Scientific Output and Impact: Relative Positions of China, Europe, India, Japan and the USA

Peder Olesen Larsen[1] Isabelle Maye[2] Markus von Ins[3]

26 August 2008

Abstract

Publication outputs and world shares of scientific publication are presented for 1981-2004 for China, Europe, India, Japan and the USA. Our results are compared with those available in the literature. The current situation whereby the main producers of scientific output statistics use different counting methods – thus, producing major differences in scientific output values – is unsatisfactory. The share in the total number of publications has been stagnating or gradually decreasing in recent years for Europe, the USA, India and Japan although there is no absolute decline in publication activities. The most dramatic trend has been the fast growth in China.

The USA is still maintaining a lead in publication impact. The impact from EU, Japan, China and India increases but is still far behind that of USA.

1 Introduction

Since the end of World War II, the USA has been the world's dominating scientific power. Europe initially ranked second, but only in a fragmented fashion. In recent decades, the situation has changed. Scientific production in European countries has increased strongly and Europe has challenged the primacy of the USA. In the 1970s Japan rose to third position after the USA and Europe. New countries, first of all China and India, have now appeared on the stage. Are Europe and the USA losing ground today and are the newcomers gaining ground? This question has been the subject of prolonged discussion, starting with the debate in the 1980s on the relative decline of British science (Martin 1994), followed by a more general approach (May 1997). There is an extensive literature on the subject but the evidence is ambiguous and no clear conclusions have been reached. The ambiguity is caused by methodological problems and unclear questions.

What is considered to be a relative decline and what does losing ground mean? What can be measured, what measurement methods give valid answers and what results permit conclusions?

We shall give a brief description of the methodological problems, followed by our own results, a comparison of our results with those available in the literature, a general discussion and a tentative conclusion.

2 Method and data

A country's scientific position can be estimated on the basis of both input and output indicators. Input indicators include, for example, total R&D expenditure, percentage of GDP spent on R&D and number of scientists. The number of scien-

[1] Marievej 10A,2, DK-2900 Hellerup, Denmark, pol at webspeed dot dk

[2] Center for Science and Technology (CEST), Effingerstrasse 43, CH-3003 Bern; Switzerland, isabelle.maye@sbf.admin.ch
Present address: State Secretariat for Education and Research SER, Hallwylstr. 4, CH-3003 Bern, Switzerland,

[3] Institute for Research Information and Quality Assurance , iFQ, Godesberger Allee 90, D-53175 Bonn, Germany, ins at forschungsinfo dot de

H. Kretschmer & F. Havemann (Eds.): Proceedings of WIS 2008, Berlin
Fourth International Conference on Webometrics, Informetrics and Scientometrics & Ninth COLLNET Meeting
Humboldt-Universität zu Berlin, Institute for Library and Information Science (IBI)

tific publications and citations are the most frequently used output indicators. We shall focus on these two indicators in our publication.

The different counting methods for publications give different results (Gauffriau and Larsen 2005; Gauffriau et al. 2007). Normalized counting methods assign one credit to each publication. If the publication has more than one author then the credit is divided by the number of objects studied. Complete and whole counting methods assign more than one credit to each publication if the publication has more than one author. Whole counting (one credit to each country contributing to a publication) gives double-counting and the results obtained by whole counting cannot be added. Therefore, the results for the European countries obtained by whole counting cannot be added to give the value for Europe. Calculation of shares based on values obtained by whole counting is problematic, if not outright impossible (Gauffriau et al. 2008).

The number of publications also depends on the scientific publication databases used. All databases include only a fraction of the literature. Whether the choice of literature and the journals included is biased (Paasi 2005) is currently subject to debate. Furthermore, the fraction of literature covered by these scientific publication databases may be decreasing over time, which creates problems for time studies (work in progress). At present, however, there is no real alternative to these databases.

There are various ways of defining Europe. Some years ago, most comparisons with the USA were performed for EU-15. This later became EU-25 followed by EU-27. EU-27+ is used in some studies where Iceland, Norway and Switzerland are included. But there are still many European countries that could be included (and many of these countries are participating in EU research framework programmes). These problems give difficulties in comparisons between different studies and in producing time series. China presents a similar problem. How should Taiwan be treated? In older studies, Hong Kong was treated separately.

The data presented in the following section are based on a database of unique codes established by the Center for Science and Technology Studies (CEST) using entries from the original databases. The CD-ROM editions of the Science Citation Index (SCI), the Social Science Citation Index (SSCI) and the Arts & Humanities Citation Index (A&HCI) from Thomson International from 1981 to 2004 were used as the original database. Only articles, notes, reviews and letters are included. The number of publications is calculated for overlapping 5-year periods. The impact is a citation indicator that is compared to the worldwide mean number of citations by scientific subfield and then normalised. The citation windows are also 5-years periods. Numbers of publications for Europe (EU-15 or EU-27) are calculated in two different ways: as the sum of the numbers for all EU member states; and as a single number obtained by considering the EU as a single country (union). In this latter case, publications are counted as EU publications when at least one EU member state contributed to producing them (Gauffriau et al. 2008). In our data contributions from Taiwan are not included in those from China.

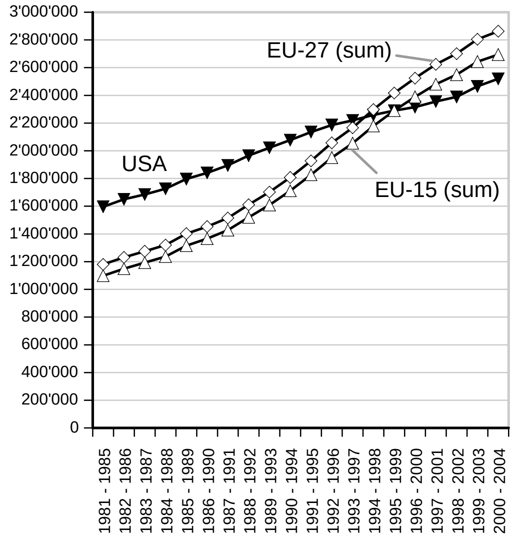
3 Results

3.1 Publication numbers for EU-27, EU-15 and the USA in the period 1981-2004

The following six graphs (Fig. 1 to Fig. 6) show the publication numbers for the USA, EU-15 (both as sum of the values for the 15 countries and as a union) and EU-27 (again both as sum of the values for the 27 countries and as a union) for the period from 1981 to 2004 according to four counting methods, complete counting (C), complete-normalized counting (CN), whole counting (W) and whole-normalized counting (WN). The values obtained by complete and complete-normalized counting are identical for EU as the sum of the values of the member countries and as a union (Gauffriau et al. 2008). Therefore, there are six, not eight graphs.

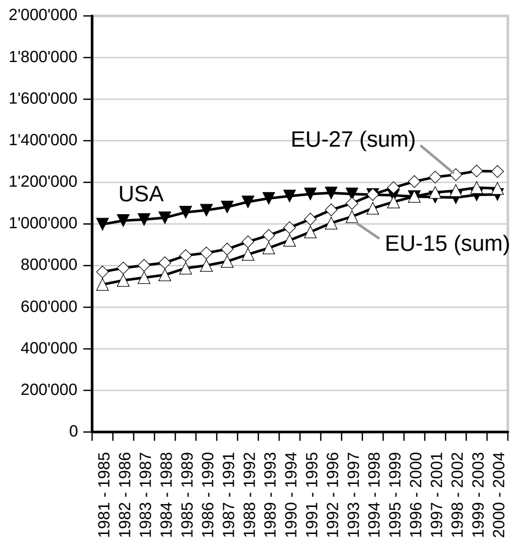
H. Kretschmer & F. Havemann (Eds.): Proceedings of WIS 2008, Berlin
Fourth International Conference on Webometrics, Informetrics and Scientometrics & Ninth COLLNET Meeting
Humboldt-Universität zu Berlin, Institute for Library and Information Science (IBI)

Fig. 1: Publications number for the USA, EU-15 (as a sum) and EU-27 (as a sum). Complete counting, C, 1981-2004.



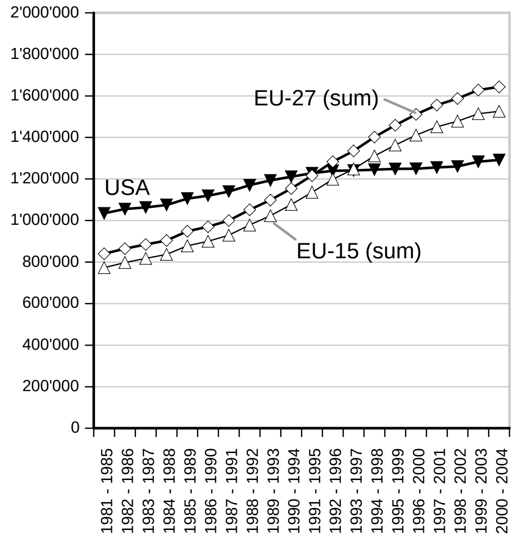
Data source: Thomson Scientific (SCI/SSCI/A&HCI)
Computation: CEST 2007

Fig. 2: Publications number for the USA, EU-15 (as a sum) and EU-27 (as a sum). Complete-normalized counting, CN, 1981-2004.



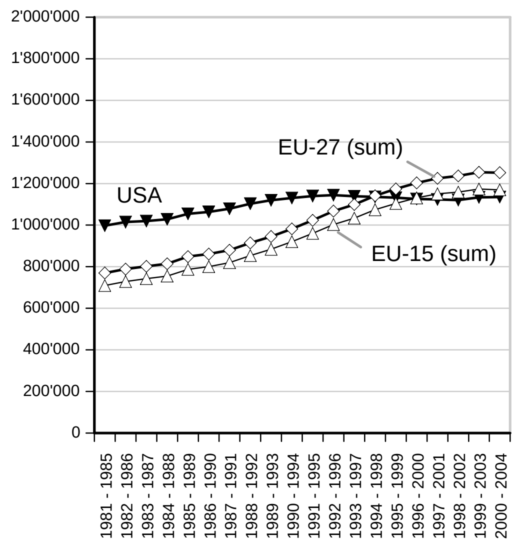
Data source: Thomson Scientific (SCI/SSCI/A&HCI)
Computation: CEST 2007

Fig. 3: Publications number for the USA, EU-15 (as a sum) and EU-27 (as a sum). Whole counting, W, 1981-2004.



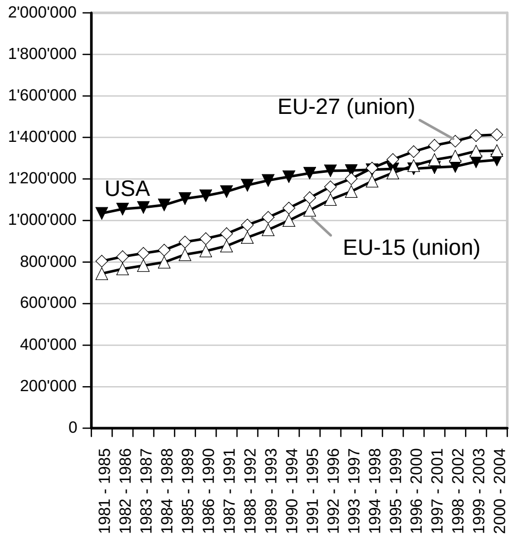
Data source: Thomson Scientific (SCI/SSCI/A&HCI)
Computation: CEST 2007

Fig. 4: Publications number for the USA, EU-15 (as a sum) and EU-27 (as a sum). Whole-normalized counting, WN, 1981-2004.



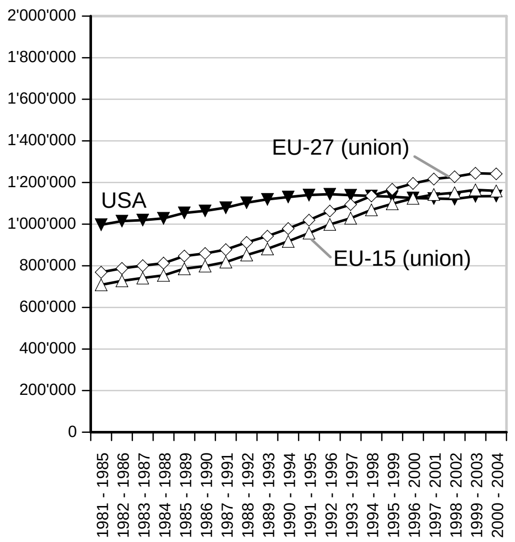
Data source: Thomson Scientific (SCI/SSCI/A&HCI)
Computation: CEST 2007

Fig. 5: Publications number for the USA, EU-15 (as a union) and EU-27 (as a union). Whole counting, W, 1981-2004.



Data source: Thomson Scientific (SCI/SSCI/A&HCI)
Computation: CEST 2007

Fig. 6: Publications number for the USA, EU-15 (as a union) and EU-27 (as a union). Whole-normalized counting, WN, 1981-2004.



Data source: Thomson Scientific (SCI/SSCI/A&HCI)
Computation: CEST 2007

In our opinion, complete-normalized counting is a sound method at the macro level for calculating absolute numbers, shares and international comparisons. There is no sound method based on whole counting, although such shares have been reported repeatedly in the literature.

The values obtained by complete-normalized counting (Fig. 2) indicate that the EU overtook the USA in the middle of the 1990s and that the growth in productivity stopped for the EU at the end of the 1990s and for the USA in the early 1990s. The complete counting (Fig. 1) and whole-normalized counting (Fig. 4) give similar conclusions. With the complete counting method, the increase for the USA in the final periods is more visible, reflecting increasing national and international cooperation. Whole counting (Fig. 3) indicates that the European productivity overtook that of the USA at an earlier time. This reflects the increasing propensity of European countries to cooperate internationally, mainly with other countries inside Europe.

The results for the EU treated as a union (Fig. 5 and Fig. 6) enable us to draw a similar conclusion: the EU overtook the USA in the middle of the 1990s.

The difference between values obtained by complete counting and by complete-normalized counting reflect the increasing national and international cooperation. The difference between EU as a sum and as a union reflects the ever-increasing cooperation within the EU (Gauffriau et al. 2008).

The difference in production between the 27 member countries of EU-27 and the 15 member countries of EU-15 is between 6 and 8% (depending on the counting method), not really important. The scientific production in Europe is mainly due to the 15 first member countries.

3.2 Publication numbers and world shares for EU-27, USA, China, Japan and India in the period 1981 - 2004.

The next four graphs (Fig. 7 to Fig. 10) show the total number and world shares of publications according to complete and complete-normalized counting for China, EU-27, India, Japan and USA for the period from 1981 to 2004.

Fig 7: World shares of publications for the USA, EU-27, Japan, China and India. Complete counting, C, 1981-2004.

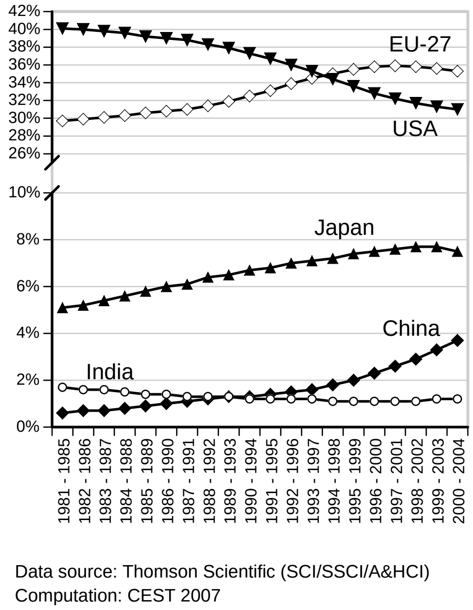


Fig 9: World shares of publications for the USA, EU-27, Japan, China and India. Complete-normalized counting, CN, 1981-2004.

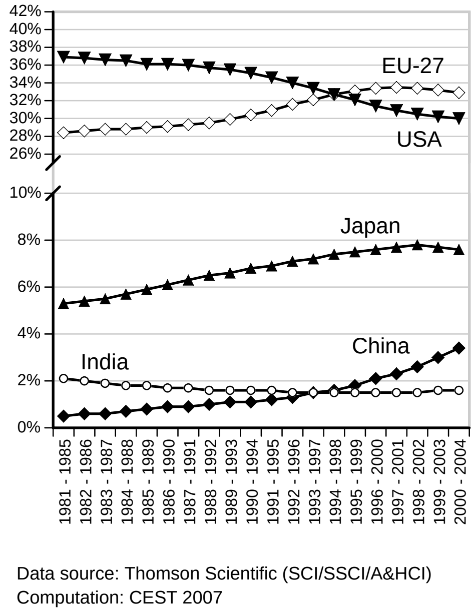


Fig 8: Publications number for the USA, EU-27, Japan, China and India. Complete counting, C, 1981-2004.

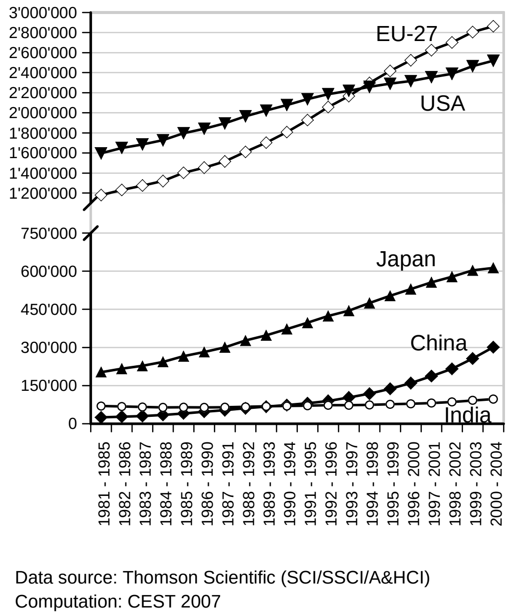
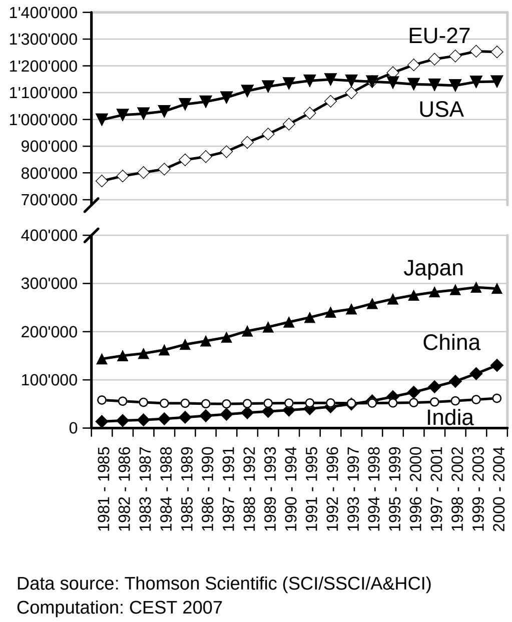


Fig 10: Publications number for the USA, EU-27, Japan, China and India. Complete-normalized counting, CN, 1981-2004



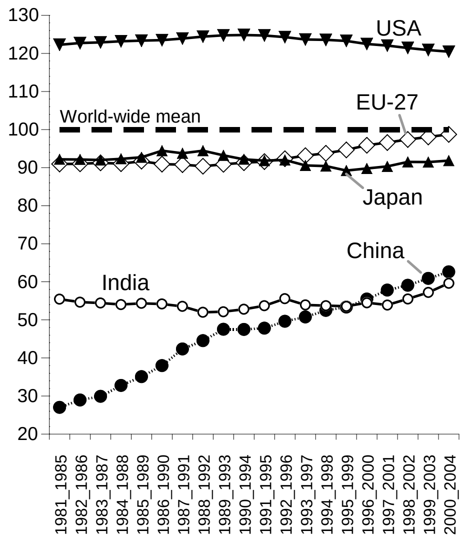
From both counting methods, it becomes clear that the USA's share declined for the entire period. The loss from 1981 to 2004, is about -9 points (complete counting, Fig. 7) and about -7 points (complete-normalized counting, Fig. 9). The USA still accounted for 1/3 of the world production in 2004. EU-27 increased its share until the end of the 1990s (approximately +6 points according to both counting methods) but has been losing share since the year 2000 (-1 point according to both counting methods). EU-27 accounts for more than 1/3 of the world production in 2004. Japan's share increased until about the year 2000 (around +3 points according to both methods) but has also been losing in share since then. Japan accounts for about 8% of the world production. Productivity in India was nearly the same for the entire period. The Indian share gradually decreased until about the year 2000, and then began to increase slightly from the year 2000 onwards, according to both methods. The Indian production represents less than 2% of the world production. The Chinese share increased for the entire period, more than 3 points and most dramatically in the last ten years. China's share of the world production was less than 1% in 1981 and about 4% in 2004.

In absolute number of publications (Fig. 8 and Fig. 10) , scientific output remained more or less stable for India for the whole period 1981-2004, with a small increase in the last periods (the growth rate between the two last periods is about +5.5% (complete counting) and +4.2% (complete-normalized counting)). The number of publications for Japan increased until the beginning of the 2000s (with a growth rate of around +4%, for both counting methods). Then growth in the absolute number of publications slowed down and stabilised towards the end with a growth rate between the two last periods of +1.6% (complete counting) and -0.9% (complete-normalized counting). For EU-27 the number of publications increased strongly in the beginning of the 1990s, with a growth rate of about +6% (complete counting) and about +4% (complete-normalized counting), then slowed down, with a growth rate of about +2.1% between the last two periods (complete counting) and -0.2% (complete-normalized counting). For the USA the number of publications has increased in the 1980s. The average growth rate is about +2.5% for the entire period and amounts to +2.2% between the last two periods (complete counting) According to complete-normalized counting the average growth rate was about +1.3% until the middle of 1990s, followed by a period of practically no growth (-0.3%), a slight increase in the beginning of 2000s (+1.2%) and then nearly no growth at all (+0.1%) between the last two periods. Chinese output increased for the entire period, slowly in the 1980s-1990s, and then fast in the last ten periods. Between 1981 and 2004, the average growth rate was +12.6% (complete-normalized counting) and +14.2% (complete counting). The growth rate from 2003 to 2004 is +15.4% (complete-normalised counting) and +17.6% (complete counting).

3.3 Impact

The following graph (Fig. 11) displays the scientific impact of publications from China, EU-27, India, Japan and the USA using the world-wide mean as a benchmark.

Fig 11: Impact for the USA, EU-27, Japan, China and India, from 1981-1985 to 2000-2004.



Data source: Thomson Scientific (SCI/SSCI/A&HCI)
Computation: CEST 2007

The USA is leading in impact for the entire period, far ahead of all other countries despite a small decline in the last years. The impact for USA increased from 122 in 1981-1985 to a peak of 125 in 1990-1994 and then slowly fell back to 120 in 2000-2004. The impact of EU-27 was around 90 in the 1980s, then at the beginning of the 1990s the impact increased regularly (+1 point per period) and came close to the worldwide mean in 2000-2004 (99). The impact for Japan was more or less stable for the entire period, around a value of 91. The impact for India was more or less stable at around 55 for the entire period with a small increase at the end, reaching 60 in 2000-2004. The impact for China was really low at the beginning but increased regularly and more quickly than for the other countries included in the comparisons. The impact for China reached and surpassed India in 1996-2000 and is now, in 2000-2004, about 63.

4 A review of the literature

A comprehensive set of data for the output in research covering the period 1995-2005 is presented in Science and Engineering Indicators 2008 from the National Science Foundation in the USA (2008). All the data presented have been produced by complete-normalized counting with institutions (or addresses) as the basic units of analysis. Data taken from SCI and SSCI are used. The data are in agreement with those described in the previous section.

Comparable data are found in OECD Science, Technology and Industry Scoreboard 2007. Data for the total publication activity in countries are derived from Science and Engineering Indicators 2006 from NSF and are therefore based on complete-normalized counting. However, in a special study of bioscience, whole counting was used.

The EU provides data for 2004 in Key Figures 2007. Again, data from SCI and SSCI are used. Whole counting is used. When calculating the EU-value for publications, double counting for the European countries is avoided. However, a comparison of the EU-data with NSF-data for 2004 indicates that in the EU-figures, the world shares for the individual EU-countries are calculated by dividing the country values obtained by whole counting with the total number of publications in the world. Using this method, the sum of shares for all countries in the world exceeds 100% by more than 20% (Gauffriau et al. 2008). Furthermore, because of double counting caused by scientific cooperation between EU-member countries and other countries, the EU-share for 2004 is stated to be 38.1% whereas the EU-share according to the NSF data is 33.5%.

UNESCO presents statistics for the world production of S&T publications for the period 1990-1997. The study is based on data from Observatoire des Sciences des Techniques (OST) in Paris Observatoire des Sciences et des Techniques – OST, Paris (2000) and produced by complete-normalized counting. Whereas the share of Science and Technology Publications in the period from 1990 to 1997 has decreased by 8% for North America and 11% for India, there has been an increase of 10% for Europe, 25% for Industrial Asia (including Japan) and 70% for China.

The relative positions of EU-15, Japan and the USA and the newcomers, Brazil, China, South Korea, Taiwan and Turkey, have been studied by Glänzel et al. (2008). Whole counting has been used to determine national publication activities. Corrections for double counting has been done to calculate publication numbers for EU-15. Calculations of shares have probably been made using a method which results in the sum of shares exceeding 100%. At least, this is the case in Table 3 of the publication. The methods used and the limitation to EU-15 make comparisons with other data difficult. However, the conclusion that Europe overtook the USA in publication activity in the 1990s and that EU-15, Japan and the USA are now at a standstill is reliable.

The relative positions of EU-15, the USA and the Asian Tigers, China, Singapore, South Korea, Taiwan, have been studied by Shelton (2008). Fractional (normalized) counting (without details) has been used to determine the national shares of publications. Although comparisons with other studies are difficult because of the use of EU-15, the conclusions are clear: USA was overtaken by EU in the 1990s, and now both EU and USA are at a standstill or slowly decreasing. The cause is the dramatic increase in the Asian Tiger countries, not a decline in investments in EU and USA.

The relative positions of EU-15 and the USA in various scientific fields have been studied by Horta and Veloso (2007). Whole counting has been used to determine shares of publications. Double counting has been avoided in determining the output of EU-15. But again, calculations of shares have been made using a method which results in the sum of shares exceeding 100%. Again, the methods used and the limitation to EU-15 makes comparisons with other data difficult.

A recent report shows that USA has a share of 50% in the core publications in science (Igami and Saka 2007).

However, the output of publications is different from the impact of publications. NSF Science and Engineering Indicators 2008 also gives the shares for EU-27 and the USA of the 1% most cited publications. This share is compared with the share of all publications to provide the index of highly cited publications. The following table is based on these data.

	Shares in per cent of the 1% most cited publications		Index of highly cited publications	
	EU-27	USA	EU-27	USA
1995	24.7	62.3	0.75	1.73
2000	25.9	59.9	0.73	1.85
2005	29.0	54.6	0.84	1.83
Data from NSF Science and Engineering Indicators 2008 depending on SCI and SSCI				

The general discussion about relative positions is not based solely on the major producers of statistics. A much cited publication on the scientific wealth of nations (May, 1997) gives data produced by whole counting. The share for the USA of the world's publications for 1981-1994 is given as 34.6%. This share has probably been calculated in a way to ensure that the shares of all countries add up to 100%. The share of citations for the USA for the period 1981-1984 is given as 49.0%. This figure again depends on whole counting. Another much cited publication on the scientific impact of nations (King 2004) again gives data produced by whole counting. National shares add up to more than 100% in the first table in this publication. Comparison with NSF-data is therefore not justified. But the shares for USA decreased from 1993 to 1997 for publications (6.9%), citations (5.5%) and the top 1% highly cited publications (4.3%). For EU-15 the comparable figures are increases of +4.8, +7.5 and +13.6%. For China the figures are +54.4, +64.2 and +125%, for India -2.7, +13.2 and +68.8% and for Japan +6.8, +11.9 and +14.4%.

5 Conclusion

It is unsatisfactory that the main producers of statistics on research output, the EU and NSF in the USA, use different counting methods giving differences in the values for the scientific output of EU and the USA. The methodological problems are even greater when it comes to citation counting (Gauffriau et al. 2007). It is problematic that Europe and EU are defined in many different ways in different studies, that many studies give national shares adding up to more than 100% and that the literature in many cases doesn't provide precise information about the counting methods (Larsen 2008).

However, there is no doubt that the share in the total number of publications is stagnating or decreasing slowly for both the EU and the USA, although they show no absolute decline in publication activities. Japan is showing stagnation in both share and absolute number of publications. There is growth in India but from a low starting point, India is not a major player in publication output. The most dramatic trend is the fast growth in China, which is now, in 2004, among the top producer countries, ranking 7th among the top 20 countries studied in CEST (complete counting method) (CEST 2007). There are many possible explanations for the growth in China, among these that the Chinese scientists have transferred their publication activity from Chinese journals to journals covered by SCI, SSCI and A&HCI, and that there has been an increase in R&D investments.

The USA is still maintaining a lead in publication impact. The impact of publications from EU-27 has increased but EU is still far behind the USA, just like Japan, India and China. The impact for Japan and India is more or less stable for the entire period. Only the impact for China really increased these last years, but the impact is still far behind the other countries. China ranked 35th in 2004 in the comparison of impact at the country level carried out by CEST (CEST 2007).

Is a relative decline a real decline? If we believe in Merton's norm of scientific results as a public good, the USA should not be seen as declining just because other countries are increasing their contributions to the common stock of scientific knowledge. The shares of publications and citations worldwide must decrease for the old scientific countries because of growth among the newcomers.

On the other hand, if the object of research policy and investment in R&D is dominance or monopoly and if scientific research is a key factor in economic development and competitiveness, then it may be appropriate to talk about decline and losers. However, because of the large citation rates and the high impact of research in the USA and the EU-countries (and the high university rankings) it is not evident that the USA and the EU-countries are losing ground and in decline and there is no evidence that this will be the case in the near future.

Acknowledgement

Support from the Carlsberg Foundation to POL is gratefully acknowledged. We are indebted to Wilfred T. Hastings for correction of the language.

References

European Commission (2007). Towards a European Research Area. Science, Technology and Innovation. Key Figures 2007. http://ec.europa.eu. Last accessed 2008-05-08.

CEST (2007). CEST Scientometrics Scoreboard: Indicateurs bibliométriques de la recherche (1981-2004) basés sur les SCI®, SSCI® et A&HCI®. isabelle.maye@sbf.admin.ch

Gauffriau, M., and P.O Larsen (2005) Counting methods are decisive for rankings based on publication and citation studies. *Scientometrics* 64 (1), 85-95.

Gauffriau, M., P.O. Larsen, I. Maye, A. Roulin-Perriard, and M. von Ins (2007). Publication, cooperation and productivity measures in scientific research. *Scientometrics* 73 (2), 175-214.

Gauffriau, M., P.O. Larsen, I. Maye, A. Roulin-Perriard, and M. von Ins (2008). Comparison of results of publication counting using different methods. *Scientometrics* 77 (1), 147-176.

Glänzel, W., K. Debackere, and M. Meyer (2008). 'Triad' or 'tetrad'? On global changes in a dynamic world. *Scientometrics* 74 (1), 71-88.

Horta, H., and F. Veloso (2007). Opening the box: comparing EU and US scientific output by scientific field. *Technological Forecasting & Social Change* 74, 1334-1356.

Igami, M., and A. Saka. (2007). Capturing the Evolving Nature of Science, the Development of New Scientific Indicators and the Mapping of Science. OECD Science, Technology and Industry Working Papers, 2007(1), OECD Publishing. DOI: 10.1787/30005636714.

King, D. (2004). The scientific impact of nations. *Nature* 430 (6997), 311-316.

Larsen P.O. (2008). The state of the art in publication counting. DOI: 10.1007/s11192-007-1991-6 (to be published in *Scientometrics* 77 (2)).

Martin, B.R. (1994). British Science in the 1980s – Has the Relative Decline continued? *Scientometrics* 29 (1), 27-56.

May, R. (1997). The Scientific Wealth of Nations. *Science* 275 (5301), 793-796.

National Science Foundation, 2008. Science and Engineering Indicators 2008. www.nsf.gov/statistics/seind08/ Last accessed 2008-05-08.

Observatoire des Sciences et des Techniques – OST, Paris (2000). Indicateurs bibliométriques des institutions publiques de recherche françaises.

OECD (2007). Science, Technology and Industry Scoreboard 2007. www.oecd.org. Last accessed 2008-05-08.

Paasi, A. (2005). Globalisation, academic capitalism, and the uneven geographies of international journal publishing spaces. *Environment and Planning A* 37, 769-789.

Shelton, R.D. (2008). Relations between national research investment and publication output: Application to an American Paradox. *Scientometrics* 74 (2), 191-205.

UNESCO Institute for Statistics (2001). The state of science and technology in the world 1996-1997. www.unesco.org/statistics. Last accessed 2008-05-08.