NUTRISS
CENTRE NUTRITION, SANTÉ ET SOCIÉTÉ

Applications of AI in Food and Nutrition Research

Benoît Lamarche, PhD

INAF INSTITUT SUR LA NUTRITION ET LES ALIMENTS FONCTIONNELS

UNIVERSITÉ LAVAL

1

Disclosures

None

NUTRISS CENTRE NUTRITION, SANTÉ ET SOCIÉTÉ

INAF

UNIVERSITÉ LAVAL

2

Application of AI in nutrition research

- Increased capacity to manage/analyse big data
  - –omics
  - precision nutrition
  - precision public health
- Dietary assessment
  - better understanding dietary patterns
  - image-based methods
  - non image-based methods
- Predicting outcomes
- Social media content analysis (NLP)

NUTRISS INAF UNIVERSITÉ LAVAL

3

Application of AI in nutrition research

- Increased capacity to manage/analyse big data
  - –omics
  - precision nutrition
  - precision public health
- Dietary assessment
  - better understanding dietary patterns
  - image-based methods
  - non image-based methods
- Predicting outcomes
- Social media content analysis (NLP)

NUTRISS INAF UNIVERSITÉ LAVAL

4

Feeding studies

NUTRISS CENTRE NUTRITION, SANTÉ ET SOCIÉTÉ

INAF UNIVERSITÉ LAVAL

5

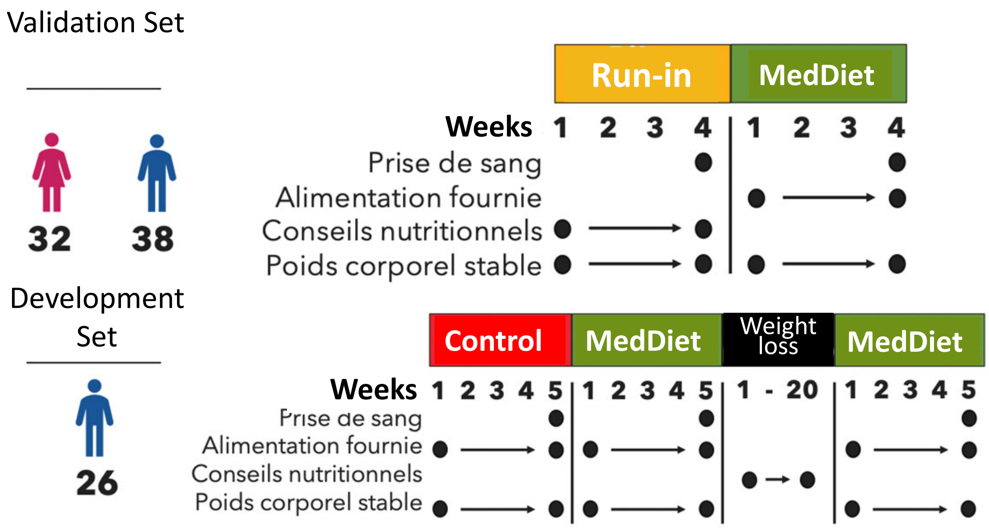
Feeding studies

NUTRISS CENTRE NUTRITION, SANTÉ ET SOCIÉTÉ

INAF UNIVERSITÉ LAVAL

6

Non targeted metabolomics to predict diet patterns



NUTRISS INAF UNIVERSITÉ LAVAL

7

Non targeted metabolomics to predict diet patterns

Performance metric	Development set		Validation set
	Train set	Test set	
Accuracy (95%CI)	0.99 (0.96-1.00)	0.97 (0.81-1.00)	0.79 (0.71-0.86)
Positive predictive value (precision) (95%CI)	1.00 (0.97-1.00)	0.98 (0.83-1.00)	0.73 (0.66-0.81)
Sensitivity (recall) 95%CI	0.99 (0.93-1.00)	0.96 (0.76-1.00)	0.91 (0.80-1.00)
F1 score (95%CI)	0.99 (0.96-1.00)	0.96 (0.81-1.00)	0.81 (0.74-0.88)

Unpublished

NUTRISS INAF UNIVERSITÉ LAVAL

8

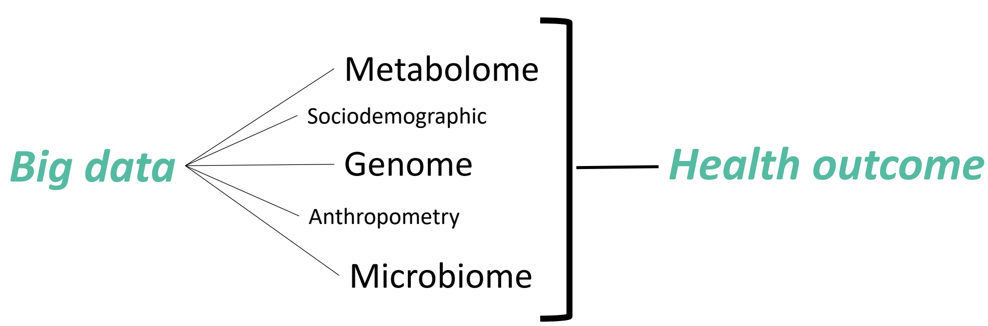
Application of AI in nutrition research

- Increased capacity to manage/analyse big data
  - –omics
  - precision nutrition
  - precision public health
- Dietary assessment
  - better understanding dietary patterns
  - image-based methods
  - non image-based methods
- Predicting outcomes
- Social media content analysis (NLP)

NUTRISS CENTRE NUTRITION, SANTÉ ET SOCIÉTÉ INAF UNIVERSITÉ LAVAL

9

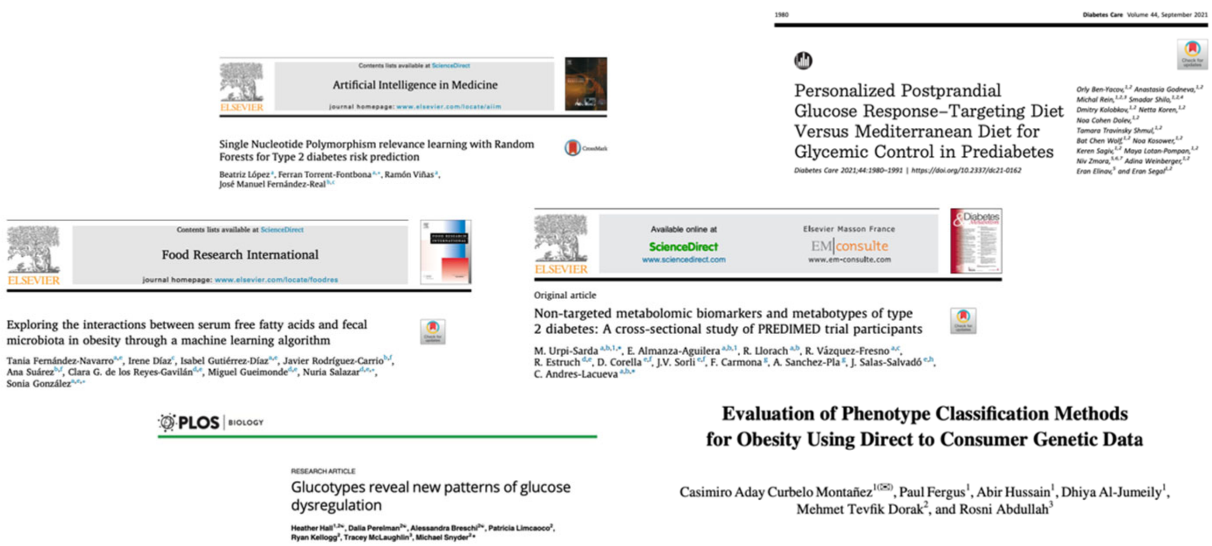
Precision nutrition



NUTRISS CENTRE NUTRITION, SANTÉ ET SOCIÉTÉ INAF UNIVERSITÉ LAVAL

10

Precision nutrition



Contents lists available at ScienceDirect
Artificial Intelligence in Medicine
journal homepage: www.elsevier.com/locate/aiim

Single Nucleotide Polymorphism relevance learning with Random Forests for Type 2 diabetes risk prediction

Beatriz López[a], Ferran Torrent-Fontbona[a,*], Ramón Viñas[a], José Manuel Fernández-Real[b,c]

1980 Diabetes Care Volume 44, September 2021

Personalized Postprandial Glucose Response–Targeting Diet Versus Mediterranean Diet for Glycemic Control in Prediabetes

Diabetes Care 2021;44:1980–1991 | https://doi.org/10.2337/dc21-0162

Orly Ben-Yacov,[1,2] Anastasia Godneva,[1,2] Michal Rein,[1,2,3] Smadar Shilo,[1,2,4] Dmitry Kolobkov,[1,2] Netta Koren,[1,2] Noa Cohen Dolev,[1,2] Tamara Travinsky Shmul,[1,2] Bat Chen Wolf,[1,2] Noa Kosower,[1,2] Keren Sagiv,[1,2] Maya Lotan-Pompan,[1,2] Niv Zmora,[5,6,7] Adina Weinberger,[1,2] Eran Elinav,[5] and Eran Segal[1,2]

Contents lists available at ScienceDirect
Food Research International
journal homepage: www.elsevier.com/locate/foodres

Exploring the interactions between serum free fatty acids and fecal microbiota in obesity through a machine learning algorithm

Tania Fernández-Navarro[a,e], Irene Díaz[c], Isabel Gutiérrez-Díaz[a,e], Javier Rodríguez-Carrio[b,f], Ana Suárez[b,f], Clara G. de los Reyes-Gavilán[d,e], Miguel Gueimonde[d,e], Nuria Salazar[d,e,*], Sonia González[a,e,*]

Available online at ScienceDirect www.sciencedirect.com
Elsevier Masson France EM|consulte www.em-consulte.com

Original article

Non-targeted metabolomic biomarkers and metabotypes of type 2 diabetes: A cross-sectional study of PREDIMED trial participants

M. Urpi-Sarda[a,b,1,*], E. Almanza-Aguilera[a,b,1], R. Llorach[a,b], R. Vázquez-Fresno[a,c], R. Estruch[d,e], D. Corella[e,f], J.V. Sorli[e,f], F. Carmona[g], A. Sanchez-Pla[g], J. Salas-Salvadó[e,h], C. Andres-Lacueva[a,b,*]

PLOS | BIOLOGY

RESEARCH ARTICLE

Glucotypes reveal new patterns of glucose dysregulation

Heather Hall[1,2☯], Dalia Perelman[2☯], Alessandra Breschi[2☯], Patricia Limcaoco[2], Ryan Kellogg[2], Tracey McLaughlin[3], Michael Snyder[2*]

Evaluation of Phenotype Classification Methods for Obesity Using Direct to Consumer Genetic Data

Casimiro Aday Curbelo Montañez[1(✉)], Paul Fergus[1], Abir Hussain[1], Dhiya Al-Jumeily[1], Mehmet Tevfik Dorak[2], and Rosni Abdullah[3]

NUTRISS CENTRE NUTRITION, SANTÉ ET SOCIÉTÉ INAF UNIVERSITÉ LAVAL

11

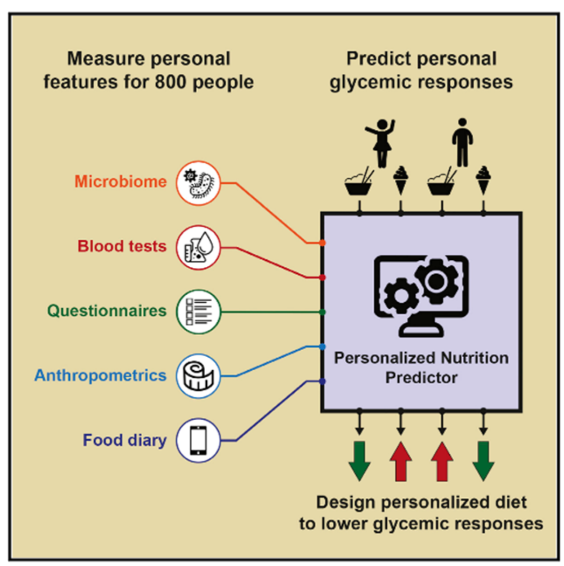
Precision nutrition

1980 Diabetes Care Volume 44, September 2021

Personalized Postprandial Glucose Response–Targeting Diet Versus Mediterranean Diet for Glycemic Control in Prediabetes

Diabetes Care 2021;44:1980–1991 | https://doi.org/10.2337/dc21-0162

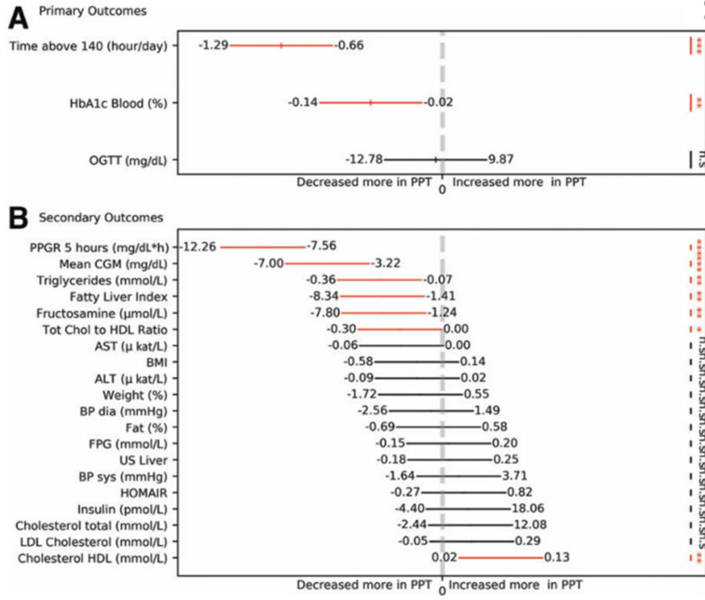
Orly Ben-Yacov,[1,2] Anastasia Godneva,[1,2] Michal Rein,[1,2,3] Smadar Shilo,[1,2,4] Dmitry Kolobkov,[1,2] Netta Koren,[1,2] Noa Cohen Dolev,[1,2] Tamara Travinsky Shmul,[1,2] Bat Chen Wolf,[1,2] Noa Kosower,[1,2] Keren Sagiv,[1,2] Maya Lotan-Pompan,[1,2] Niv Zmora,[5,6,7] Adina Weinberger,[1,2] Eran Elinav,[5] and Eran Segal[1,2]



NUTRISS CENTRE NUTRITION, SANTÉ ET SOCIÉTÉ INAF UNIVERSITÉ LAVAL

12

Precision nutrition



NUTRISS

INAF UNIVERSITÉ LAVAL

13

Application of AI in nutrition research

- Increased capacity to manage/analyse big data
  - –omics
  - precision nutrition
  - precision public health
- Dietary assessment
  - better understanding dietary patterns
  - image-based methods
  - non image-based methods
- Predicting outcomes
- Social media content analysis (NLP)

NUTRISS

INAF UNIVERSITÉ LAVAL

14

Application of AI in nutrition research

- Increased capacity to manage/analyse big data
  - –omics
  - precision nutrition
  - precision public health
- Dietary assessment
  - better understanding dietary patterns
  - image-based methods
  - non image-based methods
- Predicting outcomes
- Social media content analysis (NLP)

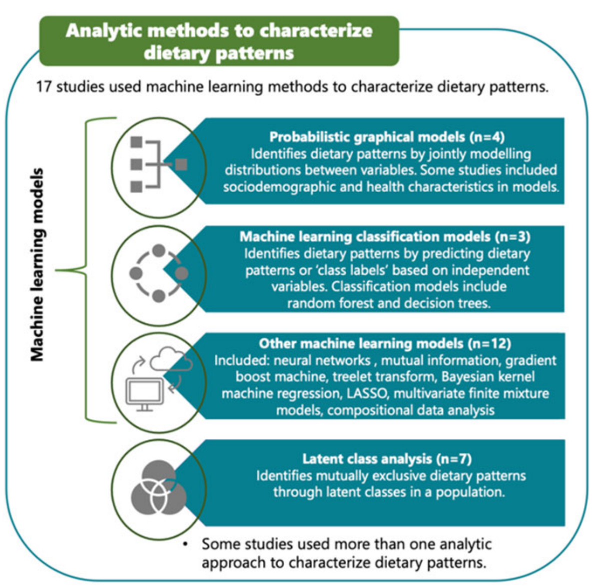
NUTRISS

INAF

UNIVERSITÉ LAVAL

15

Hutchison, Kirkpatrick et al



Probabilistic graphical models

ML (classification)

ML (other)

Latent class analysis

NUTRISS

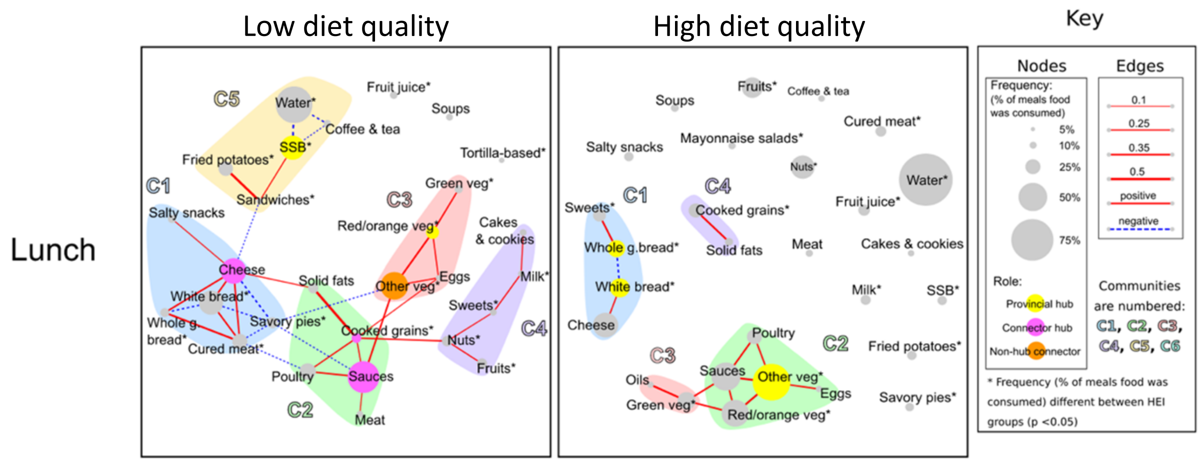
INAF

UNIVERSITÉ LAVAL

16

Food network to understand meal patterns vs HEI

Schwedhelm et al IJBNPA 2021;18:101



NUTRISS

INAF UNIVERSITÉ LAVAL

17

Application of AI in nutrition research

- Increased capacity to manage/analyse big data
  - –omics
  - precision nutrition
  - precision public health
- Dietary assessment
  - better understanding dietary patterns
  - image-based methods
  - non image-based methods
- Predicting outcomes
- Social media content analysis (NLP)

NUTRISS

INAF UNIVERSITÉ LAVAL

18

Application of AI in nutrition research

- Increased capacity to manage/analyse big data
  - –omics
  - precision nutrition
  - precision public health
- Dietary assessment
  - better understanding dietary patterns
  - image-based methods
  - non image-based methods
- Predicting outcomes
- Social media content analysis (NLP)

NUTRISS CENTRE NUTRITION, SANTÉ ET SOCIÉTÉ

INAF UNIVERSITÉ LAVAL

19

Prediction of outcomes

British Journal of Nutrition (2018), **120**, 326–334 doi:10.1017/S0007114518001150
© The Authors 2018

A comparison of statistical and machine-learning techniques in evaluating the association between dietary patterns and 10-year cardiometabolic risk (2002–2012): the ATTICA study

Dimitris Panaretos[1], Efi Koloverou[1], Alexandros C. Dimopoulos[2], Georgia-Maria Kouli[1], Malvina Vamvakari[2], George Tzavelas[3], Christos Pitsavos[4] and Demosthenes B. Panagiotakos[1]*

Are Machine Learning Algorithms More Accurate in Predicting Vegetable and Fruit Consumption Than Traditional Statistical Models? An Exploratory Analysis

Mélina Côté[1,2], Mazid Abiodoun Osseni[3,4], Didier Brassard[1,2], Élise Carbonneau[1,2], Julie Robitaille[1,2], Marie-Claude Vohl[1,2], Simone Lemieux[1,2], François Laviolette[1,3,4] and Benoît Lamarche[1,2]*

529

Canadian Science Publishing ARTICLE

Development of machine learning prediction models to explore nutrients predictive of cardiovascular disease using Canadian linked population-based data

Jason D. Morgenstern, Laura C. Rosella, Andrew P. Costa, and Laura N. Anderson

Open access Original research

BMJ Open Machine learning with sparse nutrition data to improve cardiovascular mortality risk prediction in the USA using nationally randomly sampled data

Joseph Rigdon,[1] Sanjay Basu[2]

Contents lists available at ScienceDirect

Journal of Affective Disorders

journal homepage: www.elsevier.com/locate/jad

ELSEVIER

Research paper

Identifying depression in the National Health and Nutrition Examination Survey data using a deep learning algorithm

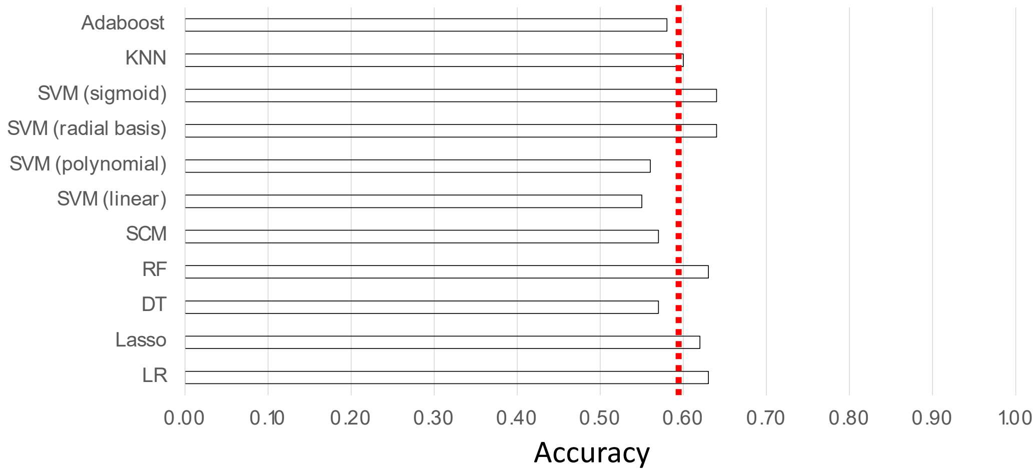
Jihoon Oh[a,1], Kyongsik Yun[b,c,1], Uri Maoz[b,d,e,f], Tae-Suk Kim[a], Jeong-Ho Chae[a,*]

NUTRISS CENTRE NUTRITION, SANTÉ ET SOCIÉTÉ

INAF UNIVERSITÉ LAVAL

20

Predicting adequate Veg/fruit consumption

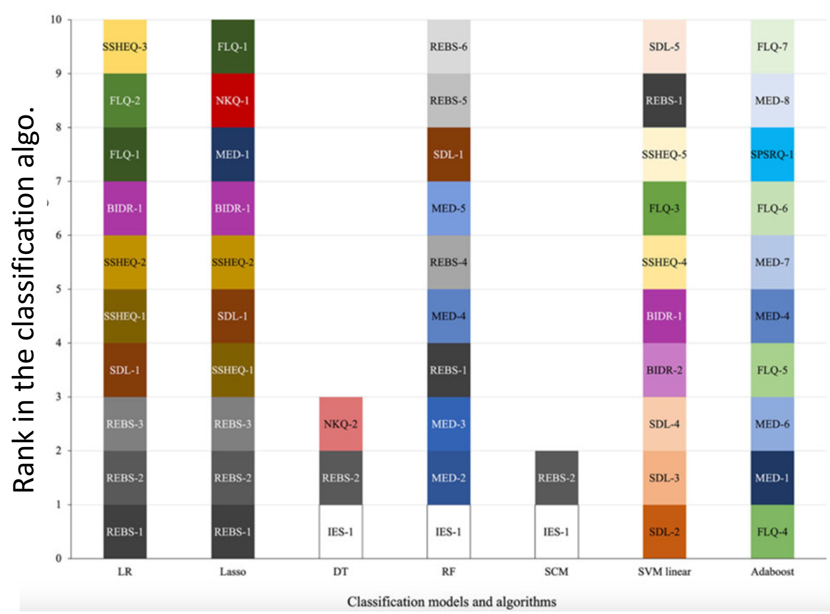


Côté et al, Front Nutr 2022

NUTRISS INAF UNIVERSITÉ LAVAL

21

Predicting adequate Veg/fruit consumption



Côté et al,
Front Nutr 2022

NUTRISS INAF UNIVERSITÉ LAVAL

22

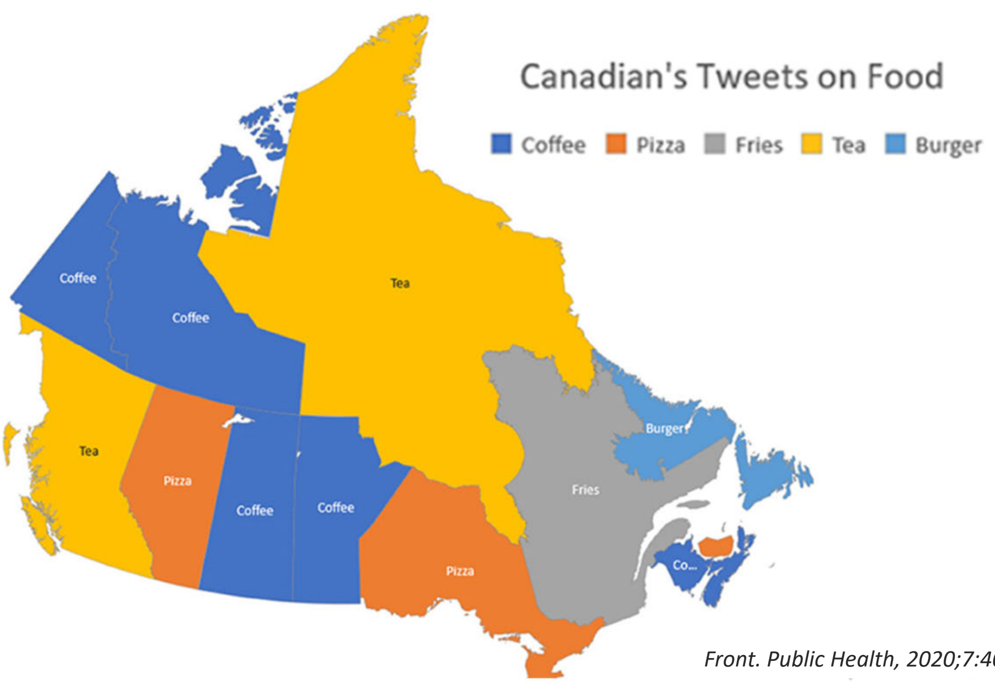
Application of AI in nutrition research

- Increased capacity to manage/analyse big data
  - –omics
  - precision nutrition
  - precision public health
- Dietary assessment
  - better understanding dietary patterns
  - image-based methods
  - non image-based methods
- Predicting outcomes
- Social media content analysis (NLP)

NUTRISS INAF UNIVERSITÉ LAVAL

23

NLP



NUTRISS

Front. Public Health, 2020;7:400+

24

Take home messages

- AI in nutrition: great promises, great challenges
  - Quality of data (garbage in...)
  - Learn a common language
  - Standardization of methods/approaches
- Keeping an open mind, no revolution here
- Training the next generation of “bilingual” researchers

NUTRISS INAF UNIVERSITÉ LAVAL

25

JEAN MAYER USDA HUMAN NUTRITION RESEARCH CENTER ON AGING

HNRCA

Designing studies to collect nutrition data for AI analysis

Sai Krupa Das, PhD
Senior Scientist & Professor
Jean Mayer USDA Human Nutrition Research Center on Aging
Friedman School of Nutrition Science and Policy
Tufts University

Tufts UNIVERSITY

1

Introduction

- The growing intersection of Artificial Intelligence (AI) and nutrition research provides exciting opportunities for streamlining conventional research protocols and revolutionizing clinical nutrition applications.
- AI can be especially useful in analyzing complex data, in informing adaptive study designs or providing personalized nutrition interventions.
- However, these advancements come with challenges that must be addressed using multidisciplinary teams.

Côté et al. (2022); Sarker (2022)

2

Precision nutrition science – the time is now!

- Personalized nutrition refers to individually "tailored nutritional recommendations aimed at the promotion, maintenance of health and prevention against diseases"

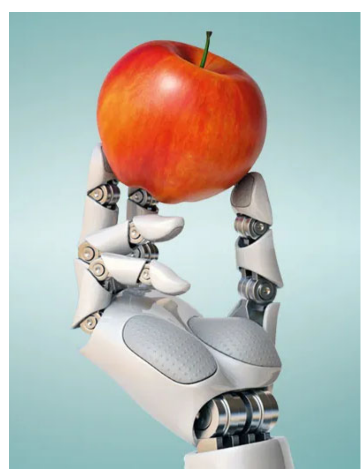
- Personalized nutrition uses diet as an intervention point to reach these favorable health outcomes.
  - These recommendations rely on the observation that **the same diet can produce variable responses across individuals.**
    - Variable responses are a product of the complex interactions among **internal** (e.g., microbiome, metabolome, genetics) and **external** (e.g., diet, physical activity) factors.

Verma et al. *Front Nutri.* 2018;5: 117.

3

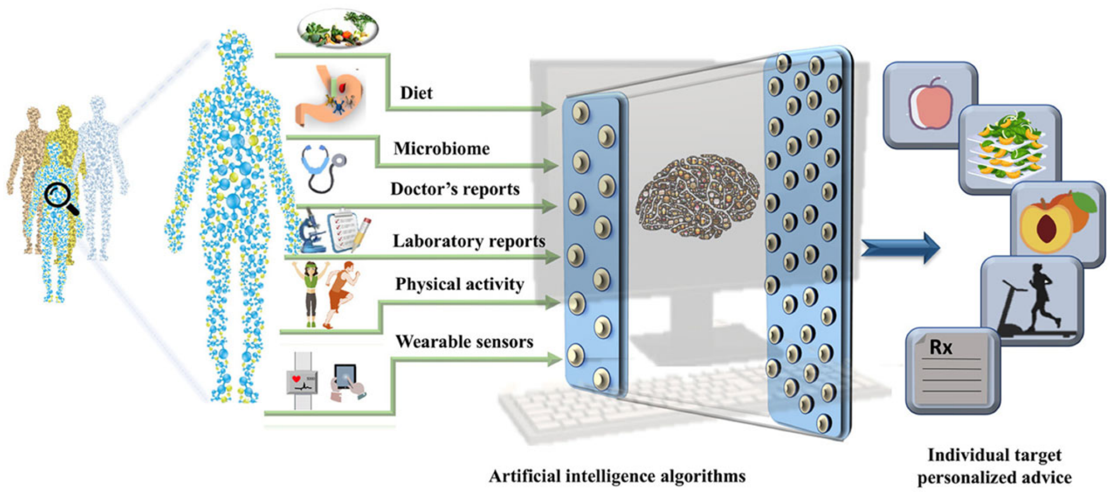
The intersection of AI and nutrition research

- Much of the difficulty in crafting personalized nutrition recommendations lies in understanding the interplay among parameters that produce **inter-individual** variation.
  - Current approaches primarily lie in fields like genomics, proteomics, and metabolomics.
  - However, a comprehensive understanding requires an integrated, "systems-wide" approach.
- In the context of disease prevention, studies investigating these variable responses to the same intervention must also quantify the unique responses from people in various disease states.



4

Multi-scale, multi-systems approach



Verma et al. *Front Nutri.* 2018;5: 117.

5

Current challenges: single outcome focus

Adapted from Verma et al. *Front Nutri.* 2018;5: 117.

6

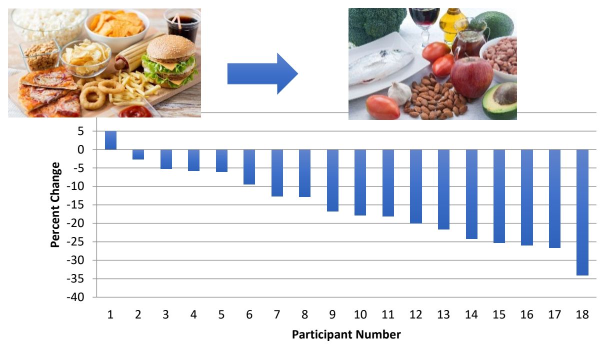
Departing from a reductionist approach

- Current objectives for nutrition data science require a departure from interrogating/examining singular biological pathways or outcomes of interest.
  - Rather, needs the investigation of dynamic interactions between pathways that are not always linear

We are different externally…

…and internally

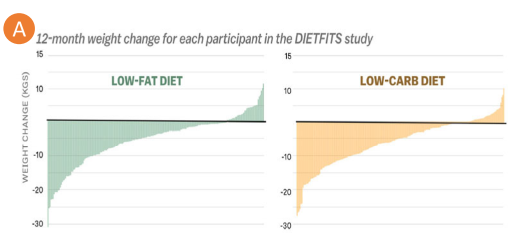
LDL-C Response to Diet Intervention



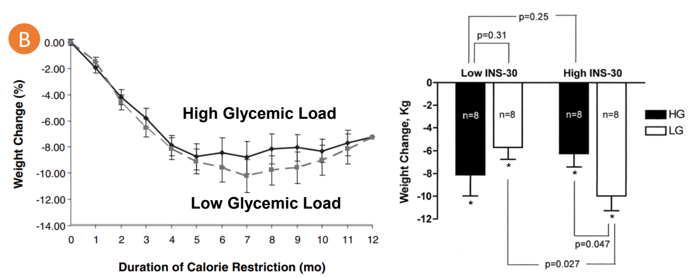
****A heart-healthy diet benefits some but not all****

7

Response variability: the current focus

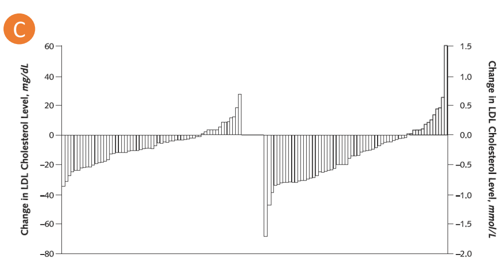


Gardner CD, et al. *JAMA.* 2018;319(7):667–679.

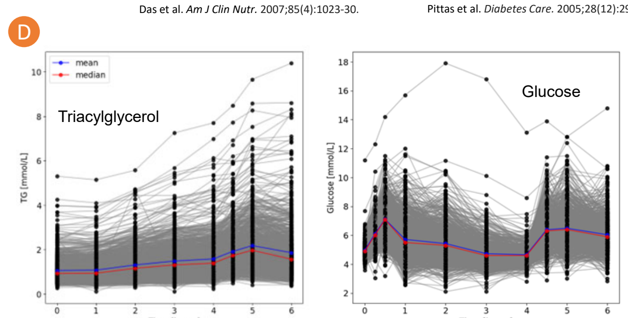


Das et al. *Am J Clin Nutr.* 2007;85(4):1023-30.

Pittas et al. *Diabetes Care.* 2005;28(12):2939-41.



Gardner CD, et al. *Ann Intern Med.* 2005;142:725-33.

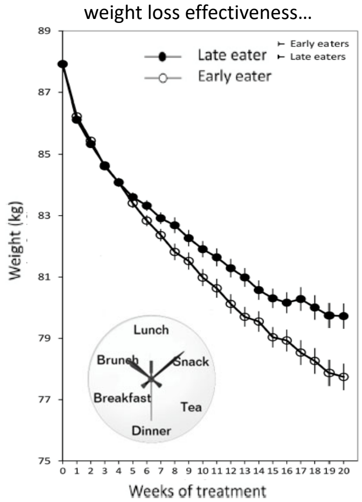


Berry et al. *Nat Med.* 2020 Jun;26(6):964-973.

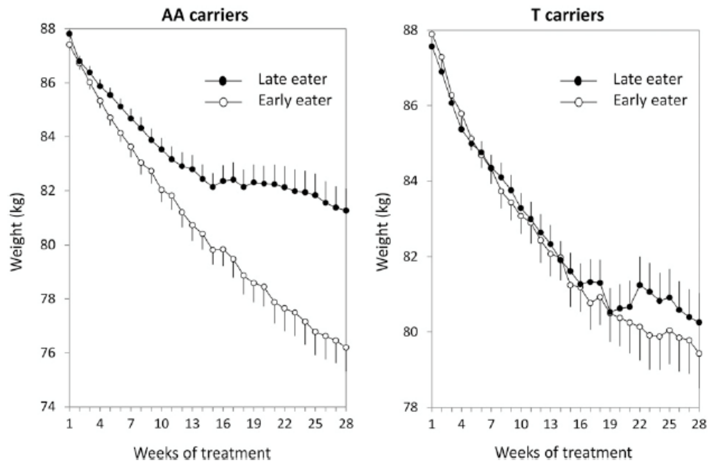
8

Response variability: eating timing x genes

Timing of food intake predicts weight loss effectiveness…



PLIN1 genetic variation **14995 A>T variants x food timing (< 3 pm OR > 3pm)**

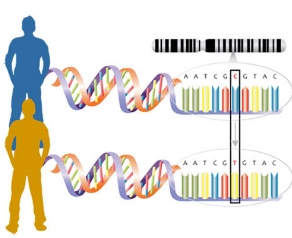


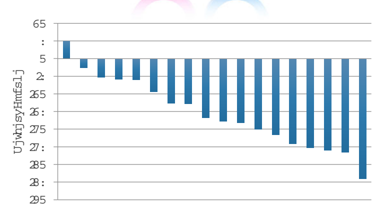
Garaulet, M, et al., *International journal of obesity (2005)*, *37*(4), 604–611.

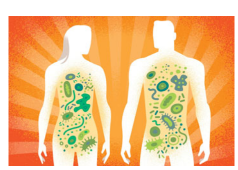
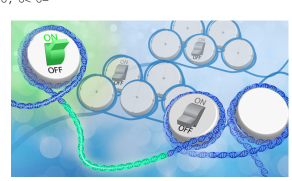
Garaulet et al. *Am J Clin Nutr.* 2016;104(4):1160-6.

9

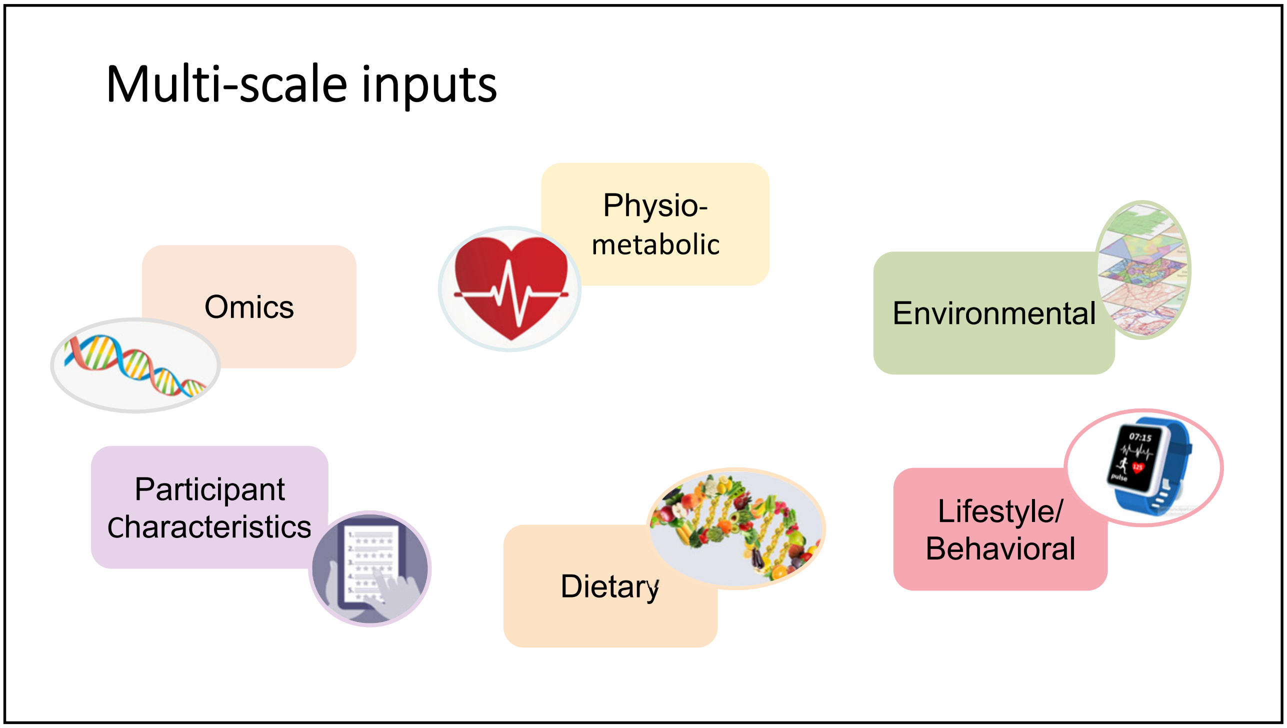
Factors associated with the inter-individual variability in response



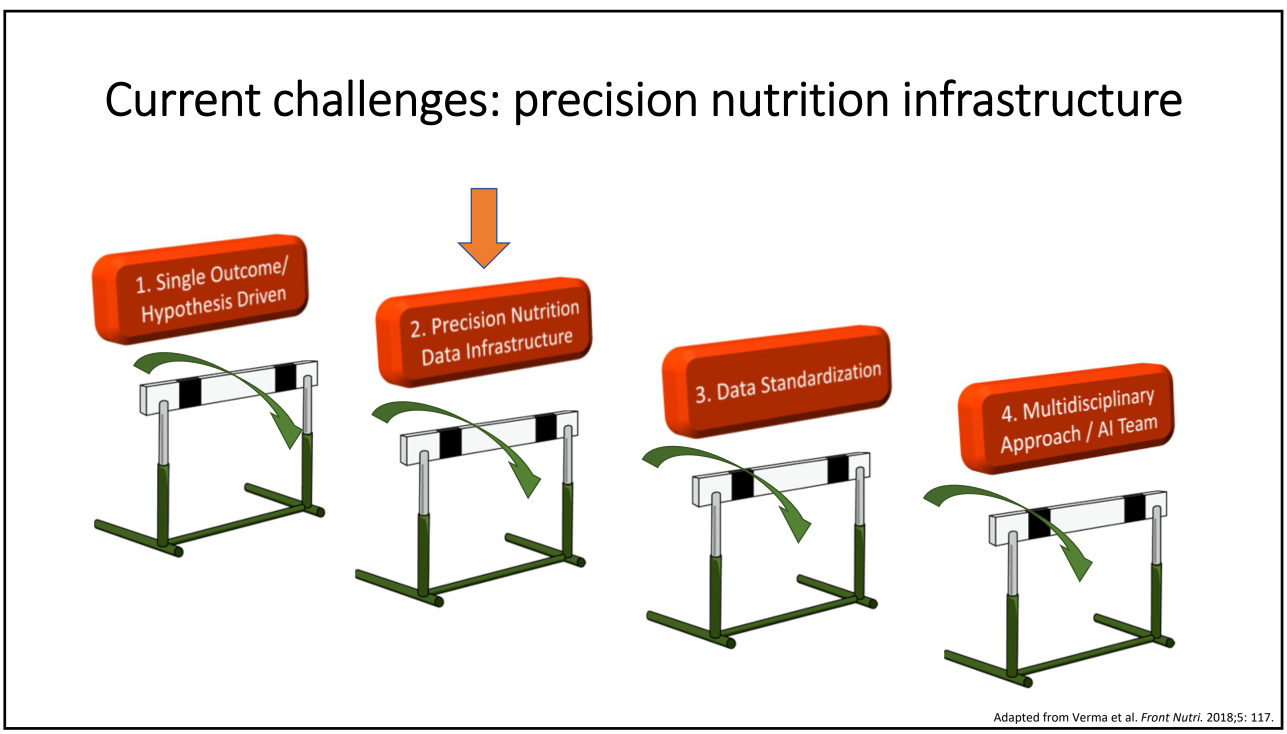


10


Multi-scale inputs
Omics
Physio-
metabolic
Environmental
Participant
Characteristics
Dietary
Lifestyle/
Behavioral

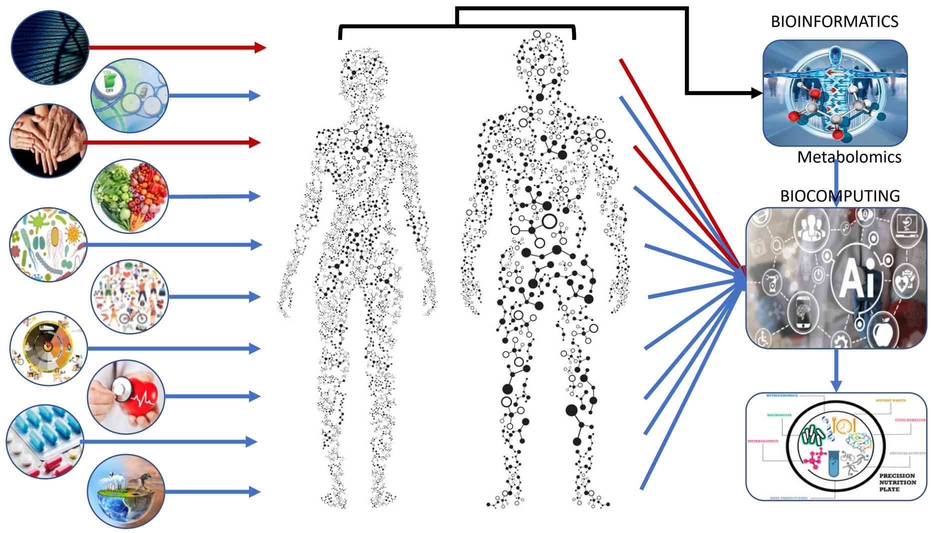
11


Current challenges: precision nutrition infrastructure
1. Single Outcome/
Hypothesis Driven
2. Precision Nutrition
Data Infrastructure
3. Data Standardization
4. Multidisciplinary
Approach / AI Team
Adapted from Verma et al. Front Nutri. 2018;5: 117.

12

Personalized nutrition computational infrastructure

INTERSECTION OF BIOLOGY AND TECHNOLOGY



Müller et al. *Eur J Clin Nutr.* 2020 Dec;74(12):1619-1629.

13

Personalized Nutrition Computational Infrastructure

- The core requirement of the infrastructure is that it needs to be classified and identified as a **food and health infrastructure (FHI).**
- The need for research infrastructures in the specific areas of food and nutrition were recently highlighted by the EuroDISH consortium's (**DISH** model).
    - **D**eterminants of food choice – key drivers of food and lifestyle choices
    - **I**ntake of foods and nutrients – past and current
    - **S**tatus and functional markers of nutritional health
    - **H**ealth and disease risk
- **Hard** Research Infrastructures – toolkits or technical equipment
- **Soft** Research Infrastructures – communication networks, methodologies and conceptual frameworks.

EuroDISH

https://www.eurofir.org/our-resources/past-projects/eurodish/

14

Hard and soft research infrastructures

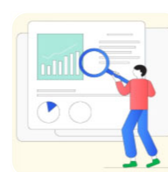
Term	Definition
Applications (apps)	Mobile applications used on a smart phone, tablet, or computer
Devices	A subset of lifestyle technology products with successful FDA approval for safety and effectiveness
Digital health	Health technology products that do not require validity or efficacy or regulatory oversight
Wearable technology	General term for body-worn sensors capable of tracking location, time, environment, motion, and certain body measures (e.g., blood glucose, etc.)
Digital medicine	Health technology products used for measurement/intervention that are supported by evidence to demonstrate quality and validity
Digital therapeutics	Evidence-based health technology products that deliver a health intervention and have been reviewed or certified by a regulatory board
Telehealth or telemedicine	Use of electronic information and telecommunication technologies to deliver and support long-distance clinical health care, patient- and professional health-related education, public health, and health administration
Health information technology (HIT)	Electronic medical records and related information systems
Web-based assessment	Tool requiring internet connection; often a "cloud-based" data source

Adapted from McClung et al 2022. *JANA*. 2022;122(1):207-18.

15

Food health infrastructure

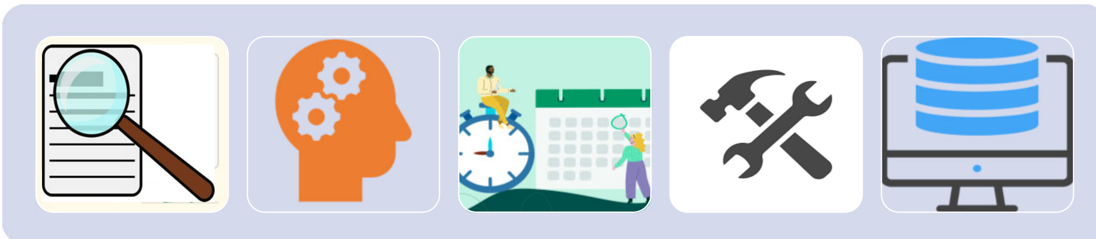
- The management and implementation of the FHIs should be driven from the user level and comprise
  - nutrition bioinformatics structures, including all biologically relevant data, preprocessed omics as well as descriptive and study participant phenotype data; prioritizing n-of-1 data;
  - data management;
  - data processing; informatics infrastructure with standardized food intake monitoring
  - data sharing capabilities; and
  - platforms for publishing the data derived from the studies to a bigger community (e.g., web portals).
- Establishing an FHI will ensure that the data related to food constituents, intake, environmental variables, determinants of health, energy expenditure, and disease risk are all in one place.
- These FHI data can help reveal the determinants of behavior, which can be used to develop prediction algorithms and nutritional interventions.



Verma et al. *Front Nutri.* 2018;5:

16

Dietary data collection



Methodology
- Food intake diary/record
- 24-hour recall
- Questionnaires
- Devices
- Apps

Biases
- Recall
- Reactivity
- Reporting
- Portion Size

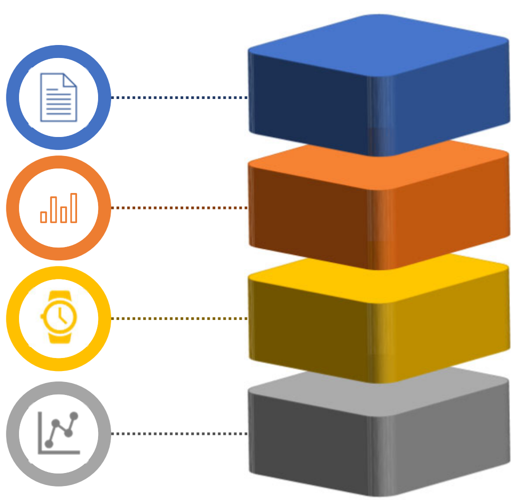
Adherence
- Definition
- Data collection
- Data reporting

Tools
- Devices
- Apps
- Measures or reference standards

Databases
- Standardized source and format
- Inclusive of ethnic foods
- Sync with other measures

17

Physical activity



Questionnaires
- Types of activity, period of capture, comprehensive list of activity types

Tracking Apps
- Phones, input/entry requiring active OR passive methods

Wearables
- Watches, sensors, trackers

Data Cleaning and Integration

18

Biosamples

Biospecimen Type
- Plasma/serum
- Tissue
- Stool
- Hair/nails
- Saliva

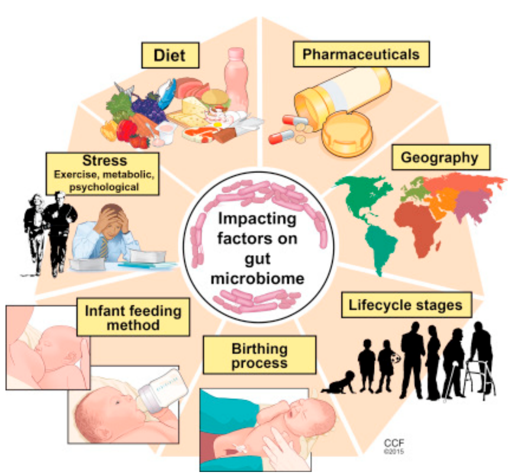
Measurement
- Continuous
- Snapshot in time

Processing
- Collection
- Storage
- Shipping
- Handling

Analysis
- Intra-study differences
- Inter-study differences

19

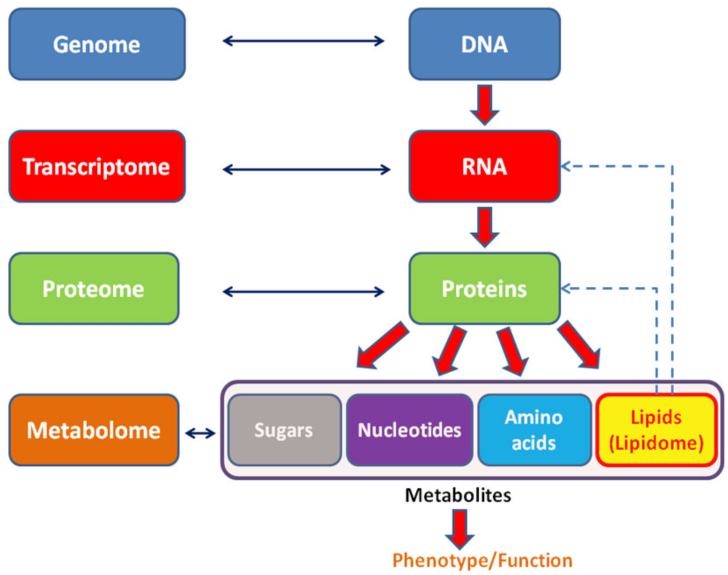
Factors influencing the gut microbiota



Nutritional, Medical, and Surgical Management, 2019, Chapter 4; gut microbiome

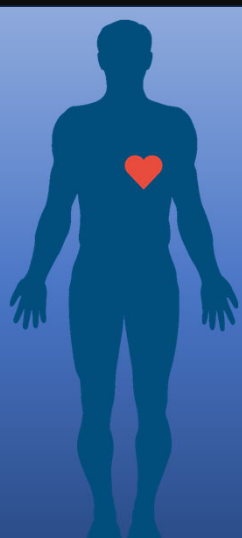
20

Analysis – omics analyses



21

Electronic Health Records

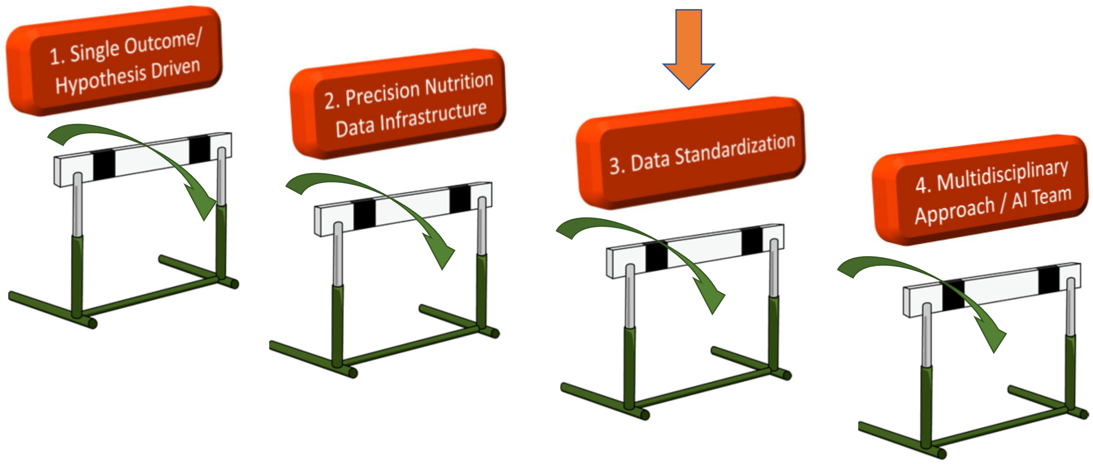


EHRs store a lot of clinical data, but the following factors must be considered:

- Data can be heterogenous, complex, and come in a mixed format of structured/unstructured forms
  - E.g., imaging, lab reports, doctor's notes
- Medication use
- PHI & sensitive information
- Release & integration with precision database

22

Current challenges: data standardization



Adapted from Verma et al. *Front Nutri.* 2018;5: 117.

23

Data Standardization

- Benefits of digitizing records include having complete relational data for each person and across participants/cases
- However, digitization can lead to challenges such as
  - Improper standardization formats
  - Lack of user interface training
  - Poorly designed technology can lead to errors in the record
- Current needs include
  - Standardizing data formats
  - *A priori* communication between AI and clinical teams from design to study setup and implementation
  - Providing sample datasets to train systems, or enhance seamless integration of multi-scale data
- Data standardization can make updating these datasets more feasible while also improving communication between clinical sites

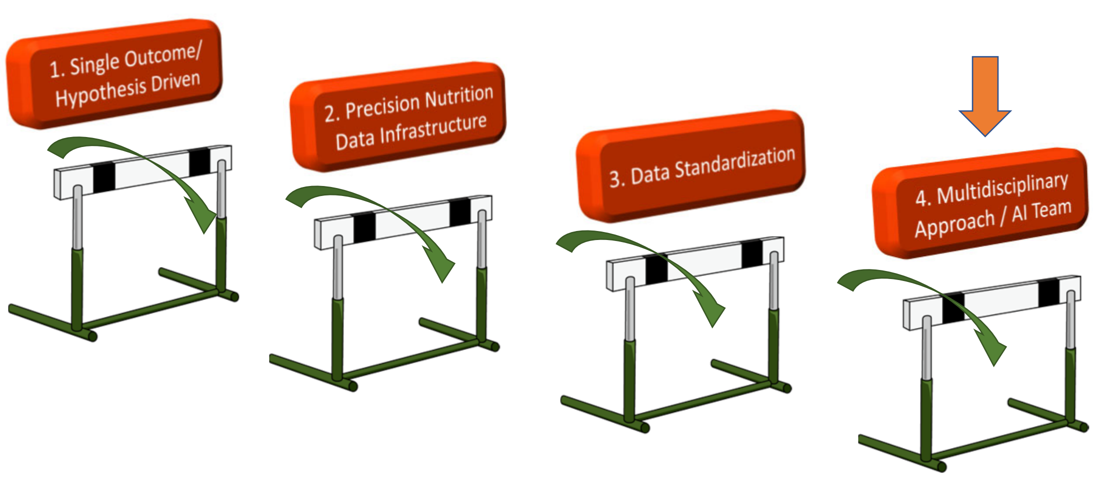
24

Missing Data

- Missing data can lead to biased and misleading results. How can we fix this?
  - Imputation methods for missing data:
    - **Mean/Median imputation** – missing values substituted for mean/median values
      - Introduces bias for extreme values
    - **K-Nearest Neighbors** – values from the grouped individuals can be averaged and assigned to the missing variable
      - May fail in cases where individuals cannot be well separated in groups based on their clinical record values
    - **Multiple imputation by chained equations (stochastic method)** – iterative process considering relationships between variables
      - Assumes normally-distributed data; excluding non-normally distributed data can lead to bias
      - Computationally intensive

25

Current challenges: team science



Adapted from Verma et al. *Front Nutri.* 2018;5: 117.

26

Team science approach for generating AI-ready data

How do we adapt to the widespread use of data-driven technologies?

- Support a team science approach and the development of professionals with clinical and computational skills.

Engagement

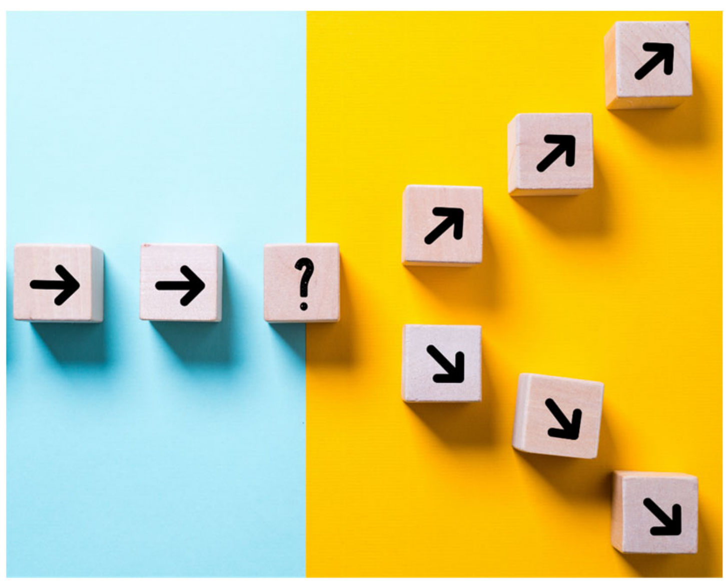
- Integrate experts in all relevant fields to provide insight from diverse perspectives
- Engage research team in all aspects of the project, from inception and planning to completion, analysis, and interpretation of findings
- Engage data scientists as integral peer collaborators

Training & Education

- Conduct periodic multidisciplinary trainings (e.g., in biology, nutrition, biomedicine, computer science, statistics, mathematics)
  - Provide an overview of the latest available technology, data standards, and methodologies
  - Consider cutting-edge knowledge informed by the change in day-to-day informatics challenges
  - Ensure users understand how to utilize technology and navigate big data to advance predictive capabilities
  - Emphasize that AI is a *tool* to facilitate decision-making, not a replacement for human experts

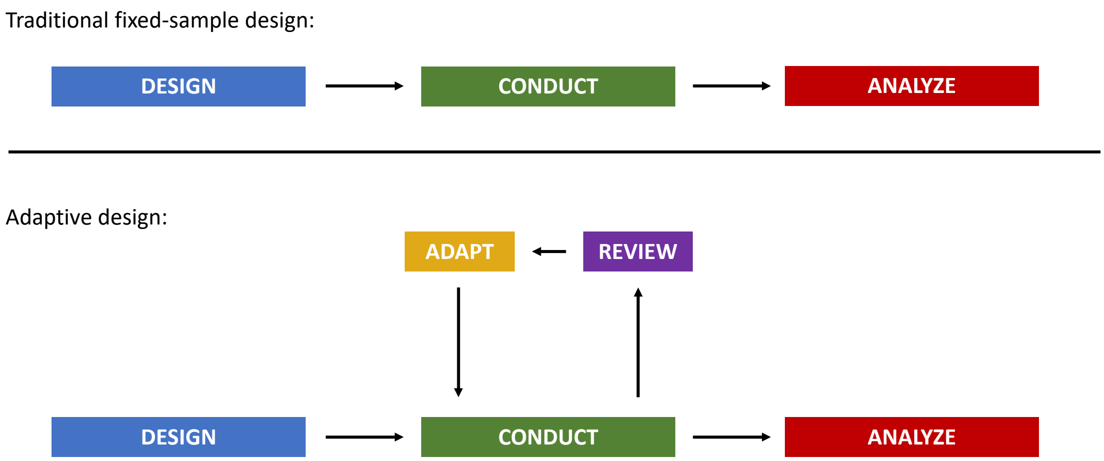
27

AI as a tool in informing study designs



28

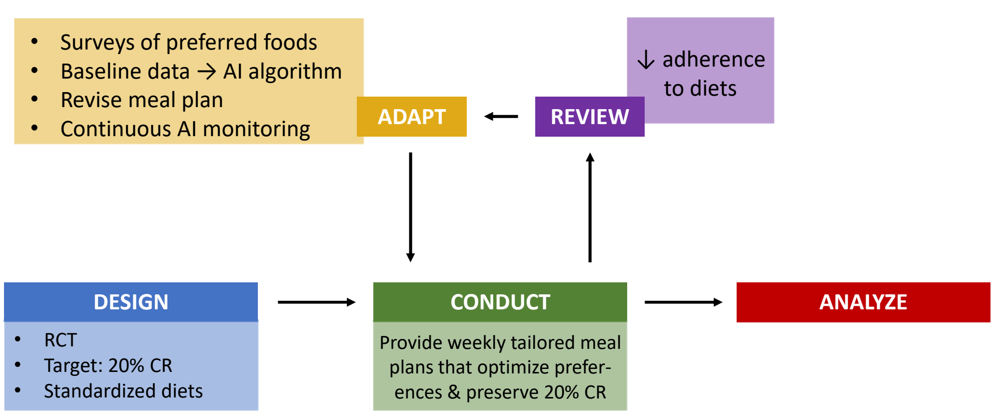
Adaptive study design



Adapted from Pallmann et al. *BMC medicine.* 2018;16(1):1-5.

29

Adaptive study example: calorie restriction



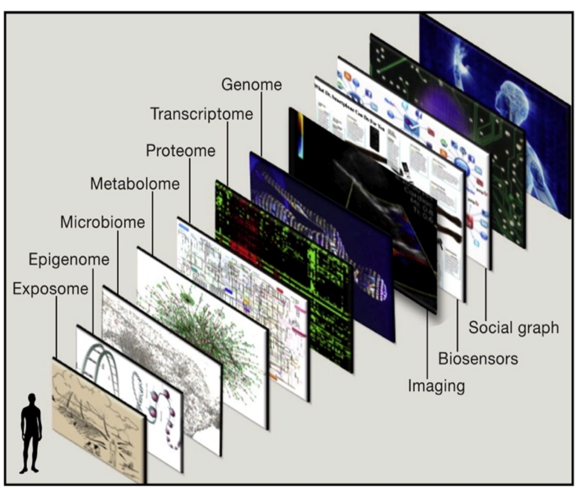
30

Integrated personalized predictive models

- The current age of big data provides exciting opportunities to integrate data on food consumption and electronic health records to create "synthetic human cohorts."
  - These cohorts can be used in responses to nutrition recommendations at the system level
- One of the primary goals for personalized nutrition is to create predictive models that draw on a history of monitored health responses.

31

Geographic information system (GIS) of a human being



The ability to digitize the medical essence of a human being is predicated on the integration of multiscale data, like a Google map, which consists of superimposed layers of data such as street, traffic, and satellite views.

For a human being, these layers include

- demographics and social graph;
- biosensors to capture the individual's metabolism, activity, and lifestyle
- imaging to depict the anatomy (often along with physiologic data);
- biology from the various omics (genome-DNA sequence, transcriptome, proteome, metabolome, microbiome, and epigenome) called **"panor-ome"**; and
- environmental exposure data, known as the **"exposome."**

Topol EJ. Cell. 2014;157(1):241-53.

32

Market ahead of the science?

Complete your tests
Test your gut, blood fat, and blood sugar responses with at-home test kit

Get your insights
About 6 weeks later receive your personalized insights report

Retrain your biology
Put your insights into action with a 4-week plan tailored to your biology

Eat for life
Get 4 months access to our app to help you sustain change and thrive
ZOE GLOBAL

https://zoe.com

33

Considerations for AI & precision health

- Address system-level challenges in the pre-competitive space, including regulatory environment, scientific and communication standards, and behavioral considerations to drive consumer adoption and effectiveness of precision health
- AI technologies must use interfaces that accommodate all levels of technological literacy.
- Promote equitable access and implementation of strategies in the health-disease continuum from wellness optimization, to functional maintenance, to disease risk reduction and disease management

34

Summary

- Nutritional factors are determined by multiple factors.
  - E.g., physiology, omics analysis, food intake, diet, metabolism, physical activity
- AI- Machine Learning, and deep learning can help provide personalized recommendations for improved health and well-being.
- Data-driven approach comes with the need to develop an integrated system to assess associated benefits and challenges.
- Longitudinal data on physiology, microbiome, and other relevant biomarkers can strengthen personalized programs
- Protection of individual data privacy and preventing discrimination is of top priority.
- Overall, a precision infrastructure framework establishes preventative and predictive guidelines for health and disease management.

35

AI and the Big Challenges in Ag and Food

- Agriculture has been hugely successful
- Big health and environmental challenges
- How AI can be a solution

Verdant Robotics weeder

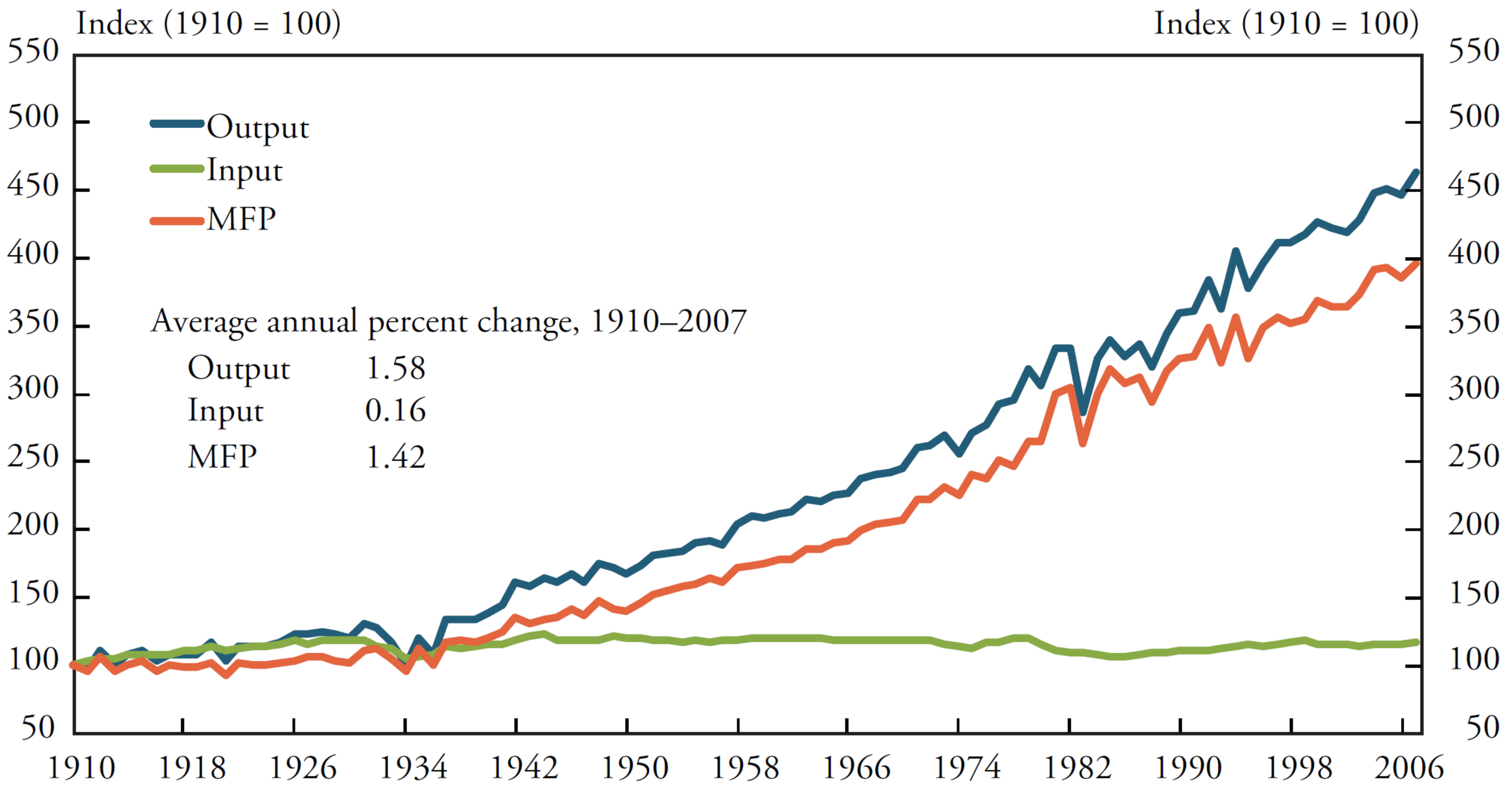
Aaron Smith
DeLoach Professor of Agricultural Economics
UC Davis
https://asmith.ucdavis.edu

AIFS AI INSTITUTE FOR FOOD SYSTEMS

UCDAVIS UNIVERSITY OF CALIFORNIA

Agricultural & Resource Economics UCDAVIS

Quantity Indexes of Output, Input, and MFP, U.S. Agriculture, 1910–2007

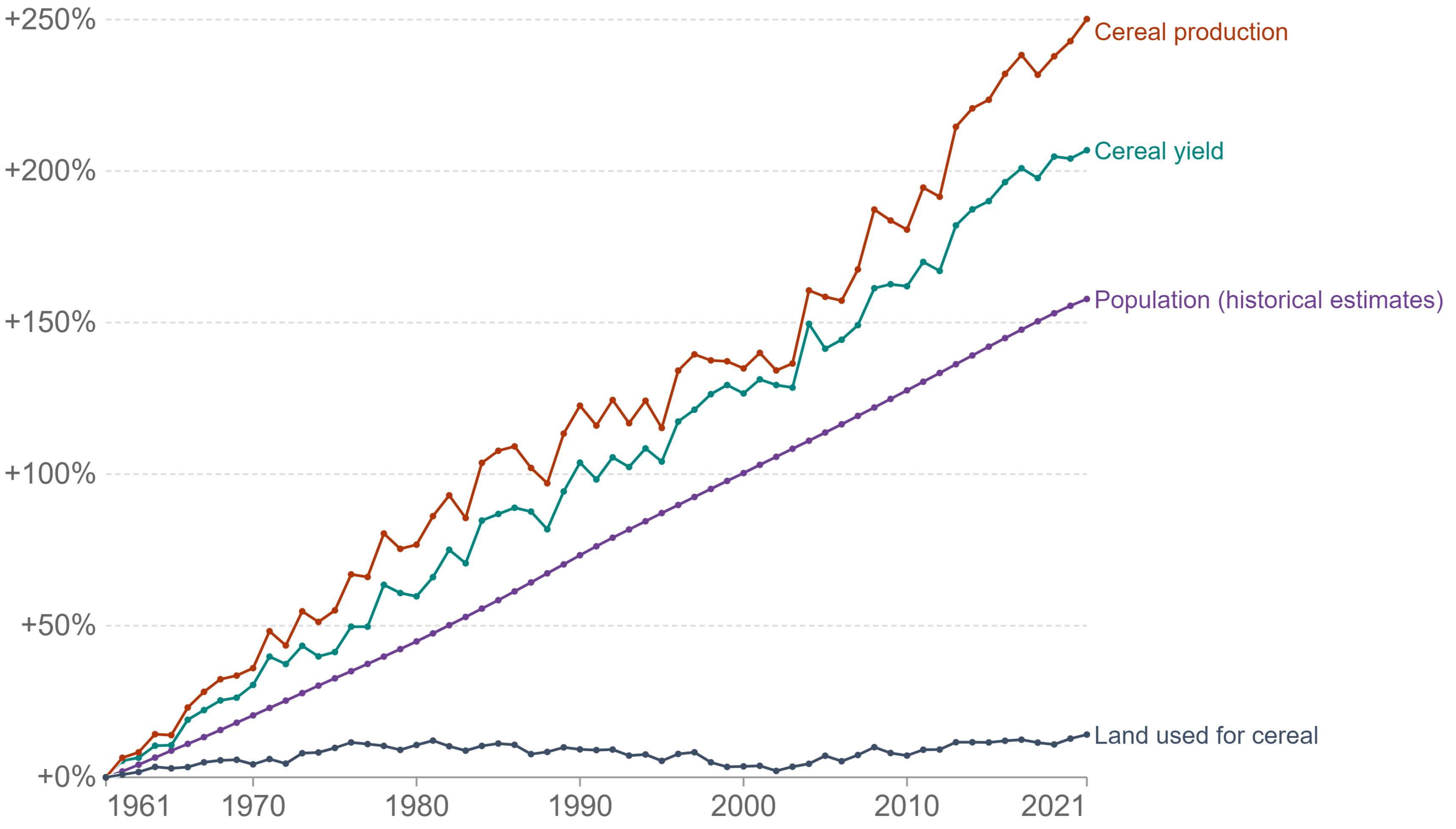


Source: Abridged version of Figure 1 in Pardey and Alston (forthcoming).

American farmers now produce four times as much with almost the same inputs

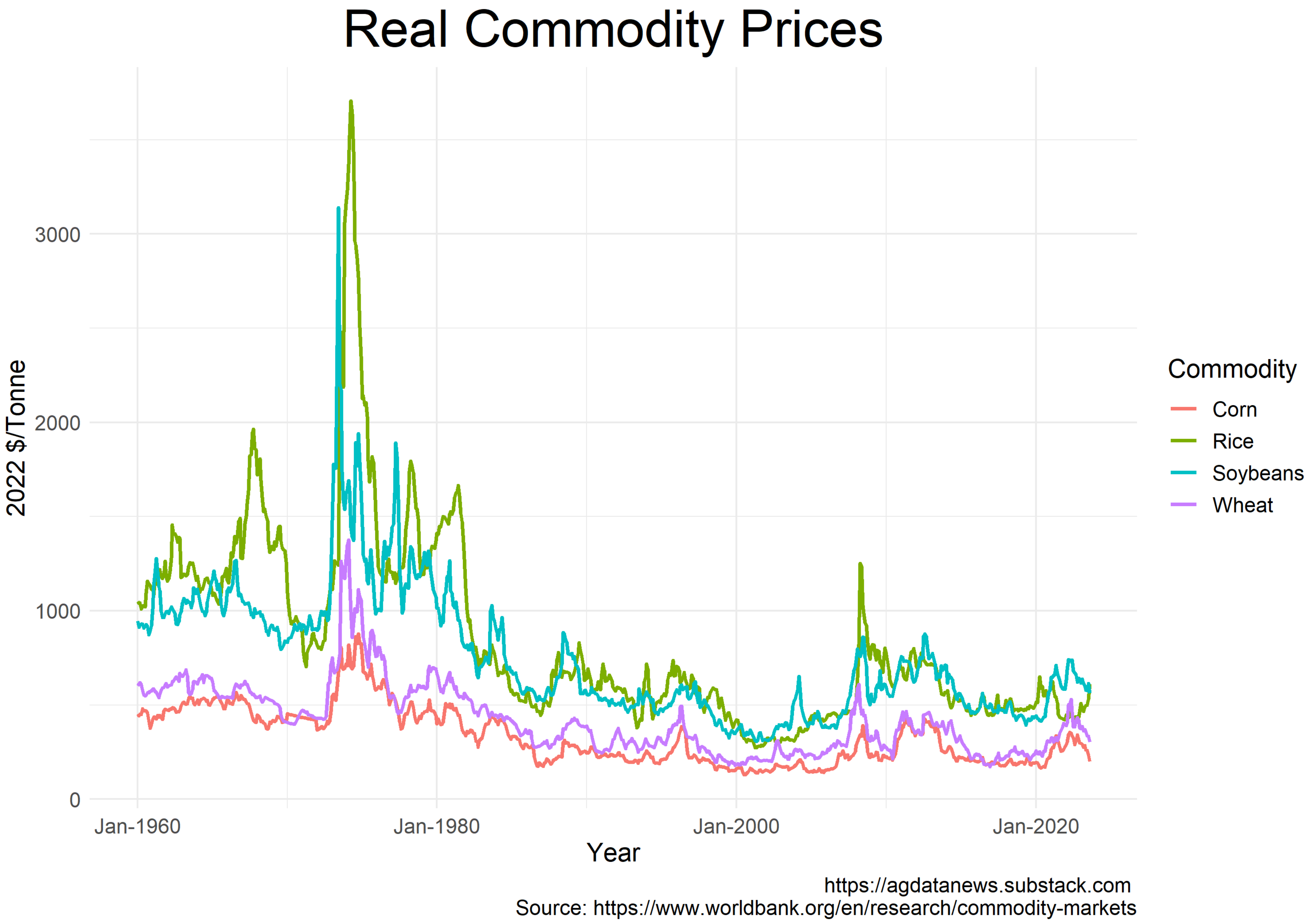
Change in cereal production, yield, land use and population, World

All figures are indexed to the start year of the timeline. This means the first year of the time-series is given the value zero.



Source: Our World in Data based on World Bank; Food and Agriculture Organization of the United Nations

The world produces 250% more cereals with only 15% more land.



Higher productivity means lower prices

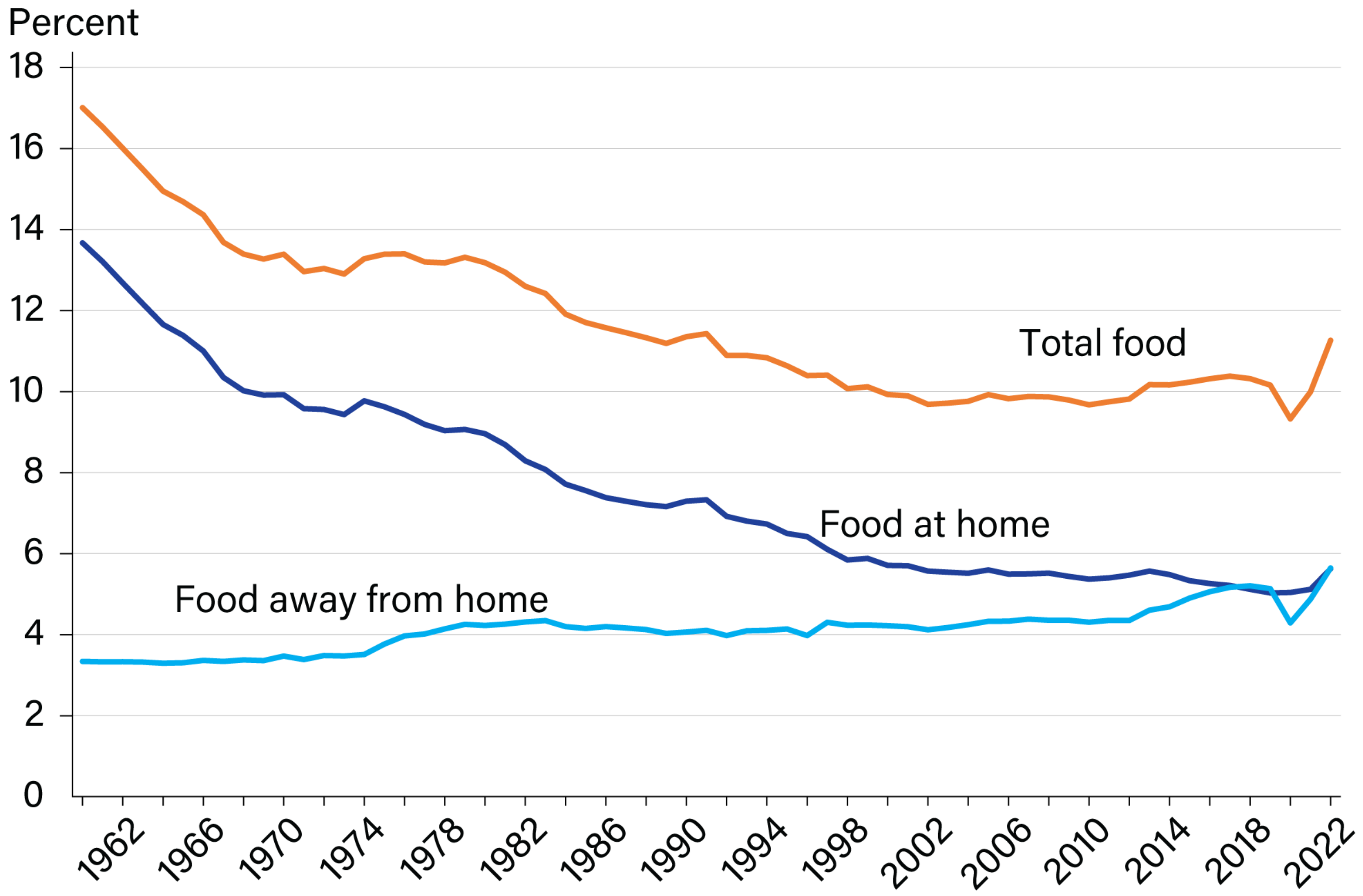
Prices down 50% since 1960 (adjusted for inflation)

AIFS AI INSTITUTE FOOD SYSTEMS

UC DAVIS
UNIVERSITY OF CALIFORNIA

Agricultural & Resource Economics UC DAVIS

Share of disposable personal income spent on food in the United States, 1960–2022

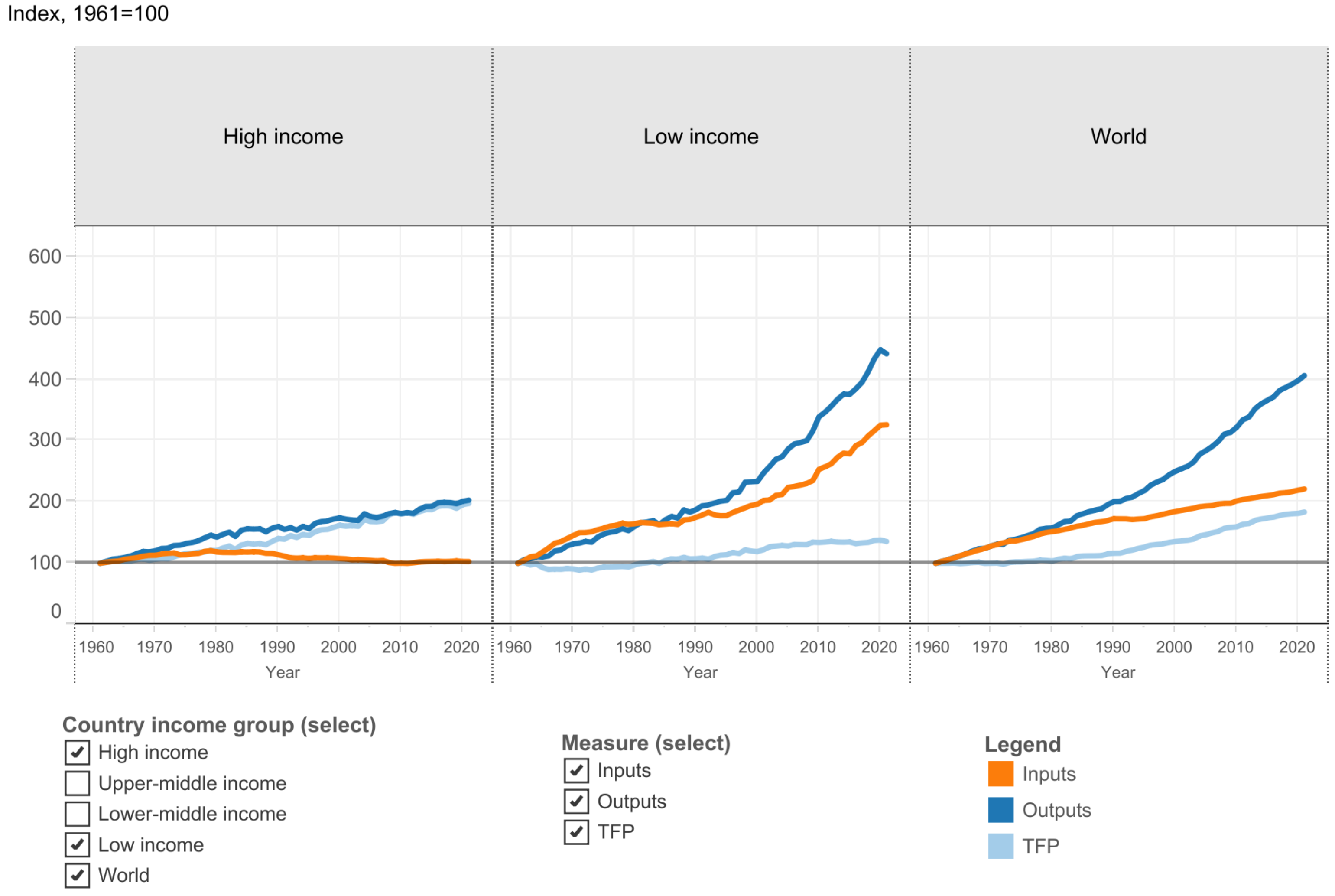


Note: Percentages are calculated using nominal values.

Source: USDA, Economic Research Service, Food Expenditure Series.

Food now a much lower percent of American household budgets

Trends in agricultural outputs, inputs, and total factor productivity (TFP) by country income group, 1961–2021



Source: USDA, Economic Research Service, *International Agricultural Productivity* data product.
Data and methods as of September 2023.

Productivity has improved the most in rich countries

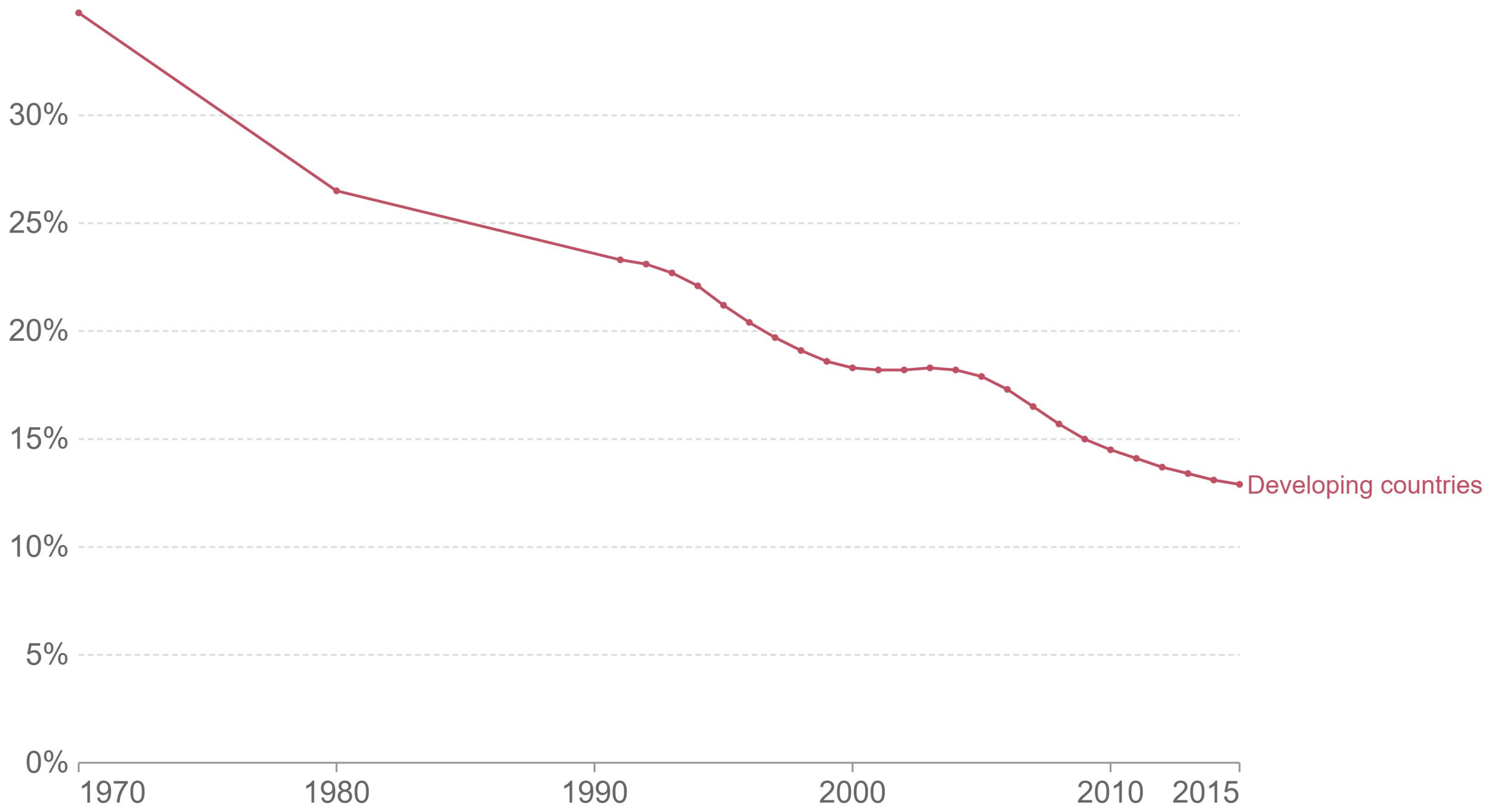
AIFS AI INSTITUTE FOOD SYSTEMS

UC DAVIS UNIVERSITY OF CALIFORNIA

Agricultural & Resource Economics UC DAVIS

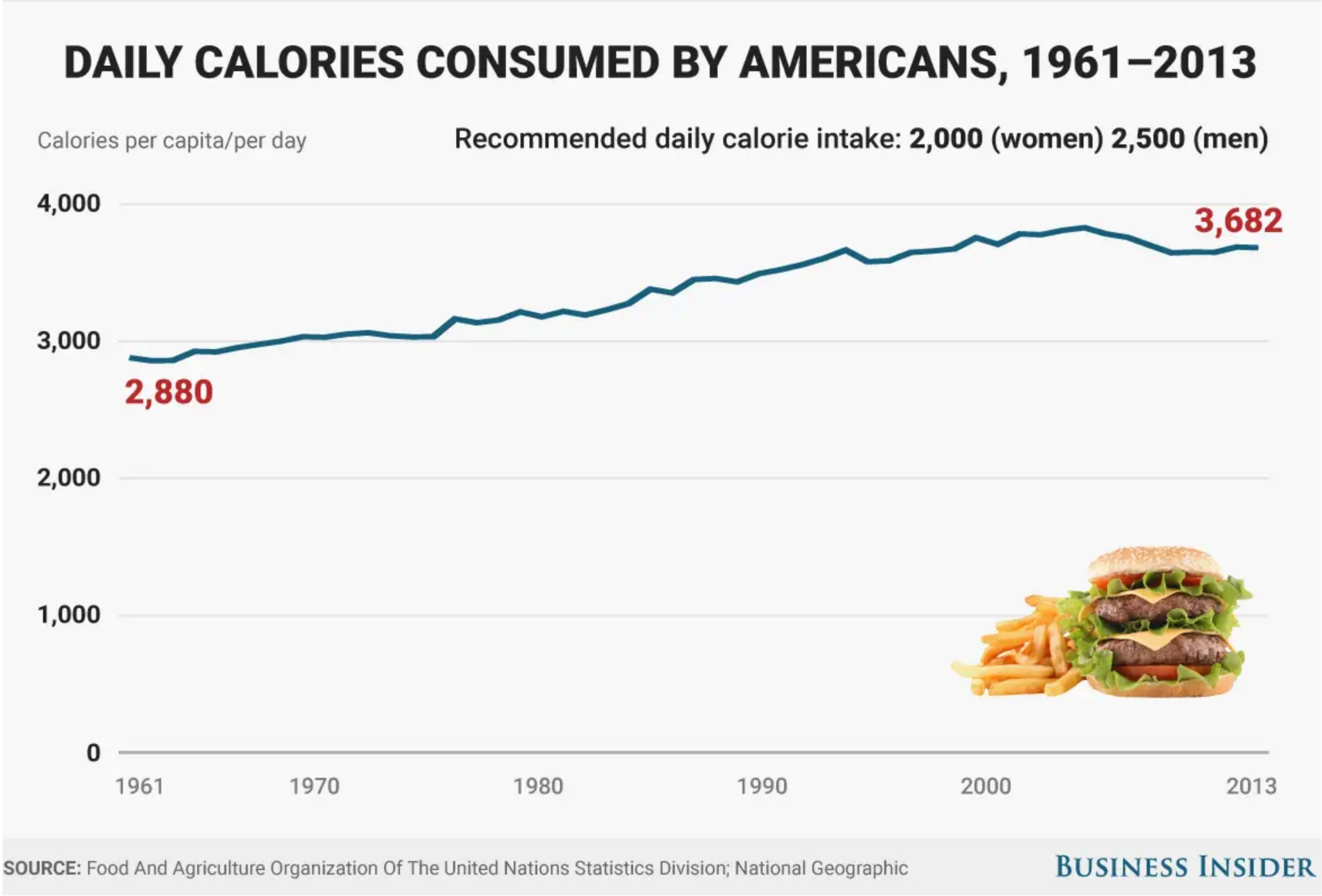
Prevalence of undernourishment in developing countries, 1970 to 2015

The share of individuals that have a daily food intake that is insufficient to provide the amount of dietary energy required to maintain a normal, active, and healthy life.

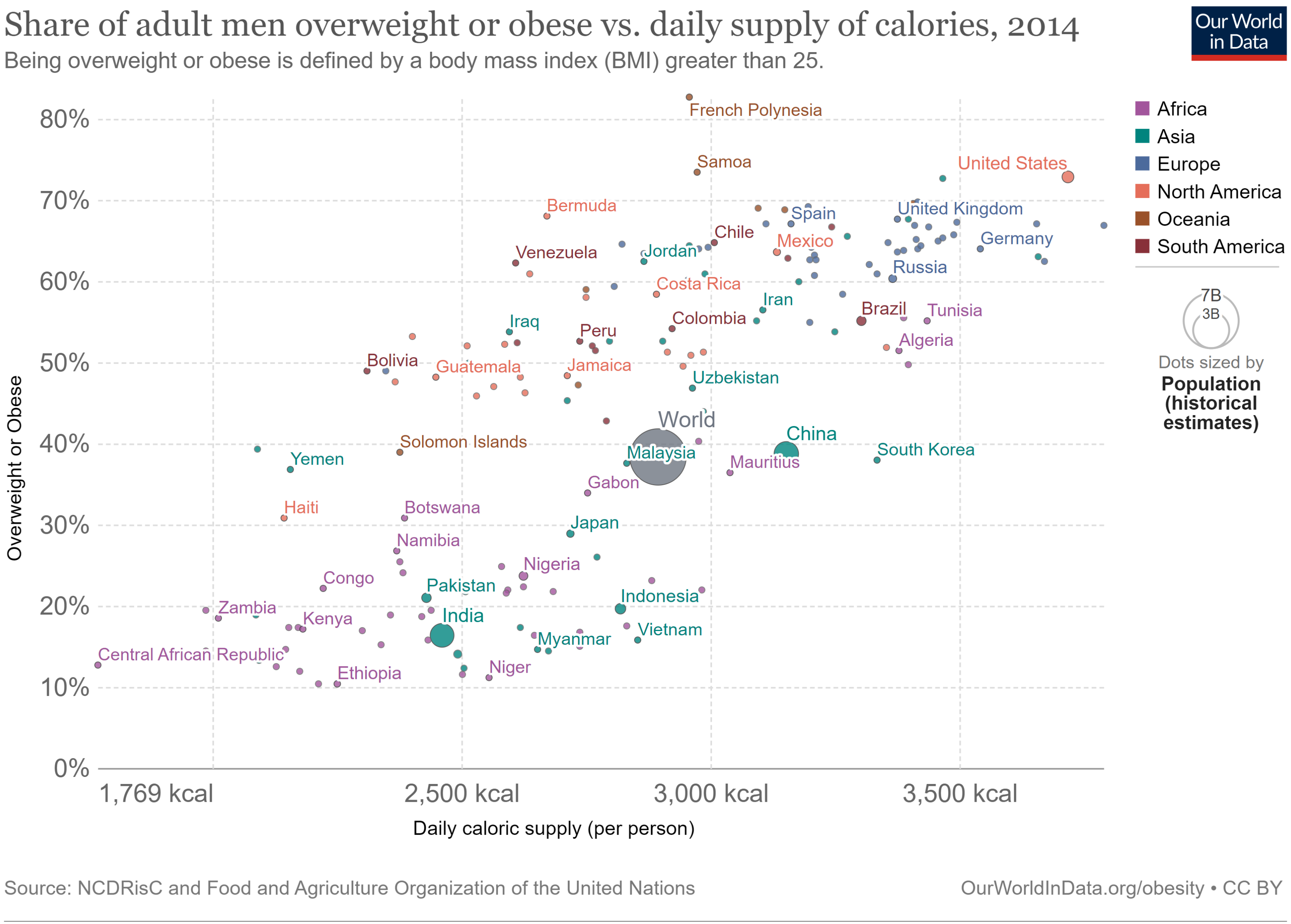


Source: Food and Agriculture Organization of the United Nations and ESS Indicators
Note: Data from 1990 onwards is well-established within FAO estimates. Earlier estimates are significantly more uncertain.
OurWorldInData.org/hunger-and-undernourishment/ • CC BY

Percent undernourished is down by 65% in developing countries



Lower prices and tasty and convenient ultra-processed food means people eat more



"Over half the population in OECD countries is overweight, with nearly 1 in 4 people considered obese."

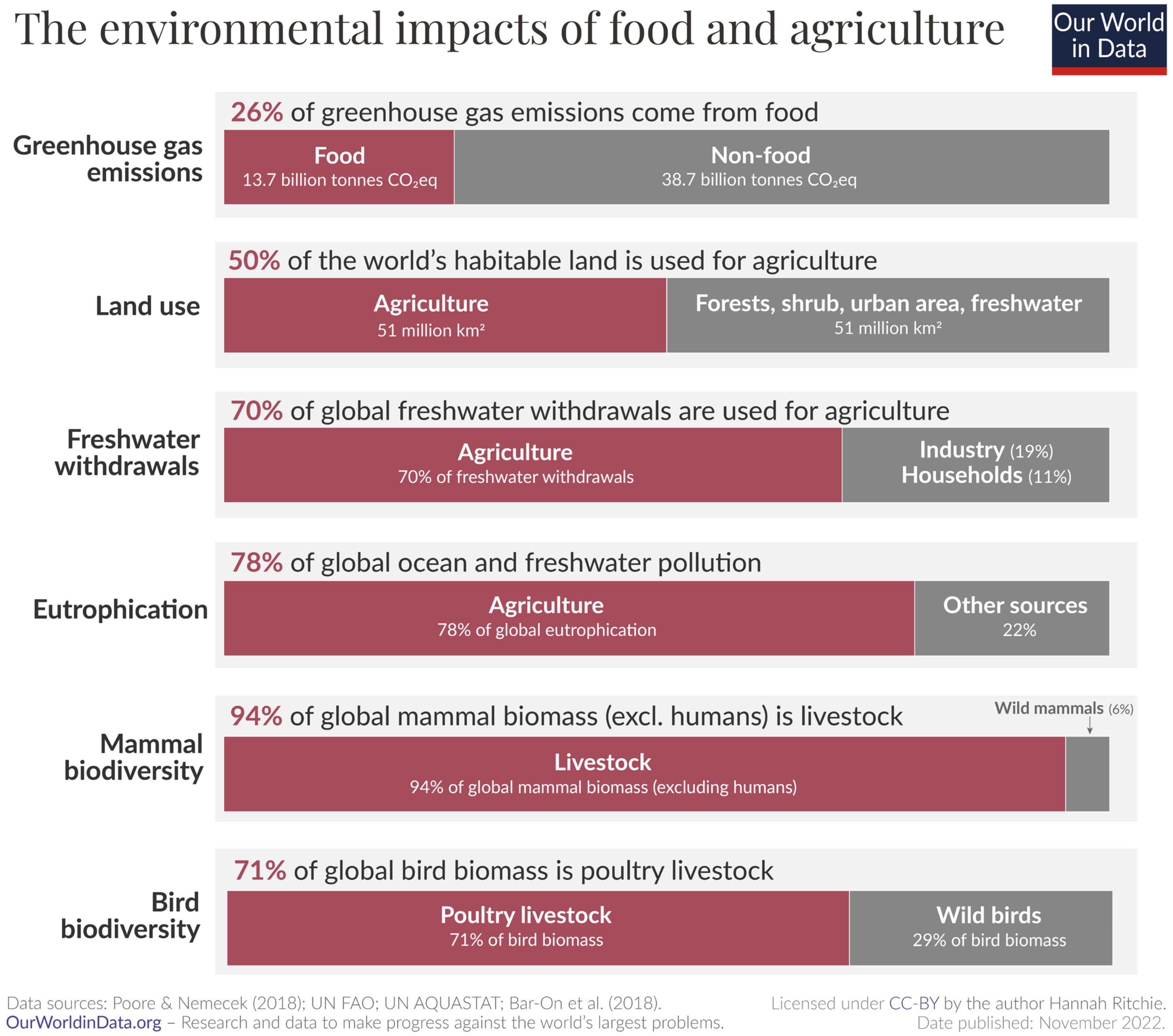
"8.4% of the health budget of OECD countries will be spent to treat the consequences of overweight over the next thirty years"

Source: *Heavy Burden of Obesity, oecd.org*

AIFS AI INSTITUTE FOR FOOD SYSTEMS

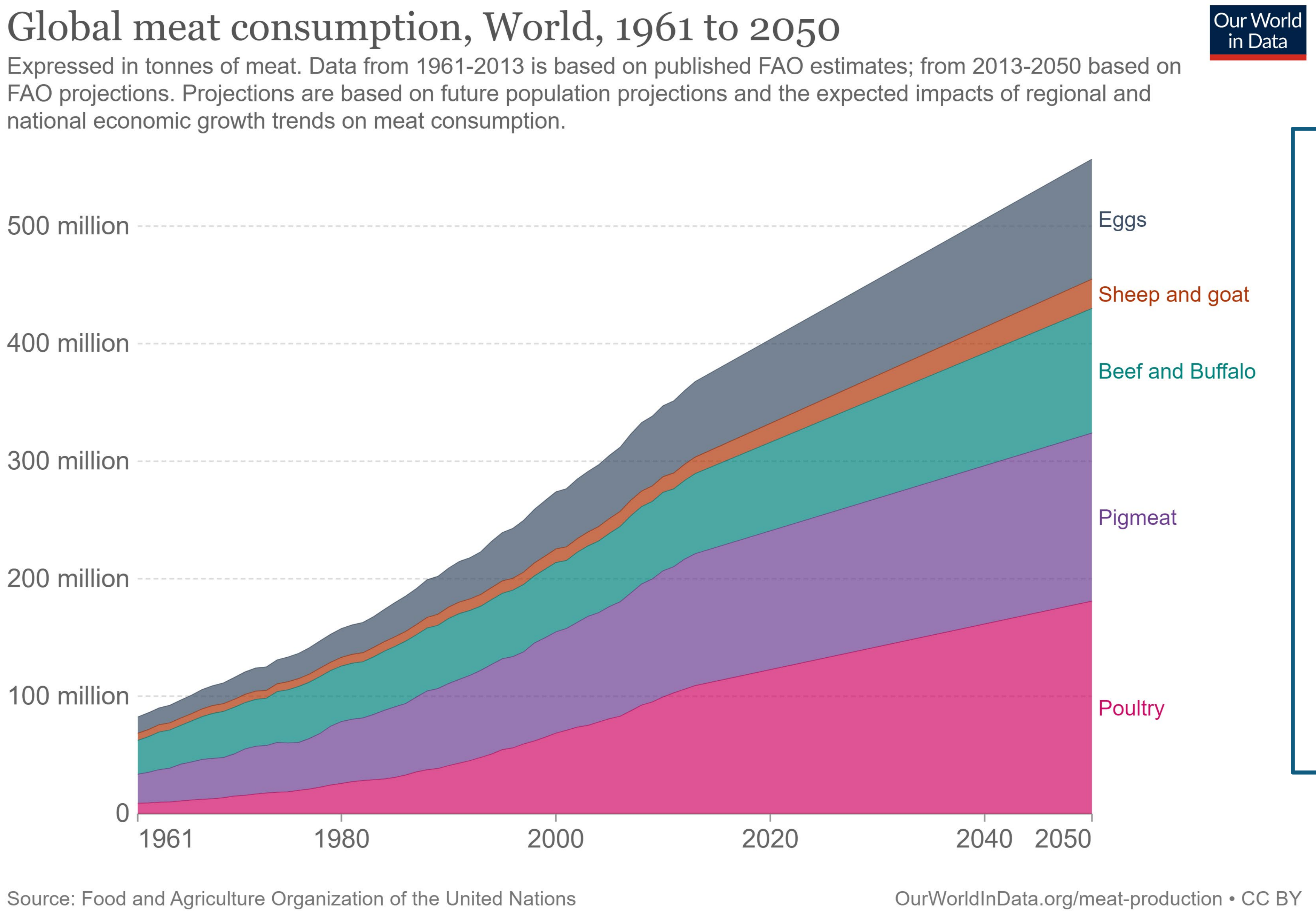
UC DAVIS
UNIVERSITY OF CALIFORNIA

Agricultural & Resource Economics UC DAVIS



Converting land to crops causes massive carbon losses

Excess fertilizer application pollutes waterways



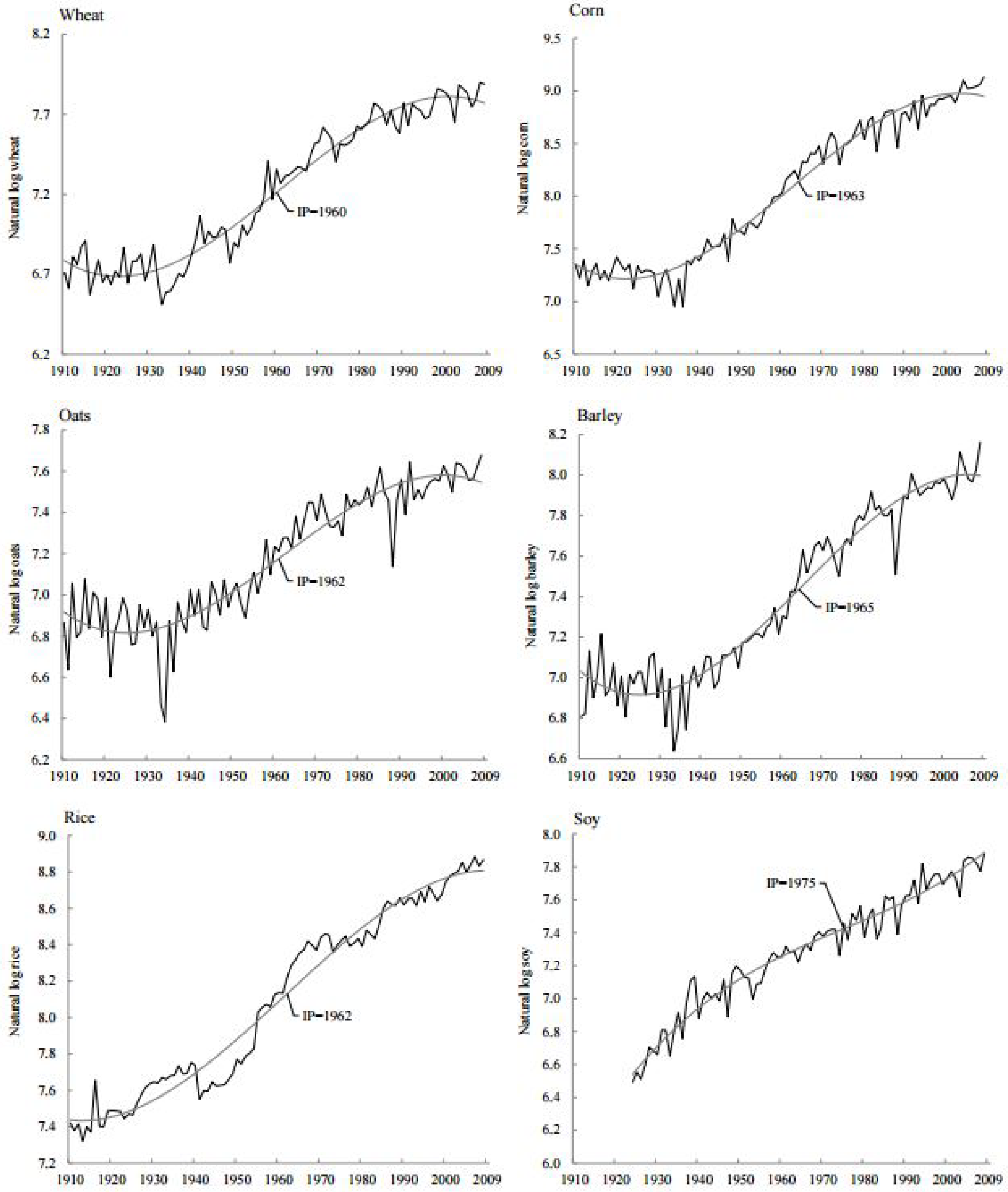
Increasing population implies more food demand

Increasing incomes mean more demand for resource-intensive food

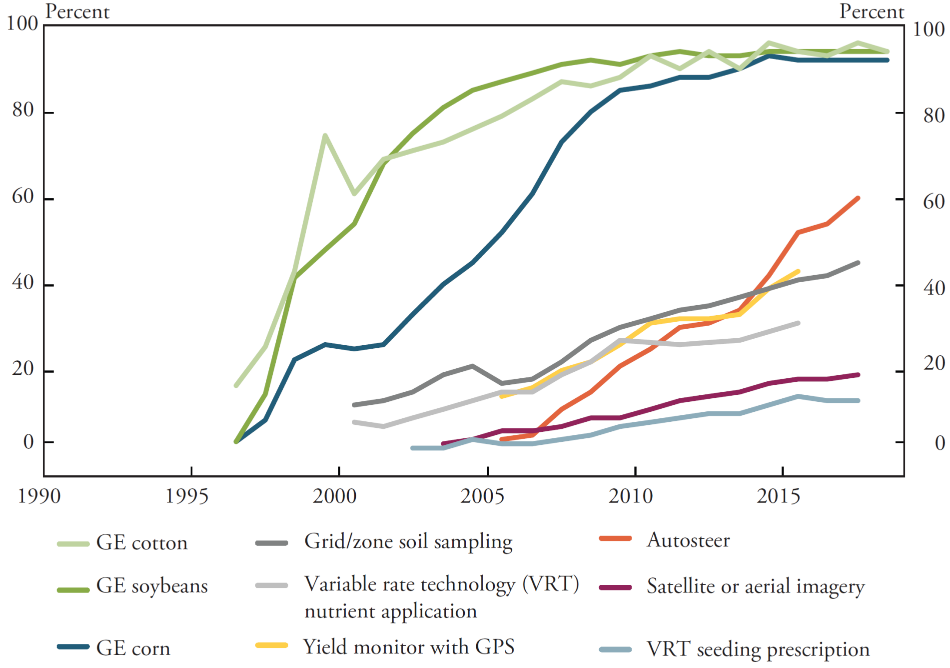
Meat demand projected to increase 50% in the next 30 years

Productivity growth has slowed

Can AI technology accelerate growth?



Panel B: Modern genetics and precision agriculture technologies



Note: Adoption rates represent shares of farms or farm area adopting.
Source: Alston and Pardey (2020).

AIFS AI INSTITUTE FOOD SYSTEMS

UC DAVIS
UNIVERSITY OF CALIFORNIA

Agricultural & Resource Economics UC DAVIS

Food is abundant and inexpensive, but

- deteriorating health
- environmental and climate challenges
- global food demand will increase

Dictionary

Definitions from Oxford Languages · Learn more

ex·ter·nal·i·ty

/ˌekˌstərˈnalədē, ˌikˌstərˈnalədē/

noun

1. **ECONOMICS**
   a side effect or consequence of an industrial or commercial activity that affects other parties without this being reflected in the cost of the goods or services involved, such as the pollination of surrounding crops by bees kept for honey.

Problem: Externalities

- Costs of eating decisions not borne by current self (“internality”)
- Costs of healthcare not borne by individual
- Costs of pollution not borne by farmer
- Benefits of innovation flow beyond innovator
- Benefits of data flow beyond provider

Solution: Public funding of ethical technologies to benefit society

- Focus on areas with substantive externalities

What do we mean by ethics?

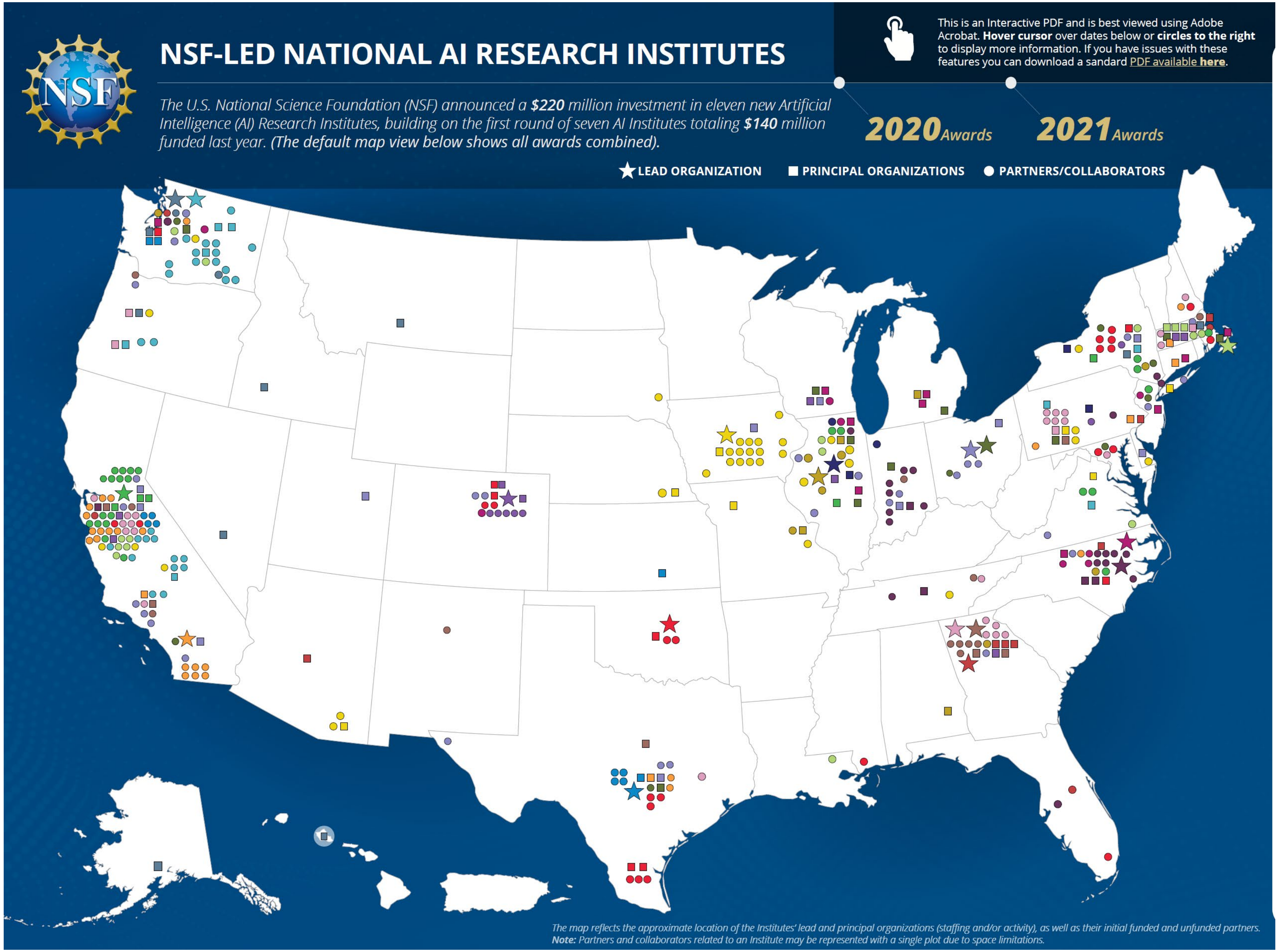
Three core questions

1. Who wins and who loses?
2. Who bears risk?
3. Who decides?

Example AI Technologies

- Autonomous weeder
- Precision seeding and fertilizer application
- Supply chain optimization
- Automatic pathogen detection in food processing
- Diet customization tools
- Engineering healthful foods

Weed Spider



NSF has funded ~25 research institutes

USDA-NIFA co-funds 5 institutes

I am part of the AI Institute for Next Generation Food Systems (AIFS)

AIFS AI INSTITUTE FOR FOOD SYSTEMS

UC DAVIS UNIVERSITY OF CALIFORNIA

Agricultural & Resource Economics UC DAVIS

Who is Responsible for Responsible AI?

Everyone is responsible

- regulations cannot govern what they do not see (e.g., micro decisions about which data to include)
- fund the AI we want!

We **interviewed** AIFS researchers

- researchers express confidence in academic research practices and outcomes; skeptical of private sector
- researchers must **navigate a complex landscape** to get data, comply with regulations, test and deploy products, and pivot quickly.
- sometimes trustworthiness makes an AI tool **less likely** to be used (“Better not to Know”).

Precision Agriculture
https://doi.org/10.1007/s11119-023-10063-3

Who is responsible for ‘responsible AI’?: Navigating challenges to build trust in AI agriculture and food system technology

Carrie S. Alexander[1] · Mark Yarborough[2] · Aaron Smith[3]

Accepted: 7 August 2023

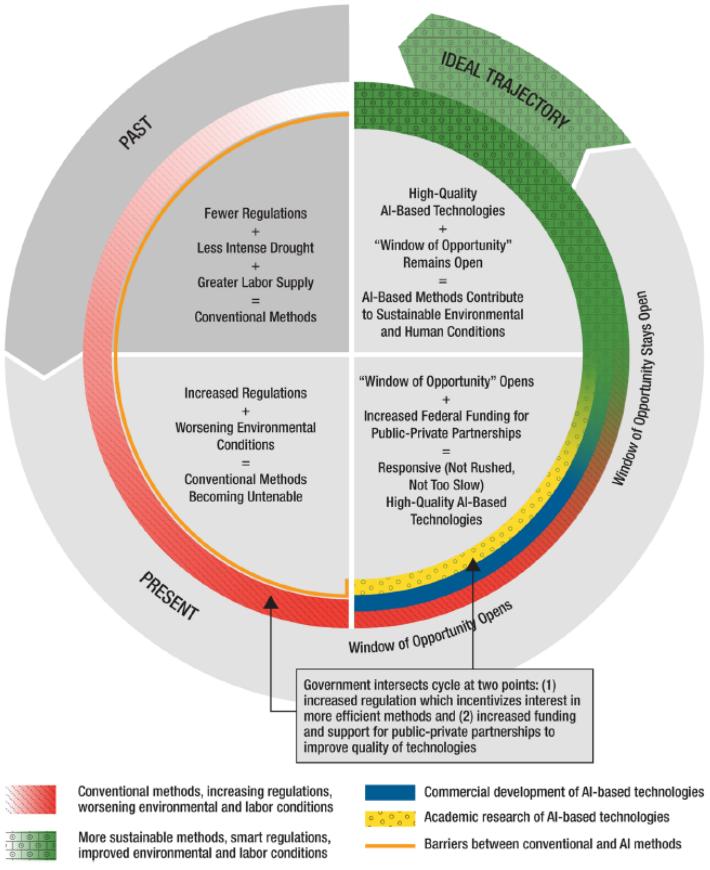


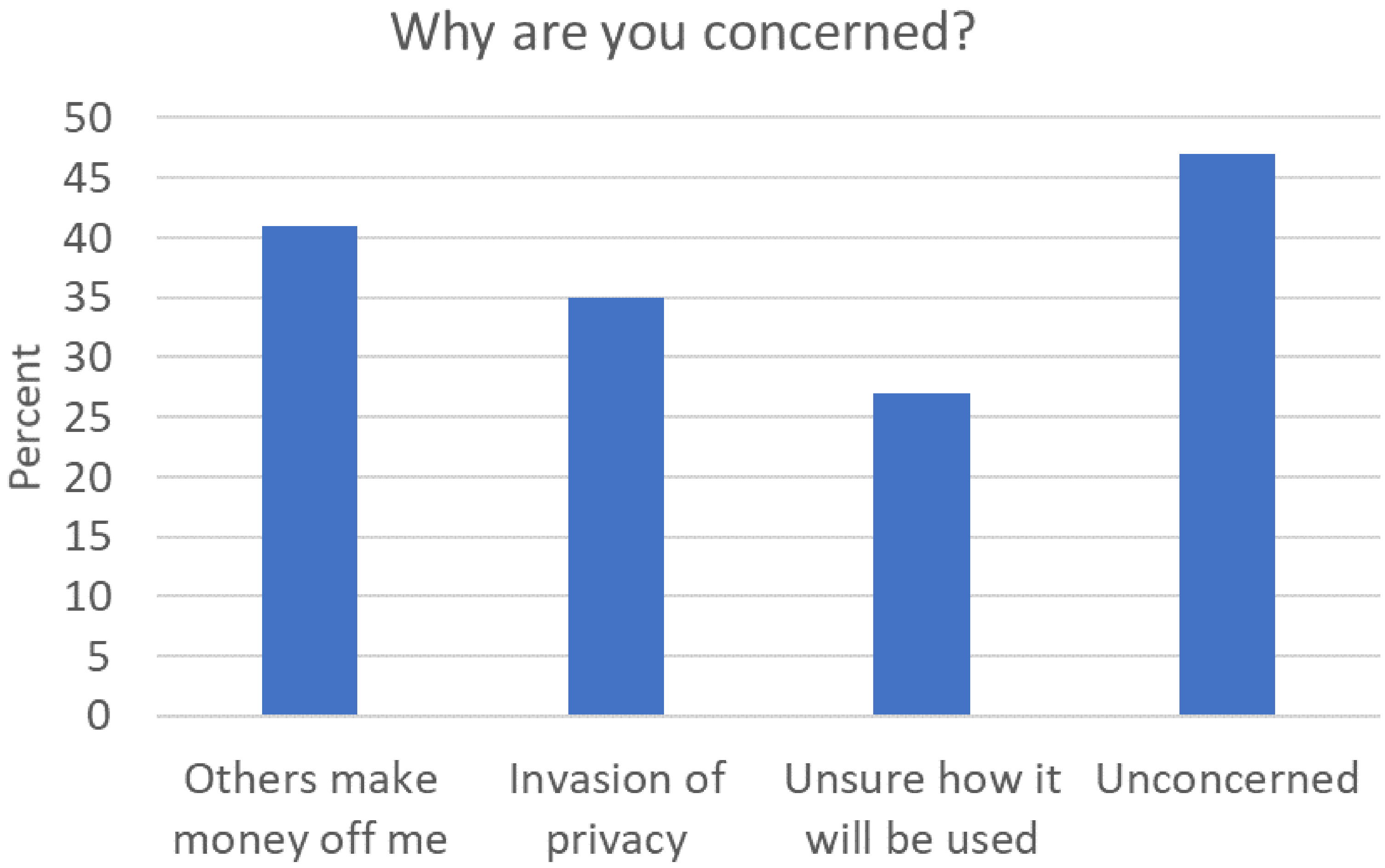
Fig. 2 Improved AI-based food system technology research, development, and adoption

AIFS AI INSTITUTE FOOD SYSTEMS

UCDAVIS
UNIVERSITY OF CALIFORNIA

Agricultural & Resource Economics UCDAVIS

Farmer Survey: Why are you concerned about sharing your data?



- **47% unconcerned**
- **Most of the concerned give multiple reasons**

Source: Author calculations from survey of Illinois corn and soybean farmers

AIFS AI INSTITUTE FOOD SYSTEMS

UC DAVIS UNIVERSITY OF CALIFORNIA

Agricultural & Resource Economics UC DAVIS

Farmer Trust and AI Technology Adoption

- modern American farmers use a lot of technology
- farmers are concerned with regulatory burden, labor scarcity, and financial pressures
- main barriers to adoption are not trust, but whether AI will solve these problems

The Challenge
Continued Productivity Gains AND Better Health and Environment

- Invest in problems beset by <u>externalities</u>
  - these are problems private firms won’t solve
- Negative externalities: health, pollution
  - for farmers, save cost <u>and</u> reduce pollution
  - foods that are more appealing <u>and</u> more healthful
- Positive externalities: systems, infrastructure, data
  - build data resources and infrastructure
- Ethical RD&D requires continuous vigilance

1. Who wins and who loses?
2. Who bears risk?
3. Who decides?

Aaron Smith
DeLoach Professor of Agricultural Economics
UC Davis

https://asmith.ucdavis.edu

https://agdatanews.substack.com

adsmith@ucdavis.edu

AIFS AI INSTITUTE FOOD SYSTEMS

The National Academies of SCIENCES ENGINEERING MEDICINE

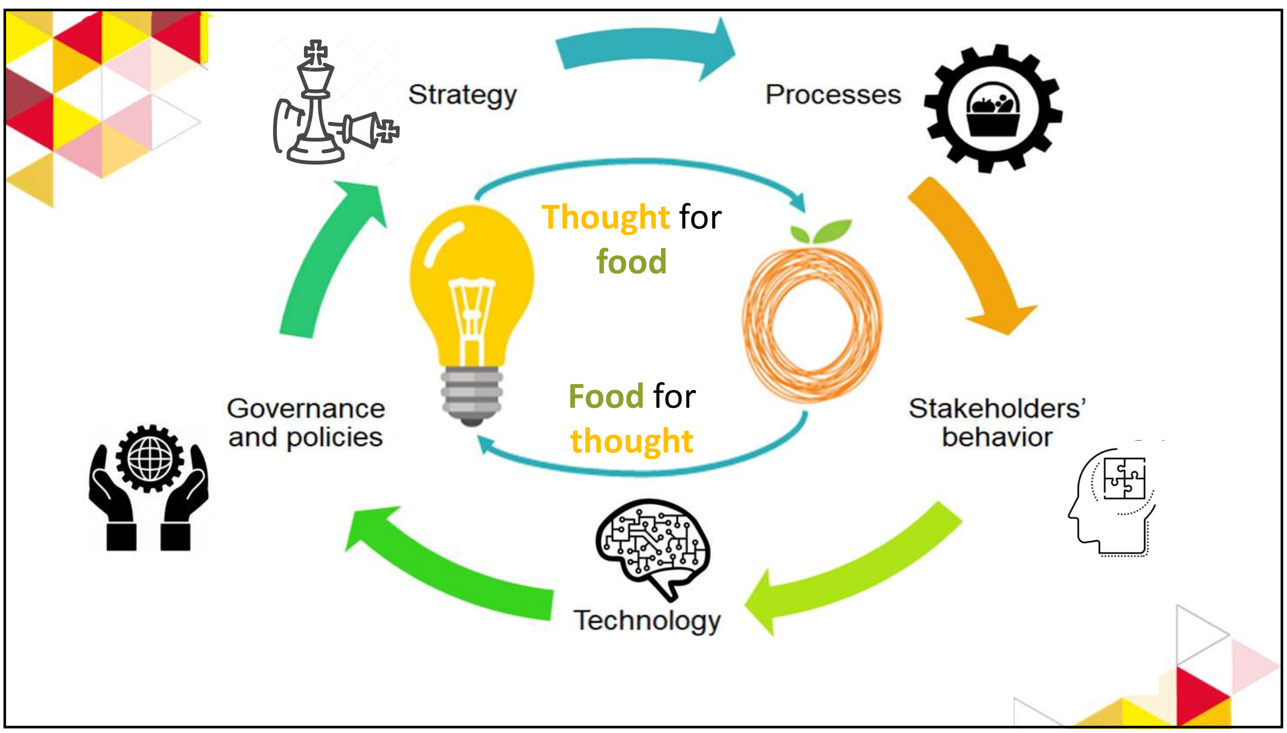
Nourish or Perish: A Research journey through Food supply chains

Christopher Mejía Argueta ● cmejia@mit.edu

Research Scientist, MIT Center for Transportation and Logistics
Founder and Director, MIT Food and Retail Operations Lab

MIT | FaROL Food and Retail Operations Lab

MIT Center for Transportation & Logistics



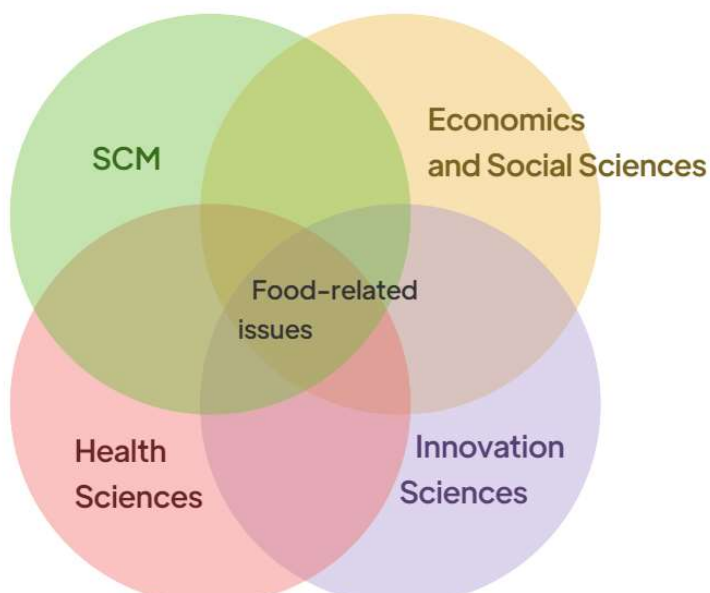
Level setting SCM and its connection to nutrition and AI/ML

What is Supply Chain Management (SCM)?

"...the **systemic, strategic coordination** of the traditional business functions and the tactics **across** these **business functions** within a particular company and across businesses within the supply chain, for the purposes of **improving the long-term performance** of the individual companies and the supply chain as a whole." - Mentzer et al. (2001)

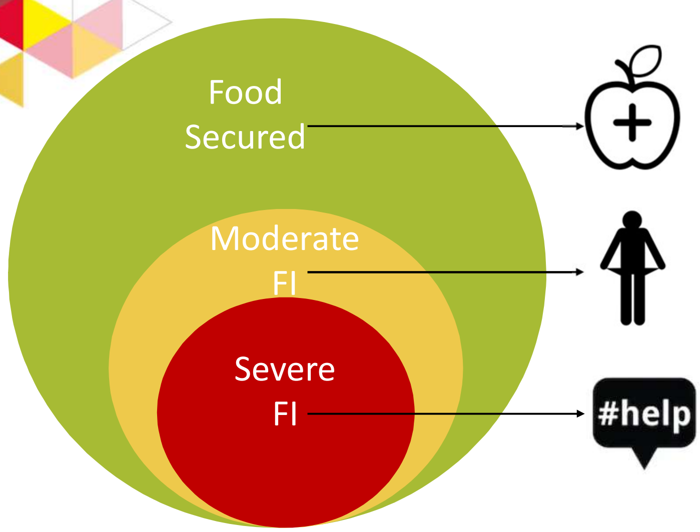
MIT Center for Transportation & Logistics

SCM for food and agribusiness is complex and multi-factorial



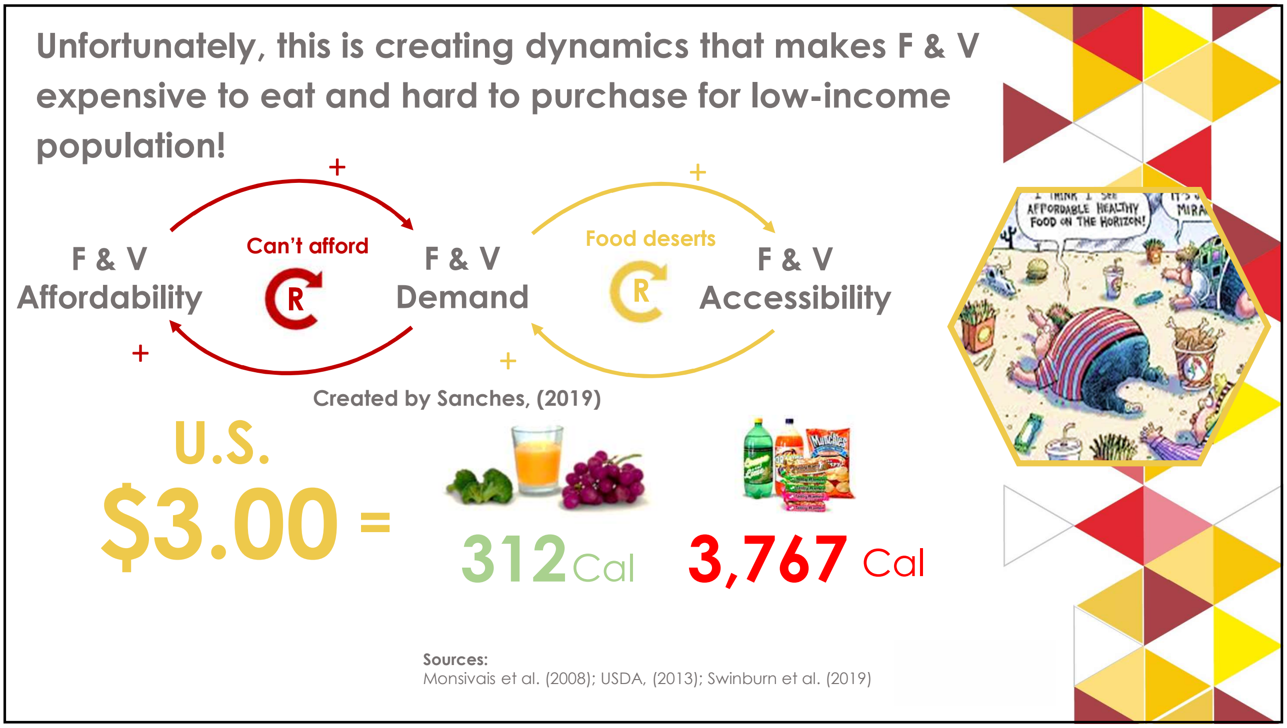
The issues

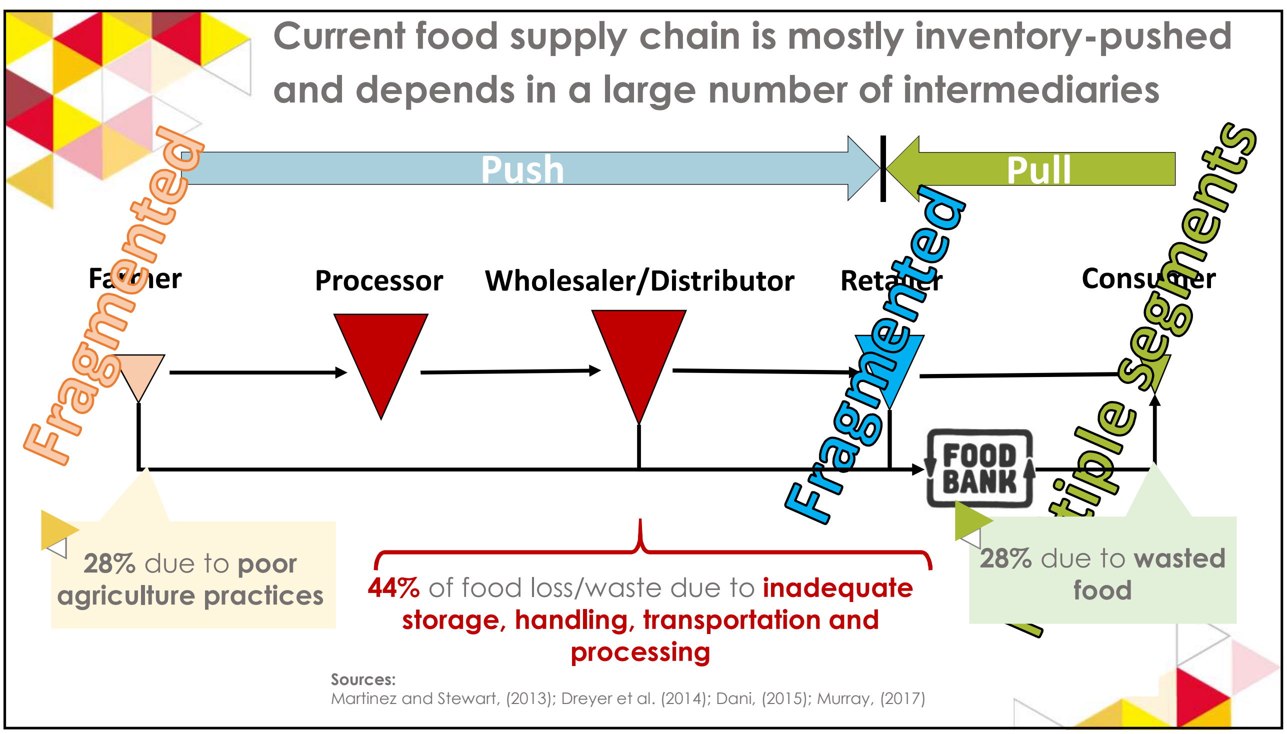
Food insecurity



5b people (446m LAC)	US$3.75/day (**US$2.98/day LAC**)
1.25b people (143m LAC)	US$2.33/day (**US$2.83/day LAC**)
746m people (62m LAC)	US$0.79/day (**US$1.06/day LAC**)

Source: FAO, (2020)


Unfortunately, this is creating dynamics that makes F & V expensive to eat and hard to purchase for low-income population!
F & V Affordability
Can't afford
F & V Demand
Food deserts
F & V Accessibility
+
+
+
+
R
R
Created by Sanches, (2019)
AFFORDABLE HEALTHY FOOD ON THE HORIZON!
U.S.
$3.00 =
312 Cal
3,767 Cal
Sources:
Monsivais et al. (2008); USDA, (2013); Swinburn et al. (2019)


Current food supply chain is mostly inventory-pushed and depends in a large number of intermediaries
Push
Pull
Fragmented
Farmer
Processor
Wholesaler/Distributor
Retailer
Consumer
Fragmented
Multiple segments
FOOD BANK
28% due to poor agriculture practices
44% of food loss/waste due to inadequate storage, handling, transportation and processing
28% due to wasted food
Sources:
Martinez and Stewart, (2013); Dreyer et al. (2014); Dani, (2015); Murray, (2017)

Agribusiness

Exploratory model

Market Size and Direct Accessibility as Mediators for Explaining Potato Prices

UTEC UNIVERSIDAD DE INGENIERÍA Y TECNOLOGÍA

Noriega, Larco, Antonini & Mejia-Argueta (2021)

Context and issue

Volume production

8% of total production of crops in Peru belongs to potato (INEI, 2017).

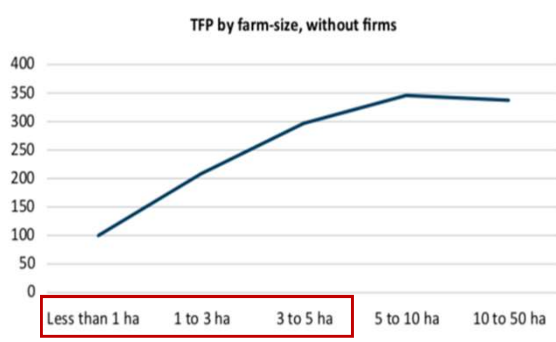
77% of sales are white potatoes in WM (CITE Papa, 2020).

Farm size distribution

~80% Small-scale producers (< 5ha).

Medium-size farms (5-50ha).

Large farms (> 50ha).

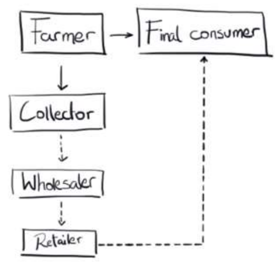


Our study

Research questions

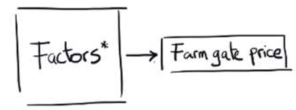
1. What factors explain better connectivity to the market channel directedness and access to larger markets?
2. How do market channel directedness or market size access have an effect on farm's gate price?
3. What is the direct or indirect influence of different factors* on farm's gate price?

Market channel directness

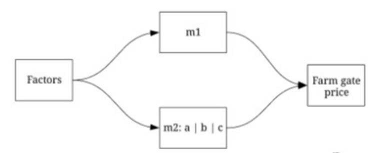


Market size access

Factors direct effect to Farm gate price



Factors indirectly affecting and mediating the farm gate price



Associativity vs intermediation

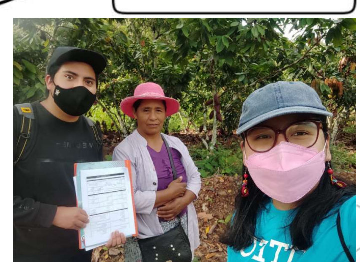
- Economies of scale related
- Transportation time to market
- Information access
- Productivity
- Associativity
- Demand outlets

Market size access

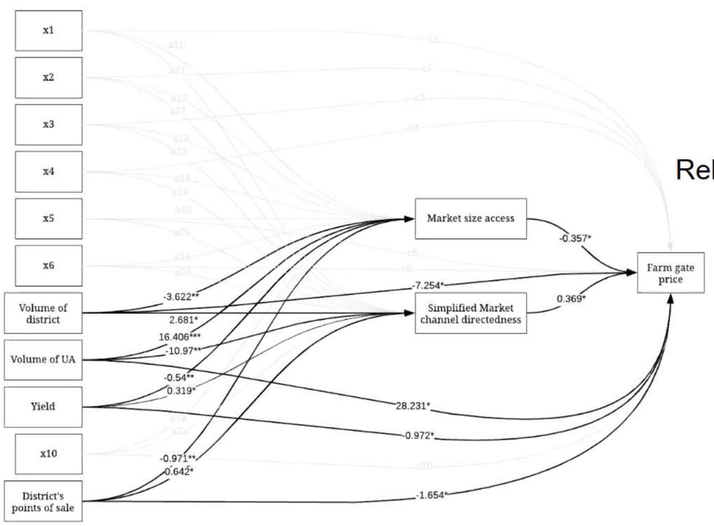
Market channel directedness

Farm gate price

UTEC
UNIVERSIDAD DE INGENIERÍA Y TECNOLOGÍA



Farm gate prices: Coupler + Wholesaler vs Retailer and end consumer



CFI: 0.99 > 0.95
RMSEA: 0.07 < 0.08
SRMR: 0.05 < 0.08

Relationships are distinct,

- Market size access is now negatively related to Farm gate price.
- Volume of district is now negatively and directly related to Farm gate price.
- District's points of sale is now negatively and directly related to Farm gate price.

Takeaways

- **Mixed effects on factors that are positively related to market directedness but negative effects of farm prices (volume & POs)**
  - This suggests that multiple price-equilibria exists and a geographical analysis is needed.
- **Importance of local demand should not overlooked**
  - Larger markets do not necessarily imply higher prices, but more local markets may do so.
- **Information is not related to market accessibility or price**
  - Having data but possibly no real choices, makes information non-actionable.

Nutrition for underserved communities

Prescriptive model

Delivering local nutritious food baskets

Supply Chain MANAGEMENT

Das, Mejia-Argueta & Sadalla (2021)

Nutrition issue in India

	rice-PDS	wheat-PDS	jowar	bajra	maize	ragi	moong	masur	urd	peas	cereals total	pulses total
count	573.00	586.00	184.00	207.00	364.00	149.00	619.00	603.00	589.00	554.00	625.00	625.00
mean	17.45	12.21	7.53	8.11	5.40	3.84	0.93	1.08	0.95	0.95	52.11	3.73
std	9.88	8.77	10.19	9.39	7.55	4.41	0.43	0.57	0.52	0.76	12.47	1.14
min	1.00	0.82	0.10	0.25	0.10	0.20	0.20	0.10	0.24	0.10	20.95	1.02
25%	10.75	4.23	2.00	2.10	1.69	1.14	0.65	0.66	0.63	0.50	42.60	2.92
50%	15.55	10.99	5.00	4.80	3.00	1.80	0.85	0.93	0.85	0.71	51.62	3.67
75%	24.19	17.48	9.81	10.00	5.95	4.66	1.11	1.41	1.14	1.03	60.97	4.54
max	53.76	68.40	80.00	52.94	80.00	20.00	4.09	4.25	6.00	4.87	93.73	7.78

Summary Statistics - Quantities in kg per household per month

ICMR Recommended Targets (Monthly Quantities)

- Cereals – 8.1 kg

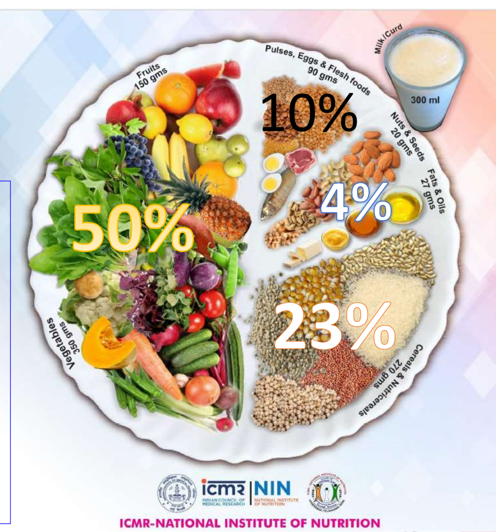
------For family of 4 -> 32.4 kg

- Pulses – 2.7 kg

-------For family of 4 -> 10.8 kg

Excess

Deficiency



Scope:
Cereals and Plant sourced Proteins

Our study

Research Question

- What are some feasible approaches for the **design of healthy, affordable, locally sourced food combinations** for Indian households?
- What supply chain mechanisms might be leveraged to **distribute the designed food baskets at scale** to vulnerable subpopulations in India?

To solve:

Availability of Nutritious Food

Affordability of food for population
- By all sections of the society

Access
- Either produced locally or through distribution, made accessible across geographies

Within Scope:

Food items – Cereals and Pulses

Distribution Channel – Government owned and operated

Targeted Beneficiaries – The poorest segment of households covered by the government run Public Distribution System in India

Methodology

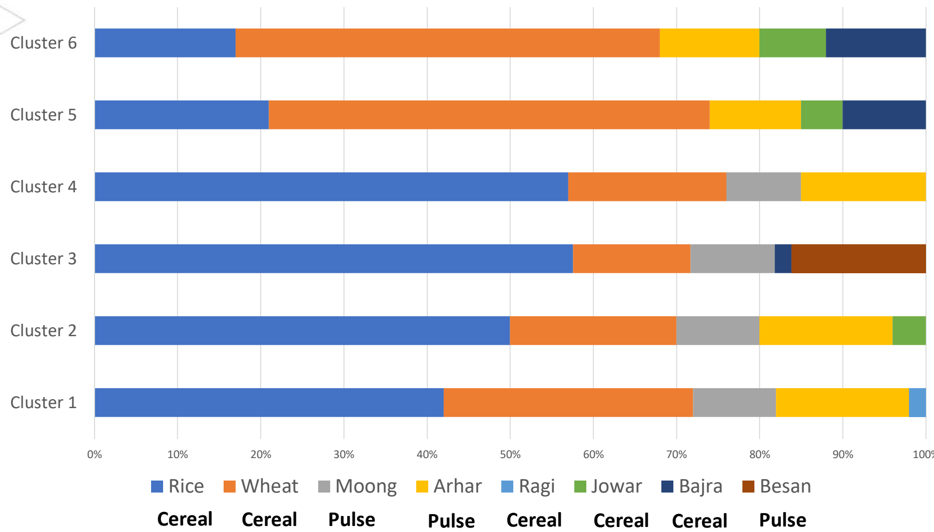
- **Segment Customers**: Characterize Customers and Clusters
  - Geography, socioeconomic features, historical consumption
  - PCA and K-means clustering
- **Design Food Basket for each cluster:**
  - Assortment of food items that meets prescribed nutritional targets
  - Bin Packing Algorithm Inspired
- **Design Distribution Network to deliver baskets to clusters (Location Allocation):**
  - Design a distribution strategy for the baskets to clusters at scale
  - Multi Commodity MILP
    - Count of storage centers is discrete
    - All other decision variables continuous

Distribution of baskets to cluster districts

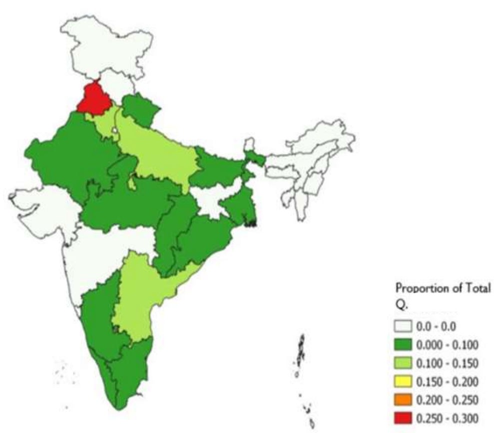
Design Food Baskets for Customer Segments

Segment Customers

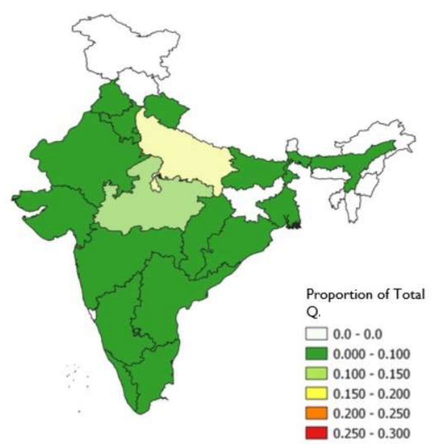
Food basket design for each cluster



Distribution



Spatial Spread of Procurement and Distribution - Now



Spatial Spread of Procurement and Distribution - Proposed

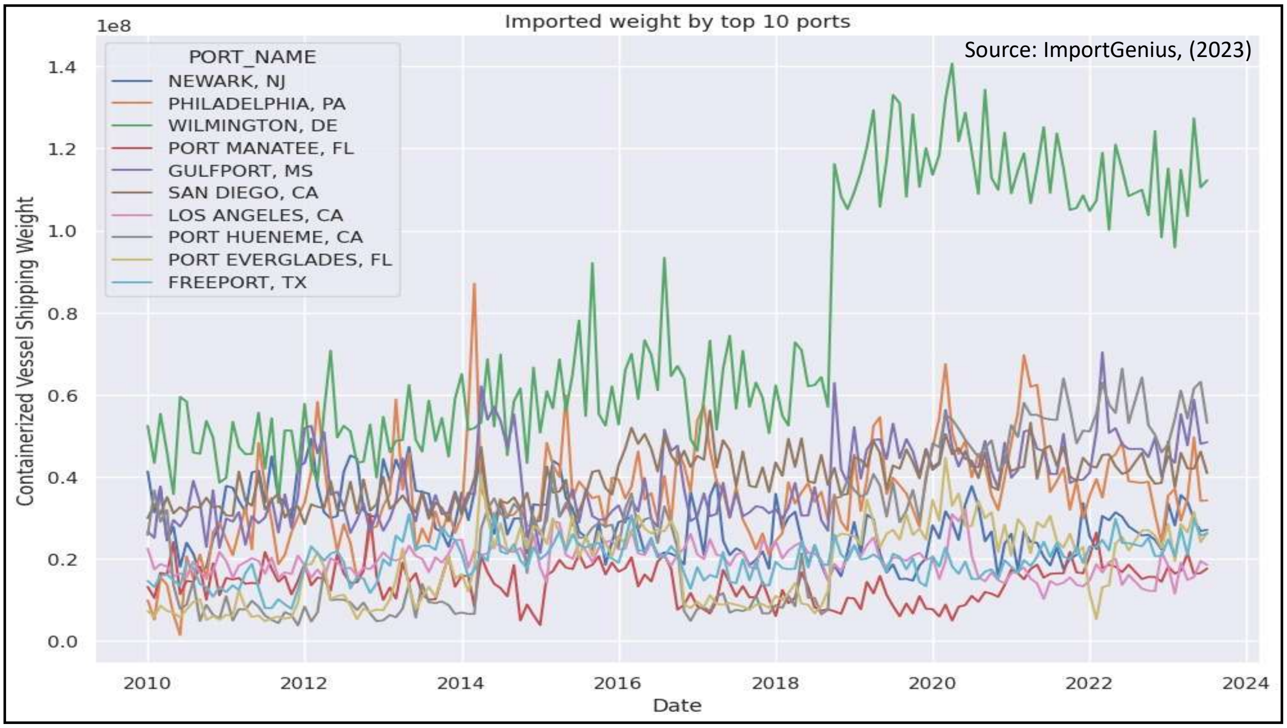
Conclusion

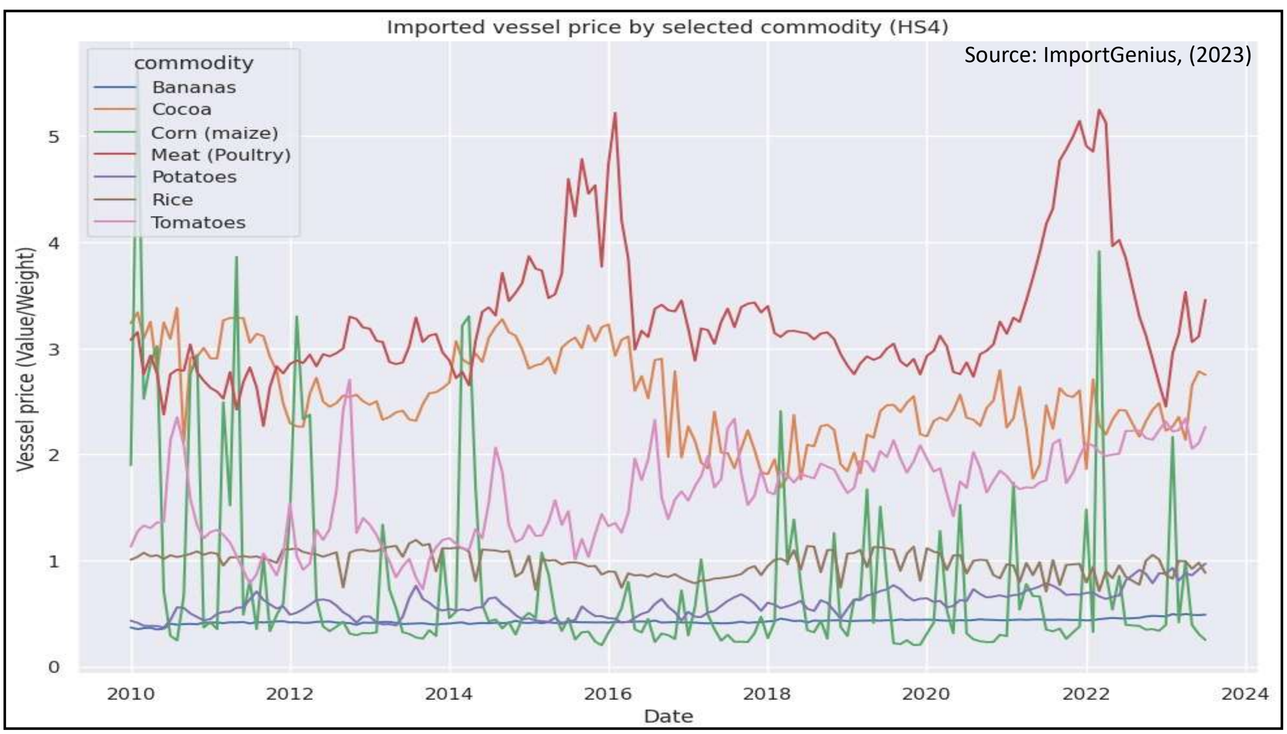
- What factors should govern **the design of healthy, affordable, locally sourced food combinations** for Indian households?
  - Identification of **locally grown pulses and cereals in different parts of India**
  - Clustering of Customers based on **Local Taste Preferences**
  - A **cost-efficient strategy for connecting** supply nodes with demand clusters
- What supply chain mechanisms might be used at a population level **to distribute the designed food baskets efficiently** in India?
  - We built on the **Government Public Distribution System adding pulses in a cost-effective way** for three reasons:
    - It has **maximum reach in terms of number of beneficiaries**, so caters to a large population base
    - It **subsidizes both production and consumption sides**
    - **Public data on** count and location of storage centers

International food flows

Criticality and resilience of global food supply chains

Dugundji, Mejia-Argueta, Koch & Gamez (2023)


Imported weight by top 10 ports
Source: ImportGenius, (2023)
1e8
PORT_NAME
NEWARK, NJ
PHILADELPHIA, PA
WILMINGTON, DE
PORT MANATEE, FL
GULFPORT, MS
SAN DIEGO, CA
LOS ANGELES, CA
PORT HUENEME, CA
PORT EVERGLADES, FL
FREEPORT, TX
Containerized Vessel Shipping Weight
0.0
0.2
0.4
0.6
0.8
1.0
1.2
1.4
2010
2012
2014
2016
2018
2020
2022
2024
Date


Imported vessel price by selected commodity (HS4)
Source: ImportGenius, (2023)
commodity
Bananas
Cocoa
Corn (maize)
Meat (Poultry)
Potatoes
Rice
Tomatoes
Vessel price (Value/Weight)
0
1
2
3
4
5
2010
2012
2014
2016
2018
2020
2022
2024
Date

Conclusions and future work

Improving food SCs will **support the productivity and growth**

Huge opportunity to link **SCs** to **nutrition intervention schemes and other disciplines**

SCs for **bottom of the pyramid** firms and consumers are still considered **unexplored** (business opportunity!!)

Long-term sustainable food ecosystems have to consider a **inter-disciplinary approach** to provide holistic solutions.


Reaching
50 Million
Nanostores
Retail Distribution in Emerging Megacities
SUPPLY CHAIN MANAGEMENT AND
LOGISTICS IN LATIN AMERICA
A Multi-Country Perspective
HUGO T.Y. YOSHIZAKI, JOSUÉ C. VELÁZQUEZ MARTÍNEZ,
AND CHRISTOPHER MEJIA ARGUETA
Supply Chain
Management
and Logistics
in Emerging
Markets
Selected Papers
from the 2018
MIT SCALE Latin
America Conference
Engineering Analytics:
Advances in Research and
Applications
edited by
Luis Rabelo
Edgar Gutierrez-Franco
Alfonso Sarmiento
Christopher Mejia-Argueta
Thank you!
Dr. Chris Mejia Argueta
cmejia@mit.edu
Food and Retail Operations Lab