

## 79. A question of quality: Criteria for the evaluation of science and medical reporting and testing their applicability

*Marcus Anhäuser and Holger Wormer, Dortmund University, Germany*

### Introduction and method

The evaluation of quality in science journalism and science communication has often been focused on the question of accuracy. But opinion on what constitutes accuracy may be different among scientists and journalists. This may be one reason why the acceptance of purely scientifically based advice for better science reporting is low among journalists (e.g. Oxman 1993). However, in recent years different monitoring projects emerged, which try to judge the quality of medical reporting on (new) treatments, tests and procedures. These attempts use a set of defined criteria which focus on questions like: Is the magnitude of the benefit reported? Are the associated risks and costs mentioned? What is the quality of the sources (studies and experts)? But also: Is there a second opinion mentioned and does the report go beyond a press release?

Mainly based on the work of Moynihan (2000), the Australian “Media Doctor” started as the first of such projects in 2004 ([www.mediadoctor.org.au](http://www.mediadoctor.org.au)) followed by a monitoring in Canada and Hong Kong as well as in the USA ([www.healthnewsreview.org](http://www.healthnewsreview.org)). In November 2010 the German Medien-Doktor – The German HealthNewsReview ([www.medien-doktor.de](http://www.medien-doktor.de)) started as the first European project in this tradition. However, the 10 criteria used in the other countries were extended by three purely journalistic criteria such as actuality and relevance of the topic, quality of presentation and journalistic accuracy. These criteria were implemented in a journalistic review process which is developed alongside scientific peer review, however, by working with reputable science journalists as reviewers (instead of mainly scientists or physicians). This pure “science journalistic peer review” may be regarded as another innovation in our project.

### Results

After the evaluation of the first 120 stories (using 10 plus 3 criteria by at least two reviewers each) there are about as many highly ranked stories as stories with poor quality (0-star-stories: 11; 1-star: 19; 2-star: 27; 3-star:27; 4-star: 28; 5-star: 8). In 35 cases the newly introduced three purely journalistic criteria were the reason for a downgrade of 1 star; in one case the new criteria justified an upgrade. Comparing the results (based on the international 10 criteria) with the US HealthNewsReview project only one major difference can be identified: in the US-project 14 percent of the evaluated stories (n=1729) got a five-star result whereas in the German project only 6 percent (n= 120) were ranked in this category.

Although the star ranking may be used as a first and rough indicator of the overall quality of different stories the individual results for each single criterion may be more interesting. According to our reviewers’ judgment the most frequent failed criterion was an appropriate discussion of the available *evidence* for a presented scientific result (75 percent of articles rated “not satisfactory”). Similar results were found for an adequate presentation of *benefits*



(74 percent “not satisfactory”) as well as for the discussion of *risk and harms* of a medical method (73 percent).

In many cases it was also criticised that the journalist did not cite independent experts (63 percent). However, journalists seem to have fewer problems with explaining the real novelty and availability of a therapy or diagnosis (29 percent each), to investigate beyond a given press release (24 percent) and to avoid “disease mongering” (12 percent “not satisfactory”). Among our newly introduced purely journalistic criteria actuality, relevance and/or originality of the chosen topics were criticised rarely (15 percent “not satisfactory”; all values rounded to full percent.)

Although these results are preliminary, some suggestions can be made on how to improve reporting on medical sciences by comparing them with the data from the US and Australia (Schwitzer 2008; Wilson 2009). In all three monitoring projects the criteria “benefits”, “harms” and “evidence” are considered to be among the worst reported objects, (with the exception of “costs” in the US (negative rank 1) and the Australian projects (negative rank 2).

Interestingly, the criterion “costs” was more often fulfilled in German health news than in stories from the other two countries. We had expected that the German journalists could consider this information as less important for the individual patient because the German health system is known for its rather extensive coverage of costs, at least in comparison to the US. However, all comparisons between the countries are limited as we imply here a similar judgment and application of the given criteria by the reviewers in the different countries.

As a means of internal quality assurance we evaluated the consistency of judgments of different reviewers in our project (already after n=70 stories with at least two reviewers each). In 77 percent the reviewers’ judgments for the same story was identical or differed by only one star. For 23 percent of the stories the rank differed by two stars which was also the maximum difference; there was no judgement of a story that differed by more than two stars.

## Conclusion and perspectives

Our method seems to be appropriate to sharpen our picture of strengths and weaknesses of medical reporting in the mass media. Therefore, it helps to specify the needed improvement in the education and permanent training of journalists reporting on science and medicine. Some cases of stories evaluated in the project are already used in our own journalism training. Furthermore, the project provides access to best practice examples for the journalistic as well as for a broader community (patients, medical doctors, science communicators).

As the project has gained remarkable acceptance among journalists and mass media (e.g., nomination for the journalistic Grime Online Award; among the awards for the “journalist of the year” of the medium-magazin; integration in the standard curriculum for (other) journalism training programmes in media houses such as the German Press Agency (dpa); wishes from several mass media to be evaluated by the Medien-Doktor team for consultancy) it also promises a sustainable impact on the self-control and reflection of journalists dealing with medical issues.

This does not necessarily mean the exact adoption of our criteria but already the awareness raising of the need or existence of any quality criteria. This may be illustrated by a reaction of a journalist to our evaluation of one of her articles:



*“I have decided for myself to establish my own small check list oriented to the Medien-Doktor criteria (however with different weighing of some criteria) and to evaluate my own articles by using them before submitting. This is my personal step for an improvement of the reporting.”*

Based on our positive experiences with the project Medien-Doktor for medical reporting we are developing now criteria for a quality monitoring in other fields of science reporting such as on environmental issues.

### **Literature and further reading**

- Arnold K (2008): Qualität im Journalismus – ein integratives Konzept. Publizistik, vol. 53 (4): 488-508.
- Moynihan R, Bero L, Ross-Degnan D, Henry D, Lee K, et al. (2000): Coverage by the news media of the benefits and risks of medications. New England Journal of Medicine 342: 1645-1650.
- Oxman AD, Guyatt GH, Cook DJ, Jaeschke R, Heddle N, Keller J (1993): An index of scientific quality for health reports in the lay press. Journal of Clinical Epidemiology 46(9): 987-1001.
- Schwitzer G (2008) How Do US Journalists Cover Treatments, Tests, Products, and Procedures? An Evaluation of 500 Stories. PLoS Med 5(5): e95.
- Wilson A, Bonevski B, Jones A, Henry D (2009) Media Reporting of Health Interventions: Signs of Improvement, but Major Problems Persist. PLoS ONE 4(3): e4831.
- Wormer H (2011): Improving Health Care journalism. In: Gigerenzer G. and Gray, JAM (Ed.): Better Doctors, Better Patients, Better Decisions: Envisioning Healthcare 2020. Strüngmann Forum Report, vol. 6. Cambridge, MA, MIT Press: 169-188.

