



An Open-Access Data Platform: Global Nutrition and Health Atlas (GNHA)

Bingjie Zhou,¹ Shiwei Liang,¹ Kyle M Monahan,² Naglaa El-Abbadi,¹ Melissa S Cruz,¹ Yutong Chen,¹ Annie DeVane,¹ Julia Reedy,¹ Jianyi Zhang,¹ Iaroslava Semenova,³ Ivan Montoliu,³ Dariush Mozaffarian,¹ Dantong Wang,³ and Elena N Naumova¹

¹Tufts University Friedman School of Nutrition Science and Policy, Boston, MA, USA; ²Data Lab, Tufts Technology Services, Tufts University, Medford, MA, USA; and ³Nestlé Institute of Health Sciences, Nestlé Research, Lausanne, Switzerland

ABSTRACT

The rapid development of nutrition science is embracing digital transformation to generate large amounts of data. Precision nutrition and “Big Data” place increasing demand for data repositories and visualization, which enhances the digital transformation. We defined the need for an integrated nutrition data platform as a web-based platform that can collect, store, track, analyze, monitor, and visually display key metrics in nutrition and health while allowing users to interact with visuals and download data provided in the platform. Interactive dashboards create new opportunities for scholars and practitioners to generate and test hypotheses. We present the development and implementation of the Global Nutrition and Health Atlas (GNHA; <https://sites.tufts.edu/gnha/>), an open-access online platform covering nutrition and health data with 26 themes and 500+ indicators from 190+ countries up to 30 y. We view GNHA as an interactive tool aiming to share information and perspectives and foster collaborations and innovations. *Curr Dev Nutr* 2022;6:nzac031.

Keywords: dashboard, data repository, food system, nutrition, visualization

© The Author(s) 2022. Published by Oxford University Press on behalf of the American Society for Nutrition. This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial License (<https://creativecommons.org/licenses/by-nc/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact journals.permissions@oup.com

Manuscript received December 10, 2021. Initial review completed February 17, 2022. Revision accepted February 23, 2022. Published online March 11, 2022.

Supported by Société des Produits Nestlé SA.

Author disclosures: The authors report no conflicts of interest. The coauthors from Tufts team are solely responsible for the design, implementation, analysis, and interpretation of the results. The coauthors employed by Nestlé provided valuable contributions for broadening the scope of the platform, assessed functionality, and shared original data and documentation.

Supplemental Figures 1–3 are available from the “Supplementary data” link in the online posting of the article and from the same link in the online table of contents at <https://academic.oup.com/cdn/>.

Address correspondence to ENN (e-mail: elena.naumova@tufts.edu).

The Global Nutrition and Health Atlas (GNHA) provides interactive data visualizations and reliable data repositories, aiming to share information and interdisciplinary perspectives and to foster collaborations and innovations.

Introduction

With the rapid development of nutrition surveillance and intensive research in public health and personalized nutrition, large volumes of data are generated from national prospective cohorts, longitudinal studies, household surveys, and other sources. Nutrition data platforms present a growing opportunity for this rapid development (1–4). Over the last decades, interactive dashboards, commonly used in business data analytics, are being frequently applied to health and nutrition domains (5, 6). In the field of food and nutrition, the Hunger Map Live dashboard, produced by the World Food Program, attempts to monitor nutrition outcomes with an emphasis on predicting hunger on a global scale (7). The Joint Child Malnutrition Estimates interactive dashboard developed by UNICEF aims to support countries in generating, analyzing, and using data to monitor malnutrition in children and women (8). These comprehensive data platforms allow a broad range of users to explore multiple data sources through interactive graphs, maps, and data summaries (9–11).

Only a few nutrition dashboards to date are equipped with all functionalities needed for effective research and communication. Moreover,

knowledge and information gaps among main stakeholders in data collection, data compilation, and data analysis are substantial. As expected, most existing nutrition data platforms are highly specialized and focus on specific nutritional topics—for example, vitamin deficiency, hunger, or food fortification. A few have attempted to cover global nutrition and health, yet the data quality, especially data granularity and completeness, were not emphasized. Interactive visualizations in modern nutrition data platforms are still limited, and visualization quality could be further improved.

Methods

Database content

We used several datasets from the Nutrition and Health Atlas (NHA) that Nestlé Research developed in 2015 as the foundation of Global Nutrition and Health Atlas (GNHA) 2021 version. We checked the accessibility of the data sources and references from the original NHA data. We updated links with new ones if they had been moved to other locations. We then reformatted the data structure by adding

gender (male, female), multiple age groups, and residential areas (urban and rural) into separate columns. This step is allowed to reduce duplication and improve data granularity. Finally, according to original sources, all compiled variables were updated to the most recent data (as of August 2021). The overall data structure is shown in **Supplemental Figure 1**.

Software implementation and query tools

We chose Harvard Dataverse, a protected and manageable research data repository, for storing and sharing. In Dataverse (Dataverse Project, dataverse.org), datasets are accompanied by well-documented metadata and a citation with a digital object identifier (DOI) for each dataset; data can be previewed and downloaded as comma-separated values (CSV) files. We created a GNHA Dataverse, uploaded all data to the GNHA Dataverse page, and embedded this specific page into the GNHA website. We also made a guestbook to monitor the downloaded entries. A sample of a Dataverse interface is shown in **Supplemental Figure 2**. We also saved data in the Box cloud service to provide high flexibility for previewing and filtering data. We connected the Box folder to an interactive data visualization software, Tableau, and used iFrame to embed the Tableau tabular dashboard to our website. In this data preview tool, filters were added to every column of a dataset and can be applied to multiple columns simultaneously. An example of Tableau query tool interface is shown in **Supplemental Figure 3**.

Visualizations

To create the visualization gallery, we connected the database to Tableau, created interactive visualization in Tableau, and then used iFrame to embed the Tableau dashboards to the GNHA platform. Visualizations are grouped into 7 categories: 6 dimensions and 1 global comparison. Each dimension has a Tableau dashboard with 2 tabs. The first tab is a uniform template across all dimensions that presents 1 indicator in all countries with available data using several plots to show different perspectives. The second tab is a country profile, offering a unique design according to specific content in each dimension. The country profile includes several key indicators within 1 dimension. Users can change the view to show a summary of this dimension in different countries. Finally, the global comparison dashboard contains a scatterplot, as an intersection in which data can be pulled from any dimension. All visualizations are interactive, supported with filters and pop-up windows, and can be downloaded in PDF or image format.

GNHA

We have developed a comprehensive nutrition and health data platform covering 6 dimensions: Demographics, Dietary Intake, Nutritional Status, Health Status, Health Economics, and Food Sustainability. Twenty-six themes and 500 indicators are organized under these 6 dimensions to cover global health and nutrition data in more than 190 countries over the past 30 y. In the present version, data are collected from more than 250 established public sources, including international institutions, governmental reports, nongovernmental organizations and academic institutions, and peer-reviewed research papers.

The GNHA platform serves 3 major missions: enabling public education, enhancing collaboration, and driving innovation. We strive to support these missions by improving data literacy and science

communication, embedded in the platform's dashboard features. The platform also allows to download data for analysis, recreate results claimed by different studies, and compare data originating from various sources (12). Therefore, we focused on broad information coverage, data credibility, visualization quality, and dashboard functionalities. Specifically, we ensured that all content on our platform—the data, visualizations, and descriptions—was accessible for various audiences, from students, nutritionists, and scientists to the public and policymakers. We produced tutorials to guide users through every element of the GNHA website, including the dashboard gallery, data portal, and data sources. We embedded detailed instructions within the visualization dashboard to help users with data and figures interpretation. Finally, we provided comprehensive metadata in an interactive and downloadable manner to help users check the reliability and credibility of the selected data sources.

Data visualizations are the key component of the GNHA platform. In the visualization gallery, we first have a Global Overview tab (**Figure 1A**) to present data on 1 variable from a specific theme. This tab is implemented a uniform design with a combination of several interconnected graphs to present 1 indicator from multiple angles. For example, information on geographic distribution, time trend, and country rankings of a presented variable can be all viewed in 1 page. Compared with Global Overview, which presents 1 indicator at a time, the Country Profile tab provides an overview in a single country, which includes several key metrics from a selected dimension (**Figure 1B**). Finally, we created a Comparison dashboard (**Figure 2**), which enables indicators from different themes and dimensions to be viewed simultaneously. The Comparison dashboard serves as a bridge that connects all GNHA data together and gives users a quick way to review the relation between any 2 variables of interest. Furthermore, we provided extensive instructions on applying filters and using interactive features to get a comprehensive understanding of the graphs. All visualizations are equipped with an explanation of why we apply those types of plots and how to correctly interpret them. One of the important goals for the GNHA visualization gallery is to serve as an easy and quick tool for students, scholars, and all other users to teach and learn essential information and generate insights through visualizations (13).

The platform is expandable and customizable at different levels. As we included data from multiple sources with varying formatting styles, we standardized them to adhere to a proper format with clearly labeled time, location, topic, and demographic characteristics. This standardization allows data to be filtered, selected with various combinations; it supports data sharing and fulfills the requirements for open data access and to promote research reproducibility. Our previous studies identified that data quality represents a major gap for many existing nutrition and health dashboards and platforms (14). Thus, we supplemented our platform by providing well-documented metadata on data completeness, a glossary of terms and their definition, the original sources, and the spatial-temporal characteristics of the data.

Overall, data literacy is highly emphasized as the main GNHA feature throughout the entire process. With the proposed organization of the compiled data and detailed guidance for data visualization, we aim to offer tools for effective data literacy programming in nutrition and health. We view GNHA as not just a data-sharing platform but, more essentially, as a visual literature review tool that connects data published

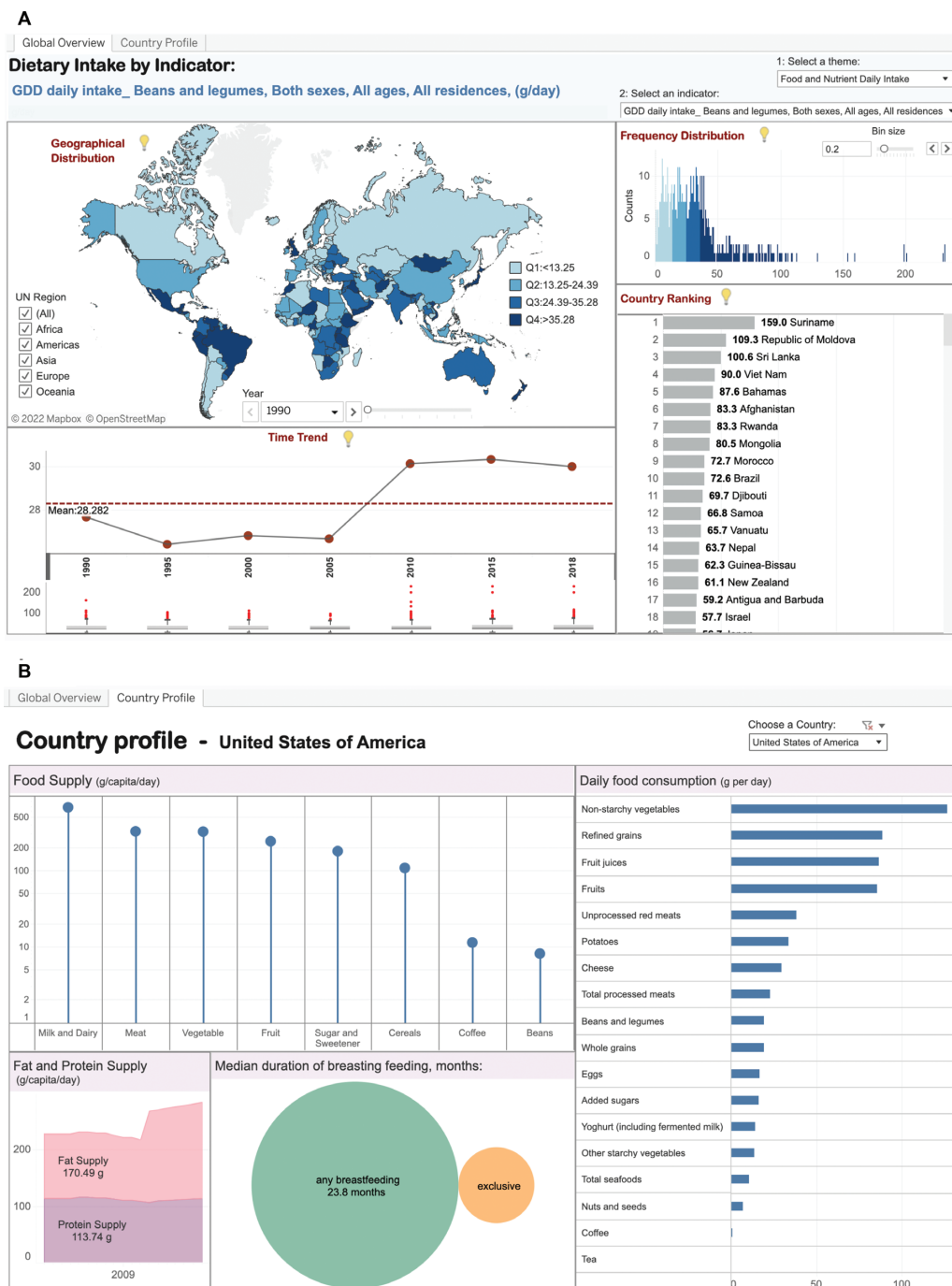


FIGURE 1 Examples of Tableau dashboard screenshots: (A) Global Overview as tab 1 and (B) Country Profile as tab 2. The Global Overview tab displays information of 1 indicator at a time from a specific dimension. In this example, any indicator within the “Dietary Intake” dimension can be selected from the dropdown menu “Choose an indicator.” This tab uses a uniform template for all dimensions with several plots to show data from different perspectives. This dashboard contains the following: 1) a choropleth map at global scale on the top left showing the geographical distribution of this indicator; 2) a histogram on the top right showing countries’ overall distributions with the color scheme representing the quartiles aligned with the color scheme in the world map; 3) a bar graph below the histogram showing the rankings and specific values of all countries for this indicator; 4) a line chart under the world map showing the available time points and the temporal trend of the chosen indicator; and 5) a boxplot at the bottom showing the countries’ distribution by year sharing the same time axis with the line chart. This organization with interconnected plots forms a cohesive and comprehensive data representation. The country-specific view provides a country profile with representative indicators for each dimension. Compared with Global Overview, which presents 1 indicator at a time, the Country Profile includes several key indicators within a selected dimension. In this example, we chose to show data from various sources on food supply, food security indicators, along with daily food consumption in the United States. Note that this dashboard allows to depict information that is limited or not available. GDD, Global Dietary Database.

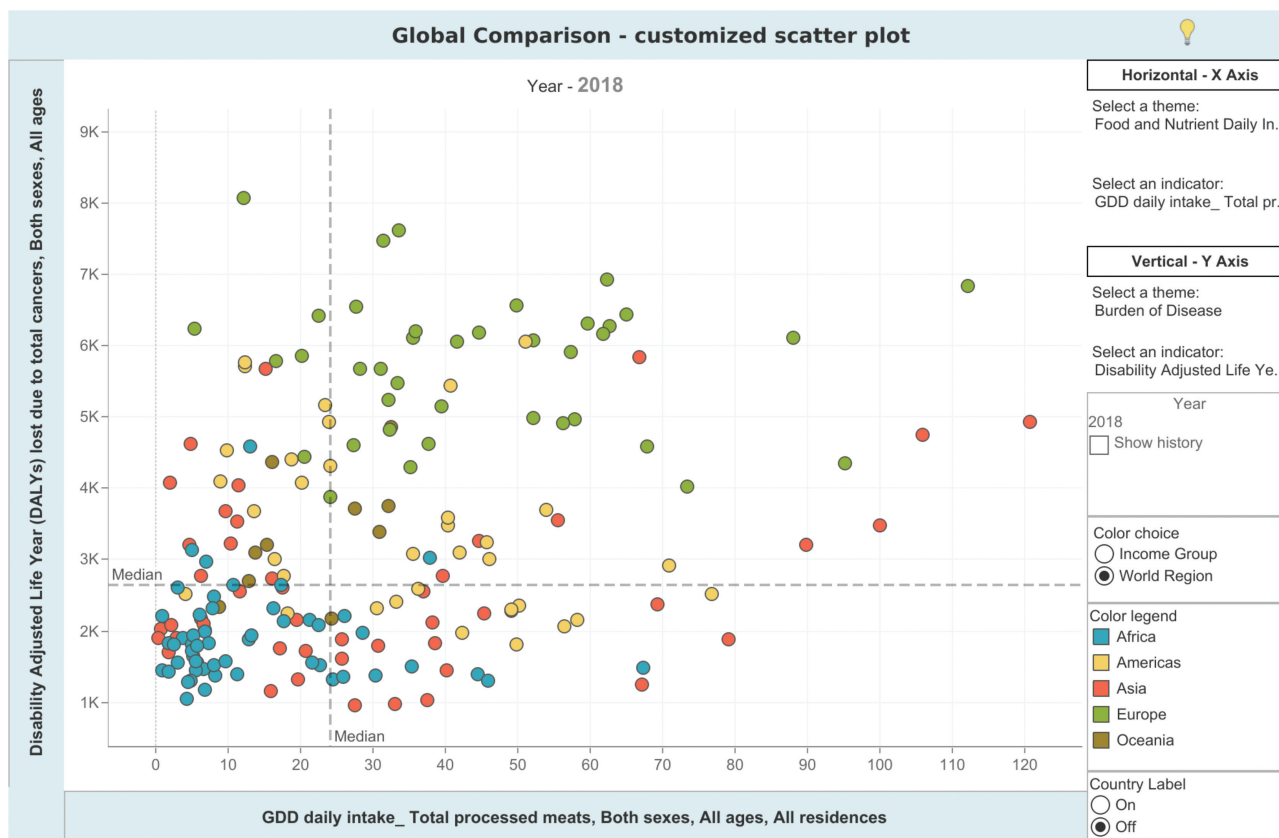


FIGURE 2 An example of the Tableau comparison dashboard—Global Comparison. The scatter plot is the central part of this dashboard screenshot. Users can choose indicators for the horizontal and vertical axes from any dimension. In this example, we chose to show the relation between total processed meats and disability-adjusted life-years lost due to cancers in 2018. We also provide the data completeness metric for each indicator by estimating the number of countries which have data available in a given year. Once indicators on 2 axes are chosen, users can also click the Year “Play” button on the right side to create an animation that shows a dynamic relation between variables over time. Points on the scatter plot can be colored by country’s income level or world region at the users’ choice. GDD, Global Dietary Database.

in different platforms, from individual studies to large-scale organizations like WHO and the World Bank.

Conclusions

The GNHA platform is hosted by the Friedman School of Nutrition Science and Policy, which aims to generate trusted science, educate future leaders, and produce a real-world impact in nutrition science and policy. We have invested heavily in data organization, standardization, and harmonization, so the data structure is flexible to incorporate data with diverse types and formats and updated regularly. Facing the growing challenges of keeping data up to date, we aim to form communities of practice and open for data contribution.

Acknowledgments

The authors’ responsibilities were as follows—BZ: contributed to conceptualization, data collection, visualization, analysis, and writing; NE-A and KMM: contributed to website design, manuscript review, and editing; SL, MSC, YC, and AD contributed to data collection and website

design; JR, JZ, IS, and IMR: contributed to manuscript review and editing; DM and DW: contributed to conceptualization, website review, and manuscript review and editing; ENN: contributed to conceptualization, methodology development, review and editing, and supervision; and all authors: reviewed, provided critical intellectual input on, and read and approved the final manuscript.

Data availability

Data described in the manuscript are from public sources. Data and analytic codes are freely available without restriction at <https://sites.tufts.edu/gnha/>.

References

1. Mason JB, Mitchell JT. Nutritional surveillance. *Bull World Health Organ* 1983;61(5):745.
2. Al Jawaldeh A, Osman D, Tawfik A; World Health Organization. Food and nutrition surveillance systems: technical guide for the development of a food

- and nutrition surveillance system for countries in the Eastern Mediterranean region. World Health Organization Eastern Mediterranean Regional Office (EMRO); 2013.
3. Ferro-Luzzi A, Leclercq C. Nutritional surveillance: an outline. *Food and Nutrition Policy in Europe* 1993;113–7.
 4. Satija A, Hu FB. Big data and systematic reviews in nutritional epidemiology. *Nutr Rev* 2014;72(12):737–40.
 5. Wu DT, Chen AT, Manning JD, Levy-Fix G, Backonja U, Borland D, Caban JJ, Dowding DW, Hochheiser H, Kagan V. Evaluating visual analytics for health informatics applications: a systematic review from the American Medical Informatics Association Visual Analytics Working Group Task Force on Evaluation. *J Am Med Inform Assoc* 2019;26(4):314–23.
 6. Preim B, Lawonn K. A survey of visual analytics for public health. *Computer Graphics Forum* 2020;39:543–80. doi:10.1111/cgf.13891.
 7. World Food Program. HungerMap LIVE. [Internet]. 2021; [cited 2021 Nov 2]. Available from: <https://hungermap.wfp.org/>.
 8. UNICEF-WHO-World Bank: Joint child malnutrition estimates 2020 edition—interactive dashboard. [Internet]. 2020; [cited 2021 Oct 14]. Available from: <https://data.unicef.org/resources/joint-child-malnutrition-estimates-interactive-dashboard-2020/>.
 9. Global Dietary Database. [Internet]. 2021; [cited 2021 Nov 1]. Available from: <https://www.globaldietarydatabase.org/>.
 10. Global Alliance for Improved Nutrition (GAIN) and Johns Hopkins University. The Food Systems Dashboard. [Internet]. 2020; [cited 2021 Oct 4]. Available from: <https://foodsystemsdashboard.org/food-system>.
 11. Global Fortification Data Exchange. [Internet]. 2021; [cited 2021 Nov 1]. Available from: <https://fortificationdata.org/>.
 12. Simpson RB, Zhou B, Alarcon Falconi TM, Naumova EN. An analecta of visualizations for foodborne illness trends and seasonality. *Scientific Data* 2020;7(1):1–15.
 13. Peng RD, Hicks SC. Reproducible research: a retrospective. *Annu Rev Public Health* 2021;42(1):79–93.
 14. Simpson RB, Gottlieb J, Zhou B, Hartwick MA, Naumova EN. Completeness of open access Flunet influenza surveillance data for Pan-America in 2005–2019. *Sci Rep* 2021;11(1):1–17.