



OPEN

DATA DESCRIPTOR

MilkyBase, a database of human milk composition as a function of maternal-, infant- and measurement conditions

Tünde Pacza¹, Mayara L. Martins¹, Maha Rockaya¹, Katalin Müller^{2,3}, Ayan Chatterjee^{4,5}, Albert-László Barabási^{4,6,7} & József Baranyi¹✉

This study describes the development of a database, called MilkyBase, of the biochemical composition of human milk. The data were selected, digitized and curated partly by machine-learning, partly manually from publications. The database can be used to find patterns in the milk composition as a function of maternal-, infant- and measurement conditions and as a platform for users to put their own data in the format shown here. The database is an Excel workbook of linked sheets, making it easy to input data by non-computationally minded nutritionists. The hierarchical organisation of the fields makes sure that statistical inference methods can be programmed to analyse the data. Uncertainty quantification and recording dynamic (time-dependent) compositions offer predictive potentials.

Background & Summary

The effect of diet on health has primarily been analysed in a descriptive way. Widely acknowledged claims, such as garlic helps preventing cardiovascular diseases, are lacking mechanistic, biochemical explanations¹. The main sources of such uncertainties are: (i) the complexity caused by thousands of chemical interactions; (ii) the inherent errors in the measurements and observations; (iii) many hitherto unknown other details¹.

Human milk (HM) is the first nutrition an infant comes across and one of our most complex foods. Ideally, mothers should breastfeed their infants, but we need to acknowledge that in many cases this is not possible, and even when mothers try their best, breastfeeding is challenging and requires a strong supportive environment.

HM has been studied extensively, still its biochemical complexity is insufficiently explored^{2,3}. It is the only food that meets all the nutritional requirements of infants and provides optimal adaptation, somatic growth, maturation, and development⁴. Beside the nutrients (carbohydrates, lipids, proteins, vitamins, and minerals), it provides bioactive components (hormones, cytokines, growth factors, antimicrobial substances, cells, etc.), which play important roles in the development of the central nervous system, metabolism, immune system, and microbiome⁵⁻⁹. Breastfeeding has been associated with improved health outcomes, including increased intelligence, reduced risks of infections and non-communicable diseases (obesity, atopic diseases, diabetes, inflammatory bowel diseases)^{6,7}. This crucial role of HM in early life nutrition gains great clinical^{5-7,10}, social and economic interest due to its impact on long-term health^{10,11}.

HM is a biological system, where both nutritional and bioactive components are in constant interactions with one another². The exact dynamics depend on characteristics related to the mother, the infant, and various environmental factors (such as the mother's diet, the gestational age, the geographic location etc.), which are also responsible for the variability of the HM composition³. Our current knowledge is largely based on studies evaluating these components, typically analysing their variability and dynamics separately^{2,3,7}. Therefore, explaining

¹Doctoral School of Food and Nutrition Science, Institute of Nutrition, University of Debrecen, Debrecen, Hungary.

²Heim Pál National Paediatric Institute, Budapest, Hungary. ³Doctoral School of Clinical Medicine, University of Debrecen, Debrecen, Hungary. ⁴Center for Complex Network Research, Northeastern University, Boston, USA.

⁵Network Science Institute, Northeastern University, Boston, USA. ⁶Department of Medicine, Brigham and Women's Hospital, Harvard Medical School, Boston, USA. ⁷Center for Network Science, Central European University, Budapest, Hungary. ✉e-mail: baranyi.jozsef@med.unideb.hu

health outcomes directly by specific components is rarely satisfactory, due to the modifying effects of the interactions between the factors in question^{2,3,7}.

As in any complex systems, the dynamics of HM cannot be predicted from the kinetics of its individual components^{2,3}. A big-data platform is needed to help. More accurately than ever, an appropriately built database could provide objective, data- and science-based guidance on the diet and lifestyle of lactating women to optimize their children's health. Besides, the development of HM substitutes could benefit enormously from the collective knowledge the database can store.

In this paper, we demonstrate that an adequately built database, combined with numerical/statistical tools, has huge potentials to unveil food complexity^{1,12} and to benefit from the stored knowledge. A key to this is the basis of our database-building principle: it considers a record as a mapping from various, possibly dynamic explanatory conditions, under which observations have been made, to the composition of HM, a truly dynamic response variable. A vital means to realize this ontology-principle is that the temporal variation of the variables is represented by tables, and pointers to these tables make sure that time-dependence is a natural attribute of the respective fields.

Methods

Food composition data have already been collected in databases, following various ontology depending on the purpose and the wanted resolution of the database. Our MilkyBase is intended to be used by academia as well as industry and regulation, therefore many compromises had to be made to find a balance between the four-V-principle of Big Data: volume, velocity, veracity, and variety.

Volume. We have set up a database that hosts published measurements of molecular components of breast milk. With its ca 10,000 datapoints, MilkyBase is far from the volume that is expectable from a Big Data project. However, we hope to initiate an ontology that would be used by researchers as well as clinicians to input their own data, so to create a “periodic table” of other important food-types, as a pool for collective knowledge¹³. Therefore, the template for inputting the data must be user-friendly enough, on commonly used platform, easily handled by the data donors. This is the reason why Microsoft Excel was chosen, as the most ubiquitous package that can link tables and be programmed via the Visual Basic for Applications (VBA) language. The VBA programs will aid both input check and data analysis (such as comparing own and others' similar observations) and serve as incentives to authors to submit relevant data. This is a kind of wiki-philosophy, which should result in a much bigger data volume than its current size.

Velocity. With its current size, the navigation and data processing are running at an acceptable speed, but the Excel platform will not be practical as the volume of the data increases, therefore, with time, it will be imported into an SQL server and the Excel sheets will serve a transit area for data donors, for initial curation.

Variety and veracity. As these are closely related, we discuss them together. Our goal was to digitize published data in a rigorously organized database, ready to be analysed by considering the milk composition as a function of various conditions. Therefore we tried to avoid changing published data, except in trivial cases, such as conversion of units for the sake of uniformity. Many times we found ambiguity or controversy in the terminology used by authors. An example for this is the concentration of a particular fatty acid molecule, which was mostly reported as a proportion relative to the total fatty acid, but sometimes proportion in the total milk mass, and sometimes even just the proportion of the total *measured* fatty acids. In such cases, we used our best knowledge and expert help to make these concepts well-defined and quantified. Such efforts admittedly bear the footprint of the database developer's judgement.

If there are trivial mistakes in the publication (such as conversion error from one unit to another one) that were easily correctable then we did so; otherwise, either we left the record out, or marked it as “suspicious”. Even so, the resultant database is inevitably imperfect. However, the discrepancies should get detected as the database is being used.

Note that the variety - veracity issue is closely related to the syntax and semantics of the fields of the database. While its syntax can be checked in an automated way, its semantics frequently reveals anomalies, affecting what data can be inputted (variety) and how can those be validated (veracity).

For compatibility, we fixed the “mass/volume milk” concentration of each biochemical component as the target response value. By “Component” we mean either a molecule or a group of molecules, such as say “linoleic acid”; or “fatty acid”. Both are “Components”, while the first is a special case of the second. Grouping like this follows a hierarchical tree structure as published data suggest (see Fig. 1). This way, not only the density of a particular molecule, but any components from the level next to the HM root, can be inputted.

Many authors only publish rescaled or derived values as components. Examples for this are the $2FL$ and $2FL/OS$ components (concentration of 2-fucosylated lactose and its proportion to that of the total oligosaccharides). To deal with such scenarios, we call a numerical value for $2FL$ as *direct*, while that for the $2FL/OS$ ratio as *indirect response*. We considered the explanatory and response variables as vectors, where each entry in the first one is a (mostly quantified) value on a specific condition that resulted in the response variables, in either direct or indirect form. Then a measurement for an indirect variable, such as $2FL/OS$, is analogous to an implicit relationship between two mathematical variables. Similarly, a variable with the name “ $C18 : 1n-9 + C18 : 3n-3$ ” indicates that the two fatty acids were measured together. So, the name of a variable may contain the “:” character to make it close to their biochemical notations as much as possible, as well as the “/” and “+” special characters, as mnemonic codes for derived variables.

The recorded values for these response variables are given in a so-called “extended numerical” format. By this, we mean that the inputted number can be supplied with its \pm standard deviation or with an interval around

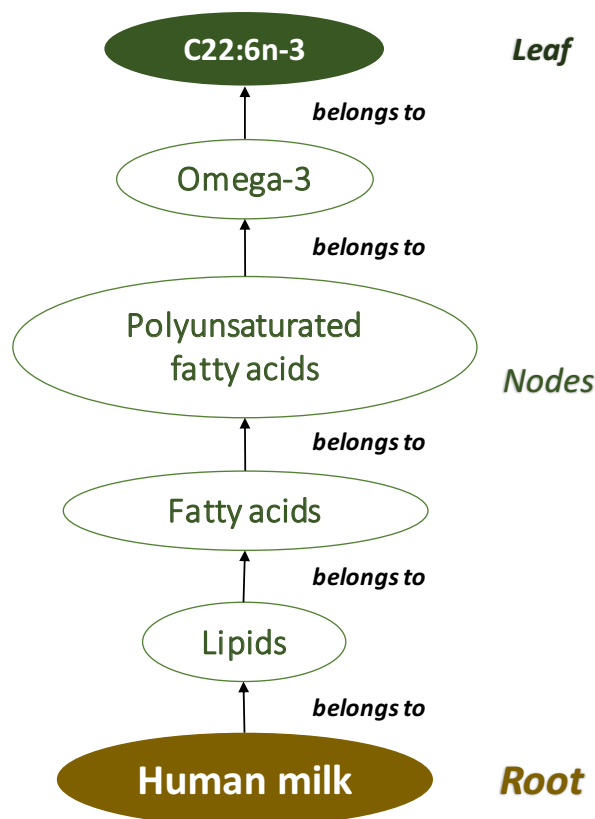


Fig. 1 Tree-structure of components.

it (like minimum-maximum, or quantile), both characterizing the uncertainty of the data. What is more, we differentiate between raw observations and estimations. Both can be inputted as response values, in the latter case with standard errors or confidence intervals. Finally, the response can be also dynamic, i.e. its temporal variation is stored in a table, and a pointer to the table is the inputted entry for the variable.

The condition fields do not necessarily hold only (extended) numerical values as above. They can be Boolean values or (a list of) categories, too. In the same way how a number belongs to an interval, a category value can belong to a group or to several groups. An example for this is the geographical region, indicating where an observation was made: the category group for China, for example can be either “Asia” or “FarEast”. Similar ambiguous definitions can occur say with Vitamin-D, by which typically we mean Vitamin-D₃, but this is not necessarily stated in the publications explicitly. Therefore, an accurate analysis of the data may introduce a probabilistic weight when characterizing the HM components at molecular level.

The variety of the data is restricted by the significance of the conditions on which the publications report. For example, the HM composition is rarely studied as a function of the sex of the new-born, so there is no separate field for that explanatory variable in the database, but the sex is included in the `cond_c` variable that contains relevant infant characteristics.

The veracity is also affected by confusions on statistical/numerical concepts. For example, sometimes the standard deviation of the measured values is mistaken with the standard error of their mean. Several publications have drawn the attention on this^{14,15}, but the mistake is still frequent. Similarly, either the publication or the person inputting the data may confuse quantiles (which is about the spread of the raw data), with confidence intervals (which is about the precision of the estimation). Whenever such errors are detected, we either correct them (if it is obvious) or mark them in the database (in less obvious situations).

Workflow.

The workflow can be overviewed as shown in Fig. 2.

Literature search. The publication search was partly manual, partly performed by FoodMine, a natural-language processing algorithm that finds papers on the chemical composition of a target food from PubMed¹⁶. The manual search used MeSH terms and Boolean operators in PubMed, with the following searching descriptor: (“human milk” OR “breast milk” OR “mothers’ milk”) AND (“nutrients” OR “components” OR “composition” OR “biochemical” OR “quantification” OR “bioactive”). The search was focused on, but not limited to, English language.

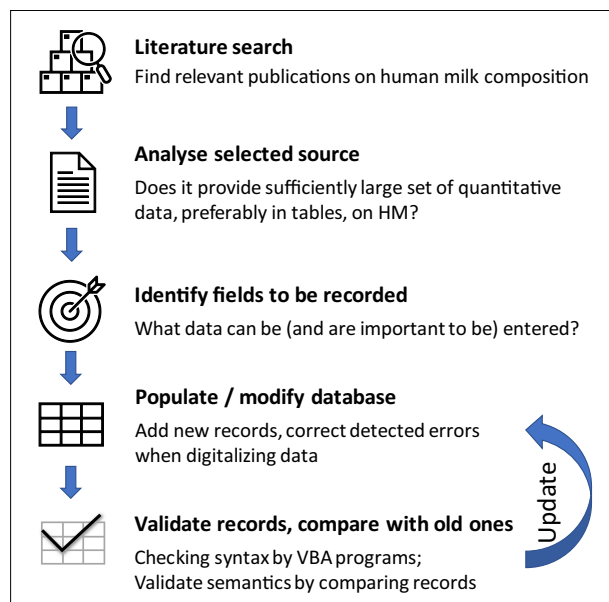


Fig. 2 Workflow of building MilkyBase.

Analyse source. The main selection criterium was quantitative data on the nutritional and/or non-nutritional components of HM. Priority was given to data (i) organized in a table format, in a systematic way; (ii) showing temporal variation (i.e., dynamic data); (iii) supplied with uncertainty quantification.

350 papers were selected by FoodMine and 201 were added from manual search. After elimination of irrelevant studies, a total of 365 potential papers were identified as suitable to enter the database. As of 1st July 2022, MilkyBase contains data from 140 papers.

Identify components. More than 600 (possibly derived) components have been identified so far, which can be either nodes or leafs of the tree-structured value set, or relationships between them. In this set, some individual molecules are represented both explicitly and implicitly (such as a specific fatty acid with unit g/litre of milk, also with a ratio to the total fatty acids, which is measured in gram. Taking out such duplicates, explicit measurements exist on ca 400 “genuine” components. Out of these, ca 50 are groups, i.e. they can be divided into either further groups or into molecules as the final leaves of the tree.

Data Records

The MilkyBase database is a system of connected tables represented by sheets in a single Microsoft Excel workbook (Fig. 3). Each record of its core (Master) sheet is identified by a unique key. Filling the source of the information, the geographic region of the measurement, the size of the cohort, the analytical method(s) measuring the component of interest in HM, as well as at least one condition and at least one response value are compulsory. The values in the Component and Condition fields can be “extended numerical” (e.g., numbers supplied with uncertainty quantification) as well as time-dependent series of numerical values, i.e., dynamic values. The syntax and the descriptions of the fields can be followed in sheets called “definition sheets”. These are also used by the “Syntax check” macro, which is part of the MBmacros.xlsm macro-enabled Excel workbook¹⁷, a collection of useful macros assigned to the database.

The relationships between the entries follow a tree-structure as before (Fig. 4). For example, the entries in the Conditions field can be numerical, just as the Component field, but also categories, which are defined in a nested way. An example for this is “Vitamin D in the diet”, which belongs to the Diet group, which in turn belongs to the mother-related “condition_m” group.

A big part of the implicit responses are proportions, mostly the concentration of a specific fatty acid molecule compared to the total fatty acid. From these, the concentration of the fatty acid molecule in question can be estimated only if the total fatty acid is known. The same holds for the situation when a molecule is measured in molecular weight; this can be converted to concentration only if the mol-weight is known; these are given in a separate field of the Master sheet. Therefore, it is possible that a certain molecule is measured in 2-3 ways. Deducing all these duplicates, the final number of explicitly recorded concentrations of molecules is currently 326. The list is expected to constantly expand as new data are coming in.

The information belonging to the CONDITION field have been organised in a similar way. 60 variables are identified and put in 6 main groups. The details are provided in the description file MBdescription.pdf¹⁷.

The MilkyBase.xlsx and its technical description MBdescription.pdf as well as the mentioned macros provided in a file called MBmacros.xlsm, were deposited in Figshare¹⁷.

| Key | Food | Source | Region | CohortSize | MeasMethod | Condition | Component |
|------------------|-----------|--------------|---------|------------|-------------------|-------------------------------------|--|
| HM-TP-Bri-86-01 | HumanMilk | Britton_86 | AZ | 70 | Lowry; CC-UV; SDS | GestAge(week)=[25,35] MilkStage(d | Prot(g/L)=!Britton_Prot AspA_mol(m |
| HM-TP-Bri-86-02 | HumanMilk | Britton_86 | AZ | 38 | Lowry; CC-UV; SDS | GestAge(week)=[38,42] MilkStage(d | Prot(g/L)=!Britton_Prot AspA_mol(m |
| HM-MM-But-84a-01 | HumanMilk | Butte_84a | TX | 8 | ABCM; KM; Colorim | GestAge(week)=33.9±2.3 Age_m(yr | Fat(g/L)=!Butte84aF_Fat Ca(g/L)=!But |
| HM-MM-But-84a-02 | HumanMilk | Butte_84a | TX | 13 | ABCM; KM; Colorim | GestAge(week)=39.2±1.4 Age_m(yr | Fat(g/L)=!Butte84aF_Fat Ca(g/L)=!But |
| HM-TP-But-84b-01 | HumanMilk | Butte_84b | TX | 45 | ABCM; KM; RGP; GL | Age_m(year)=38.0±3.1 Weight_c(g | Fat(g/L)=!Butte84b_Fat N(g/L)=!Butte |
| HM-MM-But-90-01 | HumanMilk | Butte_90 | TX | 10 | ABCM; KM; modifie | Weight_c(g)=33.45±335 GestAge(w | Fat(g/L)=33.53±7.75 N(g/L)=2.1±0.46 |
| HM-MM-But-90-02 | HumanMilk | Butte_90 | TX | 10 | ABCM; KM; modifie | Weight_c(g)=34.3±336 GestAge(w | Fat(g/L)=28.04±8.43 N(g/L)=1.73±0.14 |
| HM-TP-Cam-09-01 | HumanMilk | Campbell-Yeo | Halifax | 22 | KM; RGP; FAOM; CR | MilkStage(day)=[0,14] Domperidon | Fat(g/L)=!CampbellDomp_Fat Prot(g |

Fig. 3 The MilkyBase database is a system of 10 linked tables. The records of the core sheet are identified by a unique key and the possible values of a field are stored in the respective definition sheets with the same name.

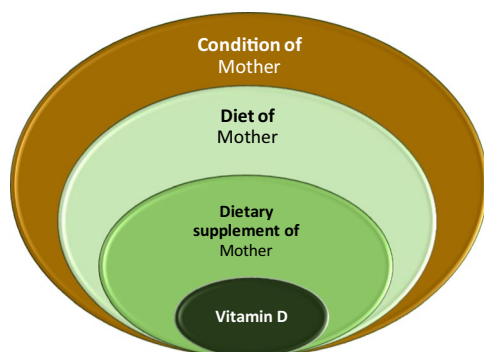


Fig. 4 Example for the nested grouping of condition values.

Technical Validation

The database validation was helped by MS Excel VBA macros. The MBmacros.xlsm file containing them is available at the Figshare repository¹⁷.

It was straightforward to develop a “Syntax check” code but semantic check would require biochemical understanding. Various comparative plots were used to identify anomalies in the publications, such as wrong units, contradictions between figures and tables or misinterpreted data-scatter and uncertainty quantifications.

Usage Notes

The presented MilkyBase database hosts records on milk composition in linked Excel tables. Its main novelty is the ontology that focusses on the effect of conditions under which the milk composition was measured, and the dynamics and uncertainty characteristics of these data, which will be entered in the explanatory and response fields. Its purpose is to provide a resource for researchers and a template for laboratories to put their own data into this format, thus initiating a knowledge-share following a kind of Wiki-philosophy.

Though the job of digitizing published data is rather laborious, as not everything can be automated, the main challenge in the development is its variety and veracity. “What to record” is a major decision and can be even biased.

It is impossible to totally automate the task of verification, either. Despite all the programming efforts, the task and responsibility must remain in the hands of the inputter and will remain dependent on human skill and expertise.

An example for the multivariate dynamic response inputted in a record is shown by Fig. 5. Such visualization is an aid to (i) recognize patterns and outliers in the data; (ii) identify data gaps; (iii) possibly identifying errors. For example this figure gives the idea, that the end of colostrum period can be defined as the time when the linear increase of fatty acid concentration is over.

Figure 6 compares the temporal variations of the concentration of Lacto-N-tetraose (LNT) in human milk as found by different authors. Here the observations of Kunz *et al.*¹⁸ show significant difference from other data, lending itself to an investigation what caused these differences.

MilkyBase demonstrates what benefits big data methods can bring for nutrition sciences. On a systematically organised database, users can run automated search and statistics that can help identifying data gaps (i.e., ideas for new research); finding mistakes in publications; and recognizing patterns, or possibly even model and optimize them for healthy infant and mother. A database like this needs to be of a relatively big volume (considering the complexity of the biochemical composition of milk), to get over a critical mass, from which we can consider the results as significant. Therefore, especially at the beginning of such database development, the amount of

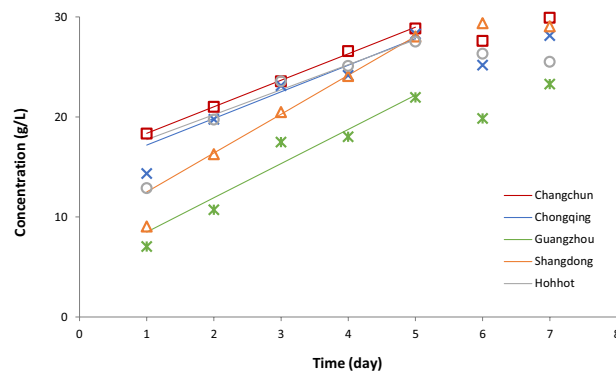


Fig. 5 Temporal variation of fatty acid concentration in colostrum breast milk in five cities of China¹⁹. After the first day, the increase of the fatty acid is remarkably linear until day 5, with similar slopes, except in Shangdong (see the continuous lines, fitted to the data shown in respective colours). The fatty acid levels of breast milk (but not the rate of its increase) in Guangzhou are significantly different from those in the other four cities.

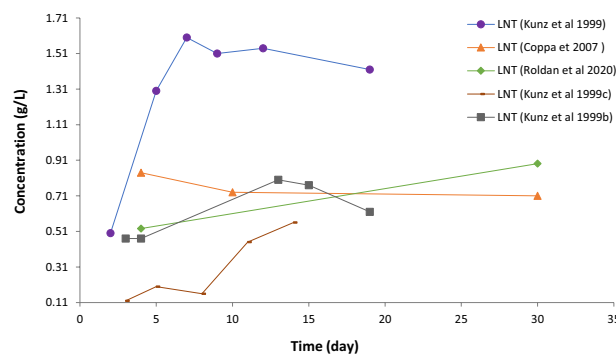


Fig. 6 Visualisation gives ideas what relationships and patterns should be checked regarding the temporal variation of components.

data that the authors make available in tables, plays a big role in the choice what papers should be digitized and recorded. Initially, the findings based on such database is inevitably more affected by what is *derivable* from the database, rather than what question is *desirable* to be solved by means of the database.

Code availability

MilkyBase.xlsx and its technical description MBdescription.pdf as well as the mentioned macros in an MBmacros.xlsm file, are available from the Figshare repository¹⁷.

Received: 6 May 2022; Accepted: 24 August 2022;

Published online: 09 September 2022

References

- Barabási, A.-L., Menichetti, G. & Loscalzo, J. The unmapped chemical complexity of our diet. *Nature Food* **1**, 33–37, <https://doi.org/10.1038/s43016-019-0005-1> (2020).
- Christian, P. *et al.* The need to study human milk as a biological system. *The American Journal of Clinical Nutrition* **113**, 1063–1072, <https://doi.org/10.1093/ajcn/nqab075> (2021).
- Samuel, T. M. *et al.* Nutritional and Non-nutritional Composition of Human Milk Is Modulated by Maternal, Infant, and Methodological Factors. *Frontiers in Nutrition* **7**, <https://doi.org/10.3389/fnut.2020.576133> (2020).
- Eidelman, A. I. *et al.* Breastfeeding and the Use of Human Milk. *Pediatrics* **129**, e827–e841, <https://doi.org/10.1542/peds.2011-3552> (2012).
- Gertosio, C., Meazza, C., Pagani, S. & Bozzola, M. Breastfeeding and its gamut of benefits. *Minerva Pediatr* **68**, 201–212 (2016).
- Carr, L. E. *et al.* Role of Human Milk Bioactives on Infants' Gut and Immune Health. *Front Immunol* **12**, 604080, <https://doi.org/10.3389/fimmu.2021.604080> (2021).
- Boix-Amorós, A. *et al.* Reviewing the evidence on breast milk composition and immunological outcomes. *Nutrition Reviews* **77**, 541–556, <https://doi.org/10.1093/nutrit/nuz019> (2019).
- Victora, C. G. *et al.* Breastfeeding in the 21st century: epidemiology, mechanisms, and lifelong effect. *Lancet* **387**, 475–490, [https://doi.org/10.1016/s0140-6736\(15\)01024-7](https://doi.org/10.1016/s0140-6736(15)01024-7) (2016).
- Patro-Gołąb, B. *et al.* Nutritional interventions or exposures in infants and children aged up to 3 years and their effects on subsequent risk of overweight, obesity and body fat: a systematic review of systematic reviews. *Obes Rev* **17**, 1245–1257, <https://doi.org/10.1111/obr.12476> (2016).
- Who. Global Strategy for Infant and Young Child Feeding. *Fifty-fourth world health assembly*, 8–8 (2003).
- Rollins, N. C. *et al.* Why invest, and what it will take to improve breastfeeding practices? *Lancet* **387**, 491–504, [https://doi.org/10.1016/s0140-6736\(15\)01044-2](https://doi.org/10.1016/s0140-6736(15)01044-2) (2016).

12. Morgenstern, J. D., Rosella, L. C., Costa, A. P., de Souza, R. J. & Anderson, L. N. Perspective: Big Data and Machine Learning Could Help Advance Nutritional Epidemiology. *Advances in Nutrition* **12**, 621–631, <https://doi.org/10.1093/advances/nmaa183> (2021)
13. PTFI. *Periodic Table of Food Initiative* <https://foodperiodictable.org/> (2021).
14. Vaux, D. L. Know when your numbers are significant. *Nature* **492**, 180–181, <https://doi.org/10.1038/492180a> (2012).
15. Chavalarias, D., Wallach, J. D., Li, A. H. T. & Ioannidis, J. P. A. Evolution of Reporting PValues in the Biomedical Literature, 1990–2015. *JAMA* **315**, 1141, <https://doi.org/10.1001/jama.2016.1952> (2016).
16. Hooton, F., Menichetti, G. & Barabási, A.-L. Exploring food contents in scientific literature with FoodMine. *Scientific Reports* **10**, <https://doi.org/10.1038/s41598-020-73105-0> (2020).
17. Pacza, T. MilkyBase, a database of human milk composition as a function of maternal-, infant- and measurement conditions, *figshare*, <https://doi.org/10.6084/m9.figshare.c.6160191.v1> (2022).
18. Kunz, C., Rudloff, S., Schad, W. & Braun, D. Lactose-derived oligosaccharides in the milk of elephants: comparison with human milk. *British Journal of Nutrition* **82**, 391–399, <https://doi.org/10.1017/s0007114599001798> (1999).
19. Liu, Y., Liu, X. & Wang, L. The investigation of fatty acid composition of breast milk and its relationship with dietary fatty acid intake in 5 regions of China. *Medicine* **98**, <https://doi.org/10.1097/md.00000000000015855> (2019).

Acknowledgements

The authors would like to thank Anna Jánosity, Gyöngyi Kirschner, Bence Pecsénye, Luis Quevedo and Chyanne Rosenbaum for their technical help. A-L B work was partially supported by American Heart Association grant no. 151708, ERC grant no. 810115-DYNASET and Rockefeller Foundation grant no. 2019 FOD 026.

Author contributions

Tünde Pacza contributed to the design of the database structure, participated in the data acquisition and validation, and contributed to the writing up of the manuscript. Mayara L. Martins contributed to the design of the database, participated in the data acquisition, worked on the validation of the database, and contributed to the writing up of the manuscript. Maha Rockaya contributed to the data acquisition and participated in its validation. Katalin Müller contributed to the writing up of the paper and to the conceptualization and design of the database. Ayan Chatterjee was responsible for the automated literature search. Albert-László Barabási contributed to the conceptualization of the study. József Baranyi contributed to the conceptualization and design of the database, coordinated the validation efforts, largely wrote and edited the manuscript. All authors reviewed and commented on the manuscript and approved the final draft.

Funding

Open access funding provided by University of Debrecen.

Competing interests

A.-L.B. is the founder of Scipher Medicine and Naring Health, companies that explore the use of network-based tools in health, and Datapolis, which focuses on urban data. The other authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to J.B.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022