Applying the Leiden Manifesto principles in practice

Commonalities and differences in interpretation

Lorna Wildgaard, Heidi Holst Madsen, Marianne Gauffriau
University of Copenhagen & Copenhagen University Library / The Royal Danish Library

UNIVERSITY OF COPENHAGEN

Agenda

Pilot study – Leiden Manifesto as a consumer label

(Presented at NWB 2017)

Analysis of interpretations of Leiden Manifesto

(Work in progress)

Template to support responsible metrics in bibliometric analysis

(Work in progress)

Pilot study – Leiden Manifesto (LM) as a consumer label

(Presented at NWB 2017)

LM as a consumer label

Nutrition Facts

Serving Size 8 fl oz
Serving Per Container 1

Amount Per Serving	
Calories 150	Calories from Fat 70
	% Daily Values*
Total Fat 8g	**12%**
Saturated Fat 5g	**25%**
Trans Fat 0g	
Cholesterol 35mg	**12%**
Sodium 120mg	**5%**
Total Carbohydrate 11g	**4%**
Dietary Fiber 0g	**0%**
Sugars 12g	
Protein 8g	**16%**

*Percent Daily Values are based on a 2,000 calorie diet.

Consumer label from: Clare, G. P. & Burghardt, K. Getting the Message: Front of Package Labeling. *Management*, 4(5): 112-122 (2014)

Leiden Manifesto: Ten principles to guide quantitative research evaluation	Assessment
1) Quantitative evaluation should support qualitative, expert assessment The bibliometric analysis is included in a CV and publication track record in a research application. The application will be evaluated by a panel of researchers from the health sciences.	☺
2) Measure performance against the research missions of the institution, group or researcher The indicators were not linked to an explicit research mission. According to the instructions for applicants, bibliometric indicators are not mandatory but may be included. No specific indicators are mentioned. The overall evaluation criterion is scientific excellence. In the health sciences, the h-index, number of publications and citations are often presented in a CV and may be seen as an implicit standard for showing impact, and together with other information indicate excellence.	😐
3) Protect excellence in locally relevant research Not relevant as the research area of Prof. NN and the application is international.	–
4) Keep data collection and analytical processes open, transparent and simple The analysis is developed in collaboration with Prof. NN and all indicators are known by the health sciences research community.	☺
5) Allow those evaluated to verify data and analysis The analysis is verified by Prof. NN.	☺
6) Account for variation by field in publication and citation practices The analysis does not support comparisons with other research fields, e.g. by showing field-normalized indicators or including indicators often used by other research fields.	☹
7) Base assessment of individual researchers on a qualitative judgement of their portfolio See principle 1.	☺
8) Avoid misplaced concreteness and false precision The analysis presents multiple indicators to give a pluralistic picture of Prof. NN's performance. The data set for the analysis is developed in collaboration with Prof. NN to ensure the best possible coverage.	☺
9) Recognize the systemic effects of assessment and indicators The analysis presents multiple indicators and not the h-index alone which is often seen in health sciences.	☺
10) Scrutinize indicators regularly and update them See principle 9.	☺

Pilot study: 2 cases

Dept. of Forensic Medicine

Analyzing department-level publication output, collaboration, and impact.

Source: http://maxpixel.freegreatpicture.com/Business-Feedback-Opinion-Group-Communication-2044702

Professor T. I. A. Sørensen

Analyzing publication output and impact of a researcher.

erc

European Research Council

Supporting top researchers from anywhere in the world

Pilot study conclusions: Does LM work as a consumer label?

- …
- Reliability of subjective interpretations of LM? 😐
- …

Analysis of interpretations of Leiden Manifesto

(Work in progress)

Three studies using the ten LM principles in practice

Evaluation of the ResearchGate Scores (ResearchGate Scores)

Orduna-Malea, E., Martín-Martín, A., Thelwall, M., & Delgado López-Cózar, E. (2017). Do ResearchGate Scores create ghost academic reputations? *Scientometrics*, *112*(1), 443–460. https://doi.org/10.1007/s11192-017-2396-9

Evaluation of the Brazilian graduate evaluation system implemented by the Federal Agency for Support and Evaluation of Graduate Education (Brazilian graduate evaluation system)

de Oliveira, T. M., & Amaral, L. (2017). Public Policies in Science and Technology in Brazil: challenges and proposals for the use of indicators in evaluation. In R. Mugnaini, A. Fujino, & N. Y. Kobashi (Eds.), *BIBLIOMETRICS AND SCIENTOMETRICS IN BRAZIL: SCIENTIFIC RESEARCH ASSESSMENT INFRASTRUCTURE IN THE ERA OF BIG DATA* (pp. 189–217). https://doi.org/10.11606/9788572051705

Evaluation of the evaluation procedures for individual researchers: The case of the Italian National Scientific Qualification (Italian National Scientific Qualification evaluation procedure)

Marzolla, M. (2016). Assessing evaluation procedures for individual researchers: The case of the Italian National Scientific Qualification. *Journal of Informetrics*, *10*(2), 408–438. https://doi.org/10.1016/j.joi.2016.01.009

Overlap, continuum, and contradictions across LM interpretations

Leiden Manifesto Principles	ResearchGate Scores vs. Brazilian graduate evaluation system	ResearchGate Scores vs. Italian National Scientific Qualification evaluation procedures	Brazilian graduate evaluation system vs. Italian National Scientific Qualification evaluation procedures
1) Quantitative evaluation should support qualitative, expert assessment			
2) Measure performance against the research missions of the institution, group or researcher			
3) Protect excellence in locally relevant research			
4) Keep data collection and analytical processes open, transparent and simple			
5) Allow those evaluated to verify data and analysis			
6) Account for variation by field in publication and citation practices			
7) Base assessment of individual researchers on a qualitative judgement of their portfolio			
8) Avoid misplaced concreteness and false precision			
9) Recognize the systemic effects of assessment and indicators			
10) Scrutinize indicators regularly and update them			

= same interpretation

= some overlap in the interpretation

= continuum but not contradicting interpretations

= contradicting interpretations

Include information on years of scientific career, experience, and activities from the users' portfolio. ≠ Selection of best publications.

3 * Indicators are updated regularly. Information on when and how is publicly available.

Analysis of statements in the documents

Negation handling

Principle 4: “RG Score is not transparent. Both indicators and weights keep under commercial secret” = Indicators and the calculation of them must be transparent for users.

General vs. detailed statements

Principle 6: “currently variations by field are not considered” vs.

“CAPES establishes the general principles for evaluation, such as the standard of the Evaluation Form, its requirements and general items that must be included in all areas. However, each of the 49 areas can customize its criteria and indicators, as long as it follows the minimum required in the regulations.
Thus, areas may give distinctive importance to intellectual products. Areas belonging to the Humanities generally give higher weight to books. Computer Science is one area that punctuates strongly scientific conferences. Biotechnology values heavily the production of patents, considering their innovative character. Indicators for each of the items also vary. In any case, all the production of the program is considered for evaluation purposes, not only scientific articles, but also books, conferences papers, technical and artistic production.
In the case of master’s professional programs there is differentiation in the evaluation items. Technical products are valued most and there is a greater variety of types of graduate work, which can be a software development, technical report, protocol, production of didactic or instructional material.
There is no a priori differentiation on weight assessment of intelelectual products as a function of language of the publication. Some areas recommend that publications be written in English, aiming for greater internationalization of the program, but this is not mandatory.
Those publications indexed in international databases such as Web of Science and Scopus are evaluated based on available bibliometric metrics. However, the committees also consider databases with greater regional coverage, such as Scielo (Scientific Electronic Library Online). Non-indexed ones are evaluated for their local or regional impact, considering the importance of the content developed and the objectives of the research.
For example, within the area of Agrarian Sciences, the development of an agricultural technique in a drought region or the planting of a specific cultivar for a particular region can have a significant impact locally and regionally, but not at a national level. The output (an article or book) from this research will probably not have a large number of citations, but that does not mean that it has not had relevant results. The same is valid for the valorization of programs in Tropical Medicine, in Brazilian public health literature, teaching of local history among others, in which the less dissemination does not mean lack of quality or prestige.”

Term interpretation

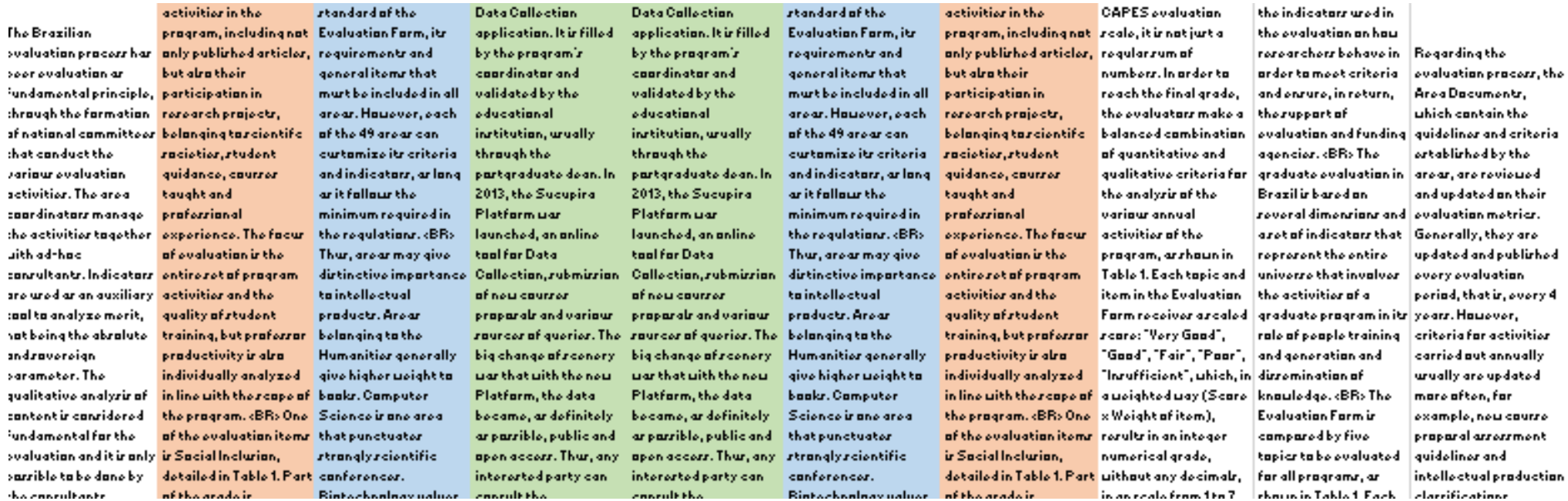
Principle 10: “algorithm”, “metrics”, “criteria for activities”, “criteria and parameters” = indicators

Next steps

2. LM interpretations' compliance with LM

3. Overlaps or contradictions in a LM interpretation:

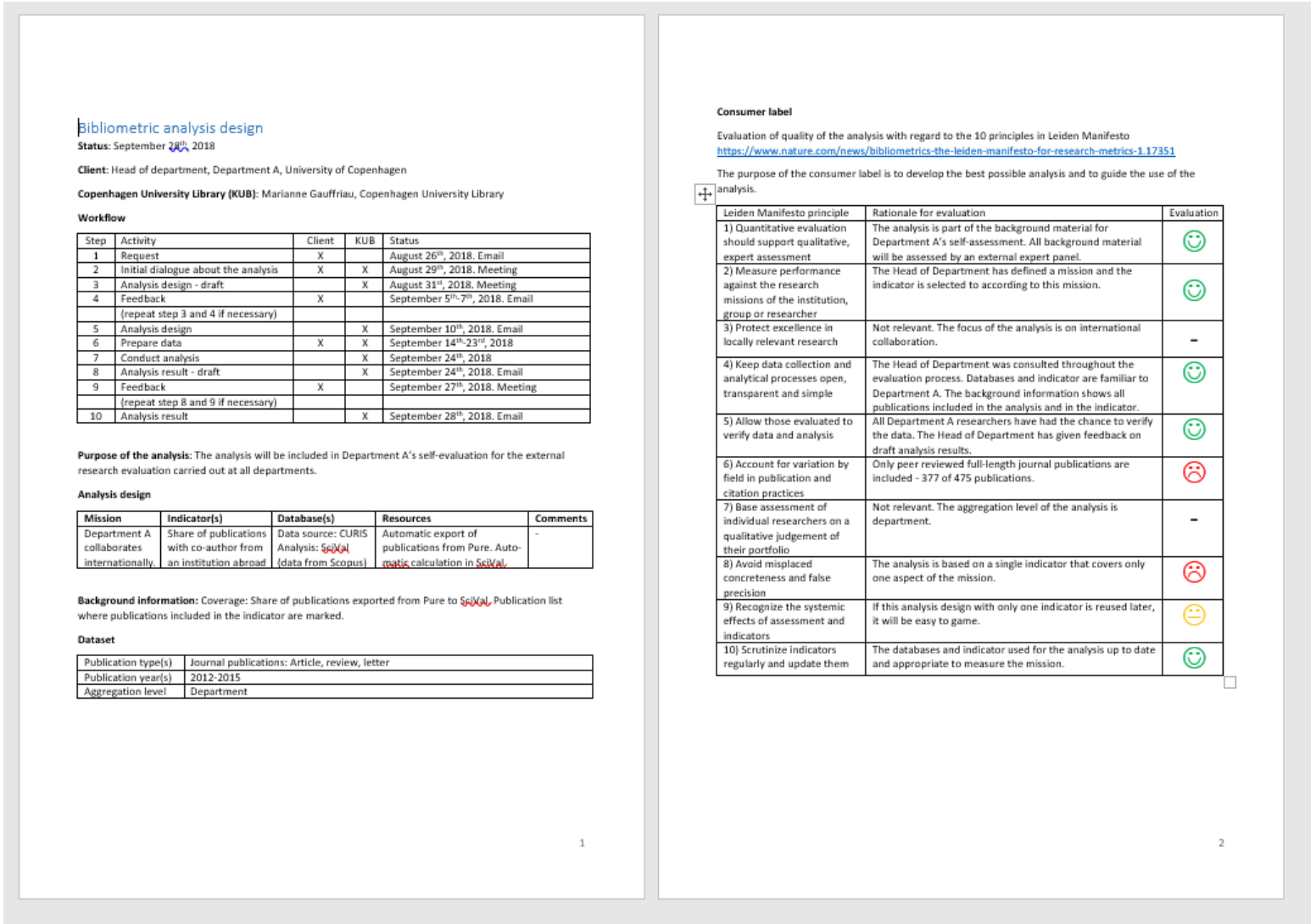
Brazilian graduate evaluation system



Template to support responsible metrics in bibliometric analysis

(Work in progress)

Template



Bibliometric analysis design

Status: September 28th, 2018

Client: Head of department, Department A, University of Copenhagen

Copenhagen University Library (KUB): Marianne Gauffriau, Copenhagen University Library

Workflow

Step	Activity	Client	KUB	Status
1	Request	X		August 26th, 2018. Email
2	Initial dialogue about the analysis	X	X	August 29th, 2018. Meeting
3	Analysis design - draft		X	August 31st, 2018. Meeting
4	Feedback	X		September 5th-7th, 2018. Email
	(repeat step 3 and 4 if necessary)			
5	Analysis design		X	September 10th, 2018. Email
6	Prepare data	X	X	September 14th-23rd, 2018
7	Conduct analysis		X	September 24th, 2018
8	Analysis result - draft		X	September 24th, 2018. Email
9	Feedback	X		September 27th, 2018. Meeting
	(repeat step 8 and 9 if necessary)			
10	Analysis result		X	September 28th, 2018. Email

Purpose of the analysis: The analysis will be included in Department A's self-evaluation for the external research evaluation carried out at all departments.

Analysis design

Mission	Indicator(s)	Database(s)	Resources	Comments
Department A collaborates internationally.	Share of publications with co-author from an institution abroad	Data source: CURIS Analysis: SciVal (data from Scopus)	Automatic export of publications from Pure. Automatic calculation in SciVal.	-

Background information: Coverage: Share of publications exported from Pure to SciVal. Publication list where publications included in the indicator are marked.

Dataset

Publication type(s)	Journal publications: Article, review, letter
Publication year(s)	2012-2015
Aggregation level	Department

1

Consumer label

Evaluation of quality of the analysis with regard to the 10 principles in Leiden Manifesto
https://www.nature.com/news/bibliometrics-the-leiden-manifesto-for-research-metrics-1.17351

The purpose of the consumer label is to develop the best possible analysis and to guide the use of the analysis.

Leiden Manifesto principle	Rationale for evaluation	Evaluation
1) Quantitative evaluation should support qualitative, expert assessment	The analysis is part of the background material for Department A's self-assessment. All background material will be assessed by an external expert panel.	☺
2) Measure performance against the research missions of the institution, group or researcher	The Head of Department has defined a mission and the indicator is selected to according to this mission.	☺
3) Protect excellence in locally relevant research	Not relevant. The focus of the analysis is on international collaboration.	-
4) Keep data collection and analytical processes open, transparent and simple	The Head of Department was consulted throughout the evaluation process. Databases and indicator are familiar to Department A. The background information shows all publications included in the analysis and in the indicator.	☺
5) Allow those evaluated to verify data and analysis	All Department A researchers have had the chance to verify the data. The Head of Department has given feedback on draft analysis results.	☺
6) Account for variation by field in publication and citation practices	Only peer reviewed full-length journal publications are included - 377 of 475 publications.	☹
7) Base assessment of individual researchers on a qualitative judgement of their portfolio	Not relevant. The aggregation level of the analysis is department.	-
8) Avoid misplaced concreteness and false precision	The analysis is based on a single indicator that covers only one aspect of the mission.	☹
9) Recognize the systemic effects of assessment and indicators	If this analysis design with only one indicator is reused later, it will be easy to game.	😐
10) Scrutinize indicators regularly and update them	The databases and indicator used for the analysis up to date and appropriate to measure the mission.	☺

2

Workflow

Bibliometric analysis design

Status: September 28th, 2018

Client: Head of department, Department A, University of Copenhagen

Copenhagen University Library (KUB): Marianne Gauffriau, Copenhagen University Library

Workflow

Step	Activity	Client	KUB	Status
1	Request	X		August 26th, 2018. Email
2	Initial dialogue about the analysis	X	X	August 29th, 2018. Meeting
3	Analysis design - draft		X	August 31st, 2018. Meeting
4	Feedback	X		September 5th-7th, 2018. Ema
	(repeat step 3 and 4 if necessary)			
5	Analysis design		X	September 10th, 2018. Email
6	Prepare data	X	X	September 14th-23rd, 2018
7	Conduct analysis		X	September 24th, 2018
8	Analysis result - draft		X	September 24th, 2018. Email
9	Feedback	X		September 27th, 2018. Meeting
	(repeat step 8 and 9 if necessary)			
10	Analysis result		X	September 28th, 2018. Email

Consumer label

Evaluation of quality of the analysis with regard to the 10 principles in Leiden Manifesto
https://www.nature.com/news/bibliometrics-the-leiden-manifesto-for-research-metrics-1.17351

The purpose of the consumer label is to develop the best possible analysis and to guide the use of the analysis.

Leiden Manifesto principle	Rationale for evaluation	Evaluation
1) Quantitative evaluation should support qualitative, expert assessment	The analysis is part of the background material for Department A's self-assessment. All background material will be assessed by an external expert panel.	☺
2) Measure performance against the research missions of the institution, group or researcher	The Head of Department has defined a mission and the indicator is selected to according to this mission.	☺
3) Protect excellence in locally relevant research	Not relevant. The focus of the analysis is on international collaboration.	-
4) Keep data collection and analytical processes open, transparent and simple	The Head of Department was consulted throughout the evaluation process. Databases and indicator are familiar to Department A. The background information shows all publications included in the analysis and in the indicator.	☺
5) Allow those evaluated to verify data and analysis	All Department A researchers have had the chance to verify the data. The Head of Department has given feedback on draft analysis results.	☺
6) Account for variation by field in publication and citation practices	Only peer reviewed full-length journal publications are included - 377 of 475 publications.	☹
7) Base assessment of individual researchers on a qualitative judgement of their portfolio	Not relevant. The aggregation level of the analysis is department.	-
8) Avoid misplaced concreteness and false precision	The analysis is based on a single indicator that covers only one aspect of the mission.	☹
9) Recognize the systemic effects of assessment and indicators	If this analysis design with only one indicator is reused later, it will be easy to game.	😐
10) Scrutinize indicators regularly and update them	The databases and indicator used for the analysis up to date and appropriate to measure the mission.	☺

Design and examples of LM principles

Purpose of the analysis: The analysis will be included in Department A's self-evaluation for the external research evaluation carried out at all departments.

Analysis design

Mission	Indicator(s)	Database(s)	Resources	
Department A collaborates internationally.	Share of publications with co-author from an institution abroad	Data source: CURIS Analysis: SciVal (data from Scopus)	Automatic export of publications from Pure. Auto-matic calculation in SciVal	-

Background information: Cove Pure to SciVal. Publication list where publications included in

Dataset

Publication type(s)	Journal publications: Article, review, letter
Publication year(s)	2012-2015
Aggregation level	Department

Principle 1: Quantitative evaluation should support qualitative, expert assessment.

Principle 2: Measure performance against the research missions of the institution, group or researcher.

Result and background information and examples of LM principles

	A	B	C	D	E	F	G	H	I
1	**Results from SciVal**								
2									
3	**Share of publications with co-author from an institution abroad**								
4	377 of 380 publications are exported from Pure to SciVal. Hereof 220 (58 %) publications have at least one co-author from an insititution abroad.								
5									
6									
7	**Background information**								
8	Exported from Pure to SciVal	Publications with at least one co-author from an insititution abroad	Title				Journal		
9	x	x	A genome-wide study			h	European Journal of Human Genetics		
10	x	x	A New Class of Amp			h	Chemistry: A European Journal		
11	x		A single bout of exer			h	P L o S One		
12	x	x	Alterations of monoca			sh	Cellular oncology (Dordrecht)		
13	x	x	Altered expression of brain monocarboxylate transp	2012	Article	English	Neurobiology of Disease		
14	x	x	An allosteric enhancer of M4 muscarinic acetylcholi	2012	Article	English	Psychopharmacology		
15	x	x	Antagonist muscle moment is increased in ACL def	2012	Article	English	Knee		
16	x	x	Assessment of a portable device for the quantitativ	2012	Article	English	Clinical Neurophysiology		
17	x	x	Association of the leucine-7 to proline-7 variation in	2012	Article	English	Acta Neuropsychiatrica		
18	x	x	Brain energy metabolism and blood flow differences	2012	Article	English	Journal of Cerebral Blood Flow and Me		
19	x		Celebrating the 30th anniversary of our journal	2012	Article	English	Journal of Cerebral Blood Flow and Me		
20	x	x	Cerebral gray and white matter changes and clinical	2012	Article	English	Neurology		
21	x	x	Cerebral oxygen metabolism in patients with early F	2012	Article	English	Journal of the Neurological Sciences		

Principle 4: Keep data collection and analytical processes open, transparent and simple.

Principle 5: Allow those evaluated to verify data and analysis.

Thank you for your attention

Nutrition Facts

Serving Size 8 fl oz
Serving Per Container 1

Amount Per Serving	
Calories 150	Calories from Fat 70
	% Daily Values*
Total Fat 8g	**12%**
Saturated Fat 5g	**25%**
Trans Fat 0g	
Cholesterol 35mg	**12%**
Sodium 120mg	**5%**
Total Carbohydrate 11g	**4%**
Dietary Fiber 0g	**0%**
Sugars 12g	
Protein 8g	**16%**

*Percent Daily Values are based on a 2,000 calorie diet.

Consumer label from: Clare, G. P. & Burghardt, K. Getting the Message: Front of Package Labeling. *Management*, 4(5): 112-122 (2014)

Leiden Manifesto: Ten principles to guide quantitative research evaluation	Assessment
1) Quantitative evaluation should support qualitative, expert assessment The bibliometric analysis is included in a CV and publication track record in a research application. The application will be evaluated by a panel of researchers from the health sciences.	☺
2) Measure performance against the research missions of the institution, group or researcher The indicators were not linked to an explicit research mission. According to the instructions for applicants, bibliometric indicators are not mandatory but may be included. No specific indicators are mentioned. The overall evaluation criterion is scientific excellence. In the health sciences, the h-index, number of publications and citations are often presented in a CV and may be seen as an implicit standard for showing impact, and together with other information indicate excellence.	😐
3) Protect excellence in locally relevant research Not relevant as the research area of Prof. NN and the application is international.	-
4) Keep data collection and analytical processes open, transparent and simple The analysis is developed in collaboration with Prof. NN and all indicators are known by the health sciences research community.	☺
5) Allow those evaluated to verify data and analysis The analysis is verified by Prof. NN.	☺
6) Account for variation by field in publication and citation practices The analysis does not support comparisons with other research fields, e.g. by showing field-normalized indicators or including indicators often used by other research fields.	☹
7) Base assessment of individual researchers on a qualitative judgement of their portfolio See principle 1.	☺
8) Avoid misplaced concreteness and false precision The analysis presents multiple indicators to give a pluralistic picture of Prof. NN's performance. The data set for the analysis is developed in collaboration with Prof. NN to ensure the best possible coverage.	☺
9) Recognize the systemic effects of assessment and indicators The analysis presents multiple indicators and not the h-index alone which is often seen in health sciences.	☺
10) Scrutinize indicators regularly and update them See principle 9.	☺