



Considering Ethics Along the Data Value Chain for Nutrition

Jessica Fanzo

Johns Hopkins University, Washington, DC, USA

Key messages

- > There are several ethical issues that need to be taken into consideration when prioritizing, collecting, analyzing and disseminating data across the *nutrition data value chain*. These include favoring public health good over gathering data for data's sake, considering individual rights over public interests, and ensuring accountability and transparency of data sharing and publication.
- > While nutrition evidence is building, the *precautionary principle* – whereby decisions to act are taken on issues even though extensive scientific knowledge on the matter is lacking or uncertain – should be carefully analyzed.
- > We are living in a time of 'big data' that is piggybacking on the digital age and economy and on social media interfaces. For those collecting, curating and using that data, there should be a form of global *Hippocratic oath* in place to “do no harm,” protect human dignity and promote wellbeing over profit interests.
- > We need to account for the *downstream uses of nutrition*, as well as dietary intake and behavior data, in ways that are consistent with the intentions and understanding of the disclosers who provide personal information to researchers.

Introduction

The 2017 Global Nutrition Report called for revolutionizing efforts across the nutrition data value chain. This chain includes data priority-setting, collection, analysis, interpretation and use of information by decision-makers working in the nutrition

space.¹ It proposed to position data as a *value product* to address malnutrition in all its forms.

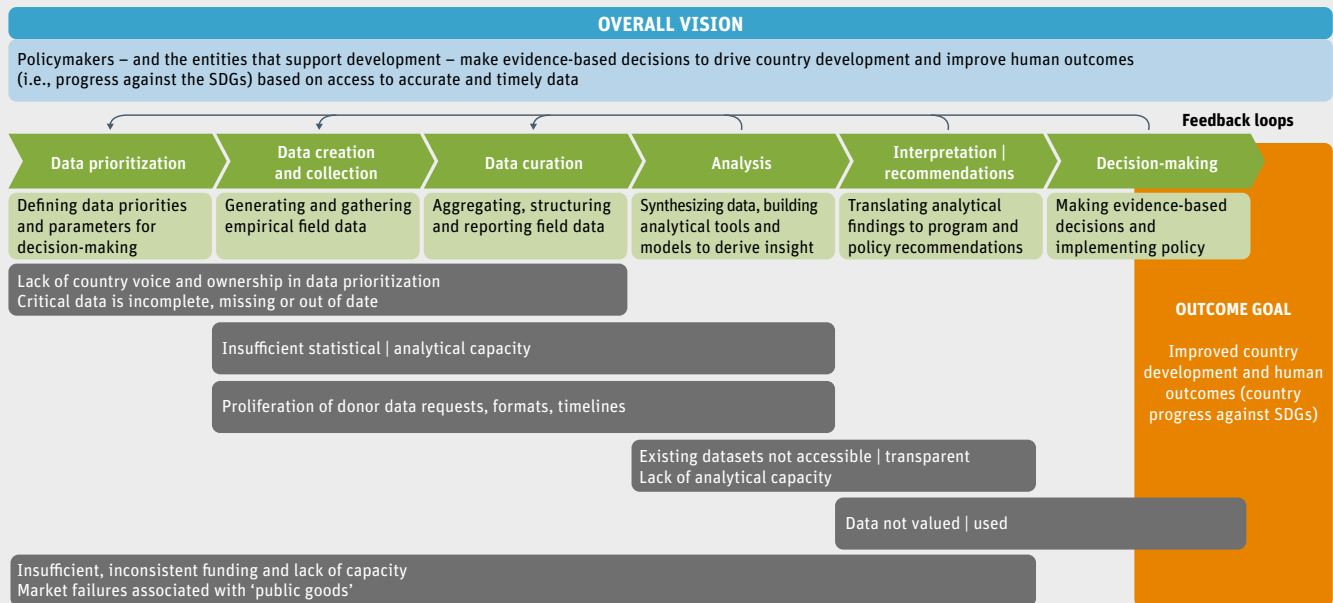
The data value chain highlighted six critical steps as shown in **Figure 1** and called for an agenda that requires: “**1**) in-country mechanisms for national priority-setting and data coordination; **2**) operational guidance for data prioritization, harmonization of indicators, and incorporation of nutrition into routine management information systems; **3**) tools for capacity development at multiple levels; **4**) costed data plans that are built into national development strategies, resourced and implemented; **5**) dissemination of tacit knowledge and experience; **6**) innovation across the value chain; and **7**) fostering a culture of data use and sharing.”

Considering ethics along the data value chain for nutrition

To undertake this revolution, there are certain *ethical* issues that should also be considered across the entirety of the nutrition data value chain. In collecting any kind of data, there are, of course, guiding ethical principles and values that should be maintained, including autonomy and confidentiality of those whose personal information is used to garner evidence that informs programs and policies. However, other ethical considerations should also be considered when prioritizing, collecting, analyzing and disseminating data, such as issues around favoring public health good over gathering data for data's sake, considering individual rights over public interests, and ensuring accountability and transparency of data sharing.

The 1978 landmark Belmont Report,² an essential reference tool for almost all institutional review boards (IRBs), provided ethical guidance on doing research involving human subjects and highlighted three core principles: respect for persons (protecting autonomy, ensuring respect and avoiding deception); beneficence of populations/persons (maximizing benefits and minimizing risks); and justice to research participants (fair and equal distribution of costs and benefits). To carry out these principles, informed consent, assessment of risks and benefits for human populations and selection criteria of human subjects should be required across research. Many of these same principles can be applied across the data value chain.

FIGURE 1: The data value chain for nutrition



Source: Development Initiatives, 2017¹

“The Belmont Report highlighted three core principles: respect for persons, beneficence of populations | persons, and justice to research participants”

In this piece, three issues that cut across the data value chain in nutrition are highlighted, including: **1)** the precautionary principle conundrum; **2)** the Hippocratic oath of big data curation; and **3)** the sensitivity of analysis and usage of data. Thereafter, a set of questions is presented that can serve as a starting point to dive deeper into specific ethical concerns at each stage of the value chain.³

The precautionary principle conundrum

In the case of nutrition, there is often an assertion that “we need more evidence” before we take action. But what if there is not enough evidence available? And when is enough, enough? When do we determine that what we have in hand is sufficient? While monitoring and evaluation are essential for guiding nutrition policies or programs, there are still interventions that have not undergone extensive evaluation to assess their impact in different contexts and in different populations. This leaves policymakers and development practitioners working somewhat in

the dark, without guidance to determine what constitutes sound evidence to justify an intervention, and what factors must be considered within such an evaluation. Interventions “should be implemented only in the face of a clear public health need and good data demonstrating effectiveness.”⁴

One cannot underestimate the complexities of the multiple forms of malnutrition, which not only have multiple causes – both immediate and underlying – but also have varying impacts at different stages of the life cycle. Because of these complexities, there are still gaps in evidence and knowledge. These gaps prevent implementation of critical nutrition interventions.⁴ This brings into consideration the *precautionary principle* in cases where decisions to act are taken on issues even though extensive scientific knowledge on the matter is lacking or uncertain. Policymakers can justify these decisions in situations where there is a social responsibility to protect the public from exposure to harm (because of a plausible risk) as opposed to not taking any action.

For example, while the nutrition community does not have all the evidence to hand of the impact of diets on health outcomes, some data suggests that certain dietary patterns promote health and that certain foods can cause deleterious health outcomes. The precautionary principle can be used in this case to draft dietary guidelines or provide dietary advice to citizens that can generally maximize health and minimize harm. For example, consumption of fruits and vegetables is most likely beneficial to human health.⁵

The Hippocratic oath of big data curation

We are living in a time of ‘big data’ that is piggybacking on the digital age and economy and the incredible breadth of information being collected across social media platforms. These big data include massive streams of information points being collected, combined and shared. Undoubtedly, big data will play an important role in delivering a deeper understanding of populations, revealing individual drivers of choice concerning what they eat, purchase and aspire to. Much of this data provides personalized nutrition insights and indicates how decisions are made, from farm to markets to individuals, and from genome to metabolome. But this personalization can create entirely new types of ethical risk concerning who is using the data and how it is disseminated. Risks include unethical or illegal use for purposes for which the original disclosers did not give their consent.⁶

There is an absolute need for oversight of these data and for sound governance, with clear transparency and accountability mechanisms. Making data open access, allowing for further analysis and interpretation of data by other parties (e.g., repeating results), and publishing ‘null’ results on research outcomes are all key to ensuring transparency and accountability in nutrition.⁷ Right now, there is a lack of authority and accompanying legislation to enforce rules of ethical practice in the use of big data. For those collecting, curating and using that data, there should perhaps be a form of global *Hippocratic oath* in place. This oath should be to “do no harm,” to protect human dignity and to promote wellbeing over profit interests. Above all, data scientists and operators of these large platforms should respect the privacy of individuals and put the people behind the data above any other objective.

.....

“For those collecting, curating and using that data, there should perhaps be form of global *Hippocratic oath* in place”

.....

For example, many people are now posting details about what they eat on Facebook, Twitter or Instagram. This could be a valuable resource in understanding what consumers like to eat, what drives their eating behaviors and what potential products they may buy. While it is enticing to use these social platforms for data purposes, most users of these social platforms share information because they want the content to reflect their online identity, they want to grow and nourish relationships, they like the feeling of having others comment on their status or they want to spread the word about an issue they care about.⁸ Most users are not posting their information for the purposes of profit

by others. Thus, it is important for these social media platforms to govern the data. Governance comes in the form of being transparent on how the data is used or sold, obtaining consent from data disclosers, instilling additional data privacy measures that data disclosers can opt in/out of, and minimizing the sale of data to third-party users.

The sensitivity of analysis and usage of data

We need to account for the downstream uses of datasets, including not only primary data collected by researchers but also secondary data analysis collected by others (e.g., using data from the Demographic and Health Surveys). Data starts as ‘raw’, but data analytics and interpretation carry a “history of human decision-making” that should be trackable and transparent regardless of whether it is primary data or secondary data analysis.^{1,9,10} This would again include accountability mechanisms for collecting data, the methods to do so, the consent for use of that data, and how data is analyzed and assessed. However, sometimes data is repurposed (e.g., secondary data analysis), which creates significant ethical risk. Researchers should strive to use data in ways that are consistent with the intentions and understanding of the data disclosers who provide personal information to the original investigators.⁹

For example, the practice of qualitative secondary analysis has been accompanied by some ethical controversy, because it involves different relationships between researchers and participants as compared with primary research data collection.¹¹ With primary data generation, “data is collected through trust interactions between researchers and participants and is shaped in ways which relate to the project design, researchers’ disciplinary assumptions, theoretical inclinations and methodological decisions.”¹² This garners concern about how the data may be used by others at a later date, and about the consent to use that data in different ways and for different purposes. Where the future use of data is not clear, it is important to secure informed consent using generic statements about how the data will be used, as well as the respective obligations or duties of research participants, primary and secondary researchers, project partners such as service providers, and also the general public.¹²

The Young Lives project – a longitudinal, international study with publicly archived survey data, exploring childhood poverty over time, including nutrition outcomes – undertook an assessment of the ethical implications of secondary analysis of qualitative data collected in the project. The authors argue that while sharing data for the purposes of secondary analysis brings many benefits, it also can be ethically complex.¹³ Issues to be addressed include countering risks of misinterpretation, proper informed consent, using data far from the intentions of the project and improper use of data representation in media dissemination. The authors argue, as does Irwin,¹¹ that devel-

TABLE 1: Ethical questions across the data value chain in nutrition

Value chain step	Ethical questions
Data prioritization, creation and collection	<ul style="list-style-type: none"> > Has the data collection procedure gone through an ethics review process? > Given the different types of data being collected, what potential harm could come from using that data? > Is there a way to minimize the volume or variety of data being collected by streamlining surveys, using secondary data and reviewing the essentiality of data in order to understand key research questions? > Are data disclosers (those whose information is being collected) aware that data has been acquired, stored or shared? > Will the data being collected, or the analysis of that data, be shared with the data disclosers?
Data curation and analysis	<ul style="list-style-type: none"> > Are there any potential harms if others access the data or the analysis of that data? > What negative consequences for the data discloser could result from the proposed analysis? What steps are being taken to mitigate these risks? > What biases have been introduced during the analysis of data and how were they addressed? > How frequently should an ethics assessment be performed on these analyses for alignment with project, product and service goals and the organization's code of ethics?
Data interpretation, use and decision-making	<ul style="list-style-type: none"> > Did the data discloser provide consent to this specific data use? Was that consent informed? > Can a data discloser discover whether data they have disclosed has been used and for what purpose? > Are there mechanisms in place for controlling access to the data? > What measures are taken to account for the risk and/or harm that could come from misusing the data? > Do data disclosers expect control, ownership, remuneration or transparency over the data they have disclosed if it is being shared? Did they provide informed consent for this action?

Source: Adapted from Accenture Labs 2016^{3,14}

oping and maintaining trusting relationships between research participants and primary and secondary researchers are key to avoiding unethical use of data.¹³

Ethical assessment across the data value chain

Introducing an ethics assessment at each stage of the data value chain can provide new perspectives of how data is being used. The questions in **Table 1** can act as a starting point for project implementers or policy analysts to dive deeper into specific ethical concerns at each stage.

“If we consider the principles of the Belmont Report, the risks can be mitigated”

Conclusion

We are living in an exciting time. Evidence is ramping up. There are more metrics and indicators to collect valuable information on diets, nutrition and other nutrition-related factors and inputs. There is more data out there – from traditional surveys to big data on social media. However, these developments also present new risks regarding the way data is collected, interpreted, used and shared. Some of these risks present ethical dilemmas. How-

ever, if we consider the three principles of the landmark 1978 Belmont Report² – beneficence, justice and respect for persons – in the data value chain in nutrition, the risks can be mitigated.

Correspondence: Jessica Fanzo PhD,

Bloomberg Distinguished Associate Professor of Global Food and Agriculture Policy and Ethics, Johns Hopkins University, Washington, DC 20036, USA
Email: jfanzo1@jhu.edu

References

01. Development Initiatives. Global Nutrition Report 2017: Nourishing the SDGs. Bristol, UK: Development Initiatives Poverty Research Ltd; 2017.
02. The National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research. The Belmont Report: Ethical Principles And Guidelines for the Protection of Human Subjects of Research. Bethesda, MD; 1978.
03. Accenture Labs. Universal principles of data ethics: 12 guidelines for developing ethics codes. 2016. Internet: www.accenture.com/t20160629T012639Z_w_/us-en/_acnmedia/PDF-24/Accenture-Universal-Principles-Data-Ethics.pdf#zoom=50
04. Hurlimann T, Peña-Rosas JP, Saxena A, Zamora G, Godard B. Ethical issues in the development and implementation of nutri-

tion-related public health policies and interventions: A scoping review. *PLOS ONE*. [serial online] 2017; Oct 26;12(10):e0186897.

05. Gonzales JF, Barnard ND, Jenkins DJ, Lanou AJ, Davis B, Saxe G, et al. Applying the precautionary principle to nutrition and cancer. *J Am Coll Nutr*. 2014 May;33(3):239–46.
06. Özdemir V, Kolker E. Precision nutrition 4.0: A big data and ethics foresight analysis—Convergence of agrigenomics, nutrigenomics, nutriproteomics, and nutrimetabolomics. *OMICS*. 2016 Feb;20(2):69–75.
07. Mayernik MS. Open data: Accountability and transparency. *Big Data & Society*. 2017 Jul 4;(2):2053951717718853.
08. Pearson E, Tindle H, Ferguson M, Ryan J, Litchfield C. Can we tweet, post, and share our way to a more sustainable society? A review of the current contributions and future potential of #Socialmediaforsustainability. *Ann Rev Environ Resour*. 2016 Nov;41:363–97.
09. Erickson LC, Harris NE, Lee, MM. It's time to talk about data ethics. 2018. Internet: www.datascience.com/blog/data-ethics-for-data-scientists
10. Teill S, Metcalf, J. Accenture Labs. Universal principles of data ethics 12 guidelines for developing ethics codes. 2016. Internet: www.accenture.com/t20160629T012639_w_/us-en/_acnmedia/PDF-24/Accenture-Universal-Principles-Data-Ethics.pdf#zoom=50
11. Mauthner, NS. Should data sharing be regulated? In: van den Hoonaard WC & Hamilton A, eds. *The Ethics Rupture: Exploring alternatives to formal research-ethics review*. University of Toronto Press; 2016:206–29.
12. Irwin S. Qualitative secondary data analysis: Ethics, epistemology and context. *Prog Dev Stud*. 2013 Oct;13(4):295–306.
13. Morrow V, Boddy J, Lamb R. *The ethics of secondary data analysis: Learning from the experience of sharing qualitative data from young people and their families in an international study of childhood poverty*. London: Thomas Coram Research Unit; 2014. Internet: http://eprints.ncrm.ac.uk/3301/1/NOVELLA_NCRM_ethics_of_secondary_analysis.pdf
14. Accenture Labs. Facilitating ethical decisions throughout the data supply chain. 2016. Internet: www.accenture.com/t20180705T112208Z_w_/us-en/_acnmedia/PDF-25/Accenture-Labs-Facilitating-Ethical-Decisions.pdf#zoom=50