

From abstract to impact in cardiovascular research: factors predicting publication and citation

Stephan Winnik^{1,2*}, Dimitri A. Raptis³, Janina H. Walker⁴, Matthias Hasun¹, Thimotheus Speer², Pierre-Alain Clavien^{3,5}, Michel Komajda⁶, Jeroen J. Bax⁷, Michal Tendera⁸, Kim Fox⁹, Frans Van de Werf¹⁰, Ciara Mundow¹¹, Thomas F. Lüscher^{1,2,5}, Frank Ruschitzka², and Christian M. Matter^{1,2,5}

¹Division of Cardiology, Department of Medicine, University Hospital Zurich, Rämistr. 100, CH-8091 Zurich, Switzerland; ²Cardiovascular Research, Institute of Physiology, University of Zurich, Winterthurerstr. 190, CH-8057 Zurich, Switzerland; ³Department of Surgery, University Hospital Zurich, Rämistr. 100, CH-8091 Zurich, Switzerland; ⁴Department of Pediatrics, Cantonal Hospital Muensterlingen, CH-8596 Muensterlingen, Switzerland; ⁵Zurich Center for Integrative Human Physiology, University of Zurich, Zurich, Switzerland; ⁶Department of Cardiology, University Pierre et Marie Curie Paris 6, Pitie Salpetriere Hospital, 47 Boulevard de l'Hopital, 75013 Paris, France; ⁷Department of Cardiology, Leiden University Medical Center, Albinusdreef 2, 2333 ZA Leiden, The Netherlands; ⁸3rd Division of Cardiology, Medical University of Silesia, Poniatowskiego 15, 40-055 Katowice, Poland; ⁹National Heart and Lung Institute, Imperial College, ICMS Royal Brompton Hospital, Sydney Street, London SW3 6NP, UK; ¹⁰Department of Cardiology, University Hospital Gasthuisberg, Herestraat 49, 3000 Leuven, Belgium; and ¹¹Department of Scientific Programme Management, European Society of Cardiology, 2035 Route des Colles – Les Templiers, Sofia Antipolis, France

Received 3 November 2011; revised 21 March 2012; accepted 13 April 2012; online publish-ahead-of-print 5 June 2012

This paper was guest edited by Prof. Dr Brahmajee K. Nallamothu, Assistant Professor of Internal Medicine, University of Michigan, USA

See page 3002 for the editorial comment on this article (doi:10.1093/eurheartj/ehs160)

Aims

Through a 4-year follow-up of the abstracts submitted to the European Society of Cardiology Congress in 2006, we aimed at identifying factors predicting high-quality research, appraising the quality of the peer review and editorial processes, and thereby revealing potential ways to improve future research, peer review, and editorial work.

Methods and results

All abstracts submitted in 2006 were assessed for acceptance, presentation format, and average reviewer rating. Accepted and rejected studies were followed for 4 years. Multivariate regression analyses of a representative selection of 10% of all abstracts ($n = 1002$) were performed to identify factors predicting acceptance, subsequent publication, and citation. A total of 10 020 abstracts were submitted, 3104 (31%) were accepted for poster, and 701 (7%) for oral presentation. At Congress level, basic research, a patient number ≥ 100 , and prospective study design were identified as independent predictors of acceptance. These factors differed from those predicting full-text publication, which included academic affiliation. The single parameter predicting frequent citation was study design with randomized controlled trials reaching the highest citation rates. The publication rate of accepted studies was 38%, whereas only 24% of rejected studies were published. Among published studies, those accepted at the Congress received higher citation rates than rejected ones.

Conclusions

Research of high quality was determined by study design and largely identified at Congress level through blinded peer review. The scientometric follow-up revealed a marked disparity between predictors of full-text publication and those predicting citation or acceptance at the Congress.

Keywords

Scientific quality • Peer review • Publication • Impact

* Corresponding author. Tel: +41 44 635 50 97, Fax: +41 44 635 68 27, Email: stephan.winnik@access.uzh.ch

Published on behalf of the European Society of Cardiology. All rights reserved. © The Author 2012.

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0/>), which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Introduction

The Congress of the European Society of Cardiology (ESC) is one of the largest international scientific meetings worldwide, yearly attracting approximately 30 000 health care professionals from more than 60 countries.^{1,2} Scientific meetings like this provide an excellent forum for distribution and discussion of the most recent scientific findings. However, the factors that determine acceptance at a scientific meeting, and whether these factors also predict subsequent full-text publication are largely unknown. Moreover, a validation of the parameters that predict acceptance at a scientific meeting or acceptance for full-text publication with respect to their later impact has not been performed.

Several studies have investigated distinct aspects such as positive outcome- or institutional bias associated with the acceptance at scientific meetings or publication in peer review journals.^{3–6} Few studies have assessed the publication fate of abstracts submitted to scientific meetings.^{7–9} In the field of cardiology, Toma *et al.*¹⁰ specifically investigated the fate of randomized controlled trials (RCTs) presented at the scientific meetings of the American College of Cardiology between 1999 and 2002. Ross *et al.* evaluated the effects of blinded peer review for abstracts submitted to the Scientific Sessions of the American Heart Association from 2000 to 2004.^{10,11} To date, no study has followed the same cohort of abstracts from their submission to a scientific meeting over their full-text publication up to their subsequent impact after publication. Without an external and neutral surrogate for scientific quality such as citation analysis, earlier studies could not appraise whether the observed effects affected later scientific impact.

Our objective was to identify distinct factors predicting acceptance at an international scientific meeting in the field of cardiology as well as parameters predicting subsequent full-text publication and citation. Beyond prediction at each of these levels, we aimed to validate the parameters predicting acceptance at Congress level or publication with those parameters predicting subsequent scientific impact, thereby appraising the quality of the peer review and editorial processes involved at each stage.

To accomplish these goals, we assessed all 10 020 abstracts submitted to the Congress of the ESC in 2006 (i.e. joint meeting of the ESC and the World Heart Foundation) and followed a representative sample of 10% ($n = 1002$ abstracts) of both accepted and rejected abstracts over a period of 4 years for full-text publication and subsequent 2-year citations as a surrogate for scientific impact.

Methods

Study design and categorization of abstracts

All 10 020 abstracts submitted to the ESC congress in 2006 were entered into a database and assessed for acceptance, presentation format (oral vs. poster presentation), and the average reviewer rating. For a precise assessment and scientometric follow-up, a computer-assisted random selection of 10% ($n = 1002$, margin of error = 0.01)¹² of all abstracts submitted was assessed in detail according to a set of pre-specified variables. These variables were origin and type (clinical vs. basic) of research, origin and type (university-affiliated vs. not university-affiliated) of institution, study

design (RCTs vs. prospective non-randomized trials vs. retrospective studies vs. systematic reviews vs. meta-analyses), the number of patients enrolled in clinical studies, the study field of clinical research, and the gender of first and last authors. The reliability of the computer-assisted representative random selection was confirmed by comparing the distribution of variables common to both the original sample of all 10 020 abstracts and the randomly selected sample of 1002 abstracts; these common variables were acceptance at the Congress (accepted vs. rejected), presentation format, and average reviewer rating. No differences were observed (see Supplementary material online, Table S4). Gender of first and last authors was assessed using a standardized algorithm based on the analysis of their first and surnames in conjunction with their country of origin. In specific cases, this approach was supplemented using online searches for photographs associated with their home institution when necessary and available. This algorithm is described in detail in the online supplementary material and was developed *a priori*. Given the need to use this empirical approach, however, the analyses that included gender should be considered exploratory.

Scientometric follow-up for publication and citation

The representative random selection of all abstracts submitted (including both accepted and rejected abstracts) was followed for the time point and journal of publication and for the number of subsequent 2-year citations, beginning from the time of publication. Abstracts were followed for 4 years. The impact factor of the journal of publication was assessed for the year of publication. To identify whether an abstract had been published, the following algorithm was used: first, the full abstract title was entered as search term in PubMed[®]. If no publication was identified, a shortened version of the abstract title (keywords only, no adjectives) was entered. If still no publication was identified, the full list of authors was entered in PubMed[®], which, in case of no match, was followed by the names of only the first and last authors. If again no publication was identified, the same algorithm was repeated first in Google Scholar, then in Google, each indexing increasing amounts of contents. Thus, articles published in journals not indexed in PubMed[®] were identified. The Cochrane library was searched for abstracts of systematic reviews. To validate that identified articles and the respective abstracts matched, the following parameters were compared for every full-text article that was identified: primary endpoint(s), the number and size of the study groups, and the conclusions. Articles assessing the same endpoints in the same number of study groups of the same or slightly increased sizes, leading to the same conclusions, and published by the same last authors compared with the respective abstract were considered 'published'. The number of 2-year citations of original articles (beginning from the time of publication) was determined using the citation indexing and search service of the ISI (Institute of Scientific Information) Web of Science (Thomson Reuters).¹³ Data were collected by four independent reviewers (S.W., D.A.R., J.H.W., M.H.) in January and February 2011, and updated in December 2011 for completion of the follow-up of 2-year citations. In a pilot study, 3% of the study samples ($n = 30$) were assessed independently by each of the four reviewers to address inter observer variances. Thereafter, one out of the four reviewers assessed each of the 1002 abstracts. The cut-off for 'high' 2-year citation rates was set to 10 or more (≥ 85 th percentile of all published studies) in order to identify the top 15% of published articles. Receiver operating characteristic (ROC) curve analyses were performed to evaluate the predictive value of average reviewer ratings for frequent 2-year citation.

Regression analyses

Univariable and multivariable binary logistic regression analyses were performed to identify factors predicting acceptance at Congress level, as well as subsequent publication and citation frequency. Multivariable regression analyses were run using backward conditional variable exclusion and list-wise case exclusion. The threshold for entry was 0.05 and for exclusion 0.10.

Statistical analyses

Metric variables were assessed for distribution using the Kolmogorov–Smirnov test and compared by the Mann–Whitney *U*, one-way ANOVA, and Kruskal–Wallis *H*-tests where appropriate. Significance levels were adjusted for multiple comparisons using conservative Bonferroni's correction where indicated. Proportional differences derived from categorical data were compared using the Fisher exact or χ^2 tests where appropriate and presented as odds ratio and 95% confidence interval. *P*-values were two-sided and considered statistically significant for $P < 0.05$, if not indicated differently (multiple comparisons). Statistical analyses were performed with IBM SPSS Statistics Version 19.

Results

From abstract at the ESC congress to publication

The 10 020 abstracts submitted to the ESC congress in 2006 from 63 different countries of all five continents gave rise to over 2900 publications and over 32 000 citations (Figures 1 and 2, Supplementary material online, Table S2). Of all abstracts submitted 90% were clinical studies and 10% were basic research. Nearly one-fifth of the research submitted had been performed in institutions not affiliated to universities (Table 1). Half of the clinical studies submitted were retrospective, 35% were designed prospectively, of which 18% were RCTs (Table 1). In a pilot study encompassing 3% of the study sample ($n = 30$), inter-observer agreement was assessed. Fleiss' kappa values for categorical variables ranged from 0.853 to 1.0, and intraclass correlation coefficients for continuous variables from 0.996 to 1.0 (see Supplementary material online, Table S5).

Peer review and editorial process at the ESC Congress 2006

All abstracts submitted were reviewed by three to eight independent reviewers, who allotted scores from 1 to 10 for each abstract. Editorial decisions were based on average reviewer ratings. Reviewer ratings showed a central tendency without being normally distributed (see Supplementary material online, Figure S1). Abstracts accepted had received significantly higher scores than those rejected ($P < 0.001$). Among accepted studies, those accepted for oral presentation had been rated significantly better than those accepted for poster presentation ($P < 0.001$), indicating a good coherency of reviewer assessment and editorial decisions. A cut-off for acceptance was identified at an average reviewer rating of 6.5 (Figure 3A). The predictive value of expert reviewer assessment for frequent citation (defined as 10 or more 2-year citations) was 70% (AUC = 0.694, Figure 3B). The best cut-off for frequent citation was identified to be an average reviewer rating of greater than 7 (AUC = 0.660, $P < 0.001$, Table 2).

Concomitantly, studies accepted at the ESC Congress received significantly more 2-year citations than rejected ones ($P = 0.004$ for oral presentations, $P = 0.008$ for poster presentations, Figure 3D).

Factors predicting acceptance at the ESC Congress

Thirty-eight per cent of all submitted studies were accepted, and only 7% of all submitted abstracts (i.e. 18.4% of all accepted abstracts) were accepted for oral presentation (Figure 1).

In univariable analyses, two factors were positively associated with acceptance at Congress level: study design ($P = 0.018$) and a patient number of 100 or more (OR = 1.93, $P < 0.001$). Within study design, specifically prospective non-randomized study design (OR = 1.90, $P = 0.016$) and randomized controlled trial design (OR 1.56, $P = 0.004$) were identified. The only factor not favouring acceptance in univariable analysis was research conducted in the field of rhythmology (OR = 0.57, $P = 0.034$) with an acceptance rate of 33% compared with 46% of other abstracts (Table 1).

When adjusting for all factors assessed, the number of patients enrolled showed the strongest association with acceptance (adjusted OR = 2.07, $P < 0.001$), followed by basic as opposed to clinical research (adjusted OR = 2.24, $P = 0.001$) with an acceptance rate of 44 vs. 37%, and study design ($P = 0.023$). Within study design, prospective data collections (i.e. prospective cohort studies and non-randomized intervention trials) (adjusted OR = 1.67, $P = 0.002$) and RCTs (adjusted OR = 1.91, $P = 0.020$) favoured acceptance (Table 1). Interestingly, none of the aspects assessed was associated with oral or poster presentation among all accepted abstracts (see Supplementary material online, Table S1).

Factors predicting full-text publication

Of all accepted studies, 38% were subsequently published, whereas only 24% of rejected studies were published ($P < 0.001$). The publication rate of studies accepted for oral presentation (37%) did not differ from the publication rate of studies accepted for poster presentation (39%, $P = 0.79$). However, significantly more of those studies that had been accepted for oral presentation were subsequently published in high impact factor journals (30%, $IF \geq 10$) compared with studies that had been accepted for poster presentation (10%, $P = 0.009$) or rejected studies (4%, $P < 0.001$) (Figure 1).

Univariable analyses revealed a significant advantage for institutions affiliated with universities to publish their research (publication rate 31%) compared with non-university-affiliated institutions (publication rate = 22%, Table 3). With a publication rate of 41%, basic research showed the strongest association with publication (OR = 1.72, $P = 0.012$), followed by study design in clinical research ($P = 0.014$) and the affiliation to a university (OR = 1.53, $P = 0.025$). Notably, neither prospective nor randomized controlled study design, but meta-analyses contributed to the significant association of study design with publication in univariable analysis (OR = 5.97, $P = 0.041$). However, the small number of meta-analyses in the studied cohort does not allow a solid statement.

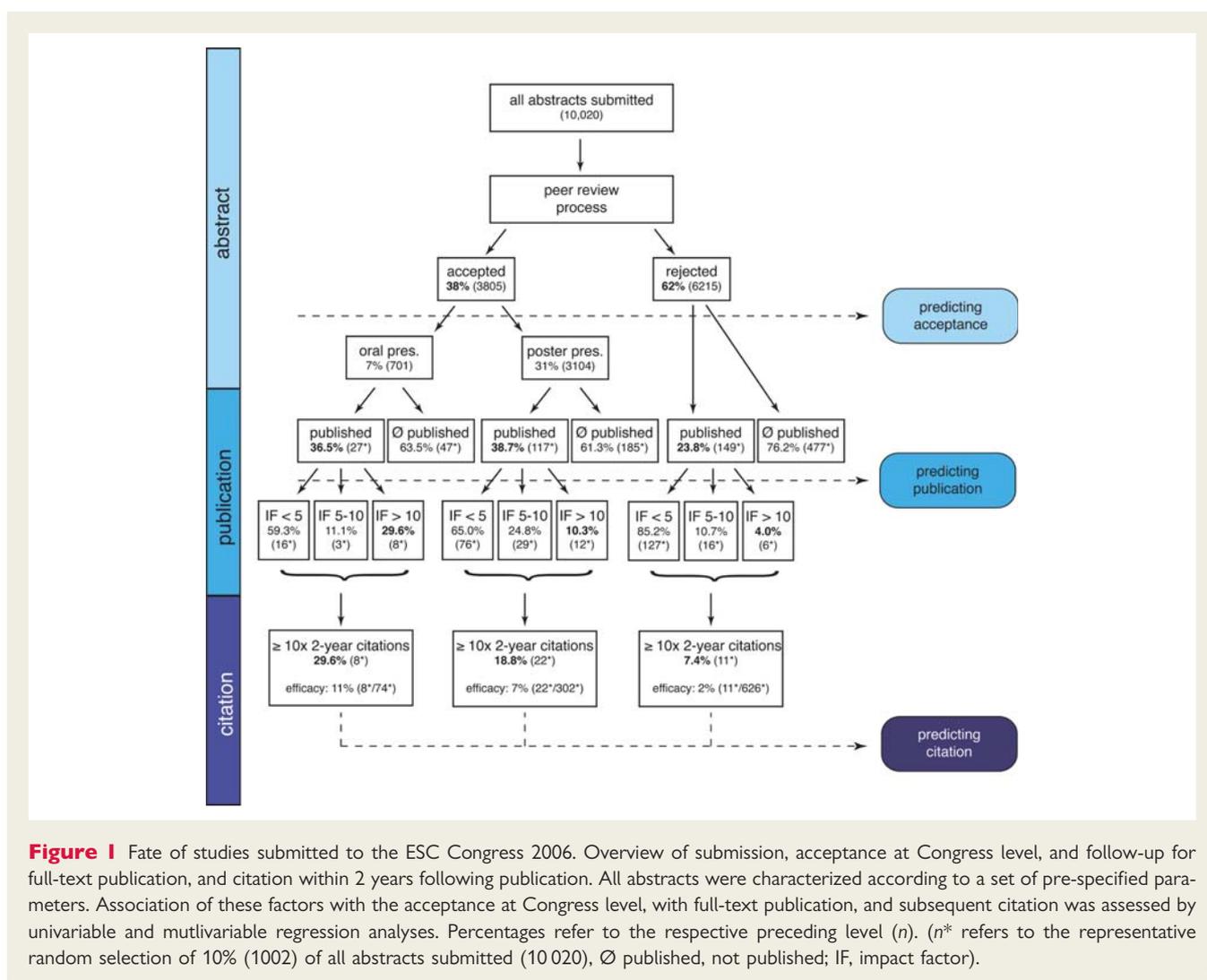


Figure 1 Fate of studies submitted to the ESC Congress 2006. Overview of submission, acceptance at Congress level, and follow-up for full-text publication, and citation within 2 years following publication. All abstracts were characterized according to a set of pre-specified parameters. Association of these factors with the acceptance at Congress level, with full-text publication, and subsequent citation was assessed by univariable and multivariable regression analyses. Percentages refer to the respective preceding level (n). (n* refers to the representative random selection of 10% (1002) of all abstracts submitted (10 020), Ø published, not published; IF, impact factor).

When adjusting for all factors assessed, prospective non-randomized study design independently predicted publication (adjusted OR = 1.44, $P = 0.032$). Moreover, multivariable analyses confirmed university affiliation to independently predict publication (adjusted OR = 1.62, $P = 0.018$). The number of patients enrolled was not associated with full-text publication, while it was with acceptance at Congress level.

Factors predicting citation

Twenty-one per cent of all studies that had been accepted and subsequently published were cited 10 times or more within the 2 years following publication. In contrast, only 7% of the published studies that had been rejected were cited 10 times or more in the 2 years following publication ($P = 0.008$). Citation rates between published studies that had been accepted for oral presentation did not differ significantly from those that had been accepted for poster presentation ($P = 0.24$). In univariable analyses, the study field ($P = 0.028$) and study design ($P = 0.001$) were associated with high citation rates (Table 4). After adjusting for the remaining factors, study design remained the only parameter independently associated with high citation rates ($P = 0.004$). In particular,

randomized controlled trial design demonstrated a marked independent association with high citation rates (adjusted OR = 6.82, $P < 0.001$), followed by prospective non-randomized studies (adjusted OR = 2.57, $P = 0.018$). Neither academic affiliation nor gender of first or last author, nor the number of enrolled patients influenced citation rates (Table 4).

At Congress level, editorial decisions correlated both with identified predictors of high citation rates, i.e. study design, and with reviewer ratings (Figure 3A). Notably, odds ratios for randomized controlled trial design differed by over three-fold between the acceptance at Congress level and high citation rates. Along this line, there was no significant difference in the number of 2-year citations between studies accepted for oral and those accepted for poster presentation (Figure 3D). Over half of the submitted RCTs were rejected at the ESC congress. However, RCTs still reached the highest acceptance rate of all factors assessed (Table 1). Over 80% of the accepted RCTs were accepted for poster presentation (see Supplementary material online, Table S1). It is consequently not surprising that a few of the most highly cited studies evolved from abstracts originally accepted for poster presentation (Figure 3D).

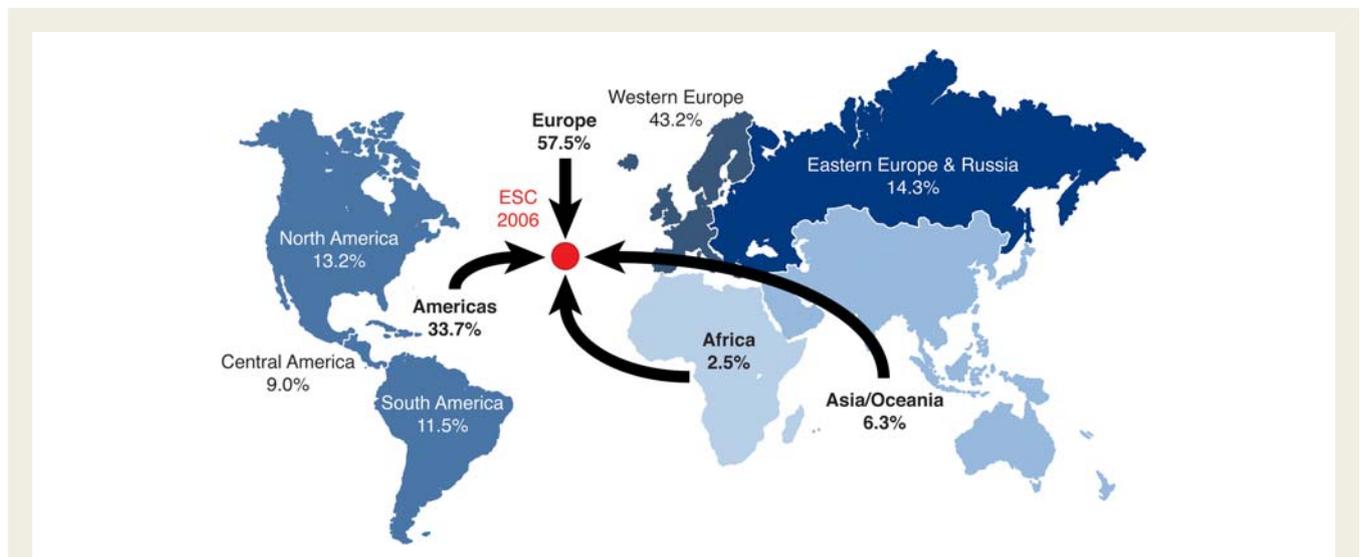


Figure 2 Origin of research submitted to the ESC Congress 2006. A total of 10 020 abstracts were submitted to the scientific sessions of the ESC 2006. Based on the assessment of a representative subsample of 10% ($n = 1002$) submitted studies originated from 63 different countries of all five continents. About one-third of all studies were submitted from the American continent, 43% from Western Europe, and 14% from Eastern Europe and Russia. Less than 10% of studies originated from Africa and Asia/Oceania.

Does female gender affect scientific success?

As noted earlier, these analyses should be considered exploratory given that information on gender was not directly available but had to be empirically determined through an examination of first and last author names. Of all abstracts submitted, 26% were conducted by female first authors (Table 1). Only 13% of senior authors were female, suggesting that both female junior and senior scientists are markedly underrepresented in the cardiovascular research community. Acceptance rates at Congress level did not differ between female and male scientists (Table 1, Supplementary material online, Table S3). Among junior researchers, publication rates did not differ markedly between male and female colleagues (30 vs. 27%, OR = 0.88, $P = 0.428$) (see Supplementary material online, Table S3). However, the publication rate of female senior scientists was markedly lower compared with the one of their male coworkers (18 vs. 31%, $P = 0.004$, Supplementary material online, Table S3). Concomitantly, female senior authorship was independently negatively associated with full-text publication in multivariable regression analysis (adjusted OR = 0.52, $P = 0.010$, Table 3). Scientific impact, as assessed by 2-year citation rates, did not differ between genders (see Supplementary material online, Table S3).

Discussion

In this study, we identified independent predictors of acceptance at the ESC Congress, subsequent full-text publication and later impact through specific assessment and scientometric follow-up of a cohort of abstracts representative for over 10 000 abstracts submitted to an international scientific meeting (Figure 4). Prospective study design determined acceptance at Congress level, full-text publication, and subsequent impact after publication. Randomized controlled trials attracted almost seven times more

2-year citations than other studies. Citation-based validation of the factors predicting acceptance at the Congress level or full-text publication reveal a potential discrimination of female senior authors at publication level and indicate superiority of blinded over open peer review.

Identification of scientific quality at an early stage

Presentation of abstracts at scientific meetings is an important part of the dissemination of knowledge. However, more than half of the studies submitted to scientific congresses are never published,¹⁴ as is also the case for the ESC 2006. Identifying high-quality research based on distinct aspects when grading large numbers of abstracts is a challenging task for reviewers and editors. Reviewers and editors at the ESC Congress successfully identified science of significant later impact, given that 29.6% of all studies accepted for oral presentation received 10 or more 2-year citations compared with only 7.4% of rejected studies. Based on the assessment of the abstracts alone, we identified prospective study design, and specifically randomized controlled study design as strong predictors of scientific impact.

Thus, key aspects of scientific quality are identifiable at the early stage of an abstract, and, even though markedly less pronounced, are taken into account in the review and editorial process of one of the largest international medical conferences.

Hurdles to overcome from abstract to impact

In the process of medical research, different stages are passed from the formulation of a well-substantiated hypothesis to a clearly proven and novel answer, and further to the acceptance of the elaborated results in the scientific community. The first hurdle to overcome is publication. Intriguingly, the factors predicting

Table 1 Factors predicting acceptance at the ESC Congress

Parameters	No. of abstracts submitted [% total (n)]	Acceptance rate [% total (n)]	Univariate reg. OR, 95% CI (P)	Multivariate reg. OR, 95% CI (P)
Overall rates	100.0 (1002)	37.5 (376)		
Type of research				
Clinical	89.7 (899)	36.8 (331)	Ref.	Ref.
Basic	10.3 (103)	43.7 (45)	1.33, 0.88–2.01 (0.174)	2.24, 1.39–3.59 (0.001)
Type of institution				
Not university-affiliated	19.2 (192)	34.9 (67)	Ref.	
University-affiliated	80.8 (810)	38.1 (309)	1.51, 0.15–1.60 (0.403)	
Study design (clinical)			(0.018)	(0.023)
Retrospective	49.3 (494)	32.4 (160)	Ref.	Ref.
Prospective non-RCT	28.9 (290)	42.8 (124)	1.90, 1.16–2.10 (0.016)	1.67, 1.22–2.29 (0.002)
RCT	6.5 (65)	47.7 (31)	1.56, 1.13–2.21 (0.004)	1.91, 1.11–3.28 (0.020)
Meta-Analysis	0.6 (6)	50.0 (3)	2.90, 0.42–10.46 (0.370)	0.0 (1.000)
Systematic Review	0.2 (2)	0.0 (0)	0.0 (1.000)	0.0 (1.000)
Other	4.2 (42)	31.0 (13)	n.a.	n.a.
Number of patients (clinical) ^a				
<100	43.4 (435)	30.3 (132)	Ref.	Ref.
≥100	41.7 (418)	45.7 (191)	1.93, 1.46–2.56 (<0.001)	2.07, 1.54–2.77 (<0.001)
Study field (clinical)			(0.373)	
Cardiac imaging, computational, acute cardiac care	10.3 (103)	44.7 (46)	Ref.	
Rhythmology	12.7 (127)	31.5 (40)	0.57, 0.34–0.96 (0.034)	
Heart failure, left ventricular function, valvular disease, pulmonary circulation	15.2 (152)	36.8 (56)	0.71, 0.44–1.16 (0.173)	
Coronary artery disease, ischaemia	13.2 (132)	31.8 (42)	0.61, 0.37–1.01 (0.055)	
Interventional cardiology, peripheral circulation, stroke	9.4 (94)	39.4 (37)	0.79, 0.47–1.36 (0.390)	
Exercise, prevention, epidemiology, pharmacology, nursing	14.6 (146)	36.3 (53)	0.68, 0.41–1.12 (0.130)	
Hypertension, myocardial and pericardial disease, cardiovascular surgery	9.5 (95)	38.9 (37)	0.83, 0.48–1.42 (0.490)	
Other	5.0 (50)	40.0 (20)	n.a.	n.a.
Gender first author ^b				
Male	72.2 (723)	39.8 (288)	Ref.	Ref.
Female	26.0 (261)	33.7 (88)	0.77, 0.57–1.03 (0.082)	0.73, 0.54–1.00 (0.051)
Gender last author ^c				
Male	84.0 (842)	38.2 (322)	Ref.	
Female	13.1 (131)	38.9 (51)	1.03, 0.71–1.50 (0.88)	

OR, odds ratio; CI, confidence interval; RCT, randomized controlled trial; ref., reference variable for odds ratio calculation; n.a., not applicable. Analyses performed on representative 10% selection ($n = 1002$), margin of error <0.01 . Backward conditional variable exclusion, c -statistic = 0.620.

Bold value indicates statistically significant findings.

^aIn 46 clinical studies, no information on the number of patients was provided.

^bIn 18 studies, no gender for the first author could be identified.

^cIn 29 studies, no gender for the last author could be identified.

publication differed substantially from those predicting high citation rates. In essence, the institution where studies were conducted and the gender of the respective principal investigator predicted publication—rather than the scientific content of a study. Such discrepancy hinders the flow of relevant findings to the scientific

community and slows clinical advancement. Once published, the scientific community itself grades the scientific importance by either spreading the word through citation or not. This second and final step of appreciation by fellow scientists in the respective field was observed to depend only and substantially on research

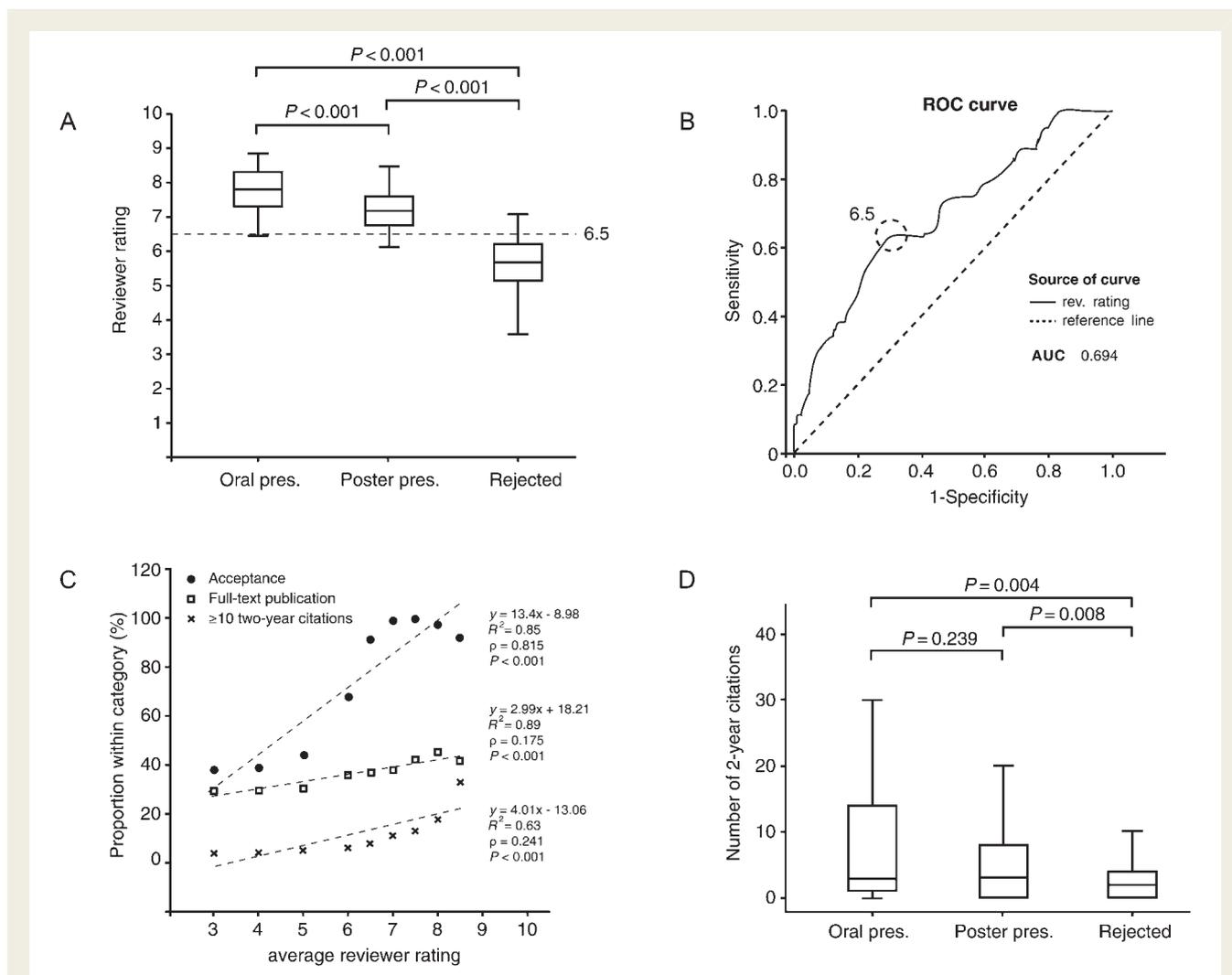


Figure 3 Peer review and editorial process of the ESC Congress 2006. All abstracts submitted were peer-reviewed in a blinded fashion by three to eight expert reviewers, and graded on a scale from 1 to 10. (A) Reviewer ratings and editorial decisions. Accepted studies had received significantly higher ratings than rejected studies; studies for oral presentation had been ranked significantly higher than those for poster presentations. The cut-off for acceptance at 6.5 was calculated using the Yuden's index (sensitivity 0.97, specificity 0.92). (B) Receiver operating characteristic (ROC) curve analyses after the scientometric follow-up of a representative subsample of 10% of all abstracts submitted ($n = 1002$) revealed a predictive value of average reviewer ratings of 69.4%. (C) Spearman's correlation of average reviewer ratings with acceptance at the Congress, subsequent full-text-publication and citation rates uncovered a significant positive correlation in all three cases. (D) Comparison of the numbers of 2-year citations of both accepted and rejected, subsequently published studies. Analyses revealed that rejected and subsequently published studies were cited significantly less frequently compared with accepted and subsequently published studies; citation frequencies between different presentation formats did not differ. AUC, area under the curve, significance level ($P < 0.017$) adjusted for multiple comparisons using Bonferroni's correction, interquartile ranges, whiskers indicate minima and maxima.

design. Considering that in current evidence-based medicine clinical guidelines are set forth based on the respective level of evidence,¹⁵ which is primarily based on study design,¹⁶ this result is encouraging. Moreover, our findings indicate that judging scientific quality of medical research based on other factors, such as the number of patients enrolled, the home institution, or the gender of the performing researchers, may not be justified.

Peer review and editorial decisions

Expert peer review is the gold standard for research evaluation when it comes to publication in peer review journals or allocation

of resources.¹⁷ Given its importance, surprisingly little research has been performed in this field.^{18,19} Abstracts submitted to the ESC Congress 2006 were subjected to ('double-') blinded peer review (neither authors knew reviewers nor vice versa). Up to eight independent reviewers graded scientific quality on a scale from 1 to 10, providing the basis for editorial decision. Information on sound criteria to specifically evaluate when grading scientific quality is scarce.²⁰ In this study, we assessed putative determinants of scientific quality, and identified four aspects that were in the focus of the peer review and editorial process of the ESC Congress 2006. External validation through citation analyses confirmed two

Table 2 Determination of the best cut-off of the average reviewer rating for frequent citation

	Frequency [n (%submitted)]	Accepted at congress [n (% within rating)]	Published [n (% within rating)]	≥10 two year cit. [n (% within rating)]	P-value (Fisher's exact)	OR (95% CI)	AUC (ROC analyses)
Av. Rev. Rating > 6	555 (55.4)	375 (67.6)	199 (35.9)	36 (6.5)	0.036	2.347 (1.04–5.27)	0.584
Av. Rev. Rating > 6.5	394 (39.3)	358 (90.9)	146 (37.1)	32 (8.1)	0.002	3.135 (1.53–6.37)	0.634
Av. Rev. Rating > 7	213 (21.3)	210 (98.6)	81 (38.0)	24 (11.3)	<0.001	4.021 (2.07–7.80)	0.660
Av. Rev. Rating > 7.5	111 (11.1)	110 (99.1)	47 (42.3)	15 (13.5)	0.001	3.491 (1.69–7.21)	0.609
Av. Rev. Rating > 8	33 (3.3)	32 (97.0)	15 (45.5)	6 (18.2)	0.015	4.193 (1.41–12.44)	0.553
Av. Rev. Rating > 8.5	12 (1.2)	11 (91.7)	5 (41.7)	4 (33.3)	0.002	24.700 (2.69–226.65)	0.543
Av. Rev. Rating > 9	Not given	–	–	–	–	–	–

P-values, odds ratios (OR), confidence intervals (CI) and area under the curve (AUC) values refer to the categorical variable of '≥ 10 two-year citations'. Random 10% selection (n = 1002), margin of error <0.01.

Bold value indicates statistically most significant finding in conjunction with the largest AUC.

out of these four aspects to be markers of scientific impact, indicating effective peer review and editorial work at this stage. Moreover, a clear cut-off for average reviewer ratings with regard to acceptance or rejection was identified. Along with significantly higher ratings of abstracts accepted for oral compared with those accepted for poster presentation, these data underscore a coherent cooperation of reviewers and editors. Scientometric follow-up assigned peer reviewer ratings a predictive value of 70% for later scientific impact, further supporting effective peer review and editorial work at Congress level.

Assessment of the predictive value of the same set of parameters for subsequent full-text publication revealed markedly different factors. External validation through citation analyses did not confirm any of the identified predicting factors to be markers of scientific impact: surprisingly, randomized controlled trials design did not predict publication. Moreover, gender and institutional affiliation of the senior author were related to whether or not an abstract led to a full-text publication. As neither gender nor institutional affiliation delineate scientific quality, a potential bias during single-blinded review cannot be excluded.

Blinded peer review and scientific quality

Most scientific and medical journals use 'anonymous' or 'single-blinded' peer review, i.e. the authors do not know the reviewers. However, only a minority applies 'blinded' or 'double-blinded' peer review, where neither authors know reviewers nor vice versa.³ Different kinds of bias are recognized as confounding factors in the effort to judge the quality of research.^{4,17,18,21,22} When the American Heart Association adopted blinded peer review for their Scientific Sessions, Ross *et al.*¹¹ demonstrated that blinded peer review at least partially reduced reviewer bias with regard to prestigious institutions or author origin. However, it remained unclear whether the reduction in bias by adopting blinded peer review would affect later scientific impact. In the present study, separate validation of blinded peer review at Congress level and presumably largely single-blinded peer review at publication level suggests substantial enhancement of scientific quality by decreasing potential bias through blinded peer review. Given these data and other reports that blinded peer

review does not only reduce bias but improve scientific quality,²³ it can be speculated that blinded peer review merits consideration on a broader scale.

Cardiovascular science and gender

The number of women in medical training and entry-level academic positions has substantially increased in recent years, whereas their representation in advanced and leading academic positions has remained comparatively low.^{24–27} One of the empirically best proven hypotheses for this disparity is that women would publish fewer papers and consequently raise less soft money compared with their male co-workers.²⁶ Among all abstracts submitted to the ESC Congress 2006, 26% originated from female first authors, and only 13% from senior female authors. Hence, women appear to be underrepresented in both the entry level and advanced academic positions in cardiovascular medicine.

There were no gender differences with regard to abstract acceptance in blinded peer review, underlining the absence of qualitative differences, as supported by others.²⁵ Interestingly, the rate of full-text publication of female senior authors was significantly lower compared with their male colleagues (19.1 vs. 30.8%). Concomitantly, senior female authorship was identified as independent negative predictor of full-text publication in multivariable regression analysis. The scientometric follow-up did not reveal any gender effects, again supporting the absence of qualitative disparities. Since we did not assess the number of submissions, but only of successful full-text publications, we cannot exclude that female senior authors have submitted their research less frequently for full-text publication. Female underrepresentation, and the fact that citation analyses in this study aimed to assess whether or not a published article was among the top 15% of published articles, limit the interpretation of female scientific performance. However, the discrepancy between blinded peer review at the Congress level and largely open peer review at the publication level in this regard may alert the scientific community to shed light on a potential bias.

Of course, all of the above findings should be interpreted cautiously and considered exploratory, given that information on the gender of first and last authors was not directly available but had

Table 3 Factors predicting publication

Parameters	No. of studies published	No. of studies not published	Publication Rate (% submitted ^a)	Univ. reg. OR, 95% CI (P)	Multiv. reg. OR, 95% CI (P)
Type of research					
Clinical	251	648	27.9	Ref.	Ref.
Basic	42	61	40.8	1.72, 1.13–2.61 (0.012)	2.11, 1.34–3.34 (0.001)
Type of institution					
Not university-affiliated	43	149	22.4	Ref.	Ref.
University-affiliated	250	560	30.9	1.53, 1.06–2.22 (0.025)	1.62, 1.09–2.43 (0.018)
Study design (clinical)					
Restrospective	124	370	25.1	Ref.	Ref.
Prospective non-RCT	90	200	31.0	1.32, 0.96–1.82 (0.090)	1.44, 1.03–2.02 (0.032)
RCT	22	43	33.8	1.53, 0.88–2.65 (0.133)	1.48, 0.84–2.61 (0.181)
Meta-Analysis	4	2	66.7	5.97, 1.08–32.98 (0.041)	0.0 (1.000)
Systematic Review	0	2	0.0	0.0 (1.000)	0.0 (1.000)
Other	11	31	26.2	n.a.	n.a.
Number of patients (clin.) ^b					
<100	120	318	27.4	Ref.	Ref.
≥100	117	298	28.2	1.08, 0.80–1.46 (0.068)	
Study field (clinical)					
Cardiac imaging, computational, acute cardiac care	32	71	31.1	Ref.	
Rhythmology	30	97	23.6	0.71, 0.41–1.25 (0.236)	
Heart failure, left ventricular function, valvular disease, pulmonary circulation	40	112	26.3	0.86, 0.51–1.47 (0.588)	
Coronary artery disease, ischaemia	39	93	29.5	0.94, 0.55–1.61 (0.821)	
Interventional cardiology, peripheral circulation, stroke	27	67	28.7	0.86, 0.47–1.55 (0.610)	
Exercise, prevention, epidemiology, pharmacology, nursing	36	110	24.7	0.76, 0.44–1.32 (0.328)	
Hypertension, myocardial and pericardial disease, cardiovascular surgery	29	66	30.5	1.12, 0.68–2.12 (0.533)	
Other	18	32	36.0	n.a.	n.a.
Gender first author ^c					
Male	217	506	30.0	Ref.	Ref.
Female	71	190	27.2	0.88, 0.63–1.18 (0.352)	0.94, 0.67–1.32 (0.723)
Gender last author ^d					
Male	259	583	30.8	Ref.	Ref.
Female	25	106	19.1	0.54, 0.34–0.85 (0.008)	0.52, 0.32–0.85 (0.010)

OR, odds ratio; CI, confidence interval; RCT, randomized controlled trial; ref., reference variable for odds ratio calculation; n.a., not applicable. Analyses performed on representative 10% selection ($n = 1002$), margin of error < 0.01 . Backward conditional variable exclusion, c -statistic = 0.757.

Bold value indicates statistically significant findings.

^aPercent submitted refers to submissions to the ESC Congress.

^bIn 46 clinical studies, no information on the number of patients was provided.

^cIn 18 studies, no gender for the first author could be identified.

^dIn 29 studies, no gender for the last author could be identified.

Table 4 Factors predicting high citation rates

Parameters	Articles cited 10× or more in 2 years [% of published (n/n)]	Univariate reg. OR, 95% CI (P)	Multivariate reg. OR, 95% CI (P)
Type of research			
Clinical	14.3 (36/251)	Ref.	
Basic	11.9 (5/42)	0.73, 0.27–1.98 (0.537)	
Type of institution			
Not university-affiliated	7.0 (3/43)	Ref.	
University-affiliated	15.2 (38/250)	2.63, 0.77–8.90 (0.121)	
Study design (clinical)			
Restrospective	6.5 (8/124)	Ref.	Ref.
Prospective non-RCT	18.9 (17/90)	3.39, 1.457.90 (0.005)	2.57, 1.20–5.68 (0.018)
RCT	40.9 (9/22)	8.77, 2.96–26.02 (<0.001)	6.82, 2.46–21.12 (0.001)
Meta Analysis	50.0 (2/4)	12.67, 1.59–100.80 (0.016)	16.30, 1.20–281.94 (0.044)
Systematic Review	—	—	—
Other	0.0 (0/11)	n.a.	n.a.
Number of patients (clinical) ^a			
<100	12.5 (15/120)	Ref.	
≥100	17.9 (21/117)	1.33, 0.76–2.67 (0.414)	
Study field (clinical)			
Cardiac imaging, computational, acute cardiac care	17.7 (5/32)	Ref.	Ref.
Rhythmology	13.3 (4/30)	0.64, 0.16–2.53 (0.528)	0.59, 0.14–2.40 (0.475)
Heart failure, left ventricular function, valvular disease, pulmonary circulation	7.5 (3/40)	0.33, 0.08–1.44 (0.142)	0.29, 0.06–1.31 (0.119)
Coronary artery disease, ischemia	30.8 (12/39)	1.96, 0.66–5.84 (0.229)	.79, 0.54–5.37 (0.331)
Interventional cardiology, peripheral circulation, stroke	25.9 (7/27)	1.86, 0.56–2.49 (0.310)	.44, 0.44–5.72 (0.582)
Exercise, prevention, epidemiology, pharmacology, nursing	5.6 (2/36)	0.69, 0.19–2.49 (0.567)	0.71, 0.19–2.95 (0.630)
Hypertension, myocardial and peri-cardial disease, cardiovascular surgery	6.9 (2/29)	0.42, 0.10–1.62 (0.254)	0.39, 0.09–1.90 (0.230)
Other	5.6 (1/18)	n.a.	n.a.
Gender first author ^b			
Male	14.3 (31/217)	Ref.	Ref.
Female	18.3 (13/71)	1.34, 0.66–2.73 (0.423)	0.0 (1.00)
Gender last author ^c			
Male	16.6 (43/259)	Ref.	
Female	4.0 (1/25)	0.22, 0.03–1.66 (0.142)	

OR, odds ratio; CI, confidence interval; RCT, randomized controlled trial; ref., reference variable for odds ratio calculation, n.a., not applicable. Analyses performed on representative 10% selection ($n = 1002$), margin of error <0.01 . Backward conditional variable exclusion, c -statistic = 0.570.

Bold value indicates statistically significant findings.

^aIn 46 clinical studies, no information on the number of patients was provided.

^bIn 18 studies, no gender for the first author could be identified.

^cIn 29 studies, no gender for the last author could be identified.

to be determined using a systematic approach. Given that this approach was applied *a priori* (i.e. before any of the analyses were conducted), however, we believe it is unlikely to lead to any systematic biases and, if anything, misclassification would result in findings shifted toward no differences. That said, the importance of understanding the role of gender in research is

critical and certainly requires further confirmation with additional work.

Limitations

In all regression analyses performed in this study, potential confounding factors that were not assessed or included in the

		<i>prediction</i>	<i>validation</i>
<i>congress</i>	Acceptance		
	1. basic research (OR 2.2, 95% C.I. 1.4 - 3.5)		−
	2. patient number ≥ 100 (OR 2.1, 95% C.I. 1.5 - 2.8)		−
	3. prospective non-randomized study design (OR 1.7, 95% C.I. 1.2 - 2.3)		+
	4. randomized controlled trial design (OR 1.9, 95% C.I. 1.1 - 3.3)		+
<i>publication</i>	Full-text publication		
	1. basic research (OR 2.1, 95% C.I. 1.3 - 3.3)		−
	2. institutional affiliation (OR 1.6, 95% C.I. 1.1 - 2.4)		−
	3. female senior authorship negative association (OR 0.5, 95% C.I. 0.3 - 0.9)		−
<i>citation</i>	Impact		
	1. randomized controlled trial design (OR 6.8, 95% C.I. 2.5 - 21.1)		
	2. prospective non-randomized study design (OR 2.6, 95% C.I. 1.2 - 5.7)		

Figure 4 Key findings: predictors of acceptance at Congress level, full-text publication, and frequent citation, and their validation with regard to scientific impact. We identified parameters predicting acceptance, subsequent full-text publication, and high citation rates (≥ 10 two-year citations). Parameters predicting scientific impact (+), no association with scientific impact (−). All findings, odds ratios, and confidence intervals (CI) refer to multivariable regression analyses based on a random 10% selection of all abstracts submitted to the ESC Congress 2006 (n = 1002), margin of error < 0.01.

respective regression analysis cannot be excluded. Therefore, extrapolations have to be done with care. For example, although no predictors of oral presentation among accepted studies were identified, other aspects such as novelty, which was not assessed in this study, may also play a role. Of note, the presentation format at the Congress level is also subject to organizational aspects, including the arrangement of consistent themes for scientific sessions.

Even though citation analysis is a widely accepted and neutral tool in evaluative bibliometrics to provide quantitative indicators for scientific performance,^{17,28–31} we are aware that it may be subject to bias. An article published in a very renowned journal or authored by a well-known expert may enhance the number of its citations compared with a similar article published in a less renowned journal or authored by a less known scientist. Citation analysis can, of course, not replace but only complement expert peer review in the evaluation of research. However, in the context of this study, we consider citation suitable as an external quality assessment in the evaluation of a large number of abstracts. Due to the extensive amount of studies, all authored by multiple

investigators, self-citation was not assessed. Importantly, no significant association between institutional affiliation and the number of citations of a published article was observed, indicating that this shortcoming is of minor importance for the respective results.

The chosen publication and citation windows, i.e. 2 years each, may not cover all full-text publications among the studies assessed. For the rapid field of biomedical research, however, a publication and citation window of 2 years each has been described to cover the majority of publications and citation peaks.^{9,17,30,31} Furthermore, the assessment of the peer review and editorial process at publication level is based on the assumption that studies that were not published in the 4 years following the Congress had been submitted for full-text publication and been rejected—which was not specifically assessed in this study. In essence, authors of those studies identified as not published were not contacted to verify that the respective studies had been submitted for publication in a peer review journal. Therefore, the results of the logistic regression analysis on the level of publication may be affected by factors that led to non-submission as opposed to reviewer ratings and editorial decisions. Consequently,

extrapolations regarding peer review and editorial process at publication level have to be interpreted with care.

The discussed potential effect of anonymized peer review at publication level is based on the work of others, stating that less than 20% of almost all English language scientific and medical journals use double-blinded peer review.³ Editorial offices of all respective journals in this study sample have not been contacted. As noted above, author gender was assessed using an algorithm based on first- and surnames in addition to photographs where necessary and available. The related findings may consequently be subject to bias based on the indirect methods used.

Generalization of the data of this study, raised in the cardiovascular field, to other medical specialties has to be handled with caution. Nevertheless, given the diversity and quantity of cardiovascular science worldwide as well as the size of the cohort in this study, our findings may be of interest to a broader audience of scientists in different fields of medical research.

Conclusions and possible room for improvement

High quality-research was determined by prospective study design that was readily identified at Congress level. Average reviewer ratings predicted frequent citation rates, indicating an effective evaluation process at Congress level. The scientometric follow-up revealed an unexpected disparity between predictors of publication and citation, suggesting a disadvantage of non-academic institutions and female senior authors at publication level. The observed discrepancy between blinded peer review at Congress level and largely single-blinded peer review at publication level indicates that blinding handling editors and reviewers at publication level may merit a closer look.

Supplementary material

Supplementary material is available at *European Heart Journal* online.

Acknowledgements

We are grateful to Florence Biazeix for administrative assistance, and to Eirini Liova for data collection and management.

Funding

This work was funded by the Foundation for Cardiovascular Research, Zurich, Switzerland, the Swiss National Science Foundation 310030-130626/1 (C.M.M.), and the University Research Priority Program 'Integrative Human Physiology' at the University of Zurich (P.-A.C., T.F.L., C.M.M.). The views expressed in this publication are those of the authors and not necessarily of the funding organizations. Funding to pay the Open Access publication charges for this article was provided by the Foundation for Cardiovascular Research (FCR) in Zurich, Switzerland.

Conflict of interest: none declared.

References

1. The European Society of Cardiology E. ESC Congresses Reports. <http://www.escardio.org/congresses/esc-2011/Pages/welcome.aspx>.
2. The European Society of Cardiology E. Past ESC Congresses. http://www.escardio.org/congresses/past_congresses/Pages/past-ESC-congresses.aspx.
3. Fisher M, Friedman SB, Strauss B. The effects of blinding on acceptance of research papers by peer review. *JAMA* 1994;**272**:143–146.
4. Garfunkel JM, Ulshen MH, Hamrick HJ, Lawson EE. Effect of institutional prestige on reviewers' recommendations and editorial decisions. *JAMA* 1994;**272**:137–138.
5. Olson CM, Rennie D, Cook D, Dickersin K, Flanagan A, Hogan JW, Zhu Q, Reiling J, Pace B. Publication bias in editorial decision making. *JAMA* 2002;**287**:2825–2828.
6. Callahan ML, Wears RL, Weber EJ, Barton C, Young G. Positive-outcome bias and other limitations in the outcome of research abstracts submitted to a scientific meeting. *JAMA* 1998;**280**:254–257.
7. Krzyzanowska MK, Pintilie M, Tannock IF. Factors associated with failure to publish large randomized trials presented at an oncology meeting. *JAMA* 2003;**290**:495–501.
8. Raptis DA, Oberkofler CE, Gouma D, Garden OJ, Bismuth H, Lerut T, Clavien PA. Fate of the peer review process at the ESA: long-term outcome of submitted studies over a 5-year period. *Ann Surg* 2010;**252**:715–725.
9. Scherer RW, Dickersin K, Langenberg P. Full publication of results initially presented in abstracts. A meta-analysis. *JAMA* 1994;**272**:158–162.
10. Toma M, McAlister FA, Bialy L, Adams D, Vandermeer B, Armstrong PW. Transition from meeting abstract to full-length journal article for randomized controlled trials. *JAMA* 2006;**295**:1281–1287.
11. Ross JS, Gross CP, Desai MM, Hong Y, Grant AO, Daniels SR, Hachinski VC, Gibbons RJ, Gardner TJ, Krumholz HM. Effect of blinded peer review on abstract acceptance. *JAMA* 2006;**295**:1675–1680.
12. Bartlett JE, Kotrlík JW, Higgins CC. Organizational research: determining appropriate sample size in survey research. *Info Technol Learn Perform J* 2001;**19**:43–50.
13. Reuters T. ISI Web of Knowledge. http://apps.webofknowledge.com/UA_GeneralSearch_input.do?product=UA&search_mode=GeneralSearch&SID=4B7Nb@C@ED49jmlB51P&preferencesSaved=.
14. von Elm E, Costanza MC, Walder B, Tramer MR. More insight into the fate of biomedical meeting abstracts: a systematic review. *BMC Med Res Methodol* 2003;**3**:12.
15. Glasziou P, Vandembroucke JP, Chalmers I. Assessing the quality of research. *BMJ* 2004;**328**:39–41.
16. Medicine TOCfE-B. *Levels of Evidence-based Medicine*. <http://www.cebm.net/index.aspx?o=1025> (last update 2009; date last accessed).
17. Pendlebury DA. The use and misuse of journal metrics and other citation indicators. *Arch Immunol Ther Exp (Warsz)* 2009;**57**:1–11.
18. Kassirer JP, Campion EW. Peer review. Crude and understudied, but indispensable. *JAMA* 1994;**272**:96–97.
19. Bailar JC 3rd, Patterson K. The need for a research agenda. *N Engl J Med* 1985;**312**:654–657.
20. Abby M, Massey MD, Galanduk S, Polk HC Jr. Peer review is an effective screening process to evaluate medical manuscripts. *JAMA* 1994;**272**:105–107.
21. Relman AS, Angell M. How good is peer review? *N Engl J Med* 1989;**321**:827–829.
22. Link AM. US and Non-US submissions: an analysis of reviewer bias. *JAMA* 1994;**272**:246–247.
23. Laband DN, Piette MJ. A citation analysis of the impact of blinded peer review. *JAMA* 1994;**272**:147–149.
24. Ash AS, Carr PL, Goldstein R, Friedman RH. Compensation and advancement of women in academic medicine: is there equity? *Ann Intern Med* 2004;**141**:205–212.
25. Tesch BJ, Wood HM, Helwig AL, Nattinger AB. Promotion of women physicians in academic medicine. Glass ceiling or sticky floor? *JAMA* 1995;**273**:1022–1025.
26. Kaplan SH, Sullivan LM, Dukes KA, Phillips CF, Kelch RP, Schaller JG. Sex differences in academic advancement. Results of a national study of pediatricians. *N Engl J Med* 1996;**335**:1282–1289.
27. Levinson W, Lurie N. When most doctors are women: what lies ahead? *Ann Intern Med* 2004;**141**:471–474.
28. Garfield E. Citation indexes for science; a new dimension in documentation through association of ideas. *Science* 1955;**122**:108–111.
29. Garfield E. Is citation a legitimate evaluation tool? *Scientometrics* 1979;**1**:359–375.
30. Haefner-Cavillon N, Graillet-Gak C. The use of bibliometric indicators to help peer-review assessment. *Arch Immunol Ther Exp (Warsz)* 2009;**57**:33–38.
31. Moed HF. New developments in the use of citation analysis in research evaluation. *Arch Immunol Ther Exp (Warsz)* 2009;**57**:13–18.