



7-31-2019

Mata-Data: Data visualization in approaching evidence-based health policy

Rodri Tanoto

Faculty of Medicine, Universitas Indonesia, rodri.tanoto89@ui.ac.id

Levina Chandra Khoe

Faculty of Medicine, Universitas Indonesia

Retno Asti Werdhani

Faculty of Medicine, Universitas Indonesia

Follow this and additional works at: <https://scholarhub.ui.ac.id/ajce>



Part of the [Social and Behavioral Sciences Commons](#)

Recommended Citation

Tanoto, Rodri; Khoe, Levina Chandra; and Werdhani, Retno Asti (2019). Mata-Data: Data visualization in approaching evidence-based health policy. *ASEAN Journal of Community Engagement*, 3(1).

Available at: <https://doi.org/10.7454/ajce.v3i1.145>

This Research Article is brought to you for free and open access by the Universitas Indonesia at UI Scholars Hub. It has been accepted for inclusion in ASEAN Journal of Community Engagement by an authorized editor of UI Scholars Hub.

Mata-Data: data visualization in approaching evidence-based health policy

Rodri Tanoto^{a*}, Levina Chandra Khoe^a, Retno Asti Werdhani^a

^a*Faculty of Medicine, Universitas Indonesia*

Received: December 4th, 2018 || Revised: December 30th, 2018 || Accepted: July 26th, 2019

Abstract

Data and information is the key to implementing evidence-based health policy, which has been problematic in Indonesia. The communication between researchers and policymakers must be further strengthened by encouraging data visualizations as the narrative tools to optimize those communications. Mata-Data (<https://mata-data.com>) aimed to present national data on health and its social determinants in dynamic data visualization, which was previously shown in tables only. This program was developed in several steps. First, data extraction and analysis from data bank and official reports. Second, website development subcontracted to external consultants. Third, trial and feedback conducted using online forms, both in quantitative and qualitative data. Fourth, evaluation and dissemination. Several significant findings were established. At first, the limitation of data types (mainly categorical and ordinal) means that the multivariate display of data in one graphic was unavailable. However, in general, the user experiences were favorable. The superiority of data visualizations was agreed by the users. Nevertheless, several concerns emerged. First, the use of a mobile device was growing. Second, the need for the user for reference points, instead of comparing between provinces. Last, the interest in geographically more in-depth data, especially on the city district level. In conclusion, it can be observed that this program has evolved into something specific for the Indonesian context. Further analysis of various concerns is needed to further how Mata-Data can improve data communication in Indonesia.

Keywords: health data; determinants of health; data visualization; evidence-based health policy; informed decision

1. Introduction

Evidence-based medicine (EBM) is an approach to improve medical practice by emphasizing and categorizing evidence from well designed and well-conducted researches to inform decision making. Inspired by this concept of evidence-based practices, various social sectors have adopted the idea previously known only within medicine, namely evidence-based education, evidence-based management, and evidence-based philanthropy, also in health sectors, evidence-based health policy and evidence-based public health programs. Baicker and Chandra (2017) have defined three characteristics required by evidence-based health policy to produce a rational approach in making policies. First, plans must be well-defined. Slogans and jargons are not enough.

*Correspondence Author: rodri.tanoto89@ui.ac.id

By avoiding specificity, slogans sidestep the importance of assessment on the effectiveness and implementation strategy required on every policy. Second, plans must be differentiated from goals. A system may have different goals for different people. Therefore, clear guidance on intended targets is essential to measure the intended policy. And lastly, plans must be based on evidence about the magnitude of the effects of the program. Conceptual frameworks may show the expected results, but never the quantity. In the same breath as Baicker's and Chandra's proposal, Jacobs et al. (2012) argued that in designing an evidence-based health policy or program, six key elements are required. First, incorporating community in assessment and making decisions. Second, utilizing data and information systems systematically. Third, making a decision based on the best available evidence. Fourth, applying the conceptual framework in planning and executing programs or policies. Fifth, concluding the plans or programs with sound evaluation. Lastly, disseminating the lesson learned. Therefore, it can be found that data and information are an essential key in implementing evidence-based health policy. Data, mainly public based data, is not only the basis for community need assessments, but also potentially screening the evidence in planning system or program.

In Indonesia, the implementation of evidence-based health policy was still problematic. Budiharsana (2017) shown in her article titled "Increasing Use of Research Findings in Improving Evidence-based Health Policy at National Level" that policy formulations in the Ministry of Health were not grounded on any studies. In designing policies, the Ministry of Health has never requested for any measurable data and information. On another hand, research teams from both within and outside of Ministry has never been asked to give any recommendation related to formulation and analysis of policies. It is no wonder that even the Ministry's research and development team have never developed researches related to any Ministry of Health's programs. Therefore, to utilize data and information as the key to evidence-based health policies, communication between researchers and policymakers should be further strengthened.

Brownson, Chriqui, and Stamatakis (2009) mentioned that in the process of previously mentioned communication between researchers and policymakers, data must be presented in the form understandable by the policymakers. If the data is well presented, it is possible for it to present global burden of diseases, to highlight health priorities, to display the relevance towards the community, to explain benefits and harms for specific

interventions, to narrate the data as the representation of impacted community, and also to calculate the cost of intervention.

Even though, as has been mentioned above, many kinds of research done by the Ministry of Health have not been related to the policies within Ministry itself, many surveys and studies portraying the health conditions and their social determinants since several years ago, such as National Health Survey, Basic Health Research, and Indonesian Health Demographics Survey have been conducted by Ministry of Health, by itself or in collaboration with Statistics Indonesia (*Badan Pusat Statistik*). Unfortunately, as has been described in the paragraph above, those data—however, limited they are in portraying health conditions, challenges and opportunities in Indonesia—has never been applied to honestly inform on how policies and programs are being planned within Ministry itself (Budiharsana, 2017). Even worse, non-governmental sectors, such as researchers from universities and research institutes, cannot acquire the raw data from the individual surveys without paying a significant amount of fund. The available data citable from the open-access reports are already aggregated in the city, provincial, or regional level of Indonesia, and they are always presented in long abstruse tables within the stories. This further complicated the utilization of health data and social determinants in measuring the condition of health in Indonesia. Therefore, innovation is needed to present the available data on health and social determinants of health, however, limited those aggregated secondary data are.

In the last few years, data visualization, both static and dynamic one, has become the solution in introducing "stories," and it has been massively implemented by different organizations, such as journalistic purposes, an activist organization, and even political parties. The development of the online world has encouraged interactive data visualizations as a narrative tool easily understandable from various social aspects (Segel & Heer, 2010). McCosker and Wilken (2014) argued that data visualization should not be perceived as the result of the problem-solving process, but instead as a method to deliver a problem that needs resolving. Therefore, data visualizations should be an innovative answer to optimize data on health and its social determinants.

The advantage of data visualization is based on the picture superiority effect, describing how image-based stimuli can be retained longer in memory compared to word-based stimuli (McBride & Anne Doshier, 2002). This phenomenon has been analyzed in hundreds of journals, but one of the persisting theoretical foundations of it is

dual-coding theory, describing how the image can be coded in memory through two pathways, verbal and image coding (Paivio, 2014). The utilization of data visualization in policymaking has been much recorded, and therefore, formed the theoretical basis that supported this program (Reuters Institute for the Study of Journalism, 2015).

One of the most famous examples of data visualization in health is Gapminder (Angotti, 2017). The idea that backed up Gapminder was that data should be utilized in storytelling, to share a fact-based worldview that everyone can understand (Tarran, 2017). Gapminder has been employed for various application, such as merging statistics education and social justice (Poling & Naresh, 2014), exploring a better paths for developing societies (either citizen engagement, or strong individual leadership) (Singh, 2014), or even as simple as challenging 'fake news' and 'alternative facts' in this era (Lee, 2017).

Therefore, considering the current limitation of available data presentation in the Indonesian context, and the superiority of data visualization, Mata-Data program (<https://mata-data.com>) was designed and aimed to present data of health conditions and social determinants in dynamic data visualization. The targeted recipients are the public, especially academia, researchers, students, and activists working in health. The end output of this program is to invite participation and increase public understanding of health condition, which hopefully, will inform the policymaking process in Indonesia.

2. Methods

Mata-Data program (<https://mata-data.com>) were developed for six months from mid-2018. The development of the Mata-Data program consisted of several steps. First, data extraction and analysis. After thorough screening through data bank and official reports of Ministry of Health and Statistics Indonesia, data were extracted from both data banks and stories from both institutions, such as National Health Survey, Basic Health Research, and Indonesian Health Demographics Survey. Because all those data have been available in the public domain, no further approval is needed. After screening all the data and all data was extracted manually by the team into spreadsheets. Those data would later be analyzed to determine the suitable types of diagrams, depending on the data types (numerical, ordinal, nominal). To ensure full integration between the spreadsheets and

the intended data visualization website, the worksheets are automatically produced by the website processing.

Second, data visualization website development. The plan was to subcontract it to the external consultant. The team gave a thorough briefing about the expected user interface and the processing behind the scene, and the consultant would provide recommendations and comments based on the feasibility of the intended interface. In summary, the input data will be uploaded as tables, and those data will be later processed by the website to produce graphs and diagrams suitable with the uploaded data. To ensure the feasibility to update the data, the site was designed so that by inputting the required parameter, it will produce a previously mentioned spreadsheet that would later be downloaded. The downloaded spreadsheet contained the needed column titles, which could then be filled with data extracted from the reports. This spreadsheet would later be reuploaded to the website altogether with the filled-in data based on the stories and data banks.

Third, trial and feedback. The experiment was conducted for a month, and feedback will be actively requested by the intended users. After the website has been completed with the intended data uploaded, it was launched, and the intended users were proactively contacted to try the data visualization website. Using social media, interest groups, academic forums, and personal contacts, intended users are contacted and invited to the trial and to give feedback. The intended users for this trial were educational professionals, researchers, governmental officials, and health practitioners that were actively interested in using health data and its social determinants.

To ensure a broad reach of audiences in the trial and feedback process, feedback for the trial was collected using the online form. The feedback was collected as quantitative and qualitative data. The online questionnaire was developed and validated by an expert team and followed by piloting with ten users with similar demographics with the intended respondents.

Fourth, evaluation and dissemination. The quantitative data were analyzed and delivered descriptively using narratives and tables. All qualitative inputs will be analyzed thematically. First, the data would be coded as it was descriptively, that all the code would be grouped thematically. All the themes would be linked towards the previously established working framework of this community engagement programs. All inputs went directly to the website improvement.

To ensure program sustainability, two approaches will be considered by the end of this program. First, this data visualization website would be self-sustainable from donation and ads. Second, this website would later be integrated as part of the UI website. The site will also receive the newest data available to ensure its relevance.

3. Result and Discussion

After data screening and extraction, it has been found that almost all variables were available in the categorical and ordinal form, which limited the possible variation of charts and correlation of variables. Moreover, the variables collected were different from year to year, and several variables had separate categories for each year, further complicated the intended annual comparison. These findings were later informed to the web developers for interface adjustments.

Mata-Data website contained several labels; Home, About Us, Disclaimer, User Guide, and Contact. The diagram on the Homepage included several groupings for the data of choice, and every chart would be interactive. The graph displayed can be downloaded as still images to be used for data dissemination. About Us page described what Mata-Data was about and the background of why it was created. The User Guide page provided guidelines and videos for using the website. The Contact contained office address, email, and telephone number. Figure 1 displayed the screenshots of the website.

In the developer settings, all data labels were editable, and new data or variables can be introduced by filling in and uploading offline spreadsheets automatically produced by the program by entering the required variable names. Various graphs and diagrams were available, such as maps, pie chart, and stacked bar chart.

The development of the data visualization website faced some difficulties because of the communication gap between IT developer and the researchers. After prolonged growth of the site, trial and feedback were conducted for a whole month, and the targeted users were contacted via social media, interest groups, academic forums, and personal contacts. Because of the broad reach of social media, it was difficult to estimate the response rate. Nevertheless, thirty-three responses were collected and analyzed.

The trial for the website reached various users from different academic backgrounds, which were medicine (53.13%), dentistry (6.25%), nursing (6.25%), midwifery (9.28%), pharmacy (3.13%), public health (6.25%), and even engineering, IT, chemistry, and

statistics. The occupational backgrounds were also varied, namely academia (18.75%), researchers (6.25%), students (15.63%), health practitioners (31.25%), government officials (21.88%), and also professional bodies (6.25%). Therefore, it can be concluded that all three main actors in policymaking (practitioners or citizens, governments or policymakers, and researchers or academia) have been involved in the trial and feedback processes. In general, the quantitative scores from the feedback questionnaire (ranged 1-10) were favorable. Further details can be viewed in Table 1. All data were not normally distributed.

Table 1. Data Visualization Website (*Mata-Data*) Trial User Experience

Questions	Parameters	Score (1-10)
Are the graphical presentations attractive enough?	Median (IQR)	7 (4)
Are the presented graphics easy to understand?	Median (IQR)	7 (2.5)
Are the available graphics informative enough?	Median (IQR)	7 (2.5)
Are the available graphics useful for your work?	Median (IQR)	8 (2)
In general, will this data visualization be useful for understanding the health situation in Indonesia?	Median (IQR)	8 (2.5)

Beside quantitative measurement, the trial feedback forms were analyzed qualitatively in a thematic approach. As expected, comments and feedback on the website technicalities, especially on the user interface, such as colors, fonts, and types of graph, emerged. However, other than technicalities, several interesting themes emerged from the thematic analysis of the qualitative answers in the questionnaire. First, the incompatibility of mobile display. This website was designed with a screen size of the personal computer as intended user experience, because of the complexities of the graphs.

Moreover, it was assumed that online researches would be mostly done from a personal computer. However, the trial feedback showed that mobile use in displaying the website was quite prominent. This may not come as a surprise, considering that mobile phones dominated the share of web traffic by a device with 69% (Kemp, 2017).

Second, the need for reference. Many users during the trial were asking for recommendations, in terms of a specific threshold within the charts or diagrams, in the form of national median or other target indicators. This showed that the users were more interested in comparing the provincial level with a specific reference, instead of comparing among provinces.

Third, the depth of data. There were specific requests for city/district level of data to be covered in this program. The arguments stated that because of the health decentralization which put the level of authority for health governance in city or regency level, the data in this governmental level became more important than national or provincial level data aggregation.

The limitation of data available means that the multivariate display of data in one graphic became unavailable in our context. The Indonesian data collected from various secondary sources were mainly categorical, meaning that most of those data can only be presented in the univariate diagram. This is unfortunate because it means that the data submitted was minimal, and correlating and comparing variables became impossible, even though correlation and comparison are much needed if the health data was to be linked with its social determinants. Therefore, further analysis of the individual variable is required in order to determine whether acquiring the raw data can push the limits.

The difficulties in communicating with external consultants, resulting in changing several web developers before finding someone familiar with the concept was also complicating the program. This technical difficulty proved to be significant because it caused many delays on the program itself. This complication may be later mitigated in the next phase of the program by collaborating with researchers with IT background, instead of using external contractors, so that the internal communication can be improved. The equal position between researchers with health-related experience and IT-related ones may ease internal communication.

Trial and feedback were run for a month, and it was safe to assume that various backgrounds have been reached both professional and academic background. Granted, a wider audience will be needed to reduce bias. However, the user experiences have so far been favorable. The users agreed on the superiority of data visualization over tables used by the original reports. This phenomenon reflected the theoretical framework that supported this program, which was the picture superiority effect (Paivio, 2014).

However, it would be interesting to examine the superiority of the interactive over the static interface in the next phase of Mata-Data, which may be demonstrated to be different, as has been observed by Wood and Badawood (2014).

From the explanation above, there was clear evidence that the optimization of data on health conditions and social determinants of health through data visualization can increase the understanding of data as evidence. Therefore, it was expected that increasing knowledge would lead to empowered general public and health workers in providing better challenges and solutions for the government's regulations and policies.

However, one of the limitations of this project is the short period of observation. More extended consideration will be needed to further analysis on how the program influences the intended users, either health communities or practitioners, researchers, and policymakers. It is expected that this program can increase their capacities to understand health conditions and social determinants of health, and therefore, increase their capabilities to make informed decisions, either in health policies or programs. Further, more extended observation with considering environmental influences is needed to measure the empowered communities, because of the prolonged effect on Outcome, even though the Output has been achieved (referring to the Input-Process-Output-Outcome-Impact cycle of program management). However, this short preliminary program has shown promising results, and further development will be more extensive engagement with the tripartite of the policymaking process, followed with more prolonged and more in-depth feedback analysis.

To ensure even better measurement against bias, the collection of feedbacks will be continued in passive setting even after the trial, to portray the authentic experiences of the users and further improvement of the program.

As have been mentioned above, three themes emerged from the qualitative thematic analysis. The first is the use of a mobile device in accessing the website. This is intriguing, considering the assumption that online researches, which utilized data visualizations such as Mata-Data, were usually done with a personal computer. However, the use of a mobile device was not surprising, considering a mobile device dominated the high percentage of internet use in Indonesia. However, bias may arise because the trial and feedback period were not done in a perfect simulation of real condition. The further

quantitative and qualitative study should be carried out on how mobile users accessing and utilizing Mata-Data.

Another theme raised was the need for reference. Several respondents requested for some forms of recommendations for the data, either in the form of the national median, or any standards. It means that there was interest in comparing the provincial results with an external standard, instead of between provinces. This may have arisen because of the respondents' background, who may have national-level-data oriented thinking.

The third theme observed the depth of data was complementary to the national-level bird-eye view of the second theme. The third theme showed the interest of respondents for more geographically specific data, especially on city/district level. This input has been thoughtfully considered from the beginning of program planning; however, those data were spread through dozens of reports, which may need more person-hour than allowed in this program. This will be implemented in the next phase.

To support the sustainability of the program, paying attention to two concerns were required. First, the durability of the program required that the data is continually updated. However, the data updates must pay attention to user preferences, as has been mentioned in one of the themes that emerged from the trial feedback. Second, the sustainability of the programs required supports from the advertisements and donations. However, for the current condition, more extensive community engagement is needed to support the possibility of using ads and donations to support this program. Another alternative under proposition is to integrate this program under the website of Universitas Indonesia so that the university can support the hosting cost. These alternatives would need further consideration and analysis to ensure the sustainability of financial support.

4. Conclusion

In conclusion, data visualizations to optimize the health data and social determinants of health has been proven to be a promising concept to be implemented and improved in Indonesia context. Even with all the program limitation, Mata-Data has opened up the possibilities to optimize the health data and its determinants. Moreover, the novelty of Mata-Data is not only on the data visualization itself but also in listening to the community voice as its users.

By engaging intended users, Mata-Data has evolved into something very different from its original design. The end product will develop further in the next phase by listening to the community voice for a better fit into the Indonesian context. Further analysis in various concerns is needed to improve more also how Mata-Data interact with and benefit the users and the general public, such as the data limitation and short period of observation.

Acknowledgments

This community engagement program was supported by the grant from Directorate of Research and Community Engagement, Universitas Indonesia. Authors would also like to extend the gratitude towards the colleagues from the Department of Community Medicine, Faculty of Medicine, Universitas Indonesia, who had provided much assistance in disseminating the invitation for trial and feedback.

References

- Angotti, R. (2017). Dynamic Displays of Data. *The Mathematics Teacher*, 110(9), 703. <https://doi.org/10.5951/mathteacher.110.9.0703>
- Baicker, K., & Chandra, A. (2017). Evidence-based Health Policy. *New England Journal of Medicine*, 377(25), 2413–2415. <https://doi.org/10.1056/NEJMp1709816>
- Brownson, R. C., Chiqui, J. F., & Stamatakis, K. A. (2009). Understanding Evidence-based Public Health Policy. *American Journal of Public Health*, 99(9), 1576–1583. <https://doi.org/10.2105/AJPH.2008.156224>
- Budiharsana, M. (2017). Increasing Use of Research Findings in Improving Evidence-based Health Policy at the National Level. *Kesmas: National Public Health Journal*, 0(0), 49–52. <https://doi.org/10.21109/kesmas.v0i0.1905>
- Jacobs, J., Jones, E., Gabella, B., Spring, B., & Brownson, R. (2012). Tools for Implementing an Evidence-based Approach in Public Health Practice. *Preventing Chronic Disease*, 9. <https://doi.org/10.5888/pcd9.110324>
- Kemp, S. (2017, February 16). Special Reports: Digital in Southeast Asia in 2017. Retrieved November 29, 2018, from We Are Social website: <https://wearesocial.com/special-reports/digital-southeast-asia-2017>

- Lee, G. (2017). The importance of facts in this 'fake news' era. *International Emergency Nursing*, 31, 1. <https://doi.org/10.1016/j.ienj.2017.03.001>
- McBride, D. M., & Anne Doshier, B. (2002). A comparison of conscious and automatic memory processes for picture and word stimuli: A process dissociation analysis. *Consciousness and Cognition*, 11(3), 423–460. [https://doi.org/10.1016/S1053-8100\(02\)00007-7](https://doi.org/10.1016/S1053-8100(02)00007-7)
- McCosker, A., & Wilken, R. (2014). Rethinking 'big data' as visual knowledge: the sublime and the diagrammatic in data visualisation. *Visual Studies*, 29(2), 155–164. <https://doi.org/10.1080/1472586X.2014.887268>
- Paivio, A. (2014). Chapter Seven: The Multimodal Dual Coding Brain. In *Mind and Its Evolution: A Dual Coding Theoretical Approach* (2013th ed.). New York: Psychology Press.
- Poling, L. L., & Naresh, N. (2014). Rethinking the intersection of statistics education and social justice. *Sustainability in Statistics Education*. Presented at the Ninth International Conference on Teaching Statistics, Arizona, USA. Retrieved from http://iase-web.org/icots/9/proceedings/pdfs/ICOTS9_C177_POLING.pdf
- Reuters Institute for the Study of Journalism. (2015). *Making Research Useful: Current Challenges and Good Practices in Data Visualisation*. Retrieved from <https://www.alliance4usefulevidence.org/assets/Making-Research-Useful-Current-Challenges-and-Good-Practices-in-Data-Visualisation.pdf>
- Segel, E., & Heer, J. (2010). Narrative Visualization: Telling Stories with Data. *IEEE Transactions on Visualization and Computer Graphics*, 16(6), 1139–1148. <https://doi.org/10.1109/TVCG.2010.179>
- Singh, S. (2014). *Choosing a Prime Ministerial Candidate - The Indian Elections 2014*. Retrieved from <https://www.youtube.com/watch?v=vpDSHIZDtH4>
- Tarran, B. (2017). "Hans was a storyteller; he turned statistics into a performing art." *Significance*, 14(2), 6–7. <https://doi.org/10.1111/j.1740-9713.2017.01011.x>
- Wood, J., & Badawood, D. (2014, June). *The Effect of Information Visualization Delivery on Narrative Construction and Development*. Presented at the Eurographics Conference on Visualization (EuroVis), Swansea, UK. Retrieved from <http://openaccess.city.ac.uk/3702/>