Influence of study design on the citation patterns of Danish, medical research

Jens Peter Andersen[1] and Jesper Wiborg Schneider[2]

jpa@iva.dk, jws@iva.dk
[1]Aalborg Hospital, Aarhus University Hospital, AHSIC, Sdr. Skovvej 15, DK-9000 Aalborg
[2]Royal School of Library and Information Science, Frederik Bajers Vej 7K, DK-9220 Aalborg East

Abstract

A recent report has shown a discrepancy between citations to Danish research publications with international collaboration and without; with a higher citation rate to international publications. This discrepancy is particularly large for the university hospitals. A possible explanation is the use of different study designs in national publications than international. Therefore, this study investigates a potential connection between study designs and citation rates in Danish, medical research.

Web of Science records of Danish, medical research from 2005-2008 are enriched with study design information from MEDLINE. Citations are item-mean-normalised with respect to publication year.

Our data confirms previous findings of correspondence between citation rates and study designs, and we observe differences in the frequencies of study designs for national and international publications. Citation rates for national and international publications vary; international publications are cited 1.50 to 3.43 times more often than national publications, depending on study design.

Our results indicate that differences in citation rates for national and international publications to some degree can be explained by differences in the study designs used.

Introduction

In health sciences, claims have been made that research designs form hierarchies of evidence. Research designs are ranked according to the perceived reliability and validity of their findings (e.g. West et al., 2002). A general ranking often suggested is: 1) Systematic reviews and meta analyses, 2) Randomized controlled trials, 3) Non-randomised clinical studies, 4) Observational studies (e.g., case-control or cohort designs), 5) Non-experimental studies (e.g., case reports or nonsystematic reviews, and 6) Expert opinion (e.g. Harbour & Miller, 2001). However, no hierarchy of evidence is unanimously accepted (Evans, 2003). A hierarchy of evidence obviously implies epistemic importance of research findings according to the perceived strength and weaknesses of the various designs. Randomized controlled trials (RCT), for example, are seen by many as the "gold standard" for clinical trials due to the employment of randomized controls. But it is a contentious issue whether RCTs in general provide different or more reliable results than non-randomized and observational studies (e.g., Black, 1996). While RCTs, for example, may in principle have methodological advantages, what seems more important is the quality of the particular study and not the design *per se*. And of course RCTs cannot be conducted for all questions of interest.

The frequency distribution of research articles according to study designs is not notably interrelated with the various hierarchies (e.g. Gnanalingham, Robinson, Hawley, & Gnanalingham, 2006). Systematic reviews and meta-analyses are resource demanding and few in numbers, whereas articles reporting results from study designs in the lower ends of the hierarchies are numerous, (e.g., Patsopoulos, Analatos, & Ioannidis, 2005).

A number of studies have investigated the citation impact of various types of research designs in medical research (e.g. Kjaergard & Gluud, 2002; Patsopoulos, Analatos, & Ioannidis, 2005). But to our knowledge only Patsopoulos, Analatos and Ioannidis (2005) have studied whether this impact is commensurate with the proposed hierarchies of evidence. Their findings suggest that the citation impact of various study designs follows the general order proposed by most current conjectural hierarchies of evidence (Patsopoulos, Analatos, & Ioannidis, 2005). On average meta-analyses receive more citations than any other study;

RCTs are the second most cited study design; and the relative impact of epidemiological research seems to lose ground. While the study seems to corroborate the assumption of hierarchies of evidence, it has some caveats. One in particular is detrimental, a short immediate citation window of two years after publication without exclusion of self-citations. Short-term impact as measured by citations is heavily influenced by self-citations (Aksnes, 2003).

Citations are an indication of use or more abstract a measure of *impact*. A high citation frequency for an individual article is no guarantee for the quality of the particular study or the importance of the results. What we know from the study of Patsopoulos, Analatos and Ioannidis (2005) is merely that on average some research study designs seems to receive a higher number of citations. While a few studies suggest that articles of the "highest quality" are on average cited more often (e.g. Patterson & Harris, 2009; West & McIlwaine, 2002), it is important to notice that extrinsic (social) factors, such as the reputation of authors, may influence such findings considerably and that "quality" is not necessarily *the* significant factor when it comes to citation counts.

The present study also explores the relation between relative citation impact and medical research study designs but from a different perspective, with a different methodology and initially for a different purpose. The basis for the study is results from a forthcoming report we are involved in on research performance indicators for Nordic research institutions. An important finding from this analysis is a higher than expected discrepancy in relative citation rates between national and international publications for university hospitals compared to universities in the Nordic countries (see Appendix 1). National publications are those where all addresses are linked to one country. While we expected that national publications on average would have a lower relative citation impact compared to international publications, it was unanticipated to observe that the discrepancy between national and international publications was especially large for university hospitals. As can be seen from Figure 3 in Appendix 1, the large discrepancy is *not* due to particularly low rates of national publications in general. National and international publications for universities and university hospitals are comparable. What stand out are the generally higher relative citation rates for university hospitals' international publications compared to those of the universities. The analysis of performance indicators for Nordic research institutions is based on address-corrected Thomson Reuters' bibliographic data from the database at the Swedish Research Council. Contrary to most analyses, we use fractionalised counting and field and item normalised relative citation indicators excluding self-citations (Lundberg, 2007).

In other words, the bibliometric validity of the Nordic analysis can be considered high and it immediately raises the question why there is a larger discrepancy between national and international publications when comparing university hospitals to universities when we have corrected for variations in citation and publication patterns between fields, self-citations publication year and publication type?

Our purpose in the present study, and a number of future ones, is to investigate why international publications in medical research coming from university hospitals apparently tend to have a higher relative citation impact. We know from a recent report on national research performance indicators for Nordic countries (Schneider et al., 2010) that Denmark has a very strong impact in medical research. We therefore use Danish publications as our data for the present study. As indicated above, in medical research, study designs and their perceived influence on scientific evidence seems to be more influential than in most other research fields. It is therefore natural to begin an investigation by investigating publication and citation patterns of study designs between national and international medical research publications. This is the aim of the present study, where Danish medical research publications from 2005-2008 coming from university hospitals are explored in order to characterize such

patterns. We study the distribution of research designs among these publications and their relative citation rates for the given period, i.e., 2005-2008 corresponding to the period in the forthcoming report. The assumption here is that the discrepancy between citations to purely national papers and those with international collaboration may partly be explained by differences in the frequencies of study designs?
The paper is organized as follows: the following section outlines the data identification, extraction and matching methods, as well as the data analyses, applied in the analysis; the following section presents the main results of the frequency and citation analyses of study designs; and the final section discusses the findings. Appendix 1 presents the main findings from the unpublished report upon which this analysis is based.

Materials and method

Data for the analysis were gathered from Thomson Reuters Web of Science (WoS), using Science Citation Index-Expanded (SCI-E), Social Sciences Citation Index (SSCI), Conference Proceedings Citation Index – Science (CPCI-S) and Conference Proceedings Citation Index – Social Science & Humanities (CPCI-SSH). The databases were queried for "Denmark" OR "Danmark" in the address field and delimited to items published from 2005 to 2008 inclusive (54,060 items). This dataset was further delimited to selected document types (Articles, Proceedings Papers, Reviews and Letters), and subject areas (All WoS-subject areas related to medicine and biomedicine, following the classification used in Schneider et al. (2010)). The data were imported to a database, skipping those items which did not contain ISSN-numbers, as these were necessary for the further analysis, leaving 23,026 items. The principal loss in that particular case is conference publications published only in non-periodical proceedings. WoS document types are not related to study designs; however, this is registered in the MEDLINE database. By matching records from WoS to MEDLINE records, it was possible to enrich our records with study design-related document types. This was achieved in a two-step process, matching first those items which contained DOIs followed by a combination of ISSN, publication year and first page (most medical journals employ consecutive pagination, leaving volume and issue largely irrelevant). Only exact, unique matches were included, resulting in a total of 21,444 items. These were enriched with publication type and MeSH-descriptor data from MEDLINE (Some study designs are coded as publication type, while others as MeSH-descriptors).
The original unpublished analysis focused on university hospitals and national versus international publications. We therefore need to categorize publications according to institution and nationality. Author affiliations from WoS were categorized into Non-danish affiliations, Danish university hospitals and other Danish affiliations (e.g. universities, pharmaceutical industry, (non-)governmental organisations and general physicians). Some items did not contain Danish affiliations after all, and some did not involve Danish university hospitals and were excluded.
MEDLINE publication types and MeSH descriptors were used to categorise items into nine study designs (here ranked according to evidence level, descending):

1. Meta-analysis (Publication type)
2. Systematic review (Publication type = review AND "systematic review" in title words)
3. Randomised Controlled Trial (Publication type)
4. Cohort study (MeSH descriptor)

5. Multicenter study (Publication type)[4]
6. Clinical trial (Publication types = Clinical trial and Clinical trial phase I-IV)
7. Case-control study (MeSH descriptor)
8. Case report (Publication type)
9. Nonsystematic review (Publication type = review NOT "systematic review" in title words)

As not all MEDLINE items contained these publication types (many are merely classified as "journal article") or MeSH descriptors, the final analysis was limited to 5,816 items (25.3 % of the initial data).
Citations to the items were included from WoS and normalised with respect to publication year and study design in order to increase comparability across time. No fixed citation window was used, as items were normalised according to publication year, instead the complete citation count until time of collection was used. Field-normalisation was not applied, as many of the strata would become too small for sensible normalisations. We used item-normalised-mean-citation-scores to normalise each item (Lundberg, 2007).

Results

The absolute number of publications after the filtering processes is listed in Table 1. Publications are separated into national (Danish) and international publications. National publications are those affiliated to at least one Danish university hospital and no non-Danish addresses, while international publications are affiliated to at least one Danish university hospital and at least one non-Danish address.
Nonsystematic reviews are the most frequent study design in the data set for national publications, whereas clinical trials are the most frequent international study design. Interestingly, there are twice as many international clinical trials as national. This is the greatest absolute difference in publication counts, but also the second highest relative difference. The highest relative difference is found in case reports, with almost three times as many national as international publications. All study designs are higher cited when international collaboration is involved, ranging from factor 1.50 to 3.43. This tendency follows previous research and the findings from the unpublished report outlined in Appendix 1. Notice the relative citation scores are obviously different for the two studies, as the data sets are different, we only include those publications we could identify in MEDLINE, some are not identified others are not indexed, and because we use a slightly different calculation of citation scores. Nevertheless, the tendencies are clearly the same, and the present findings are reliable for studying publication and citation patterns for the current purpose.

Table 1. Publication counts and mean normalised citation impact for the eight study designs, divided into Danish and international publications.

	Publications		Mean normalised citation impact		
Study design	Danish	International	Danish	International	Factor difference
Systematic review	36	37	1.07	1.65	1.54
Meta-analysis	65	89	1.01	2.52	2.50
RCT	611	674	0.61	2.09	3.43

[4] Multicenter study is a subheading to clinical trials, eventually we merged this study design with clinical trials.

Clinical trial	565	1187	0.60	1.78	2.97
Cohort	460	445	0.57	1.04	1.82
Case-control	356	380	0.54	1.04	1.93
Nonsystematic review	900	744	0.94	1.41	1.50
Case report	291	118	0.22	0.37	1.68

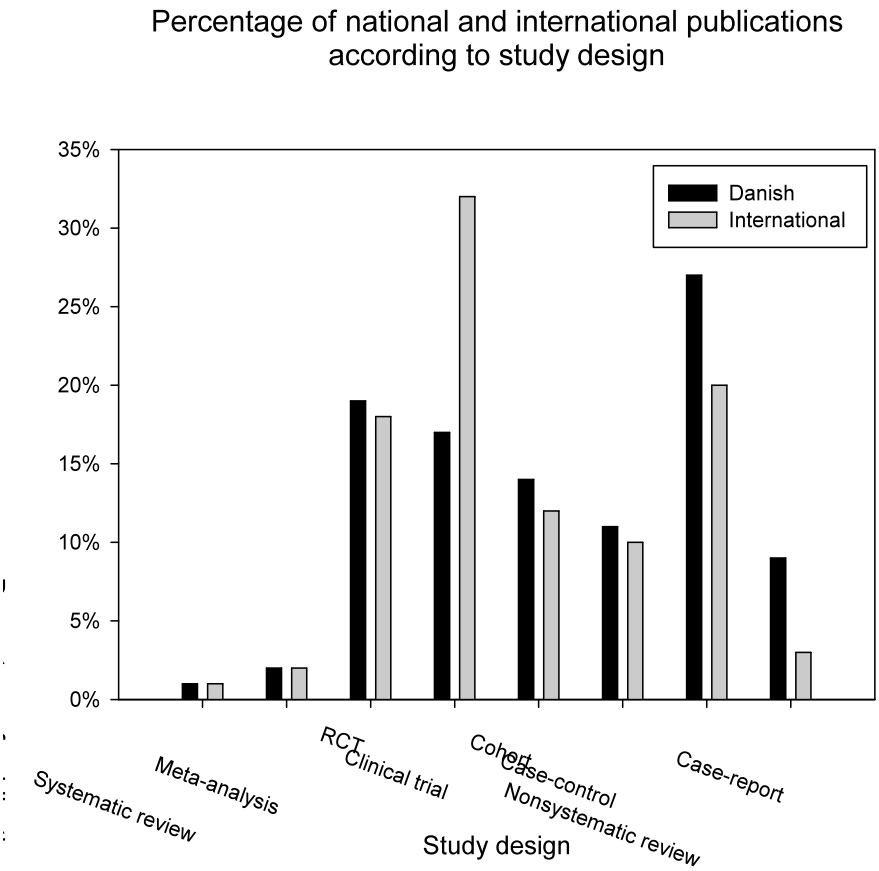

Figure 1. Relative publication frequencies for national and international study designs.

Figure 1 shows the relative distribution of study designs for national and international publications. It is evident that for international publications one-third of publications are clinical trials, whereas non-systematic reviews constitute more than 25% of national publications in the set. If we compare the patterns between national and international publications we see clear similarities in the relative distributions except for clinical trials and non-systematic reviews mentioned above, and case-reports. The distribution of the latter is markedly larger for national publications. What is perhaps interesting in this respect is that high profiled (in regard to the hierarchy) study designs such as systematic reviews, meta-analyses and RCTs seem to be more or less evenly distributed between national and international publications.

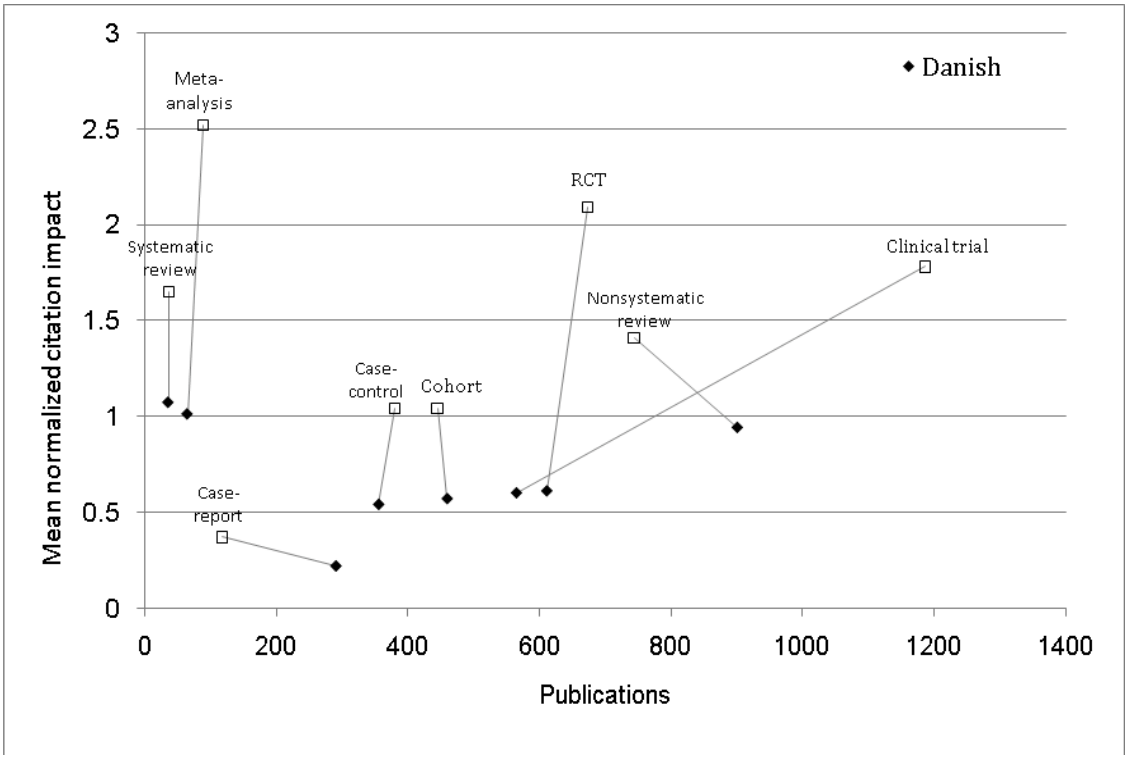

Figure 2. This figure is an illustration of impact, productivity and their differences for the various study designs for national and international publications.

Figure 2 plots publication frequency as a function of relative citation impact for national and international publications. It is based on the information in Table 1. Figure 2 clearly illustrates the intense discrepancy in citation impact for meta-analyses and RCTs but also perhaps surprisingly clinical trials, and to a lesser degree non-systematic reviews. With almost similar publication frequencies for national and international publications, it is interesting to observe the very large discrepancy in impact for meta-analyses and RCTs. Since RCT's are regarded as the "gold standard" for clinical trials, it is surprising to see how low the impact of national RCT's is – it is barely higher than that of clinical trials in general, cohort studies and case-control studies. Most interestingly, international clinical trials deliver the highest publication frequency combined with a very high impact. Also non-systematic reviews deliver robust impact and publication frequencies, which might be surprising, compared to previous research. This is possibly due to the simple, algorithmic approach to discerning systematic and non-systematic reviews (i.e., "systematic review" present in titles). Figure 2 shows that meta-analyses, RCTs, non-systematic reviews and clinical trials receive significantly more citations for its international publications. It is among these study designs that the explanations for the general discrepancies in citation impact should be found.

Discussion & conclusion

Our results show huge discrepancies in the impact of national and international publications, thereby confirming previous research. It was our aim to investigate whether some of these discrepancies are due to differences in the study designs used in national and international publications.

The two largest differences in frequencies of study designs were found for case-reports and clinical trials, with a national publication focus for the former and an international publication focus for the latter. Given the very low impact of case reports and the equally high impact of clinical trials, this gives us an indication that certain study designs influence the discrepancy in citation impact between national and international publications. Notice, the importance is

pattern of publication. RCT's receive more than three times as many citations on average when at least one author is international. This of course influences the overall difference between national and international publication impact especially when considerable numbers of publications are international. It is very clear from the findings that the prime movers for the very high discrepancy in citation impact between national and international publications for Denmark is clinical trials and RCTs, the discrepancy in impact is large and in the case of the former the international publication frequency is largest overall. Interestingly, the rather mixed category of "non-systematic reviews" drops in frequency from national to international though impact rises. The latter is not surprising as most other designs do the same: the interesting findings however are that these reviews have a higher than expected impact and most importantly, a very high publication frequency. They contribute considerably to the high overall impact for international publications.

More detailed data, including language, sources and subjects would be useful in giving a clearer image of the actual influence of study design on impact. One problem of the current data set is the size – we were only able to discover the study types of 25% of all relevant publications, which leaves a large set of important data untouched. Further analysis of the remaining publications is crucial for the thorough investigation of the problem.

Finally, and very interestingly, our results seem to support the findings in Patsopoulos, Analatos and Ioannidis (2005) in regard to whether impact is commensurate with the proposed hierarchies of evidence, at least for study designs in international publications (see Table 1; systematic reviews are not included in Patsopoulos, Analatos and Ioannidis (2005) and seem in general to have lower citation rates than meta-analyses). But we recapitulate that such a finding perhaps says more about social habits and paradigmatic thinking than research quality?

References

Aksnes, D. W. (2003). A macro study of self-citation. *Scientometrics*, *56*(2), 235-246.

Evans, D. (2003). Hierarchy of evidence: a framework for ranking evidence evaluating healthcare interventions. *Journal of Clinical Nursing*, *12*(1), 77-84. doi: 10.1046/j.1365-2702.2003.00662.x.

Gnanalingham, M. G., Robinson, S. G., Hawley, D. P., & Gnanalingham, K. K. (2006). A 30 year perspective of the quality of evidence published in 25 clinical journals: signs of change? *Postgraduate medical journal*, *82*(968), 397-9. doi: 10.1136/pgmj.2005.041251.

Harbour, R., & Miller, J. (2001). A new system for grading recommendations in evidence based guidelines. *BMJ*, *323*(7308), 334-336. doi: 10.1136/bmj.323.7308.334.

Kjaergard, L. L., & Gluud, C. (2002). Citation bias of hepato-biliary randomized clinical trials. *Journal of clinical epidemiology*, *55*(4), 407-10.

Lundberg, J. (2007). Lifting the crown - citation z-score. *Journal of Informetrics*, *1*(2), 145-154. doi: 10.1016/j.joi.2006.09.007.

Patsopoulos, N. A., Analatos, A. A., & Ioannidis, J. P. A. (2005). Relative citation impact of various study designs in the health sciences. *JAMA : the journal of the American Medical Association*, *293*(19), 2362-6. doi: 10.1001/jama.293.19.2362.

Patterson, M. S., & Harris, S. (2009). The relationship between reviewers' quality-scores and number of citations for papers published in the journal Physics in Medicine and Biology from 2003–2005. *Scientometrics*, *80*(2), 343-349.

Schneider, J. W., Aksnes, D. W., Sivertsen, G., Faurbæk, L., Finnbjörnsson, T., Morthens, S. G. S., et al. (2010). *Bibliometric Research Performance Indicators for the Nordic Countries* (p. 56). A report from NoriaNet: "the use of bibliometrics in research policy and evaluation activities", NordForsk: Oslo. Retrieved from http://www.nordforsk.org/_img/bibliometri_2.pdf.

West, R., & McIlwaine, A. (2002). What do citation counts count for in the field of addiction? An empirical evaluation of citation counts and their link with peer ratings of quality. *Addiction*, *97*(5), 501-504. doi: 10.1046/j.1360-0443.2002.00104.x.

West, S., King, V., Carey, T. S., Lohr, K. N., McKoy, N., Sutton, S. F., et al. (2002). *Systems to Rate the Strength of Scientific Evidence* (p. 204). Rockville, MD: Agency for Healthcare Research and Quality: 64-88. AHRQ publication 02-E016.

Appendix 1

The figures summarize some specific findings from a forthcoming report on performance indicators for Nordic research institution which is the basis for the present analysis.

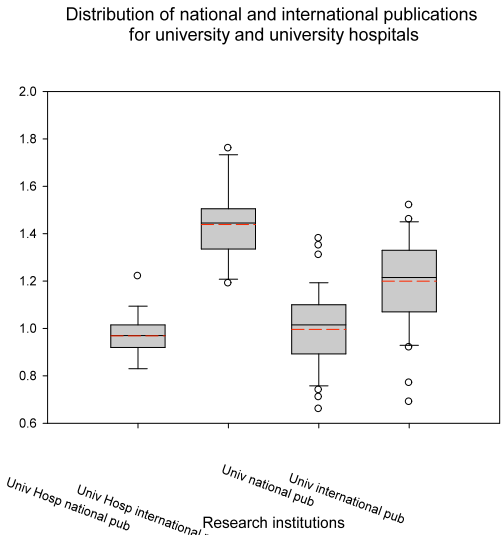

Figure 3: Field normalised citation impact for national and international publications for Nordic research institutions in the period 2005-2008. Research institutions are separated into universities and university hospitals.

Comment: National publications from universities and university hospitals have essential the same basis, whereas they differ markedly for international publications. University hospitals have a markedly higher relative citation impact for their international publications.

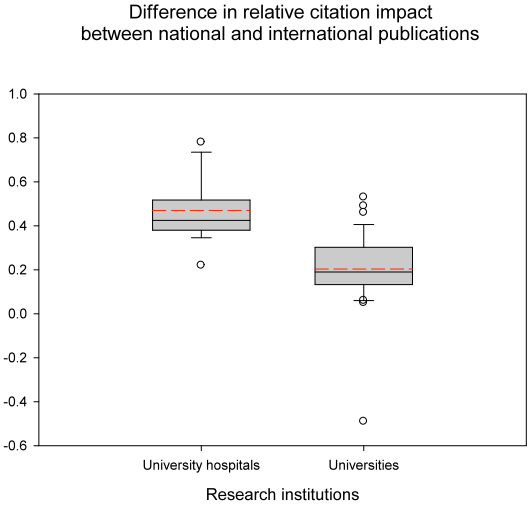

Figure 4: Difference in field normalised citation impact for national and international publications for Nordic research institutions in the period 2005-2008.

Comment: Figure 4 illustrates the distribution of actual differences in impact between national and international publications. We clearly see a larger difference for university hospitals in accordance with the previous boxplot.