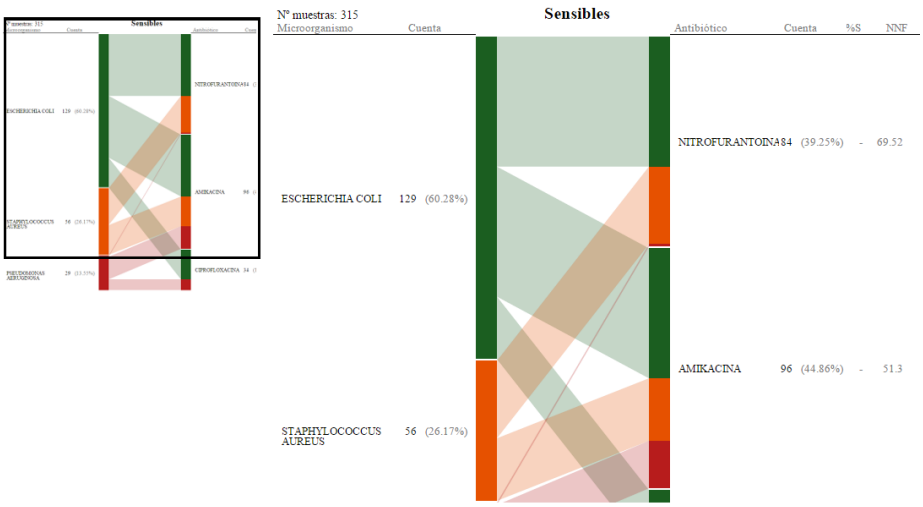

Fig. 2. Bipartite model with multiple microorganisms and antibiotics.

The microorganisms are sorted from higher to lower prevalence, while the antibiotics are ordered by efficacy. In this way, the clinic can focus rapidly on the most relevant elements because they appear in the upper zone of the model.

We use the length of the microorganism's bar to represent and the prevalence. The color codification in the microorganisms is also associated to their prevalence: The green, yellow and red tones belong to the 50% accumulated prevalence, the 80% and the 100% respectively.

In the case of the antibiotics, the length of the bar shows the amount of appearances in the antibiograms, whereas the colors allow us to identify rapidly if it is active against the most prevalent organism.

This model allows to select dynamically a unique element to focus the attention on it, being a microorganism or an antibiotic, and the information is

recalculated for that specific element. For instance, with an antibiotic, its individual efficacy instead of the cumulated one is shown. In this model we have no identified restrictions from the technical point of view.

2.3 Tree model extended for empiric treatment

The Tree model, as it is showed in the Figure 3, represents a hierarchical structure. In the first level of the Tree all associated microorganisms to a type of sample appear. The second level represents the antibiotics which are susceptible to the superior level microorganism. Dynamically a third level associated to the antibiotic shows the resistant microorganisms for that antibiotic. The user can interactively expand each level of the tree.

The color code used in the first and third level are exactly the same ones as those described before. We included textual information with all the samples, both the sensitive and the resistant ones of the microorganisms. The microorganisms and antibiotics follow the same sorting criteria and leading to the same consequences that we previously commented.

Moreover, the same level uses the area to represent the prevalence of the microorganisms, while in the second level it is used to represent the resistance to an antibiotic with an inner red circle.

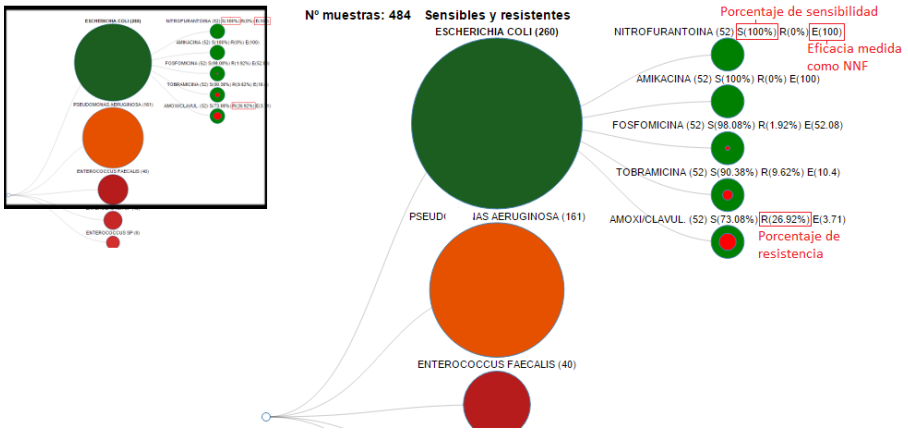

Fig. 3. Tree model with second expanded level.

A limitation of this model is that it is not able of showing a unique graphic with the relationships between all microorganisms and all antibiotics.

3 Evaluation of the visualization models

The evaluation of models consists of two different parts. On the one hand, the capability to represent the clinical concepts needed for decision making in the

empiric treatment. On the other hand, the UP and the UE of the clinical professionals. This study is done using a web-based development over D3.js.

3.1 Representation capability

Table 1 summarizes all the items defined in the Section 2. The Bipartite model does not present any technical limitation for representing all concepts and for solving the clinical tasks identified.

The second best model according to the items would be the Tree model. This model is able to represent the relationships of each microorganism with each antibiotic, as well as being able to use the same chart for the empiric treatment and the directed one. The hierarchical structure is not as expressive as the Bipartite model for our tasks. Nevertheless, the visualization of proportions is more comfortable because of comparison of the area of a circle is easier than the length of a bar.

Finally, Sunburst model is ranked in last position because it does not accomplish several items. Despite of being visually attractive, this model has serious limitations for the decision making in the empiric treatment.

	Sunburst	**Bipartite**	**Tree**
Item 1	No	Yes	Yes
Item 2	Yes	Yes	Yes
Item 3	No	Yes	Yes
Item 4	Yes	Yes	Yes
Item 5	No	Yes	Yes
Item 6	Yes	Yes	No
Item 7	No	Yes	No
Item 8	Only susceptibilities	Yes	Yes

Table 1. Comparative table of models' representation capabilities.

3.2 User's performance and experience

We have evaluated the UP and UE on empiric treatment with a survey [6]. The survey is structured according to the following parameters: task solving (it measures model's efficacy through accuracy, miss rate, specificity and sensitivity), usability (for detecting user's trouble by means of subjective comments), domain (we modify the quantity of data shown), correctness and time employed to solve the questionnaire. We have selected Bipartite and Tree visualization models to do the evaluation, excluding Sunburst model on account of its limitations for this problem.

The survey has been verified by clinic staff to assure the language used is the usual one of clinicians and it does not contain ambiguous terms. Besides, users have to follow a 10 minutes video tutorial before completing the survey. The survey contains 22 mixed questions in total. Questions related to models are closed with policotomic single selection answers, and they include the option

"don't know" (DK) for those cases where the model is not able to communicate all the necessary information to make correct decisions. The answers have no scale, because the aim is to identify the correct answer.

What is more, the survey is organised in blocks related to distinct tasks and clinical concepts: identification of prevalence, identification of activity, and identification of resistance. Questions are ordered by a funnel-like technique from general to specific. Also, the survey has two parts: the first one evaluates both models individually (Bipartite and Tree); the second one evaluates both models in a compared way, that is, one sentence is presented to the user and he or she has to decide which of them reflects such affirmation.

Survey's population consisted of 9 participants, 7 of them were physicians and 2 computer science engineers, with an average age of 26.5 years old. The participants needed, on average, 18 minutes (6 minimum and 32 maximum) to finish the survey. Google Forms was used to design and fill the survey.

Figure 4 shows the results for the questions about the individual models. The average hit rate is 64.84%. *"Don't know"* option was selected, on average, just the 6.25% of times. Task A, about microorganism prevalence both models have similar behaviour. Task B, about the most active antibiotic identification, the Tree model was easier for users. Questions 5 and 6, about the identification of the set of microorganism with highest prevalence for a set of antibiotics, were specially difficult. However, task C, about the identification of the most effective antibiotic, the Bipartite model was the most useful for respondents. Note that task D about identification of resistances was only defined for the Bipartite model since it was not possible to represent the same concept with the Tree model. For that task, the model showed to be useful.

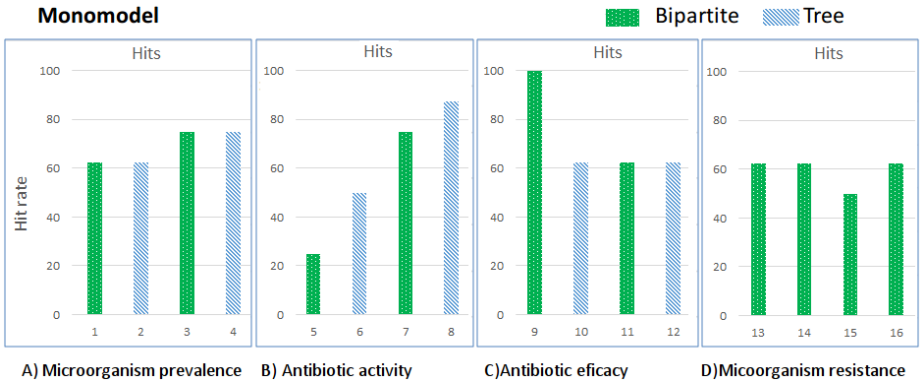

Fig. 4. Hit rates for the tasks in monomodel questions.

Figure 5 shows hit rates and UE for the compared analysis of the two models. These questions presented the same scenario in both models simultaneously but only one had the correct answer. Users showed more confidence comparing the

two models simultaneously than for the individual model questions since the rate of correct answers is higher and they did not select the DK option.

For the task A, about identifying the prevalent microorganisms, users did not show preference between the models and both models enabled to solve the task correctly in 90% of times. For the task B, about identifying the antibiotic activity, users preferred clearly the Tree model although hit rate was high in both. Although task C, about identifying the most effective antibiotic, its turn out to be more difficult and users did not show a clear preference between models.

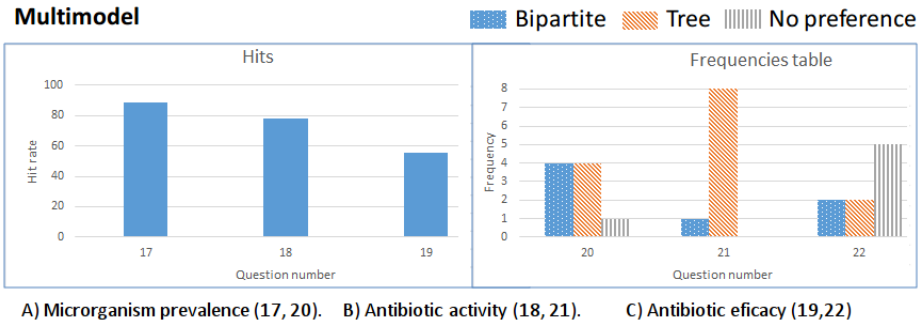

Fig. 5. Hit rate and user's experience for tasks in multimodel questions.

Regarding to user's comments, we highlight that they preferred a voice guided instruction and subtitles contained in the video tutorial.

We found two limitations in the evaluation. Firstly, clinic users do not have a high level of experience in the application of antibiotic treatment, however, they are used to work with new technologies. Secondly, although clinical situations shown are based on reality, we did not include questions with a different volume of data, so it is not possible to identify which model is the preferred one for simpler situations or more complex ones.

4 Conclusions

The goal of this work was to improve the clinical decision support system for bacterial infection management by means of visualization models to help during the prescription of empiric treatment. The empiric treatment is based on expert rules that need to include local epidemiology of each hospital. This epidemiology is presented annually in the hospital on a static way through the cumulative antibiogram given in a tabular form and with limited capability regarding to decision making.

In this work we have proposed three visual models, Sunburst, Bipartite and Tree, that we extended and made use of visual elements such as areas, lengths, colours and associations in order to represent concepts and to solve clinical tasks needed during selection of adequate empiric treatment.

We evaluated the models in two ways: we analysed their representation capabilities and user's performance and experience. Regarding the intrinsic characteristics, Bipartite model is the only one that accomplish all items identified. However, users preferred the Tree model for the task of identifying the most active antibiotic. It could be because users are used to hierarchical representation. For the rest of the tasks, users have not shown preferences for any model.

The current lines of work are based on developing a stratification system more complex that enables parameters selection in a more accurate way. We are also working in the elaboration of dual cross antibiogram visualization.

Acknowledgement This work was partially funded by the Spanish Ministry of Economy and Competitiveness under project WASPSS (Wise Antimicrobial Stewardship Program Support System, ref TIN2013-45491-R), and by European Fund for Regional Development (EFRD).

References

1. Policy statement on antimicrobial stewardship by the Society for Healthcare Epidemiology of America (SHEA), the Infectious Diseases Society of America (IDSA) and Pediatrics Infectious Disease Society (PIDS), 2012.
2. FJ Candel, M Matesanz, I Candel, C Betriu, and JJ Picazo. ¿ qué hacer frente al aumento de la resitencia a los antibióticos. *JANO*, 64(1462):35–40, 2003.
3. Luca Chittaro, Carlo Combi, and Giampaolo Trapasso. Data mining on temporal data: a visual approach and its clinical application to hemodialysis. *Journal of Visual Languages & Computing*, 14(6):591–620, 2003.
4. Carlo Combi, Elpida Keravnou-Papailiou, and Yuval Shahar. *Temporal information systems in medicine*. Springer Science & Business Media, 2010.
5. Denis Klimov, Yuval Shahar, Meirav Taieb-Maimon, et al. Intelligent interactive visual exploration of temporal associations among multiple time-oriented patient records. *Methods Inf Med*, 48(3):254–262, 2009.
6. Heidi Lam, Enrico Bertini, Petra Isenberg, Catherine Plaisant, and Sheelagh Carpendale. Empirical studies in information visualization: Seven scenarios. *Visualization and Computer Graphics, IEEE Transactions on*, 18(9):1520–1536, 2012.
7. Andreas Laupacis, David L Sackett, and Robin S Roberts. An assessment of clinically useful measures of the consequences of treatment. *The New England journal of medicine*, 318(26):1728, 1988.
8. Francesco Pinciroli, Luisa Portoni, Carlo Combi, and Francesco Fabio Violante. WWW-based access to object-oriented clinical databases: the KHOSPAD project. *Computers in biology and medicine*, 28(5):531–552, 1998.
9. Yuval Shahar, Dina Goren-Bar, David Boaz, and Gil Tahan. Distributed, intelligent, interactive visualization and exploration of time-oriented clinical data and their abstractions. *Artificial intelligence in medicine*, 38(2):115–135, 2006.
10. Antonia Zapantis, Melinda K Lacy, Rebecca T Horvat, Dennis Grauer, Brian J Barnes, Brian O'Neal, and Rick Couldry. Nationwide antibiogram analysis using nccls m39-a guidelines. *Journal of clinical microbiology*, 43(6):2629–2634, 2005.
11. Yongnan Zhu, Liang Sun, Alexander Garbarino, Carl Schmidt, Jinglong Fang, and Jian Chen. Pathrings: a web-based tool for exploration of ortholog and expression data in biological pathways. *BMC bioinformatics*, 16(1):165, 2015.