BIBLIOMETRIC INDICATORS – DEFINITIONS AND USAGE AT KAROLINSKA INSTITUTET

UNIVERSITY LIBRARY BIBLIOMETRIC TEAM 2014

CATHARINA REHN, DANIEL WADSKOG, CARL GORNITZKI & AGNE LARSSON

This appendix to the *Bibliometric handbook for Karolinska Institutet* lists indicators used, or considered for use, at Karolinska Institutet together with their definitions, some comments on advantages and shortcomings of the different indicators, and how/if they are implemented at Karolinska Institutet.

First, some general notes on the definitions and the calculation of indicators in the appendix:

- Inclusion or exclusion of self citations – see the handbook for more information – might affect the resulting indicator values, but not how the indicators are calculated. Self citations are therefore noted as a separate indicator, but not in the context of any of the other indicators. At Karolinska Institutet, we do not presently remove self citations when calculating our indicator values.
- Fractionalization or any other form of weighting of publications between the contributing authors – see the handbook for more information – will affect most indicators. It will, however, not affect the basic calculation principles, and, for reasons of clarity, this aspect has been left out in the indicator descriptions. At Karolinska Institutet, we do not currently use any fractionalization or weighting when calculating our indicator values.
- The validity of several of the indicators improves if the authors themselves validate or supply information about their publications before the indicator values are calculated. If the analysis is done on anything below university level it is particularly important.
- CWTS indicators and denotations are included in this indicator definition list where appropriate, since these are well known in the bibliometric community.

Note: The word *unit* is here to be interpreted as “unit of analysis”, unless in the context of “research unit”.

TABLE OF CONTENTS

DENOTATION INDEX

P	Total number of publications
P_{WoS}	Number of publications in Thomson Reuters indices
P_{TJ}	Number of publications in top journals
$P_{f5\%}$	Number of articles among the top 5% most cited in the field, of the same age and article type
p	Relative share of publications
$p_{f5\%}$	Top 5% – share of articles among top 5% most cited in the field, of the same type and age
p_u	Uncitedness – share of uncited publications
p_x	Co-authoring – share of publications co-authored with another unit
p_f	A unit's world share of publications in field f
p_w	A unit's world share of publications in all fields
C	Total number of citations
c_i	Number of citations to a single publication *i*
$\bar{c}$	Average number of citations per publication
c_s	Self citedness – share of citations from the own unit
F_{ij}	The cumulative distribution function for subject area ij
μ_f	Field reference value (field citation score) for articles of the same type, age and in the same field of research
$\bar{\mu}_f$	Mean field reference value (mean field citation score)
μ_j	Journal reference value
RIP	Raw citations per paper; i.e. the average number of times the journal's publications in the three preceding years were cited in the year of analysis
DCP	Database citation potential; i.e. a harmonic average number of active references in the publications belonging to the journals subject field

ABBREVIATIONS

CWTS	Center for Science and Technology Studies, Leiden University
WoS	Web of Science® prepared by Thomson Reuters®, Inc.

1. PUBLICATIONS

1.1 NUMBER OF PUBLICATIONS

Designation	Total number of publications
Denotation	P
Description	The number of scientific publications produced by the analyzed unit during the analyzed time span. Sometimes results are also presented separately per document type.
Calculation	Count the full number of scientific publications produced at the analyzed unit during the analyzed time span.
Formula	-
Data Requirements	Publication data from a local publication database or from an external database (such as the Thomson Reuters indices) complemented by self-reported publications from the analyzed unit.
Advantages	Relatively easy to produce.
Disadvantages	Does not take the size of the analyzed unit into account and does not say anything about the impact of the publications.
KI usage	At Karolinska Institutet we currently give every contributing unit full credit for the publication, i.e. no fractionalization or weighting between authors or institutions is used.
Reference	-

1.2 NUMBER OF WOS PUBLICATIONS

Designation	Number of publications in Thomson Reuters indices
Denotation	P_{WoS}
Description	The number of scientific publications produced by the analyzed unit during the analyzed time span, found in the Thomson Reuters indices.
Calculation	Count the full number of publications in Thomson Reuters indices, produced at the analyzed unit during the analyzed time span.
Formula	-
Data Requirements	Publication data from Thomson Reuters indices.
Advantages	Easy to retrieve from the Thomson Reuters Web of Science.
Disadvantages	Does not take the size of the analyzed unit into account and does not say anything about the impact of the publications. Does not take into account publications not present in the Thomson

	Reuters indices.
KI usage	At Karolinska Institutet, publications where the analyzed unit is only a part contributor is fully accounted to the unit, i.e. no fractionalization or weighting between authors is currently used.
Reference	-

1.3 NUMBER OF PUBLICATIONS IN TOP JOURNALS

Designation	Number of publications in top-ranked journals
Denotation	P_{TJ}
Description	The number of publications the analyzed unit has published in a selected number of journals during the analyzed time span.
Calculation	Select journals according to a suitable criterion. Check how many of the unit's publications that are published in these journals during the analyzed time span.
Formula	-
Data Requirements	A bibliographic database (for instance Thomson Reuters Web of Science or a local publication database) to count publications and addition of publications not present in the database.
Advantages	Does reflect the potential impact of the unit's articles more than a mere publication count.
Disadvantages	Does not take the size of the analyzed unit into account.
KI usage	At Karolinska Institutet, journals classified as being focused on other subjects than life sciences are sometimes excluded from the journal list to make the indicator more relevant for assessments of life science research. No fractionalization or weighting between authors is currently used.
Reference	-

1.6 RELATIVE ACTIVITY INDEX

Designation	Relative activity index
Denotation	RAI
Description	The relative effort a unit of analysis devotes to a specific field measured in publications.
Calculation	The analysed unit's world share of publications in a given field divided by the unit's world share of publications overall.
Formula	$$\text{RAI} = \frac{WS_f}{\text{WS}}$$ p_f = The unit's world share of publications in a given field p_w = The unit's world share of publications in all fields
Data Requirements	Requires data from a comprehensive bibliographic database such as the Thomson Reuters citation indices.
Advantages	-
Disadvantages	The indicator is not normalized with regard to document type or publication year. The classification used for domains and subdomains is the journal classification scheme supplied by Thomson Reuters.
KI usage	At Karolinska Institutet this indicator is not used at present.
Reference	Frame, J. D (1977). Mainstream research in Latin America and the Caribbean. Interciencia, 2, 143. Read more about the method: http://link.springer.com/article/10.1007/BF02017249

1.5 RELATIVE SPECIALIZATION INDEX

Designation	Relative Specialization index
Denotation	RSI
Description	Indicates how active an analysed unit is in a certain field. A value of -1 indicates a completely idle research field and a value of 1 if all publications from the unit are in one field.
Calculation	Divide the (activity index -1) with the (activity index +1).
Formula	$$RSI = \frac{RAI - 1}{RAI + 1}$$ RAI = Relative activity index, defined above
Data Requirements	Requires data from a comprehensive bibliographic database such as the Thomson Reuters citation indices.
Advantages	
Disadvantages	This indicator is not normalized with regard to document type or publication year. The classification used for domains and subdomains is the journal classification scheme supplied by Thomson Reuters.
KI usage	At Karolinska Institutet this indicator is not used at present.
Reference	Schubert, A., Braun T. (1986). Relative indicators and relational charts for comparative assessment of publication output and citation impact. Scientometrics, 9(5–6), 281–291 http://link.springer.com/article/10.1007/BF02017249

2. CITATIONS

2.1 NUMBER OF CITATIONS

Designation	Total number of citations
Denotation	C
Description	The total number of citations to articles published by an analyzed unit during the analyzed time span.
Calculation	Find all articles published by the analyzed unit during the analyzed time span and sum their citation values (usually retrieved from the Thomson Reuters indices via Web of Science or another source to the indices).
Formula	$C = \sum_{i=1}^{P} c_i$ where: P = number of publications c_i = number of citations for publication i
Data Requirements	Requires data from a comprehensive citation database such as the Thomson Reuters citation indices.
Advantages	Gives an indication of the scientific impact the unit's published articles as a whole.
Disadvantages	Does not take into account that older articles usually are more cited and that citation rates vary between document types and subject areas. Does not compensate for the size of the unit.
KI usage	At Karolinska Institutet, publications where the analyzed unit is only a part contributor of the publication is fully accounted to the unit, and thus also the corresponding citation count.
Reference	-

2.2 CITATIONS PER PUBLICATION

Designation	Average number of citations per publication
Denotation	Karolinska Institutet: $\bar{c}$ CWTS: CPP
Description	The average number of citations to articles published by an analyzed unit during the analyzed time span.
Calculation	Find all articles published by the analyzed unit during the analyzed time span in a citation index, sum up the citations and divide by the number of publications.

Formula	$\bar{c} = \frac{1}{P}\sum_{i=1}^{P} c_i$ where: c_i = number of citations for publication i P = number of publications
Data Requirements	A comprehensive citation index as Thomson Reuters citation indices.
Advantages	Gives an indication of the average scientific impact of the unit's published articles.
Disadvantages	Does not take into account that older articles usually are more cited if a variable, cumulative citation time window is used, and that citation rates vary between document types and subject areas.
KI usage	At Karolinska Institutet, citations to publications where the analyzed unit is only a part contributor of the publication is fully accounted to the unit. $\bar{c}$ is usually used as background information and complement to other indicators.
Reference	-

2.3 FIELD NORMALIZED CITATION SCORE

Designation	Item oriented field normalized citation score average
Denotation	Karolinska Institutet: $\bar{c}_f$ CWTS: MNCS
Description	This indicator corresponds to the relative number of citations to publications from a specific unit, compared to the world average of citations to publications of the same document type, age and subject area. As an example, 0.9 means that a unit's publications are cited 10% below average and 1.2 that they are cited 20% above average. The term "item oriented" indicates that the normalization of the citation values is done on an individual article level. (CWTS previously used a version called "crown indicator"* which closely resembles this one.)
Calculation	The number of citations to each of the unit's publications is normalized by dividing it with the world average of citations to publications of the same document type, publication year and subject area, which is called the field reference value (μ_f). The indicator is the mean value of all the normalized citation scores for the unit's publications.

Formula	$\bar{c}_f = \frac{1}{P}\sum_{i=1}^{P}\frac{1}{N_i}\sum_{j=1}^{N_i}\frac{c_i}{\left[\bar{\mu}_f\right]_{ij}}$ where: c_i = number of citations to publication i $\left[\bar{\mu}_f\right]_{ij}$ = the average value of citations to publications of the same type, published the same year as article i in the subject area j P = the unit's number of publications N_i = the number of subject areas that the publication i belongs to
Data Requirements	Requires data from a comprehensive citation database such as the Thomson Reuters citation indices.
Advantages	As the normalization takes place on the level of the individual publication the indicator gives each publication equal weight in the final value.
Disadvantages	If the normalization is done on an article level, a few highly cited articles in a moderately cited research area may contribute unproportionately to the value of the field normalized citation score.
KI usage	At Karolinska Institutet, the item oriented field normalized citation score average is only used for units of publications greater than 50 and where the publication year <= the year of analysis-2. Citations to publications where the analyzed unit is only a part contributor of the publication is fully accounted to the unit.
References	$\bar{c}_f$: Lundberg, J. (2007). Lifting the crown – citation *z*-score. Journal of Informetrics, 1(2), 145–154. MNCS: http://www.sciencedirect.com/science/article/pii/S1751157710000817 * Moed, H. F., Debruin, R. E., & Vanleeuwen, T. N. (1995). New bibliometric tools for the assessment of national research performance: database description, overview of indicators and first applications. Scientometrics, 33(3), 381–422.

2.4 TOTAL FIELD NORMALIZED CITATION SCORE

Designation	Total item oriented field normalized citation score
Denotation	Karolinska Institutet: C_f, sum cf CWTS: TNCS (formerly Brute force)
Description	This indicator gives an indication of the combined impact of the production volume and field normalized citation score of the analyzed unit. It can be interpreted as the number of "normal cited" publications required to achieve the Cf value.
Calculation	Add together the item oriented field normalized citation scores for all the publications of the analyzed unit.
Formula	$c_f = \sum_{i=1}^{P} \left[c_f \right]_i$ where: $\left[c_f \right]_i$ = item oriented field normalized citation score for publication i P = total number of publications for the analyzed unit
Data Requirements	Requires data from a comprehensive citation database such as the Thomson Reuters citation indices.
Advantages	Gives and indication on both the volume and the impact of the publications from the analyzed unit. Assesses the total impact on the scientific community.
Disadvantages	Does not compensate for the size of the analyzed unit.
KI usage	Cf is usually used when comparing units of similar size such as two different universities.
Reference	Karolinska Institutet: Lundberg, J. (2007). Lifting the crown – citation *z*-score. Journal of Informetrics, 1(2), 145–154. TNCS: Waltman, L., van Eck, J. N., van Leeuwen, T. N., Visser, M. S., van Raan, A. F. J. (2011). Towards a new crown indicator: some theoretical considerations. Journal of Informetrics. 5(1), 37–47. http://www.sciencedirect.com/science/article/pii/S1751157710000817

2.5 LOGARITHM-BASED CITATION Z-SCORE

Designation	Item oriented field normalized logarithm-based citation z-score average
Denotation	$\bar{c}_{fz[\ln]}$
Description	The logarithm-based citation z-score relates the logarithm of the number of citations that a publication has received with to the mean and the standard deviation for the logarithms of the

	citation rates for all the corresponding reference publications of the same type, age and subject area.
Calculation	The average of the logarithms of the number of citations (plus 1 to avoid the value 0) to publications of the same document type, publication year and subject area, which is called the logarithm-based field reference value ($\mu_{f[\ln]}$), is subtracted from the logarithms of the citation counts (plus 1) for each article produced by the analyzed unit during the analyzed time span. If an article is classified as belonging to several subject areas, a mean value of the areas is used as $\mu_{f[\ln]}$. The resulting value is then divided by the standard deviation for the logarithm of the citation count plus one of the population of articles that constitutes the logarithm-based field reference value. Finally, the mean value of all values calculated as mentioned above is calculated by dividing the values with the number of analyzed publications, and this gives the logarithm-based citation z-score indicator for the unit.
Formula	$\bar{c}_{fz[\ln]} = \frac{1}{P}\sum_{i=1}^{P}\frac{\ln(c_i+1)-\lfloor\mu_{f[\ln]}\rfloor_i}{\lceil\sigma_{f[\ln]}\rceil_i}$ where: c_i = number of citations to publication i $[\mu_{f[\ln]}]_i$ = the logarithm-based field reference value; the average value of the logarithms of the number of citations plus one to publications of the same type, published the same year in the same research area as article i $[\sigma_{f[\ln]}]_i$ = the standard deviation of the $[\mu_{f[\ln]}]_i$ distribution P = the unit's number of publications
Data Requirements	Requires data from a comprehensive citation database such as the Thomson Reuters citation indices and verification of articles as belonging to the analyzed unit.
Advantages	Since the distribution of citation rates differs between research fields, publication years and document types it can be argued that using a z-score in the normalization procedure would be more appropriate than a field normalized citation score. The z-score indicator gives information both if the citation value of the publication is lower (negative z-score) or higher (positive z-score) than the field score, and how far from the mean the value is, measured in a normalized way by using the standard deviation for the field citation score as a measuring unit.
Disadvantages	A complex calculation and therefore not easily explained to a client.
KI usage	This indicator is rarely used at Karolinska Institutet.

Reference	Lundberg, J. (2007). Lifting the crown – citation *z*-score. Journal of Informetrics, 1(2), 145–154.

2.7 AVERAGE PERCENTILE

Designation	Average field normalized citation percentile
Denotation	$\bar{c}_{perc}$
Description	The average of shares of publications within the same types, ages and subject areas that have fewer citations than the analyzed publications.
Calculation	For each publication in the analysed unit, the percentile rank, $F(c_i\text{-}1)$, for the number of received citations is calculated for c_i as the rank position among publications within the same type, age and subject area. The rank position is thus the number of publications that have a lower citation score than c_i. The rank position for each publication is then divided with the total number of publications within the same type, age and subject area to get $c_{perc}(i)$. The average percentile for an analysed unit is the average of the individual $c_{perc}(i)$.
Formula	$$\bar{c}_{perc} = \frac{1}{P}\sum_{i=1}^{P}\frac{1}{N_i}\sum_{j=1}^{N_i}\frac{F_{ij}(c_i-1)}{M_{ij}}$$ P = the unit's number of publications c_i = The number of citations to publication i N_i = the number of subject areas that the publication i belongs to M_{ij} = The number of publications within the same type, age and subject areas as publication i where j denotes the specific subject area. F_{ij} = The cumulative distribution function for subject area ij
Data Requirements	Requires data from a comprehensive citation database such as the Thomson Reuters citation indices.
Advantages	Compared to $\bar{c}_f$, the c_{perc} is less sensitive to outliers and extreme citation scores.
Disadvantages	The distribution of c_{perc} is less affected by actual differences in citation rates than $\bar{c}_f$.
KI usage	At Karolinska Institutet we usually use the percentile as a complement to $\bar{c}_f$.
Reference	Different variants of defining the percentile in bibliometrics: Bornmann, L., Leydesdorff, L., Mutz, R. (2013). The use of percentiles and percentile rank classes in the analysis of bibliometric data: opportunities and limits. Journal of

	Informetrics, 7(1), 158–165. http://www.sciencedirect.com/science/article/pii/S1751157712000831#

2.7 TOP 5%

Designation	Item oriented field normalized top 5%
Denotation	Karolinska Institutet: $p_{f5\%}$ CWTS: PPtop (Ptop, or sometimes HCP or A/E)
Description	Top 5% shows the share of publications attributed to a unit that belong to the 5% most highly cited publications in the world from the same year, in the same subject and of the same document type. Other top values, as top 1% and top 10% are also used, and calculated in the same way as top 5%. (Some variants of this indicator, such as HCP or A/E, is written as a decimal number that shows the relation to the world average. A value over 1 shows that the analyzed unit has more of its publications among the top 5% than the world average, a value below 1 that it has less.)
Calculation	Calculate the citation percentile for each publication and count the number of publications with a citation percentile above 0.95 (if the article is classified as belonging to several subject areas, a mean value between subject areas is calculated). The indicator can either be used as a number of highly cited articles, as described above, or as a share. To get the share, divide the number of top 5% publications with the unit's total number of publications
Formula	$p_{f5\%} = \frac{P_{f5\%}}{P}$ where: $P_{f5\%}$ = publications above citation threshold for 5% most cited for the same article type, year and field P = total number of publications for the analyzed unit during the analyzed time span
Data Requirements	Requires data from a comprehensive citation database such as the Thomson Reuters citation indices.
Advantages	Compared to $\bar{c}_f$, the $p_{f5\%}$ is less sensitive to outliers and extreme citation scores.
Disadvantages	Only uses a small share of the citation distribution.
KI usage	At Karolinska Institutet we usually use the top 5% as a complement to $\bar{c}_f$.
Reference	Lutz Bornmann, L., Anegón, F. M., Leydesdorff, L. (2012). The new Excellence Indicator in the World Report of the SCImago

	Institutions Rankings 2011. Journal of Informetrics, 6(2), 333–335. http://www.sciencedirect.com/science/article/pii/S1751157711001052

2.11 JOURNAL NORMALIZED CITATION SCORE

Designation	Item oriented journal normalized citation score average
Denotation	Karolinska Institutet: $\bar{c}_j$ CWTS: CPP/JCSm
Description	This indicator corresponds to the number of citations to publications from a specific unit during an analyzed time span, compared to the world average of citations to publications of the same document types, ages and in the same journals. As an example, 0.9 means that the unit's publications are cited 10% below average and 1.2 that they are cited 20% above average. The term "item oriented" indicates that the normalization of the citation values is done on an individual article level. A high indicator value suggests that a group is highly cited within the journals they choose to publish in.
Calculation	The number of citations to each of the unit's publications is normalized by dividing it with the world average of citations to publications of the same document type, published the same year in the same journal. The indicator is the mean value of all the normalized citation counts for the unit's publications.
Formula	$\bar{c}_j = \frac{1}{P}\sum_{i=1}^{P}\frac{c_i}{\mu_j}$ c_i = number of citations to publication i μ_j = the average number of citations to publications of the same type, published the same year and in the same journal as article i P = the unit's number of publications
Data Requirements	Requires data from a comprehensive citation database such as the Thomson Reuters citation indices.
Advantages	
Disadvantages	
KI usage	At Karolinska Institutet, this indicator is rarely used at present.
References	CWTS: Moed, H. F., Debruin, R. E., & Vanleeuwen, T. N. (1995). New bibliometric tools for the assessment of national research performance: database description, overview of indicators and first applications. Scientometrics, 33(3), 381–422.

2.13 H-INDEX

Designation	Hirsch index (h-index)
Denotation	h
Description	The h-index is the number of publications (h), attributed to the analyzed unit during the analyzed time span that have at least h citations.
Calculation	Find the unit's published articles in a citation index and sort them in descending order by number of citations. Count articles from the top of the list and downwards, and when the number of an article rises above the citation count for that very article, the number of the preceding article is to be counted as the h-index.
Formula	See Hirsch's original article, referenced below.
Data Requirements	Requires data from a comprehensive citation database such as the Thomson Reuters citation indices.
Advantages	Very easy to calculate in Web of Science.
Disadvantages	h-index gives positive bias to senior researchers with older articles, since these have had more time to be cited, though the demand that new articles with comparable citation levels has to be added has a certain damping effect on that bias. The indicator is not field normalized which makes it unsuitable for comparisons between researchers in different research fields. The value is static for researchers that have stopped publishing or no longer produce highly cited work.
KI usage	h-index is presently not used by the Karolinska Institutet bibliometrics group.
Reference	Hirsch, J. E. (2005). An index to quantify an individual's scientific research output. Proceedings of the National Academy of Sciences of the United States of America, 102(46), 16569–16572.

2.14 UNCITEDNESS

Designation	Uncitedness
Denotation	Karolinska Institutet: p_u CWTS: %Pnc
Description	The share of a unit's publications that that remain uncited after a certain time period. Self-citations should be removed from the citation count.
Calculation	Count the number of publications that have never been cited during a specified time period, excluding self-citations. Divide with the total number of publications from the same unit during the same time period.

Formula	$p_u = P_u / P$ where: P_u = the unit's number of publications which has received no citations P = the unit's total number of publications
Data Requirements	Requires data from a comprehensive citation database such as the Thomson Reuters citation indices and validation of the unit's publications.
Advantages	-
Disadvantages	-
KI usage	Seldom used at Karolinska Institutet
Reference	-

2.15 SELF CITEDNESS

Indicator	Self citedness
Denotation	Karolinska Institutet: c_s CWTS: %SELFCIT
Description	The share of a unit's received citations where authors refer to their own papers.
Calculation	Count the total number of citations to the unit's publications during the analyzed time span. Check where citations are coming from and count the number coming from the unit itself. Divide the second number with the first to get share of self citedness.
Formula	$c_s = C_S / C$ where: C_S = citation to the unit's publications emanating from the unit itself C = the total number of citations to the unit's publications
Data Requirements	Requires data from a comprehensive citation database such as the Thomson Reuters citation indices, validation of publications and analysis of citing articles, which can be done in the Web of Science.
Advantages	-
Disadvantages	-
KI usage	Seldom used at Karolinska Institutet
References	-

3. COOPERATION

3.1 CO-AUTHORING

Denomination	Share of articles co-authored with another unit
Denotation	p_x
Description	This group of indicators is used to show to what extent an analyzed unit cooperates with other units in the production of articles. International collaboration – share of publications with co-authors from organizations in at least two different countries. National collaboration – share of publications with co-authors from at least two organizations within the same country. Department collaboration – share of publications with co-authors from at least two departments within the same organization. Unit collaboration – share of publications with co-authors from two or more research units.
Calculation	Count the number of articles published by the analyzed unit during the analyzed time span and check how many that were co-authored together with a selected other unit. Divide the second number by the first one to get the share of articles co-authored between the units.
Formula	$p_x = P_x / P$ where: p_x = share of publications co-authored with a certain unit P_x = number of publications co-authored with the selected unit P = total number of publications produced at the analyzed unit during the analyzed time
Usage	
Data Requirements	Full addresses to all participating units.
Advantages	
Disadvantages	
KI Usage	Often used to assess Karolinska Institutet's degree of copublication in different geographical areas.
Reference	

4. JOURNALS

4.1 THOMSON REUTERS JOURNAL IMPACT FACTOR

Designation	Thomson Reuters Impact Factor
Denotation	Karolinska Institutet: I_{WoS}, JIF CWTS: IF
Usage	Used to measure the impact of scientific journals.
Description	The impact factor is a number that corresponds to the average number of citations a publication in a specific journal has received during the two years following the year of publication.
Calculation	The impact factor for a specific journal (J), one specific year (Y) is calculated by counting the number of citations to articles in that journal the two preceding years (Y-1 and Y-2) from publications in year Y and dividing this with the number of publications defined by Thomson Reuters as "citeable" in journal J the two preceding years (Y-1 and Y-2).
Formula	$I_{WoS} = C / P$ where: I_{WoS} = the impact factor for journal J in year Y C = the number of citations from publications in year Y to publications in journal J published Y-2 and Y-1 P = total number of citeable publications in journal J in year Y-2 and Y-1
Data Requirements	No own data is required; Thomson Reuters journal impact factor is available through the service Journal Citation Reports.
KI Usage	The JIF is a regular part of analyses at Karolinska Institutet since this indicator is well known within the Medical scientific community.
Reference	The Thomson Reuters Impact Factor: http://wokinfo.com/essays/impact-factor/

4.2 NORMALIZED JOURNAL IMPACT

Designation	Normalized journal impact
Denotation	$\bar{c}_f$ or J_{cf}
Description	This indicator corresponds to the relative number of citations to publications in one specific journal, compared to the world average of citations to publications of the same document type, age and subject area. As an example, 0.9 means that publications in this journal are cited 10% below average and 1.2 that they are cited 20% above average.

Calculation	The number of citations to each of the journal's publications is normalized by dividing it with the world average of citations to publications of the same document type, publication year and subject area, which is called the field citation score (μ_i). If an article is classified as belonging to several subject areas, the mean value of the field citation scores is used. The indicator is the mean value of all the normalized citation counts for publications in this journal.
Formula	$\bar{c}_f = \frac{1}{P}\sum_{i=1}^{P}\left[\frac{c_i}{\bar{\mu}_f}\right]_i$ c_i = number of citations to publication i $\left[\bar{\mu}_f\right]_i$ = the average value of citations to publications of the same type, published the same year in the same research area as article i P = the number of publications in the journal during the selected time period
Data Requirements	Requires data from a comprehensive citation database such as the Thomson Reuters citation indices and calculation of field normalized citation scores for normalization of citation values.
Advantages	Compared to JIF, Jcf can be used when comparing journals across different fields of research.
Disadvantages	
KI usage	At Karolinska Institutet we use publications from the publication year and the two preceeding years (Y-1 and Y-2) in order to mimic the data set used in the calculation of IF.
Reference	

4.3 SOURCE NORMALIZED IMPACT PER PAPER

Designation	Source normalized impact per paper
Denotations	SNIP
Usage	Used to measure the relative impact of scientific journals.
Description	The incoming citations to a journal is normalized based on the characteristics of the citing journal.
Calculation	The average number of *citations* per publication in a journal, divided by the average number of *references* per publication in the journals subject field(s). At present a harmonic average is used for references and the time window is the latest three years.
Formula	$SNIP = \frac{RIP}{DCP}$ Where

	$RIP = \frac{C(y \rightarrow (y-3:y-1))}{P(y-1:y-3)}$ $DCP = \frac{1}{3} * \frac{n}{\sum_{i=1}^{n} \frac{1}{p_i r_i}}$ RIP = Raw impact per paper. The number of citations from publications in year Y to publications in the journal published between Y-3 and Y-1. DCP = Database citation potential. A weighted harmonic average of the number of available citations in the field of the journal. n = the number of publications in the subject field of the journal. r_i = the number of active references in publication i in the journal's subject field. p_i = the proportion of publications in publication i:s journal that have at least one active reference (active means references to records in the same database).
Data Requirements	Requires data from a comprehensive citation database such as the Thomson Reuters citation indices. The original SNIP wad based on Scopus citation database.
KI Usage	Not currently used at Karolinska Institutet
Reference	Waltman, L., van Eck, J. N., van Leeuwen, T. N., Visser, M. S. (2013). Some modifications to the SNIP journal impact indicator. Journal of Informetrics, 7(2), 272–285. http://www.sciencedirect.com/science/article/pii/S1751157712001010

5. CITATION REFERENCE VALUES

5.1 FIELD CITATION REFERENCE VALUE

Designation	Field citation reference value
Denotation	Karolinska Institutet: μ_f CWTS: FCS (field citation score)
Description	The world average of citations to publications of the same document types, ages and subject areas.
Calculation	All documents are divided into groups where the items have the same document type, age and subject area. The mean value of the citations to all publications within the same group is the international field reference value for that particular group.
Formula	$\mu_f = \frac{1}{P}\sum_{i=1}^{P} c_i$ where: c_i = number of citations to publication i P = number of publications in the normalization group
Data Requirements	Requires data from a comprehensive citation database such as the Thomson Reuters citation indices.
KI usage	At Karolinska Institutet the field reference value is used to normalize citation rates for calculation of the more advanced citation indicators. Presently, we use the Thomson Reuters subject classification of the journals where the articles were published as a basis for grouping articles by subject.
Reference	-

5.3 JOURNAL CITATION REFERENCE VALUE

Designation	Journal citation reference value
Denotation	Karolinska Institutet: μ_j CWTS: JCS (journal citation score)
Description	The world average of citations to publications in the same journal and of the same document type and age.
Calculation	All documents are divided into groups consisting of items published in the same journal, having the same document type and age. The mean value of the citations to all publications within the same group is the journal reference value for that particular group.

Formula	$\mu_j = \frac{1}{P}\sum_{i=1}^{P} c_i$ where: c_i = number of citations to article i, belonging to the selected group of articles P = number of publications in the selected group of articles
Data Requirements	Requires data from a comprehensive citation database such as the Thomson Reuters citation indices.
KI usage	Used to calculate the Normalized journal impact
Reference	