Media Representations of Genetic Discoveries: Hype in the Headlines?

Timothy Caulfield & Tania Bubela

Many commentators have expressed concern regarding the sensationalistic reporting of biomedical stories by the popular press.[1] It has been suggested that inaccurate or exaggerated reporting can have an adverse impact on public understanding, creating unwarranted hope or fears, and the development of informed policies.[2]

Readers get their first or only impressions from headlines. Unfortunately, there are reasons to believe that headlines may be particularly inaccurate or "hyped." Science and medical stories have to compete with other news stories and, as such, headlines must be constructed to catch the attention of both the potential reader and editors who make publishing decisions.[3] As a result, even when a media report is circumspect, headlines may be sensationalized.[4]

"Headline sensationalism" has been associated with a variety of specific social concerns. For example, in the context of genetic discoveries, repeated exposure of the lay public to such headlines may lead to heightened genetic determinism.[5] That is, the public will come to develop an inaccurate belief that there is a tight causal linkage between a gene and a given human trait or disease. It has also been noted that headlines can influence how the reader interprets the information presented within the body of the full article.[6]

Sensationalized headlines that bear little resemblance to the article may generate antipathy or disappointment among some readers[7], creating a degree of bad will with a portion of the potential audience that should give editors pause.[8] They may also alienate sources, including those in the scientific community.

While there is a growing body of research on the accuracy and nature of newspaper stories[9], there is little available data on the accuracy of headlines in the context of genetic research. This study builds on the results of a previous paper examining the accuracy of newspaper stories.[10] We examine the degree and nature of the "hype" present in newspaper headlines associated with stories on genetic discoveries.

Methods

We describe only those methods that are specific to this study on headlines. The selection of scientific papers and newspaper articles, the general coding frame and coding, and a detailed explanation of the statistical analysis using CART were presented in Bubela and Caulfield.[11]

Parts of the coding frame were specific to the headline study. The coders were asked a series of questions with standardized categorical responses on the theme; the source of information (voice) beside the scientific paper; assessment of risk, benefits, or controversy for headlines, newspaper articles, and the scientific paper that generated the press coverage (Table 1).

Three coders, who all had scientific backgrounds, were asked to subjectively assess the technical accuracy of the

headline compared to the scientific paper. The coders also subjectively assessed whether the claims made in the headline were exaggerated (1) relative to the newspaper article and (2) relative to the scientific journal article. The coders assigned the headline to one of three categories: no exaggerated claims, slightly exaggerated claims, and moderately-highly exaggerated claims with respect to both the contents of the newspaper article and the scientific journal article. When compared to the newspaper and the scientific journal article, the headline was not considered to have been exaggerated if its claims had first been made there.

Statistical Analysis

All three coders assessed a random selection of 84/627 (13.7%) headlines, newspaper articles and their associated papers to assess inter-coder reliability. We calculated Intraclass Correlation Coefficients (Model 2) for questions on the coding frame for which there was only one answer, and in all cases the coefficient was greater than 0.75, indicating good agreement (Table 2).[12]

A classification tree analysis determined which variables from the coding frame contributed to the assignment of the headline to one of the three categories of exaggerated claims using CART 4.0 (Classification and Regression Tree) software when compared to the contents of the newspaper article and the scientific paper, respectively.[13] The relative contribution of variables as primary splitting variables in the classification tree are listed in Table 3. These indicate the variables that were most important in determining the category of exaggerated claims of newspaper articles.

Results

We examined 590 headlines from 627 newspaper articles reporting on 111 scientific papers from 24 scientific and medical journals. The majority of headlines were generated by articles in Science (32.6%), Nature (18.8%), Cell (16.3%), and Nature Genetics (16.0%). The theme of the headline closely reflected the main theme of the newspaper and scientific journal articles (Table 1). The dominant themes were genetic diseases or disease genes and basic research. This was not surprising given our search criteria. However, behavioural or neuro-genetics were the main theme in over 15% of headlines and may indicate an over-representation of this research field in newspaper coverage of genetics.[14] Similarly, the lack of scientific papers on pharmacogenetics and proteonics when compared with headlines and newspaper articles on this topic suggests that press stories are emphasising this potential application of basic research and gene discoveries. These themes were secondary in the scientific papers.

Most headlines were not framed as a controversy (93.2%) and this reflected the lack of controversy in the newspaper articles and the scientific journal articles (Table 1). The majority of headlines mentioned the benefits of the research (93.2%), again reflecting the emphasis on benefits in both the newspaper and scientific journal articles (Table 1). Only 7.1% of headlines mentioned risks, even lower than the 15% of newspaper articles that discussed risks. Risks were discussed in only 5.4% scientific journal articles. This result is consistent with a 1991 study that found a preference for positive messages for headlines.[15]

Most headlines were categorized as having no exaggerated claims (64.2%) or slightly exaggerated claims (24.9%) when compared to the newspaper article (Table 2). About one tenth (10.8%) of headlines were moderately-highly exaggerated when compared to the newspaper article. However, twice as many headlines (20.9%) had moderately-highly exaggerated claims when compared to the scientific journal article. By comparison, 11% of newspaper articles were categorized as having moderate-highly exaggerated claims when compared to their scientific journal article sources.[16]

CART Analysis

The CART analysis determined the relative contribution of each descriptive variable as a primary splitting variable (Table 3). The misclassification cost for the learn data comparing the headline with the newspaper article indicated a good fit of the model to the data: 0.06 for no exaggerated claims, 0.22 for slightly exaggerated claims and 0.29 for moderately to highly exaggerated claims. The misclassification cost for the learn data comparing the headline with the scientific journal article also indicated a good fit of the model to the data: 0.21 for no exaggerated claims, 0.22 for slightly exaggerated claims and 0.32 for moderately to highly exaggerated claims.

The likelihood of risks was a primary splitting variable that determined the category of exaggeration when the headline was compared to the newspaper article. Both risks and the likelihood of benefits associated with the research were pri-

mary splitting variables that determined the category of exaggeration when the headline was compared to the scientific journal article. This again reflects the emphasis placed on benefits and the paucity of either scientific or media coverage of risks of scientific research.

The theme of the headline, the newspaper article, and the scientific journal article were also primary splitting variables. Headlines on behavioural and neuro-genetics, genetically modified organisms, reproductive technologies, and pharmacogenetics were exaggerated while headlines on life threatening and prevalent diseases such as cancer, stroke, and heart disease were more circumspect.

Finally, the scientific journal was a primary splitting variable when the headline was compared to the scientific journal article (Table 3, Fig. 1). Headlines on articles published in the generic science journals such as Nature and Science were less exaggerated than those published in the specialist medical journals (Fig. 1). However, this result may be an artifact of the difference in sample size, given the dominance of media coverage of articles published in Science and Nature.

Comment

Our study has a number of limitations. We only surveyed the print media and limited our analysis to the more respected newspaper publications that do not necessarily have as high a circulation as some tabloids. In addition, the study is limited to newspaper stories that are directly related to peer reviewed articles and published abstracts. As a result, stories that flow from other sources, such as abstracts from scientific meetings that may never be published[17], may be under-represented. Also, our study did not include speculative or editorial pieces that may be about a particular biotechnology controversy and such stories may be given more sensationalistic titles. Indeed, given that our study is limited to stories about genetic discoveries and not stories about specific scientific controversies (such as human cloning, genetic discrimination or embryonic stem cell research) one could argue that sensationalistic headlines are under-represented.

Nevertheless, our results show a moderate degree of exaggeration in headlines about genetic discoveries when compared to both the newspaper story and the scientific source article. The fact that the headlines were twice as likely as newspaper stories to moderately or highly exaggerate the claims made in the source science article (21% as compared to 11%) supports the general impression that headlines are more sensational and should be viewed with circumspection by the public. Nonetheless, we also found that 62.3% had no exaggerated claims, thus, the majority of headlines reasonably reflect scientific information. Overall, headlines amplify the reporting trends found in the newspaper articles. For example, there is slightly more exaggeration and speculation in headlines as compared to newspaper articles. In addition, headlines perpetuate the overemphasis of benefits over risks found in both the newspaper stories and the scientific article.

Timothy Caulfield is Canada Research Chair in Health Law & Policy, Research Director, and Associate Professor at the Health Law Institute, University of Alberta, Edmonton, Alberta. Tania Bubela is a student-at-law at Field Law, Edmonton, Alberta, and a former research assistant with the Health Law Institute.

Acknowledgements: The study was supported by Genome Prairie, SSHRC and AHFMR. The authors wish to thank Edna Einsiedel PhD and Usher Fleising PhD for study design, Tim Karels PhD for statistical advice, David Hik PhD for statistical advice and editorial comments, Jordan Ward MSc, Christine Caulfield MSc, and Qusair Mohammedbai BSc for coding and Nina Hawkins at the Health Law Institute for administrative assistance.

Notes

1. Miriam Schuchman & Michael S. Wilkes, "Medical scientists and health news reporting: A case of miscommunication" (1997) 126 Annals of Internal Medicine 976; Dorothy Nelkin, "Beyond risk: Reporting about genetics in post-Asilomar press" (2001) 44 Perspectives Bio. & Med. 199.
2. Timothy Caulfield "Genetics services and market forces" (2002) 3 Nature Reviews Genetics 734; Ian Hargreaves, Justin Lewis & Tammy Speers, *Toward a better map: science, the public and the media* (UK, Economic and Social Research Council, 2003), online: ESRC <http://www.esrc.ac.uk/esrccontent/DownloadDocs/Mapdocfinal.pdf>.
3. Vladimir de Semir, "What is newsworthy" (1996) 347 Lancet 1163.

4. Jerome P. Kassirer & Marcia Angell, "Prepublication release of journal articles" (1997) 337 New Eng. J. Med. 1762.
5. Dorothy Nelkin & Susan M. Lindee, "Elvis' DNA: The gene as cultural icon" (1995) 55 Humanist 10; Celeste M. Condit *et al*., "An exploratory study of the impact of news headlines on genetic determinism" (2001) 22 Science Communication 379.
6. P.H. Tannenbaum, "The effects of headlines on the interpretation of news studies" (1953) 30 Journalism Quarterly 189.
7. Jose A. Leǿn, "The effects of headlines and summaries on news comprehension and recall" (1997) 9 Reading and Writing 85.
8. Condit *et al.*, *supra* note 5.
9. Tania Bubela & Timothy Caulfield, "Does the print media hype genetic research?: A comparison of newspaper stories and peer reviewed research papers" Can. Med. Assn. J. [forthcoming in 2004]; M.S. Wilkes & R.L. Kravitz, "Medical researchers and the media: attitudes toward public dissemination of research" (1992) 268 J. Am. Med. Assn. 999.
10. Bubela & Caulfield, *ibid.*
11. *Ibid.*
12. Leslie Gross Portney & Mary P. Watkins, *Foundations of clinical research applications to practice*, 2d ed. (Upper Saddle River, NJ: Prentice-Hall Inc., 2000).
13. Tannenbaum, *supra* note 6; D. Steinberg & P. Colla, *CART- Classification and regression trees* (San Diego, CA: Salford Systems, 1997); Leo Breiman *et al.*, *Classification and regression trees* (Belmont, CA: Wadsworth International, 1984); Salford Systems, "CART- Product overview", online: Salford Systems <http://www.salford-systems.com/products-cart.html>.
14. Bubela & Caulfield, *supra* note 9.
15. G. Koren & N. Klein, "Bias against negative studies in newspaper reports of medical research" (1991) 266 J. Am. Med. Assn 1824.
16. Bubela & Caulfield, *supra* note 9.
17. L. Schwartz, S. Woloshin & L. Baczek, "Media coverage of scientific meetings" (2002) 287 J. Am. Med. Assn. 2859.

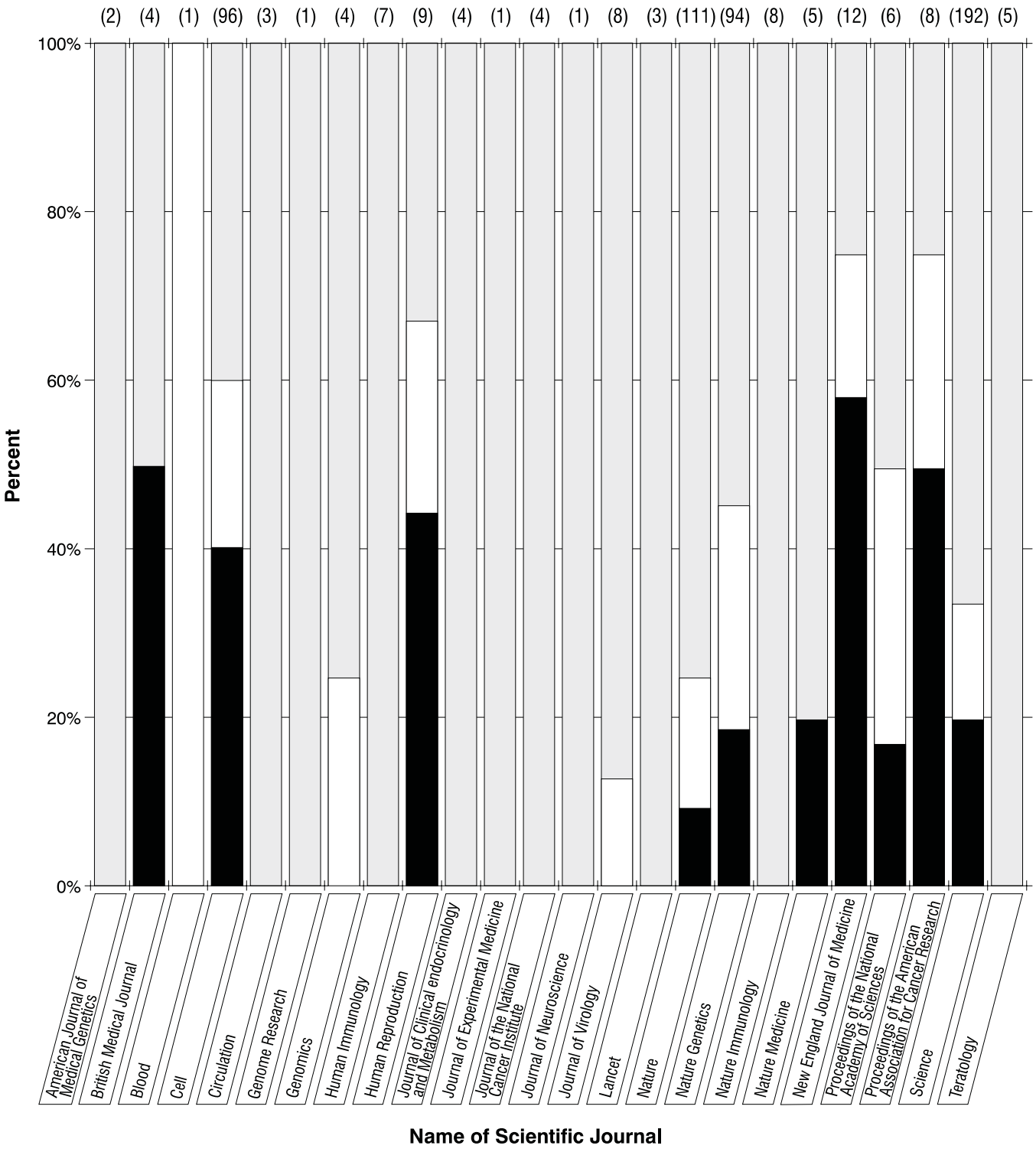

Fig. 1. The percent of newspaper headlines with moderately-highly exaggerated claims (white), slightly exaggerated claims (grey), or no exaggerated (black) claims when compared to their source in a scientific journal article. The number of headlines is presented at the top of the graph.

Table 1: Results of the Coding Frame for Headlines, Newspaper Articles, and Scientific Journal Articles				
Coding Frame	**Options**	Headline Percent (number)	Newspaper Article Percent (number)	Scientific Journal Article Percent (number)
Main Theme	**Science/Medicine**			
	1. Basic Research	19.4 (105)	14.5 (94)	27.9 (31)
	2. Disease/ Disease Gene	40.4 (219)	47.5 (298)	50.5 (56)
	3. Gene Therapy	0.9 (5)	0.2 (1)	0.9 (1)
	4. Behavioural/Neuro Genetics	15.3 (83)	15.6 (98)	12.6 (14)
	5. Genomics	0.7 (4)	1.1 (7)	1.8 (2)
	6. Pharmacogenetics	4.6 (25)	2.2 (14)	
	7. Proteonics	2.2 (12)	2.9 (18)	
	8. Genetically Modified Animals	6.3 (34)	4.1 (26)	1.8 (2)
	9. Genetically Modified Humans	0.9 (5)	0.3 (2)	
	10. Cloning Animals	1.8 (10)	1.9 (12)	
	11. Stem Cells/Therapeutic cloning	2.1 (11)	2.2 (14)	2.7 (3)
	12. Reproductive Technology (not Cloning)	0.7 (4)	1.0 (6)	0.9 (1)
	Safety/Risks			
	13. Health	0.6 (3)	0.3 (2)	
	Other Issues			
	14. Diagnosis, genetic testing, predictive medicine (in adults or post birth)	1.1 (6)	0.6 (4)	0.9 (1)
	15. Patenting/Property Rights	0.2 (1)		
	16. Economic Prospects, opportunities		1.9 (12)	
	17. Biopharmaceutical Industry	0.9 (5)	1.3 (8)	
	18. Legal/Regulatory	1.1 (6)	5.9 (37)	
	19. Education/Genetic Literacy	0.2 (1)		
	20. Ethical Issues	0.6 (3)	0.6 (4)	
	Total	**100 (542)**	**100 (627)**	**100 (111)**

Table 1: Results of the Coding Frame for Headlines, Newspaper Articles, and Scientific Journal Articles				
Coding Frame	**Options**	Headline Percent (number)	Newspaper Article Percent (number)	Scientific Journal Article Percent (number)
Main Voice (who/what is the main spokesperson/group/ institution quoted or described in the headline and newspaper article)	1. Not applicable/ unknown	80.8 (476)	0.8 (5)	
	Public Sector			
	2. Parliament/Congress	0.2 (1)		
	3. Government Research Institutions/scientists	2.5 (15)	14.5 (9)1	
	4. University or hospital research scientists	9.7 (57)	75.4 (472)	
	5. The Public, public opinion (e.g., surveys)	0.2 (1)	0.3 (2)	
	6. The media, published opinion	3.1 (18)	0.3 (2)	
	7. Celebrity (sports, film, TV)	0.5 (3)	0.5 (3)	
	Private Sector- Business			
	8. Scientists in Private laboratories	1.0 (6)	5.3 (33)	
	9. Biotechnology Company/ Spokesperson	0.5 (3)	1.1 (7)	
	10. CEO or upper management		0.6 (4)	
	11. Stock Exchange	0.5 (3)	0.5 (3)	
	Private Sector- Other			
	12. Patient Groups/Lobbies		0.5 (3)	
	13. Professional Organizations	0.2 (1)	0.2 (1)	
	Total	100 (589)	100 (626)	
Is the headline/ article framed as a controversy?	**No**	93.2 (550)	77.0 (483)	96.4 (107)
	Yes- balanced	3.2 (19)	13.6 (85)	1.8 (2)
	Yes- imbalanced	3.6 (21)	9.4 (59)	1.8 (2)
	Total	100 (590)	100 (627)	100 (111)
Are benefits mentioned	**Yes**	93.2 (550)	96.8 (605)	**98.2 (109)**
	No	3.2 (19)	3.2 (20)	**1.8 (2)**
	Total	100 (590)	100 (625)	100 (111)
Are risks mentioned	**Yes**	7.1 (42)	15.0 (94)	**5.4 (6)**
	No	92.9 (548)	85.0 (532)	**94.6 (105)**
	Total	100 (590)	100 (626)	100 (111)

Table 2: Questions on the Coding Frame Comparing Newspaper Headlines with (111?) Scientific Journal Articles and the Contents of 627 Newspaper Articles. The Intraclass Correlation Coefficient Measures Inter-coder Reliability of the Three Coders.

Question Comparing Newspaper and Scientific Journal Articles	Options	Percentage (number)	Intraclass Correlation Coefficient
What research model does the scientific journal article use?	**Human**	41.4 (46)	0.98
	Non-human	36.0 (40)	
	Both	22.6 (25)	
Does the scientific journal article discuss the research in terms of humans?	**Yes**	87.4 (97)	1.0
	No	12.6 (14)	
Does the headline discuss the non-human model?	**Yes**	20.5 (121)	1.0
	No	79.5 (470)	
Does the headline discuss the research in terms of humans?	**Yes**	87.8 (518)	1.0
	No	12.2 (72)	
	Total	100 (590)	
Are there any significant technical/scientific errors in the Headline?	None	83.2 (492)	0.76
	1-3	16.8 (99)	
	>3	0	
Overall, do the main claims made in the headline article reflect the contents of the newspaper article?	**Moderately-highly exaggerated claims**	10.8 (64)	0.77
	Slightly exaggerated claims	24.9 (147)	
	No exaggerated claims	64.2 (379)	
	Total	100 (590)	
Overall, do the main claims made in the headline reflect the contents of the scientific journal article?	**Moderately-highly exaggerated claims**	20.9 (123)	0.77
	Slightly exaggerated claims	16.8 (99)	
	No exaggerated claims	62.3 (367)	
	Total	100 (589)	

Table 3: Relative Contribution of the Variables that were the Main (50.00) Primary Splitting Variables in a Classification Tree Analysis Determining Whether Newspaper Articles Were Classified Into One of Three Categories: No Exaggerated Claims, Slightly Exaggerated Claims, or Moderately to Highly Exaggerated Claims.	
Variable	Variable Importance Score
Overall, do the main claims made in the headline article reflect the contents of the newspaper article?	
Main theme of the newspaper article	100.00
Main theme represented in the headline	97.03
Likelihood of risk discussed in the newspaper article	76.44
Overall, do the main claims made in the headline reflect the contents of the scientific journal article?	
Whether risks are discussed in the scientific journal article	100.11
The likelihood of the secondary benefit discussed in the scientific journal article	97.72
The main theme of the scientific journal article	96.00
The main theme represented in the headline	90.99
Likelihood of main benefit discussed in the scientific journal article	70.45
Whether the scientific journal article has a positive valuation tone	65.91
The name of the scientific journal	57.25