Applying bibliometrics in research assessment and management ...

The real deal !

Dr. Thed van Leeuwen
Presentation at the NARMA Meeting, 29th march 2017

Outline

- *CWTS and Bibliometrics*

- *Detail and accuracy in bibliometric applications*

- *Normalization in bibliometrics*

- *Coverage in bibliometric studies*

- *Infamous bibliometric indicators – What to avoid*

- *CWTS methodology – basic indicators*

- *Advantages and disadvantages in bibliometric analysis*

CWTS and Bibliometrics

Introduction of bibliometrics

- Quantitative analysis + the cognitive and organizational structure of science and technology

- Scientific communication - journal publications

- Output and Impact, as measured through publications and citations

- Scientists express, through citations, a certain degree of influence of others on their own work

- Citations indicate influence or (inter)national visibility
    → Does not equal *'quality'*

CWTS

CWTS data system

- CWTS has a full bibliometric license from Thomson Reuters to conduct evaluation studies using the Web of Science.

- Our database covers the period 1981-2015/6.

- Some characteristics:
    - Over 46.000.000 publications.
    - Over 700.000.000 citation relations between source papers.
    - Author disambiguation tools are applied, linked with acquired experience
    - Various bases for field normalization
    - Continuous address cleaning tools being developed, related to the *Leiden Ranking*.
    - Contains reference sets for journal and field citation data.

CWTS

A less neutral approach …

- Bibliometric measures tend to shape what they measure

- Bibliometrics has some serious shortcomings

- Better not be used as a stand-alone tool

- There is a lot of academic debate on the meaning of citations

- However, we still consider bibliometric techniques helpful tools in the assessment of research performance and everything that comes with it

CWTS

Coverage in bibliometric studies

Appropriateness of bibliometric analysis

Introduction

- The use of evaluative bibliometrics can only become meaningful when used in a the right context.

- Publication culture of the unit(s) under assessment are shaping that context.

- As such, any bibliometric study should start with an assessment of the adequacy of metrics in that particular context.

- Therefore, CWTS has developed methods to assess that fit of metrics in a certain context.

How to define adequate coverage ?

- In order to determine whether metrics applied in an assessment context are meaningful, one needs to know what is represented through the metrics.

- We distinguish two types of coverage:
  - *Internal*   (from <u>inside</u> the perspective of the WoS)
  - *External*   (from the perspective of a total output set)

CWTS

Assessing the adequacy of WoS for bibliometrics: The *Internal* coverage method

- Look at publications in WoS across fields,

- Use the references given by the authors of the publications,

- Analyze the communication channels referred to,

- Usage of WoS journals as share of the total number of references is an indication of the relevance for the authors involved,

- Thereby constituting a basis for the usage of bibliometrics as evaluation tool !

Assessing the adequacy of WoS for bibliometrics: The *External* coverage method
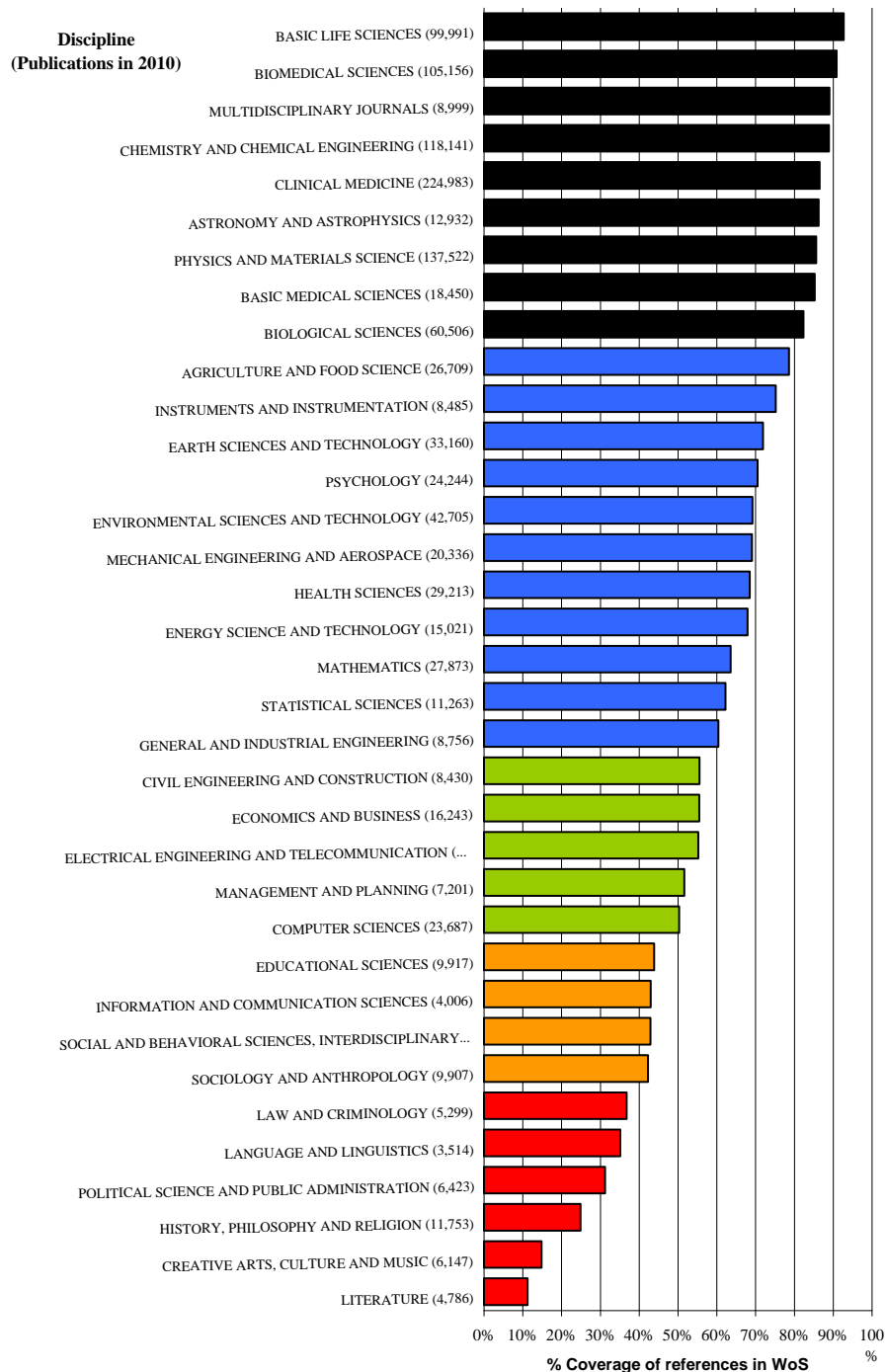
– Use the list of publications of an organization, subject of a bibliometric analysis => here in Norway, one could use Cristin

– Match the submitted list with the WoS.

– Degrees of covered scientific outlets indicate the relevance of WoS journals.

– Thereby constituting a basis for the usage of bibliometrics as an evaluation tool !

CWTS

Internal coverage in bibliometric studies

		in WoS
AU	Moed, HF; Garfield, E.	
TI	In basic science the percentage of 'authoritative' references decreases as bibliographies become shorter	
SO	**SCIENTOMETRICS** 60 (3): 295-303, 2004	Y
RF	**ABT HA**, J AM SOC INF SCI T, v 53, p 1106, 2004	Y
	GAR... ...TION INDEXING, 1979 (BOOK!)	N
	GAR... ...S INFORMATION S, v 8, p 403, 1985	N
	GILBERT GN, SOC STUDIES SCI, v 7, p 113, 1977	Y
	MERTON...	Y
	RO...	Y
	ZUC...	Y

Not in WoS

WoS Coverage = 5/7 = 71%

Discipline (Publications in 2010)

Bar chart: % Coverage of references in WoS

Discipline	
BASIC LIFE SCIENCES (99,991)	(black)
BIOMEDICAL SCIENCES (105,156)	(black)
MULTIDISCIPLINARY JOURNALS (8,999)	(black)
CHEMISTRY AND CHEMICAL ENGINEERING (118,141)	(black)
CLINICAL MEDICINE (224,983)	(black)
ASTRONOMY AND ASTROPHYSICS (12,932)	(black)
PHYSICS AND MATERIALS SCIENCE (137,522)	(black)
BASIC MEDICAL SCIENCES (18,450)	(black)
BIOLOGICAL SCIENCES (60,506)	(black)
AGRICULTURE AND FOOD SCIENCE (26,709)	(blue)
INSTRUMENTS AND INSTRUMENTATION (8,485)	(blue)
EARTH SCIENCES AND TECHNOLOGY (33,160)	(blue)
PSYCHOLOGY (24,244)	(blue)
ENVIRONMENTAL SCIENCES AND TECHNOLOGY (42,705)	(blue)
MECHANICAL ENGINEERING AND AEROSPACE (20,336)	(blue)
HEALTH SCIENCES (29,213)	(blue)
ENERGY SCIENCE AND TECHNOLOGY (15,021)	(blue)
MATHEMATICS (27,873)	(blue)
STATISTICAL SCIENCES (11,263)	(blue)
GENERAL AND INDUSTRIAL ENGINEERING (8,756)	(blue)
CIVIL ENGINEERING AND CONSTRUCTION (8,430)	(green)
ECONOMICS AND BUSINESS (16,243)	(green)
ELECTRICAL ENGINEERING AND TELECOMMUNICATION (...	(green)
MANAGEMENT AND PLANNING (7,201)	(green)
COMPUTER SCIENCES (23,687)	(green)
EDUCATIONAL SCIENCES (9,917)	(orange)
INFORMATION AND COMMUNICATION SCIENCES (4,006)	(orange)
SOCIAL AND BEHAVIORAL SCIENCES, INTERDISCIPLINARY...	(orange)
SOCIOLOGY AND ANTHROPOLOGY (9,907)	(orange)
LAW AND CRIMINOLOGY (5,299)	(red)
LANGUAGE AND LINGUISTICS (3,514)	(red)
POLITICAL SCIENCE AND PUBLIC ADMINISTRATION (6,423)	(red)
HISTORY, PHILOSOPHY AND RELIGION (11,753)	(red)
CREATIVE ARTS, CULTURE AND MUSIC (6,147)	(red)
LITERATURE (4,786)	(red)

0% 10% 20% 30% 40% 50% 60% 70% 80% 90% 100%

% Coverage of references in WoS
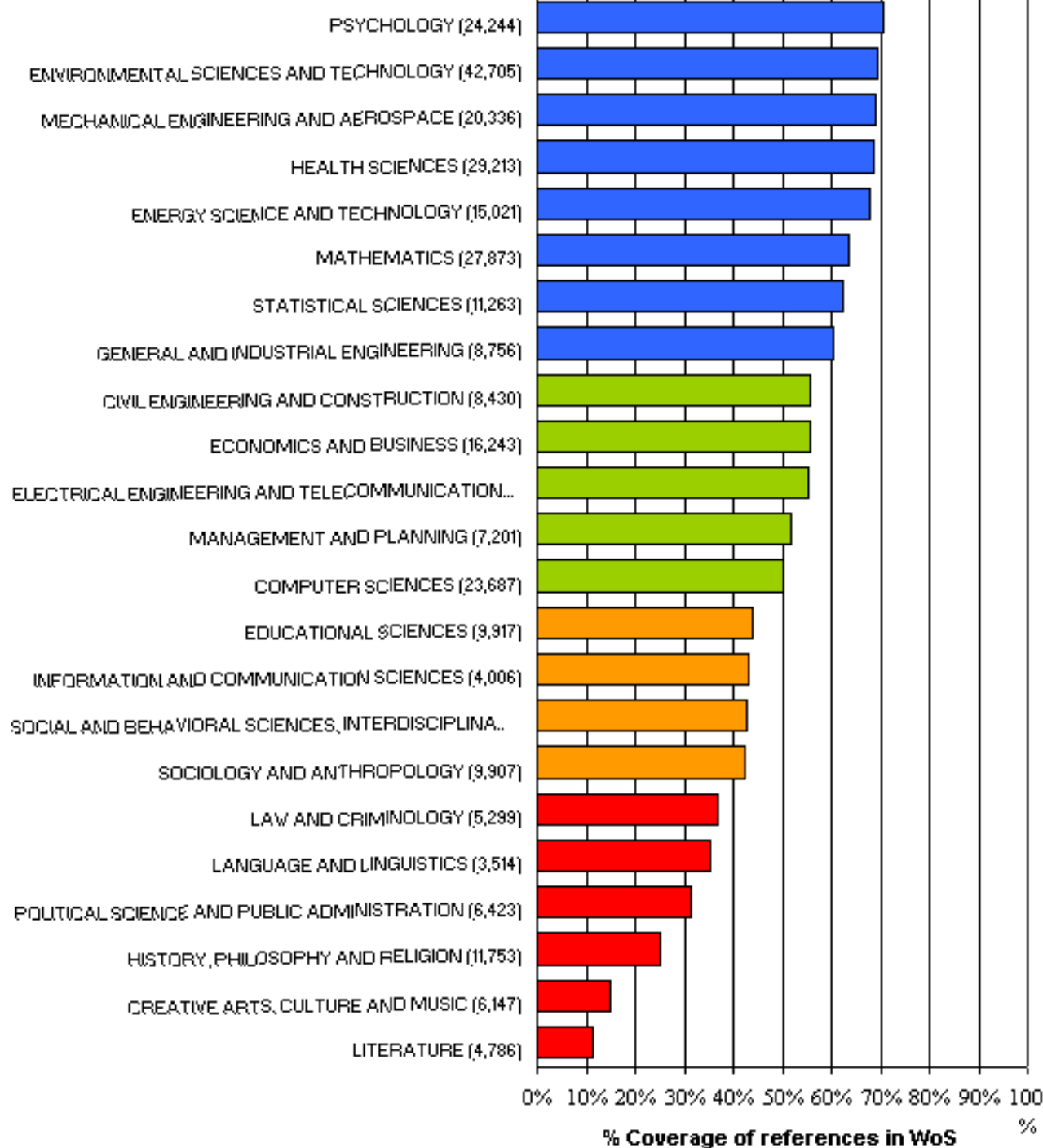
# WoS Coverage in 2010 across disciplines

- **Black=Excellent coverage** (>80%)

- **Blue= Good coverage** (between 60-80%)

- **Green= Moderate coverage** (but above 50%)

- **Orange= Moderate coverage** (below 50%, but above 40%)

- **Red= Poor coverage** (highly problematic, below 40%)

PSYCHOLOGY (24,244)
ENVIRONMENTAL SCIENCES AND TECHNOLOGY (42,705)
MECHANICAL ENGINEERING AND AEROSPACE (20,336)
HEALTH SCIENCES (29,213)
ENERGY SCIENCE AND TECHNOLOGY (15,021)
MATHEMATICS (27,873)
STATISTICAL SCIENCES (11,263)
GENERAL AND INDUSTRIAL ENGINEERING (8,756)
CIVIL ENGINEERING AND CONSTRUCTION (8,430)
ECONOMICS AND BUSINESS (16,243)
ELECTRICAL ENGINEERING AND TELECOMMUNICATION...
MANAGEMENT AND PLANNING (7,201)
COMPUTER SCIENCES (23,687)
EDUCATIONAL SCIENCES (9,917)
INFORMATION AND COMMUNICATION SCIENCES (4,006)
SOCIAL AND BEHAVIORAL SCIENCES, INTERDISCIPLINA...
SOCIOLOGY AND ANTHROPOLOGY (9,907)
LAW AND CRIMINOLOGY (5,299)
LANGUAGE AND LINGUISTICS (3,514)
POLITICAL SCIENCE AND PUBLIC ADMINISTRATION (6,423)
HISTORY, PHILOSOPHY AND RELIGION (11,753)
CREATIVE ARTS, CULTURE AND MUSIC (6,147)
LITERATURE (4,786)

0% 10% 20% 30% 40% 50% 60% 70% 80% 90% 100 %

% Coverage of references in WoS

External coverage in bibliometric studies

Difference between the internal registration system & representation WoS
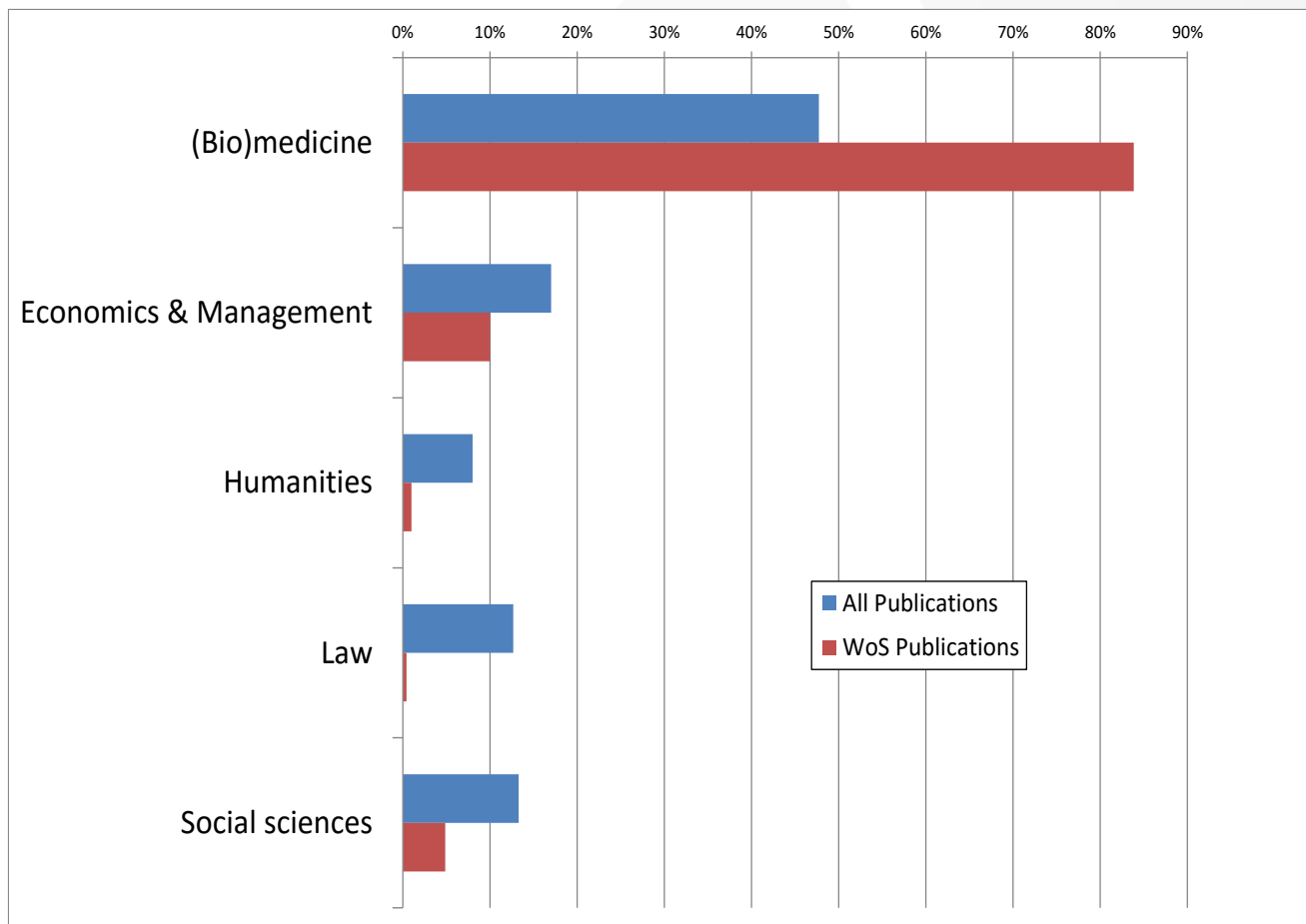
- Dominance university hospital in WoS realm extremely visible
- Law and Humanities 'disappear' in WoS realm

Composition of the output of the university in METIS
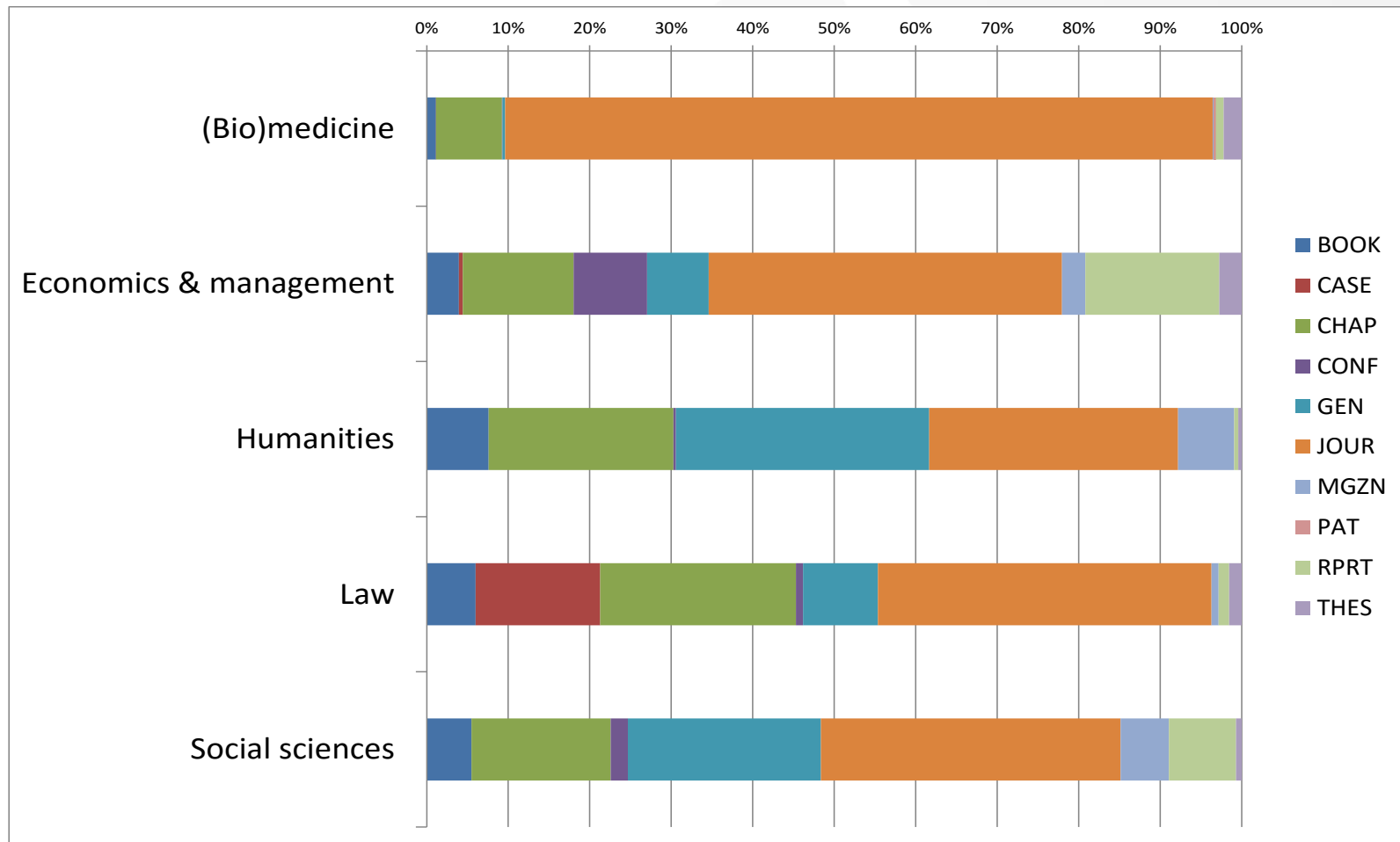
- The category General is in some cases voluminous

- All units **do** have journal publications !

Let us get started: Selection of indicators

What indicators are considered as valid in research assessment contexts?

- **Absolute numbers: publications**

    – Too little specific, only focus on productivity

- **Absolute numbers: citations**

    – Too little specific, as well as too much dependent on field

- **Average numbers: publications**

    – Related to the number of staff involved, in combination with field specific publication culture

- **Average numbers: citations**

    – Combining the disadvantages of the two previous options, namely field specific production and reference cultures.

CWTS

Dutch evaluation system: SEP protocol

- **System approved by VSNU-KNAW-NWO**

  – Focus on Institute/Department

  – Stay away from productivity as indicator

  – Include Societal Relevance as dimension

  – Peer review is central

- **Applies also on non-academic research**

- **Periodical/disciplinary by nature**

CWTS

The landscape of Dutch Psychology research

Vrije Univ Amsterdam

1614 Publications in 11-15

Univ Groningen

1164 Publications in 11-15

UvA

1260 Publications in 11-15

Univ Leiden

1245 Publications in 11-15

Univ Utrecht

1563 Publications in 11-15

Erasmus Univ

608 Publications in 11-15

Open Univ

298 Publications in 11-15

Univ Tilburg

1335 Publications in 11-15

Univ Maastricht

1543 Publications in 11-15

CWTS
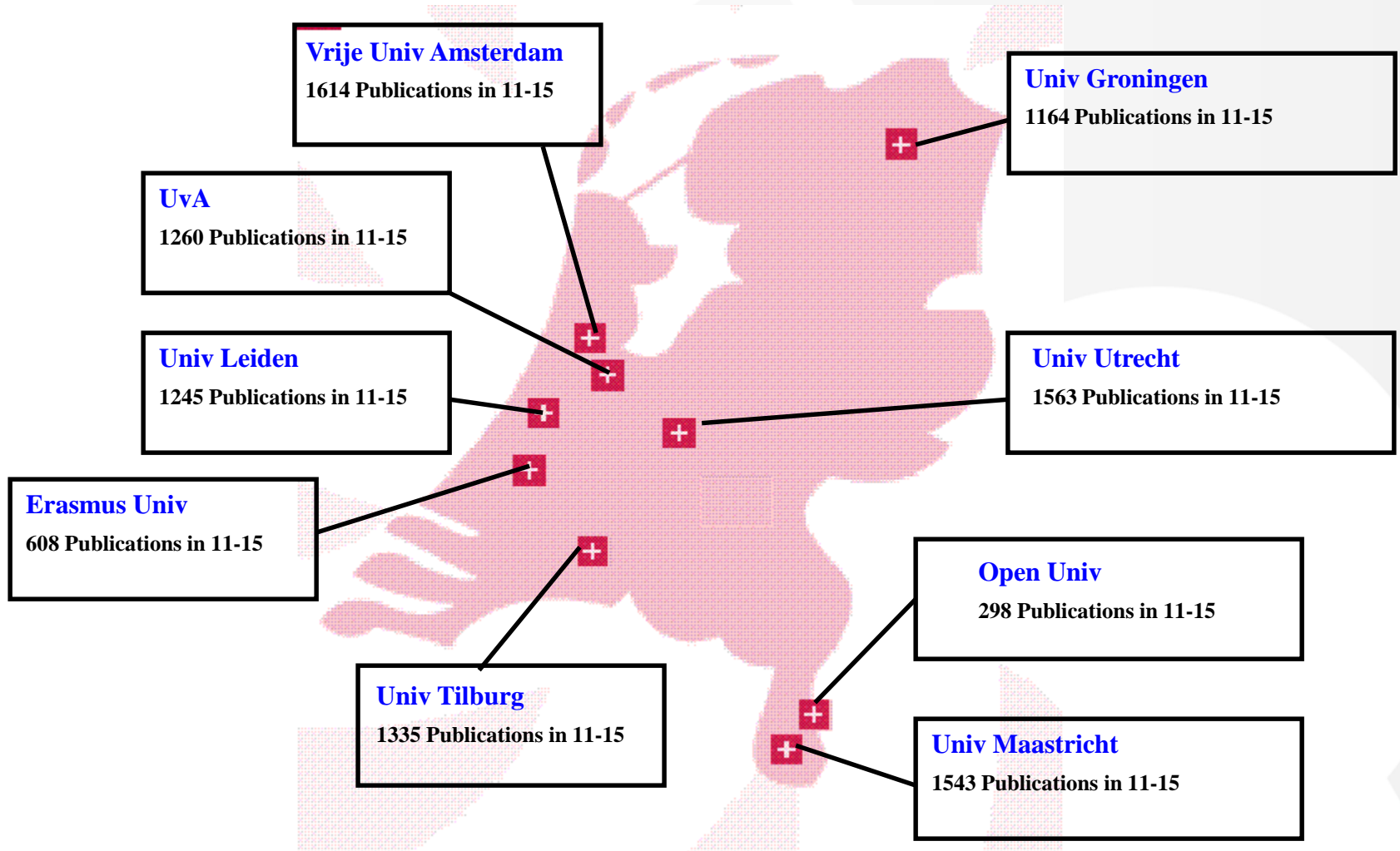
What if ... ?

- **When we use the Journal Impact Factor (JIF) ?**
  - ....

- **When we use the h-index ?**
  - ....

CWTS

Definitions of Journal Impact Factor & Hirsch Index

- **Definition of JIF:**
  - The mean citation score of a journal, determined by dividing all citations in year T by all citable documents in years T-1 and T-2.

- **Definition of h-index:**
  - The 'impact' of a researcher, determined by the number of received citations of an oeuvre, sorted by descending order, where the number of received citations on that single paper equals the rank position.

CWTS

Departments sorted by FTe

Sum of JIF values

Mean of JIF values

	Pubs	tcs	mcs	t_JIFs	m_JIFs
Psy Dept A	303	2741	9,05	882,75	2,91
Psy Dept B	607	6252	10,30	1659,93	2,73
Psy Dept C	1177	12358	10,50	3759,63	3,19
Psy Dept D	1245	14851	11,93	4168,19	3,35
Psy Dept E	1268	18945	14,94	4830,89	3,81
Psy Dept F	1359	13686	10,07	4081,37	3,00
Psy Dept G	1554	17595	11,32	5281,18	3,40
Psy Dept H	1574	16940	10,76	5062,70	3,22
Psy Dept I	1632	28359	17,38	7412,37	4,54

CWTS

… but what does the *Mean of JIF values* really mean ?

Problems with JIF

- **Methodological issues**

  - Was/is calculated erroneously (Moed & van Leeuwen, 1996)

  - Not field normalized

  - Not document type normalized

  - Underlying citation distributions are highly skewed (Seglen, 1994)

- **Conceptual/general issues**

  - Inflation (van Leeuwen & Moed, 2002)

  - Availability promotes journal publishing

  - Is based on expected values only

  - Stimulates one-indicator thinking

  - Ignores other scholarly virtues

CWTS

A policy related question

- What is the status of our current work force, compared to a previous situation ?

- People move, so what happened in time with new staff members coming in, and others move out ?

- Therefore, two analyses are made:
  1) consisting of all staff appointed previously, that left/retired, etc.
     - Output of the institute alone, nothing more

  2) consisting of staff that is currently appointed
     - Output from elsewhere as well

CWTS

Mobility analysis and h-index values

	H-index PastPerf	n_cits	H-index ResPot	n_cits
Psy Dept A	22	23	25	26
Psy Dept B	34	34	35	34
Psy Dept C	42	42	42	42
Psy Dept D	43	44	47	48
Psy Dept E	58	58	55	56
Psy Dept F	43	43	44	44
Psy Dept G	48	49	50	50
Psy Dept H	45	45	47	48
Psy Dept I	70	71	74	74
All	101	101	103	104

… but how to interpret the h-index values for a department, against the national score ?

CWTS

Problems with H-index

- **Bibliometric-mathematical issues**

  - mathematically inconsistent (Waltman & van Eck, 2012)

  - conservative

  - Not field normalized (van Leeuwen, 2008)

- **Bibliometric-methodological issues**

  - How to define an author?

  - In which bibliographic/metric environment?

- **Conceptual/general issues**

  - Favors age, experience, and high productivity (Costas & Bordons, 2006)

  - No relationship with research quality

  - Ignores other elements of scholarly activity

  - Promotes one-indicator thinking

CWTS

CWTS methodology: basic indicators

Indicators suitable for assessment (1)

p: the number of publications of a unit, in a certain period.

tcs: The total number of citations received in a certain period.

mcs: the mean citation score of the oeuvre of a unit.

% not cited: the share of that oeuvre that is not cited.

% self citations: the share of citations given by the (co-)authors.

Indicators suitable for assessment (2)

mncs: the comparison of the actual impact with expected field average impact scores.

mnjs: comparison of the journals in which the unit published, with the field average impact in which the output was published.

internal coverage: indicates relevance of the bibliometric analysis, based on reference behavior of units themselves.

Top 10%: The share of the output that belongs to the top 10% most highly cited in the fields the unit is active in.

CWTS

Various additional types of analysis focus on …

- *Research profiles:* a break down of the output over various fields of science.

- *Scientific cooperation analysis:* a break down of the output over various types of scientific collaboration.

- *Knowledge user analysis:* a break down of the 'responding' output into citing fields, countries or institutions.

- *Network analysis:* how is the network of partners composed, based on scientific cooperation?

CWTS

	p	tcs	mcs	% not cited	% selfcits
Psy Dept A	298,75	1933,75	6,47	18%	29%
Psy Dept B	608,25	4867,25	8,00	13%	23%
Psy Dept C	1164,50	9448,50	8,11	15%	23%
Psy Dept D	1245,00	11761,00	9,45	13%	22%
Psy Dept E	1260,50	15009,75	11,91	11%	21%
Psy Dept F	1335,50	10163,50	7,61	15%	26%
Psy Dept G	1543,00	13556,25	8,79	14%	23%
Psy Dept H	1563,25	12970,75	8,30	15%	23%
Psy Dept I	1614,50	20913,75	12,95	13%	26%

CWTS

	mncs	mnjs	Internal coverage	% collab	% int collab
Psy Dept A	1,12	1,05	76%	86%	42%
Psy Dept B	1,44	1,24	80%	79%	46%
Psy Dept C	1,37	1,28	79%	72%	42%
Psy Dept D	1,30	1,25	85%	78%	36%
Psy Dept E	1,64	1,44	85%	78%	49%
Psy Dept F	1,24	1,25	79%	83%	42%
Psy Dept G	1,33	1,24	84%	80%	53%
Psy Dept H	1,40	1,28	80%	79%	41%
Psy Dept I	1,84	1,69	84%	86%	52%

CWTS

National analysis of academic medical centers

Annual monitoring of research performance of Dutch university medical centers (UMCs)

- Integration of medical faculty with the academic hospital

- Analysis on internal structure, combined with a national perspective.

- National comparison is standard, local analysis is custom made

- Data delivery by own formats

CWTS

The landscape of Dutch UMC's

VUmc

22.405 publications in 98-14

UMC Groningen

21.833 Publications in 98-14

AMC

31.335 publications in 98-14

LUMC

23.572 publications in 98-14

UMC Utrecht

24.724 publications in 98-14

Erasmus MC

32.338 publications in 98-14

UMC Radboud

24.826 publications in 98-14

UMC Maastricht

22.548 publications in 98-14

CWTS
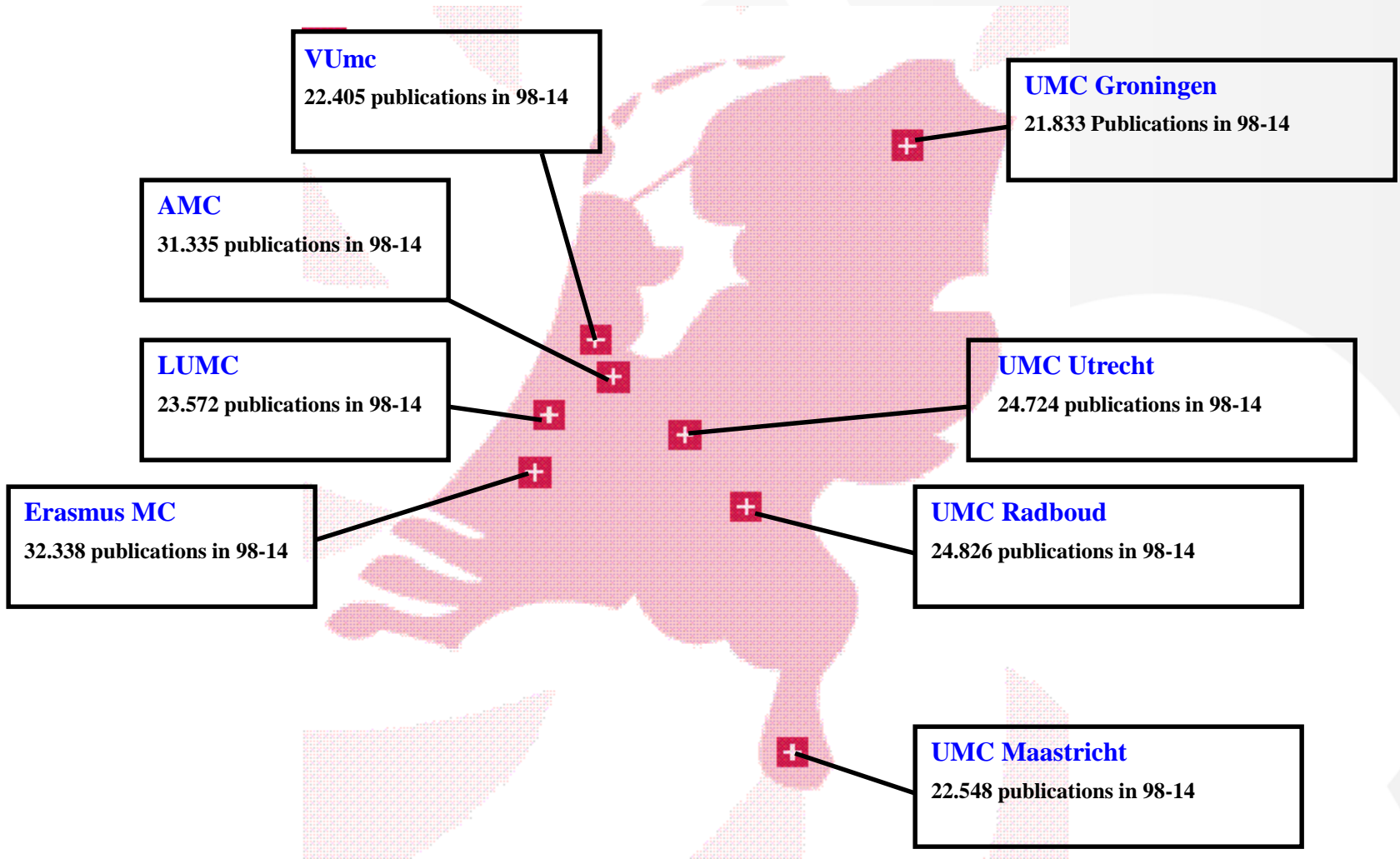
Overall tables and trend analysis (1998-2014/2015)

	p	tcs	mcs	mncs	mnjs	pp_top_perc	pp_uncited	prop_self_cits	int_cov
Erasmus MC	32338	1052533	32,55	1,65	1,42	18%	5%	16%	89%
LU MC	23572	724565	30,74	1,52	1,38	17%	5%	17%	92%
Radboud UMC	24826	655694	26,41	1,47	1,33	16%	5%	17%	90%
UMC Maastricht	22548	662294	29,37	1,54	1,28	16%	5%	15%	87%
UMCG	21833	534729	24,49	1,44	1,36	16%	6%	17%	90%
UU UMC	24724	765568	30,96	1,59	1,43	18%	5%	15%	91%
UvA AMC	31335	868131	27,70	1,51	1,36	17%	6%	16%	90%
VUmc	22405	689691	30,78	1,66	1,36	19%	6%	16%	89%

CWTS

Landscaping: mapping the situation for the UMCs

- Showing positions of UMCs, combining output and impact (mncs) and journal impact (mnjs)

  – Overall

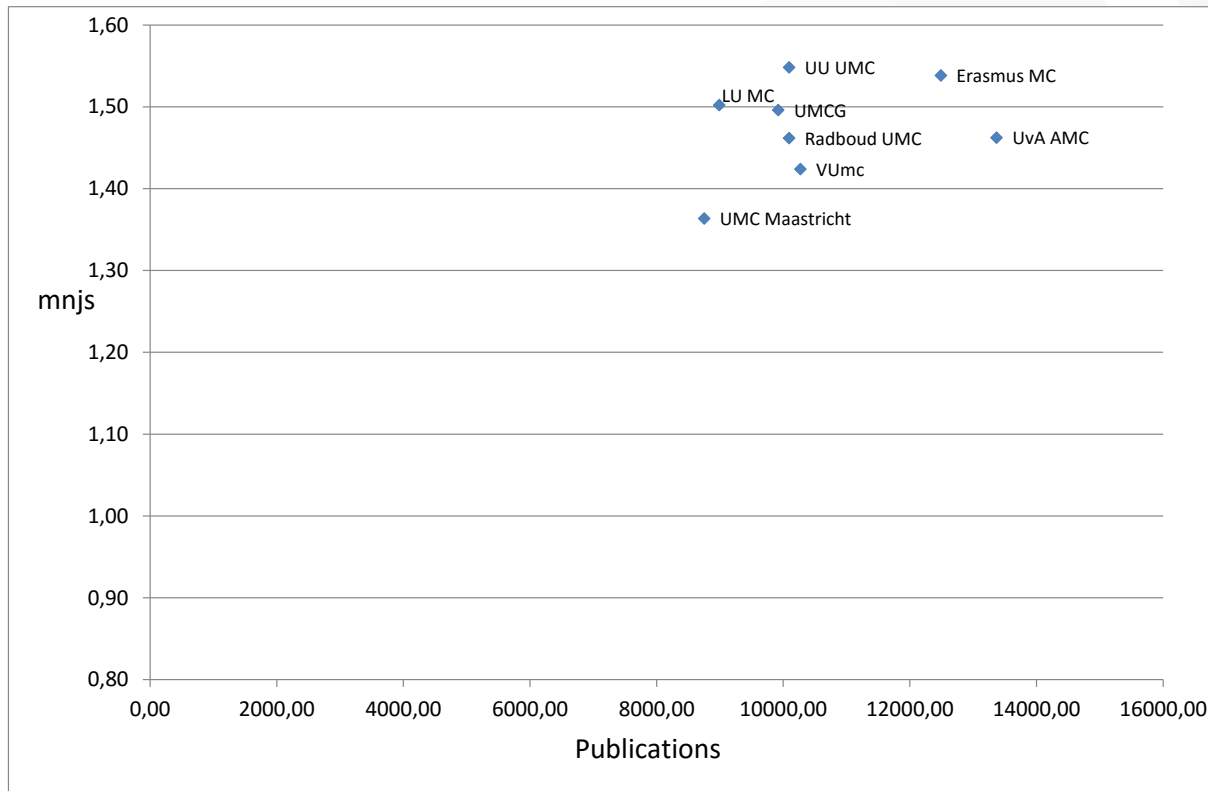 – Scientific cooperation analysis

  – Academic leadership

CWTS

Landscaping: Overall output and field impact (2010-2014/15)
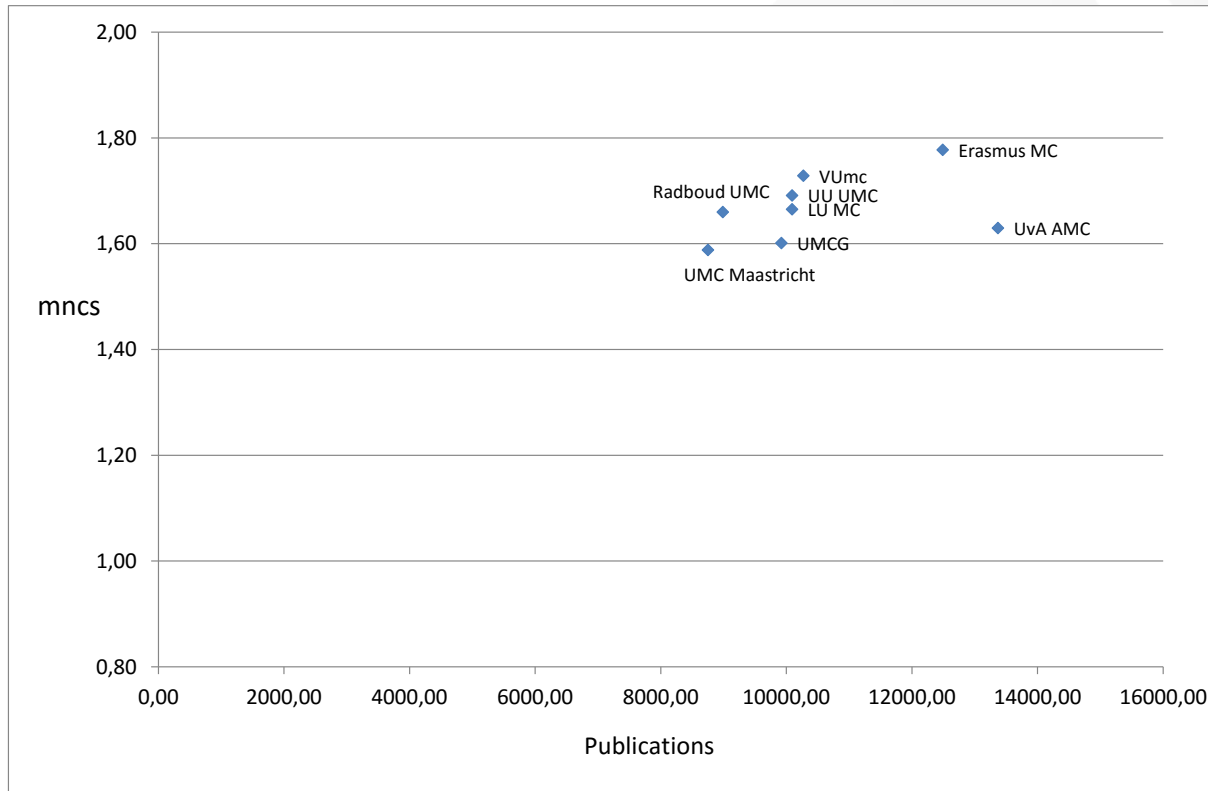
Conclusions:

- 6 produce between 8.000-10.000 papers, 2 stand out

- Impact varies between 60-80% above world average
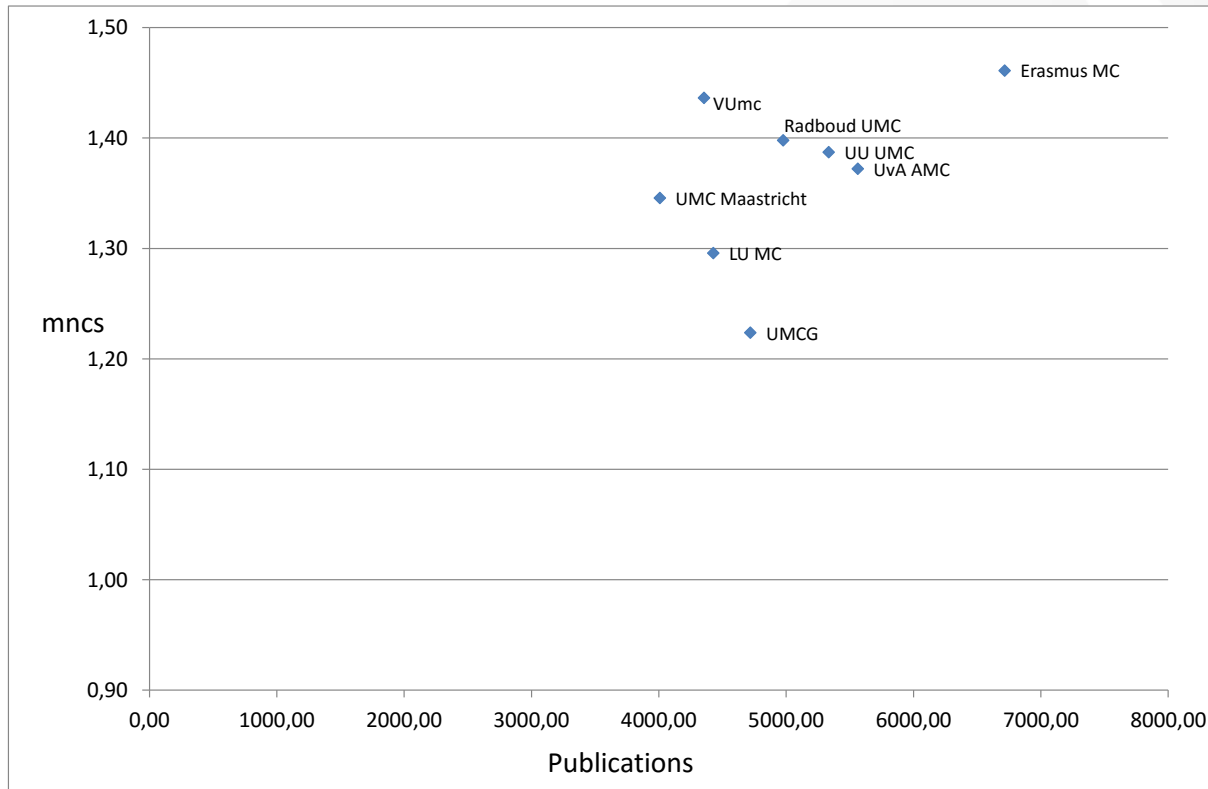
- 2 behave 'counter intuitive'!

Landscaping: Overall output and journal impact (2010-2014/15)

- Conclusions:

  – Choice for high impact journals

  – Positions of journals varies between 40-55% above field average

  – 3 publish in top journals

Landscaping: Overall output and field impact, first authorships, (2010-2014/15)

- Conclusions:
  - Output and impact decreases
  - Partial dependence on contributions from elsewhere
 - Still a strong position, far above world field average

Landscaping: Overall output and field impact, single institute output, (2010-2014/15)
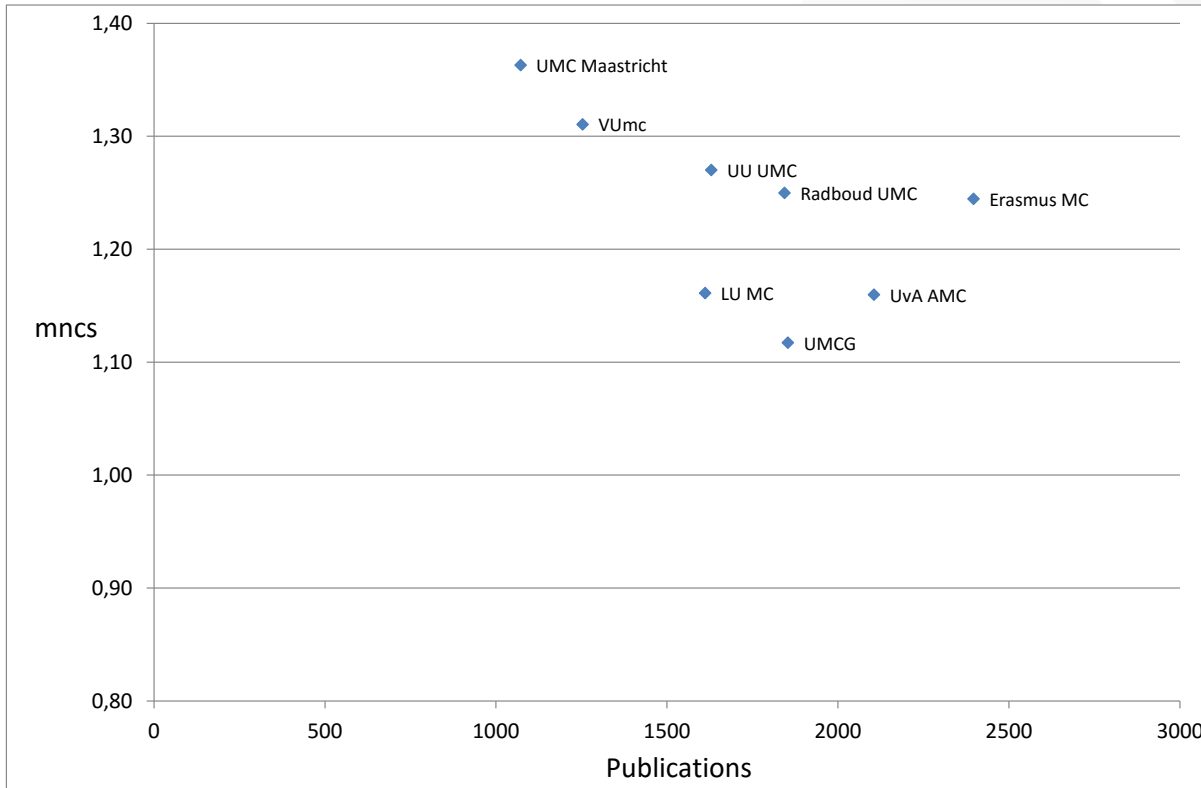
- Conclusions:
  - Small part of the output of UMCs
  - Academic leadership is visible, as this indicates the strength in the house
  - Impact is still above world average impact level

Landscaping: Overall output and field impact, international collaboration, (2010-2014/15)
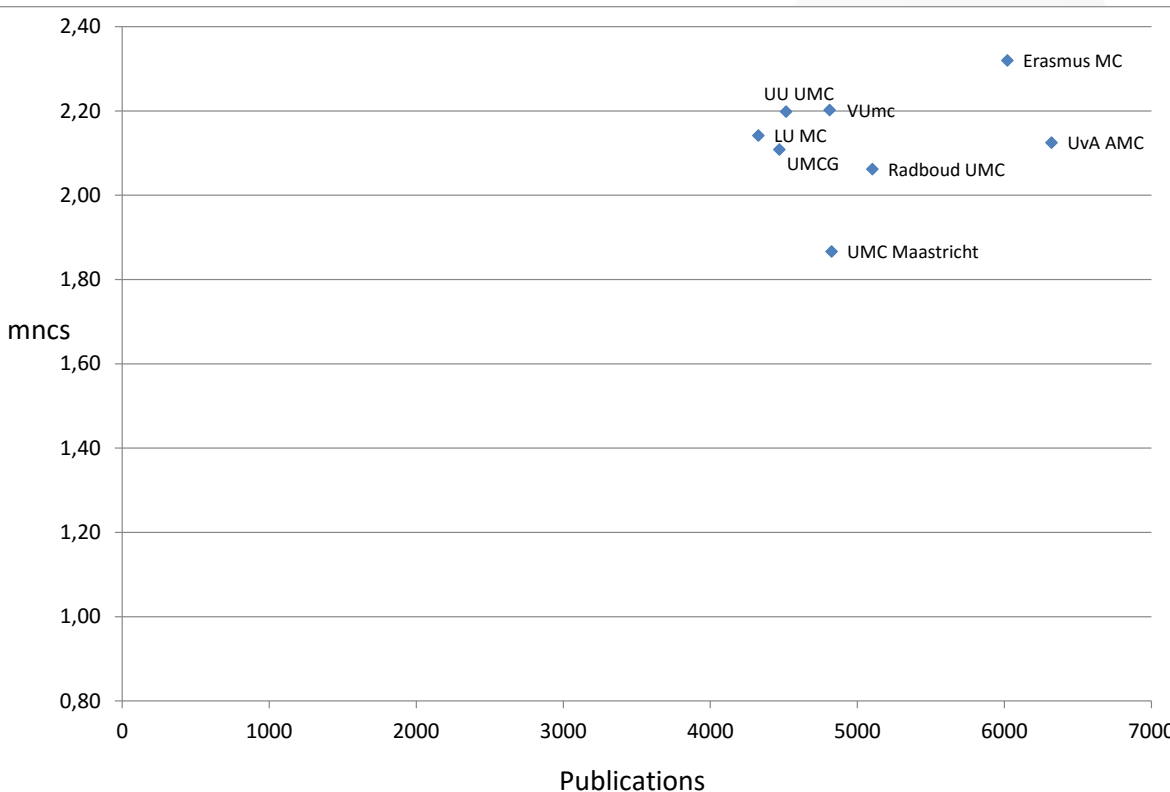
- Conclusions:
  - Large parts of the output result from international cooperation
  - Impact levels are very high
  - Dutch UMCs are attractive partners !

Research profiles, output and impact displayed

- Based upon output distribution over fields (WoS JSCs).

- Impact indicators are mncs and mnjs.

- We now apply WoS JSCs for normalization

CWTS

On normalization in bibliometric analysis

- The use of normalization is **conditio sine qua non** in applying bibliometric techniques.

- The most used system is the one based upon Web of Science Journal Subject Categories, which fits the multidisciplinary nature of the Web of Science.

- However, this most applied system, that of Journal Subject Categories, has serious drawbacks *

* Van Eck, N.J., et al (2013). Citation analysis may severely underestimate the impact of clinical research as compared to basic research. PLoS ONE, 8(4), e62395. arXiv:1210.0442

46

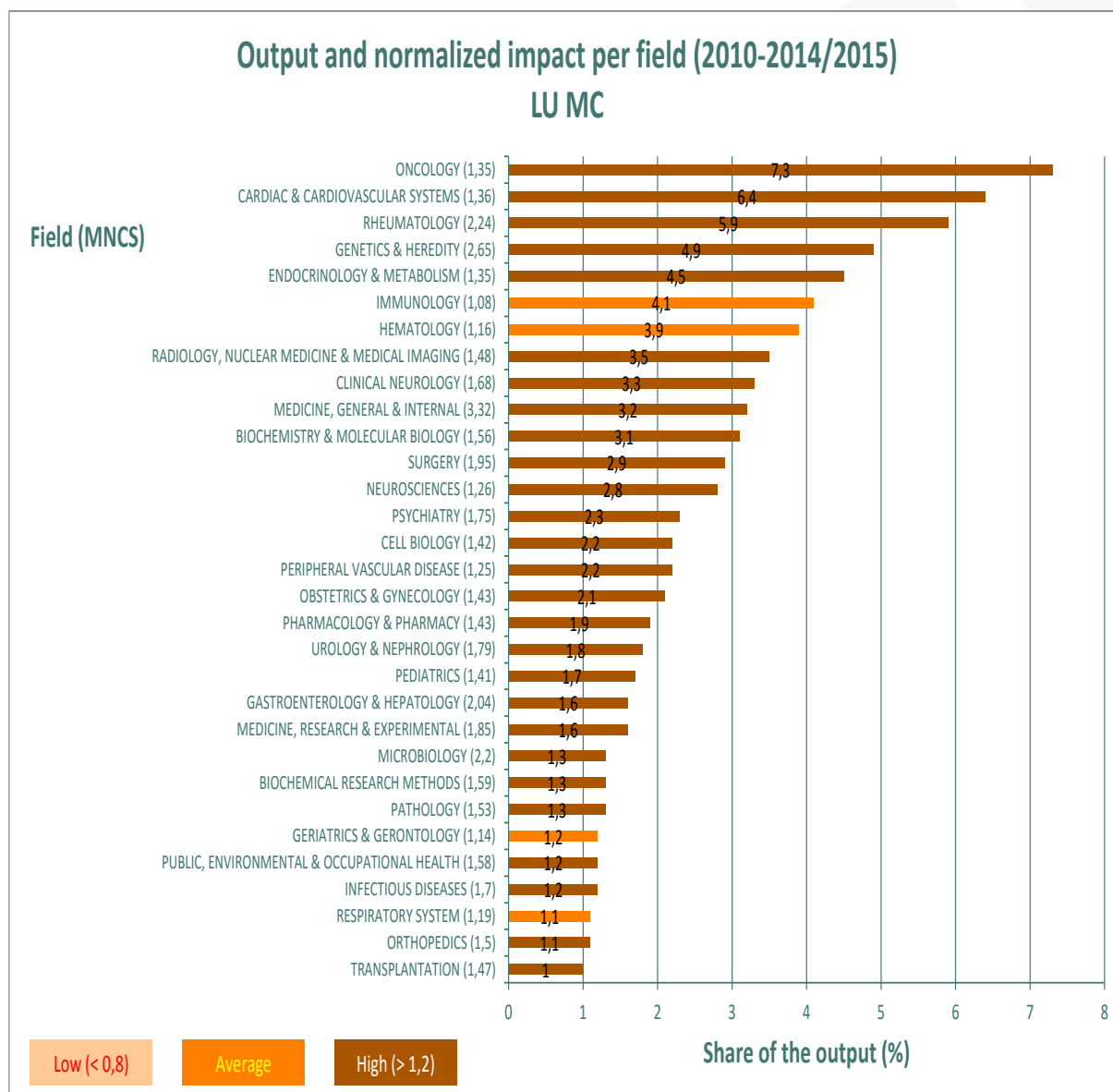# Journal Subject Category *"Clinical Neurology"*

VOSviewer

Some conclusions on normalization

- Therefore, CWTS has developed methods to normalize in a different way, avoiding these problems.

- Preferred is the CWTS Publication Cluster dataset.

- However, normalization and level of aggregation remain in a complex relationship.

- We have to remain aware of the other meaning of the word normalization, and avoid that this becomes a straight jacket.

Research profile focused on overall impact level

Output and normalized impact per field (2010-2014/2015)
LU MC

Field (MNCS)

Field	Share of the output (%)
ONCOLOGY (1,35)	7,3
CARDIAC & CARDIOVASCULAR SYSTEMS (1,36)	6,4
RHEUMATOLOGY (2,24)	5,9
GENETICS & HEREDITY (2,65)	4,9
ENDOCRINOLOGY & METABOLISM (1,35)	4,5
IMMUNOLOGY (1,08)	4,1
HEMATOLOGY (1,16)	3,9
RADIOLOGY, NUCLEAR MEDICINE & MEDICAL IMAGING (1,48)	3,5
CLINICAL NEUROLOGY (1,68)	3,3
MEDICINE, GENERAL & INTERNAL (3,32)	3,2
BIOCHEMISTRY & MOLECULAR BIOLOGY (1,56)	3,1
SURGERY (1,95)	2,9
NEUROSCIENCES (1,26)	2,8
PSYCHIATRY (1,75)	2,3
CELL BIOLOGY (1,42)	2,2
PERIPHERAL VASCULAR DISEASE (1,25)	2,2
OBSTETRICS & GYNECOLOGY (1,43)	2,1
PHARMACOLOGY & PHARMACY (1,43)	1,9
UROLOGY & NEPHROLOGY (1,79)	1,8
PEDIATRICS (1,41)	1,7
GASTROENTEROLOGY & HEPATOLOGY (2,04)	1,6
MEDICINE, RESEARCH & EXPERIMENTAL (1,85)	1,6
MICROBIOLOGY (2,2)	1,3
BIOCHEMICAL RESEARCH METHODS (1,59)	1,3
PATHOLOGY (1,53)	1,3
GERIATRICS & GERONTOLOGY (1,14)	1,2
PUBLIC, ENVIRONMENTAL & OCCUPATIONAL HEALTH (1,58)	1,2
INFECTIOUS DISEASES (1,7)	1,2
RESPIRATORY SYSTEM (1,19)	1,1
ORTHOPEDICS (1,5)	1,1
TRANSPLANTATION (1,47)	1

Share of the output (%)

Low (< 0,8) Average High (> 1,2)

- Conclusions
  - Easy way to view the most prolific activities
  - Output shares and impact are viewed in one glance
  - Also smaller fields (>1%) become visible

# Research profile indicating journal impact levels

Output and journal impact per field (2010-2014/2015)
LU MC
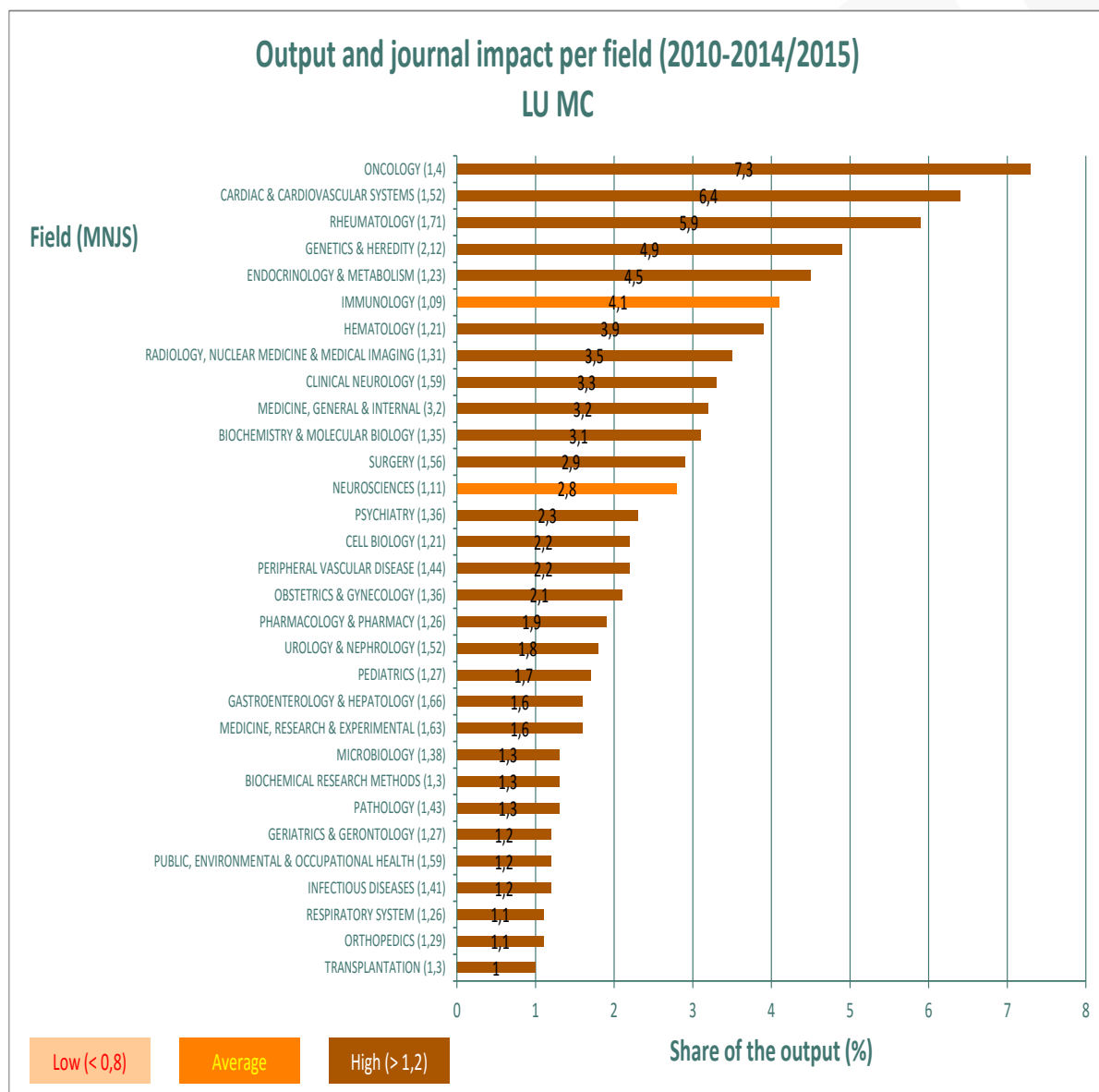
- Conclusions
  - Here journal impact is the impact indicator
  - In a glance, one observes selectivity and success in publication strategies
  - Again, also in less prolific fields

Top paper analysis, output and impact displayed

- Based upon output distribution over fields (WoS JSCs).

- Impact indicators are mncs and mnjs.

- We now apply WoS JSCs for normalization

- Preferred is the CWTS Publication Cluster dataset.

CWTS

Top paper analysis: visibility among the top x% in the field
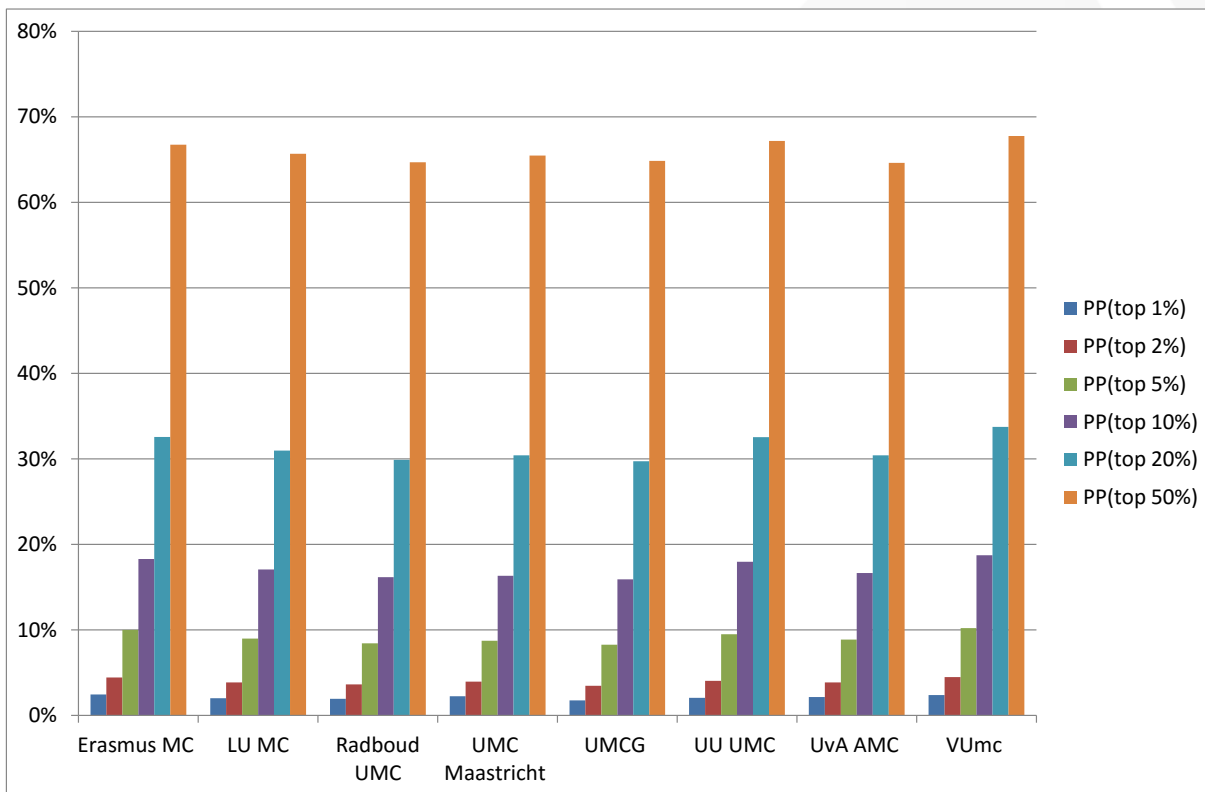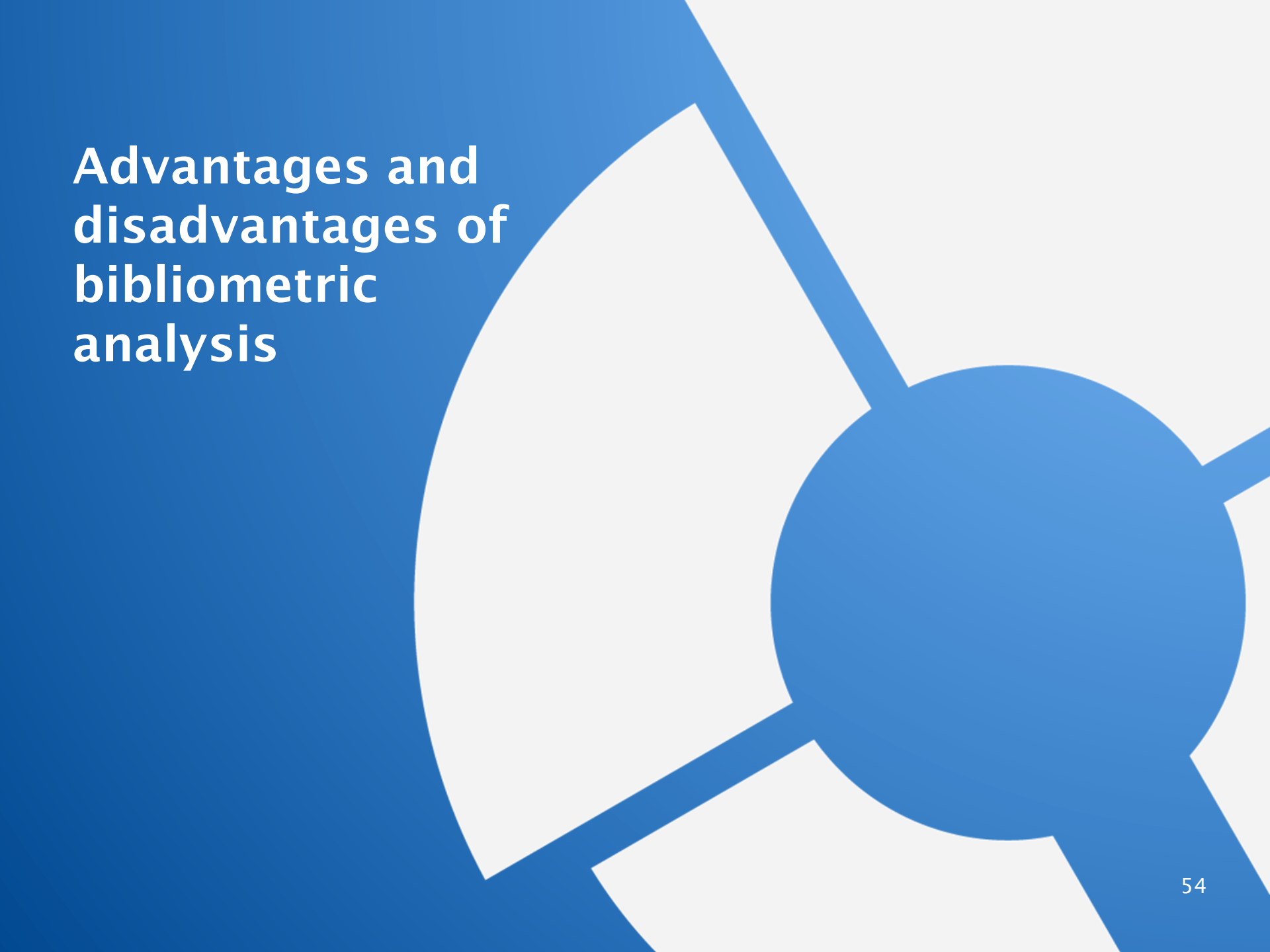
- Conclusions
  - Dutch academic medical research is very visible
  - This supports the usage of mncs as an indicator !

CWTS

Top paper analysis: activity and impact in top General Medicine journals
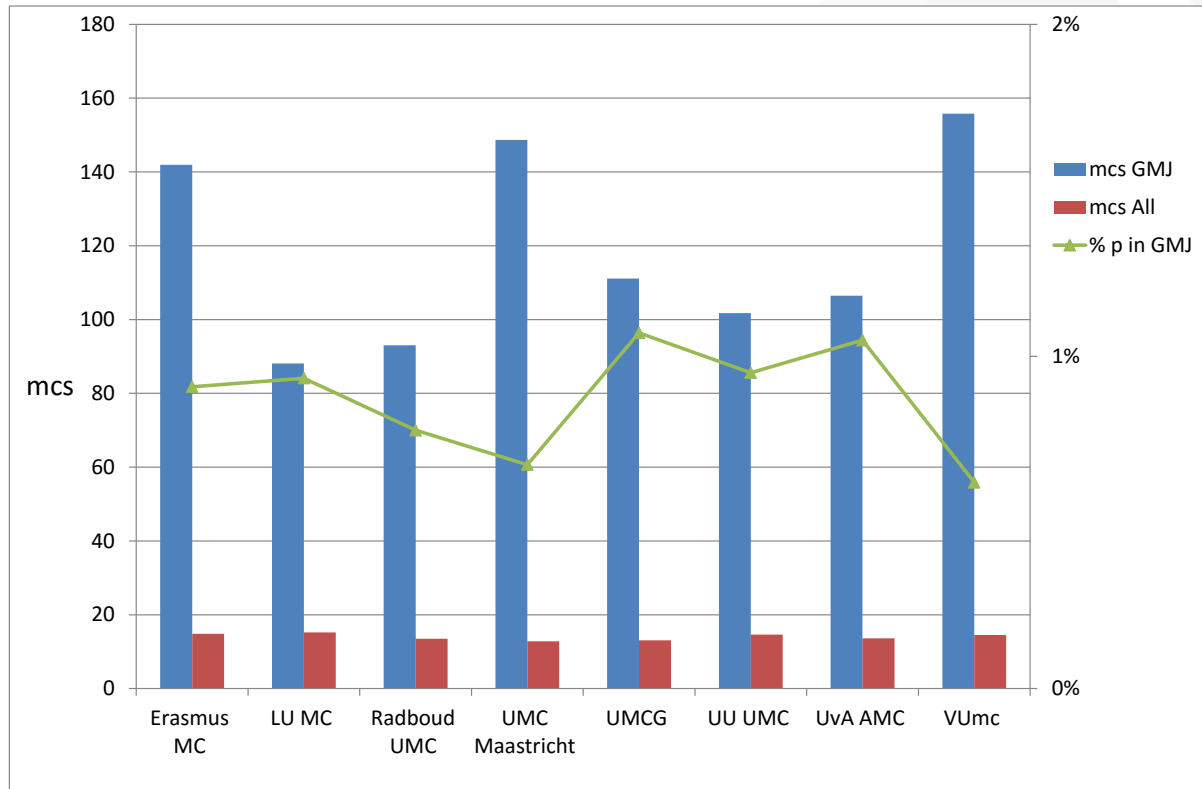
- Conclusions
  - Impact is very high
  - Here mcs is a valid indicator
  - A small output can generate a large audience

CWTS

Advantages and disadvantages of bibliometric analysis

Some disadvantages of applying bibliometrics …

- Steers away from more qualitative considerations.

- Metrics shape as much as they measure scientific activity.

- People tend to forget we are talking about *'indicators'*.

- Tends to stimulate one-dimensional thinking.

- It requires skills to calculate and interpret results.

- ….

CWTS

Some advantages of applying bibliometrics …

- Metrics tend to offer insights into underlying structures and patterns.

- Metrics tend to be a strong complementary tool to peer review.

- Metrics tend to be relatively stable in time.

- ….

CWTS

Some conclusions …

- Bibliometrics should always be combined with peer review,

- … and preferably conducted by skilled experts !

- Always contextualize the bibliometric scores !

- One better avoids the '*Quick & Dirty*' indicators !

- Advanced bibliometrics can be very helpful in research management, at various levels.

CWTS

Thank you for your attention!

Any questions?
Ask me now, or mail me
Leeuwen@cwts.nl