NRIN Research conference 2016

Fostering Responsible Research Practices

Abstract book

Plenary session

P1 Is Dutch science at greater risk?

Daniele Fanelli

Meta-Research Innovation Center at Stanford (METRICS), Stanford University, USA

The Netherlands are leading the way in Research Integrity. In addition to important educational initiatives and research activities, they have taken important policy stances, for example by dropping the "productivity" parameter from the Protocol for Research Assessments. Such laudable attention to integrity has been spurred by the uncovering of prominent cases of scientific misconduct and perhaps by a general perception that pressures to publish in this country are excessive.

Are the Netherlands at greater risk from bias and misconduct compared to other countries? I will present and discuss multiple lines of evidence that may help answer this question. In addition to published data about productivity and retraction patterns, I will present results of a 3-year project funded by the Office of Research Integrity, which analyzed data from published meta-analyses to identify factors that push researchers to select, embellish and falsify their findings. Results are based on a sample of over 3000 meta-analyses from the biomedical and social sciences, from which the reported effect sizes of over 60,000 primary studies were extracted and standardized to measure the amount of over-estimation of individual findings. Data on primary studies was then matched with multiple characteristics of the paper and its authors, including country, collaboration distance, career stage, gender, average publication rate, average impact, etc. These and other sources of evidence will be reviewed to test many hypotheses about risk factors for misconduct and to check whether and why studies by Dutch researchers might have been, in the recent past, more likely to make exaggerated or false claims.

P2 Integrity challenges in Ian McEwan's Solar – on the use of novels in research and education concerning research integrity in the era of big science

Hub Zwart

Radboud University Nijmegen, Faculty of Science, Department of Philosophy, Institute for science, Innovation and Society (ISIS)

Research Integrity has become an important item on the agenda, for research communities themselves, but also for managers, funders and publishers of research. Integrity challenges are often seen as symptomatic for various ongoing transformations concerning the way in which research is conducted and knowledge is produced. This includes the emergence of 'big science' (programmatic, large-scale research, conducted by heavily funded, international research consortia using big technology: big computers, automation, robotics,

etc.). Other important developments are: increased entanglement of research with policy agenda's (so that research outcomes becomes politically sensitive) and of research with industry (public-private partnerships, intellectual property rights, etc.). In other words, scientific research has become a complex endeavour, a challenged profession. Our EU H2020 project PRINTEGER (*Promoting Integrity as an Integral dimension of Excellence in Research*), which was launched in 2015, aims to address emerging integrity challenges in contemporary science. In my own research, I use 'genres of the imagination' (novels, cinema, etc.) as a window into what is happening in science, and in my NRIN presentation I will focus on a particular novel, namely *Solar* by Ian McEwan (2010), about a Nobel laureate who committed plagiarism. In this novel, some of the integrity challenges entailed in contemporary research are casted and enacted in a very telling and credible way. Therefore, it may serve as a magnifying glass for highlighting some of the key aspects of research integrity. I will indicate how novels such as *Solar* can be used for research and education (notably ethics courses) on research integrity.

P3 Non-publication of clinical drug trials is common among phase 1 and single center trials. Results of an inception cohort study

Cornelis A. van den Bogert, Patrick C. Souverein, Cecile T.M. Brekelmans, Susan W.J. Janssen, Gerard H. Koëter, Hubert G.M. Leufkens, Lex M. Bouter

Objective: to investigate the occurrence and determinants of non-publication of clinical drug trials in the Netherlands.
Design: Inception cohort study.
Setting: Clinical drug trials carried out in the Netherlands
Population: Clinical drug trials reviewed by the 28 Institutional Review Boards in the Netherlands in 2007.
Determinants: Trial phase, applicant, centers, drug type, type of study, participant category, prospective registration, sponsor, therapeutic/non-therapeutic, approval status of the drug(s) in the trial, and therapeutic area. The risk ratio (RR; bivariate analysis), crude and adjusted odds ratio (OR; multivariable logistic regression), and 95% confidence interval (CI) were used to investigate the associations between determinants and non-publication.
Main outcome measures: Non-publication as peer reviewed article.
Results: In 2007, 622 clinical drug trials were reviewed by the IRBs in the Netherlands. By the end of observation (January/February 2016), 19 of these were rejected by the IRB, another 19 never started inclusion, 10 were still running, and 102 were terminated early. Of the 472 trials remaining in the analysis, 169 (35.8%) were non-published as peer-reviewed article. The multivariable logistic regression model identified the following determinants with a robust, statistically significant lower likelihood of non-publication: phase 2 (adjusted OR 0.4, 95% CI 0.2-0.9), phase 3 (adjusted OR 0.3, 95% CI 0.2-0.7), and trials not belonging to phase 1-4 (adjusted OR 0.3, 95% CI 0.1-0.6) compared to phase 1 trials; multicenter trials also conducted outside the European Union compared to single-center trials (adjusted OR 0.4, 95% CI 0.2-0.9). Trials that were not prospectively registered had a higher likelihood of non-publication compared to prospectively registered trials (adjusted OR 1.9, 95% CI 1.0-3.4).
Conclusions: Overall, non-publication is common. Phase 1 trials, single-center trials, and no prospective registration are strong determinants of non-publication

P4 Integrity and sustainability in the digital age: Storing, sharing, and documenting digital data by sociologists in the Netherlands

Beau Oldenburg
Department of Sociology, Faculty of Behavioural and Social Sciences, University of Groningen

The digital revolution brought lasting benefits for scientific research, but it also brought new challenges. Perhaps one of the biggest current challenges is how digital data can be handled in an integer and sustainable way. In this small-scale study we conducted in-depth interviews with fourteen sociologists – six

research directors and eight PhD students – of six different Dutch universities and asked them how they store, share, and document their digital data.
Based on these fourteen interviews we conclude that there have been several positive changes in past years, which have been accelerated by the Stapel case. One positive development is that in nearly all sociology departments researchers are requested to make so-called publication packages: digital packages that contain all data, syntaxes, and other important information needed to replicate results that are published in a scientific journal or book (chapter). However, there is still quite some room for improvement. Most respondents store their data on vulnerable and unreliable media (e.g., Dropbox and USB sticks). Moreover, it was unclear how publication packages are checked and what the consequences are when these packages are not in order. With respect to sharing data we found that most researchers are reluctant to share their data with other researchers, even though they are aware that it is a good scientific habit. Finally, only a few of the researchers were familiar with the term Data Management Plan. Respondents did discuss Data Management related topics with their colleagues but did not write them down.
Practical solutions for these problems are discussed. In addition, we discuss changes that recently have been implemented in the sociology department of the University of Groningen based on this report.

Parallel session A1 – Reporting

A1.1 Do trialists hedge their claims sufficiently?: Towards automated detection of overstatement and spin

Gerben ter Riet[1], Sufia Amini[2], Lotty Hooft[3], Halil Kilicoglu[4]

[1] Department of General Practice, Academical Medical Center Amsterdam, The Netherlands

[2] Dept General Practice, Academic Medical Center, Amsterdam, The Netherlands

[3] Dutch Cochrane Center, University Medical Center, Utrecht, The Netherlands

[4] Lister Hill National Center for Biomedical Communications, U.S. National Library of Medicine, USA

Introduction: Linguistic spin overstates claims and misleads. Hedges are phrases that lessen the impact of claims and may change readers' interpretation of a text. For scientific studies, ideally, the weaker a study (given a particular objective), the weaker the justifiable knowledge claims and the stronger the hedging required.
Methods: We sampled 100 Cochrane-reviewed randomized clinical trials and extracted data on risk of bias (ROB) and intervention effect size. Linguistic software generated word-count corrected hedging scores. Hedges were assigned a weight between 1 (weak hedge) and 5 (strong). We calculated 10th and 90th centile reference values for hedging scores and their relation to ROB and decisiveness of results.
Results: 98 RCTs were analyzed. Hedging scores varied between 2.1% (2.1 hedging points per 100 words) and 7.5% (mean 4.5). The proportion of fulfilled ROB items varied between 0 and 100 (mean 56%). Hedging was *not* associated with ROB. Hedging score variability was slightly greater among trials of low quality. Magnitude of effects and their statistical precision were not associated with hedging. The 10th centile of the hedging scores, at ROB score of 100, was 3.2%.
Discussion and Limitations: The absence of an association between study quality, strength of findings and hedging suggests that authors may insufficiently temper their claims given the rigor of trial methodology. We studied randomized trials only and focused solely on the primary outcomes. We cannot exclude that authors of excellent trials hedged too much. Automated detection of overstatement and spin may be useful for users of manuscripts.

A1.2 Reporting Bias in observational epidemiologic research on phthalates

Gerard Swaen, Miriam Urlings, Maurice Zeegers

Research institute CAPHRI Maastricht University, The Netherlands

Background: Observational epidemiology studies on phthalates are diverse, have yielded contradictory results and consequently are difficult to assess. In order to better evaluate these studies we conducted a systematic review including a request for the underlying study protocols.
Methods: A literature search and additional searches yielded 158 eligible journal articles. Corresponding authors were asked to participate in a short telephone interview and to provide a copy of their study protocol. Study characteristics were scored and their association with protocol provision examined.
Results: 47 (29.7%) Corresponding authors agreed to be interviewed and ultimately 22 (14%) provided a copy of the study protocol. Of the 43 publications for which we received information about whether a protocol existed, 16 confirmed there was none and three protocols had been lost. Corresponding authors reporting their study as being positive were three times less likely to provide a copy of their protocol and to participate in the interview (OR=0.31 95% CI: 0.11-0.86). Concordance between the protocol and the publication could not be assessed because of lack of detail in nearly all protocols.
Interpretation: Epidemiology studies on phthalates often lack a protocol and transparency. Given these results we are uncertain whether a formal systematic review of this body of literature will provide reliable risk estimates and we refrain from conducting one. We recommend researchers conducting systematic reviews on observational epidemiology studies to obtain a copy of the underlying protocols prior to conducting the systematic review.
Funding: The European Council of Plasticizers and Intermediates (ECPI)

A1.3 Determinants of selective reporting: a review and content analysis of a random selection of the literature

Jenny T. van der Steen, Cornelis A. van den Bogert, Mirjam C. van Soest-Poortvliet, Soulmaz Fazelifarsani, René H.J. Otten, Gerben ter Riet, Lex M. Bouter

Selective reporting distorts the body of scientific evidence, wastes resources and can harm patients' health and the credibility of science. We assess the nature of possible determinants of selective reporting in the scientific domain using principles of systematic review methods along with principles of qualitative content analysis.
We searched four databases with terms for bias and selection combined with terms for reporting and publication. We reviewed a 25% random selection of eligible records for inclusion. Examining the content of the entire article, we extracted phrases mentioning putative determinants of selective reporting. We documented article and study characteristics, such as academic discipline and study design, and characteristics of the determinant, such as empirical evidence or opinion (viewpoint), and if there was any association with selective reporting. We categorized content of the determinants in an iterative procedure. All steps were performed with at least two researchers.
We included 64 articles listing a total of 502 determinants. Half of the determinants incriminated an actor (most often the researcher or editor) and most could not be interpreted in terms of a single cause (e.g., sample size). The central and most frequently stipulated determinant was a "focus on preferred or significant findings." Other important categories referred to, e.g., conflict of interest and area/specialty.
Our review provides a taxonomy of determinants of selective reporting. It inspires research on mechanisms and informs preventive measures or their study. It may assist in assessing risk of selective reporting which may help inform policy development on responsible research conduct.

A1.4 Impact of reporting bias on estimates of test performance: an empirical study of uterine artery Doppler testing

Gerben ter Riet[1], Nicholas Bodmer[2], Janneke 't Hooft[3], Khalid Khan[4]

[1] Department of General Practice, Academical Medical Center Amsterdam, The Netherlands

[2] Department of Gynecology, Academical Medical Center Amsterdam, The Netherlands

[3] Medignition AG, Verena Conzett-Strasse 9, Post Box 9628 CH-8036 Zürich, Switzerland

[4] The London School of Medicine and Dentistry, London, United Kingdom

Introduction – Biomedical research is plagued by selective reporting of (statistically significant) outcomes. This practice jeopardizes patient care. In diagnostic accuracy and prediction research on a single indicator or predictor, statistical significance plays a lesser role than in trials. However, other preferences may influence authors' inclination to report particular outcomes.
In a review of Doppler ultrasonography for the prediction of pre-eclampsia and intra-uterine growth restriction, we noticed large variation in the numbers of predictive indices reported on in a single paper. Since these Doppler indices are constructed from a few basic measures, all authors could have reported on many of over 15 different indices. We investigated if the strengths of the indices were associated with the number of other indices that were reported on in the same paper.
Methods – Systematic review. We linearly regressed the weighted natural logarithm of the (diagnostic) odds ratio (DOR) of the respective indices against the number of co-reported Doppler indices in the same paper, accounting for confounders and within-study clustering.
Results – 257 odds ratios were calculable from 70 studies. In total, 13 indices were investigated. The median DOR was 7. The median number of co-reported Doppler indices was 2 (range 0-7). The DOR across all indices decreased by 8.3 percent for each additional index reported on (95%CI 0.7 to 15.6%).
Discussion and conclusion(s) – Reporting the results of greater numbers of predictive indices in the same publication was associated with lower predictive performances of these indices. This may reflect reporting bias in prediction research.

A1.5 The impact of 10 years of STARD on the reporting of diagnostic accuracy studies[1]

Daniël A. Korevaar
Department of Clinical Epidemiology, Biostatistics and Bioinformatics, Academic Medical Centre, University of Amsterdam, Amsterdam, The Netherlands

Introduction: In response to increasing evidence of suboptimal reporting of diagnostic accuracy studies, the STARD (STAndards for Reporting of Diagnostic accuracy) statement was launched in 2003. We evaluated the impact of STARD on reporting completeness.
Methods/Results: First, we performed a systematic review and searched Medline, Embase and the Cochrane Library Methodology Register (inception to August 2013) for studies that had evaluated adherence of published diagnostic accuracy studies to the 25-item STARD checklist. We included 16 evaluations, together analyzing the reporting of 1496 diagnostic accuracy studies. The mean number of STARD items reported varied from 9.1 to 14.3 out of 25. Six evaluations compared post-STARD with pre-STARD diagnostic accuracy studies; random-effects meta-analysis revealed a modest but significant increase in adherence after STARD's launch (mean difference 1.41 items (95%CI 0.65-2.18)). Because all evaluations included in the systematic review had been performed in the first few years after STARD's launch, we subsequently performed our own evaluation of adherence to STARD among 112 diagnostic accuracy studies that were published more recently (in 2012) in 12 high-impact-factor journals. The mean number of items reported was 15.3 (range 6.0-23.5). Compared to the findings of two previous evaluations of adherence among diagnostic accuracy studies published in the same 12 journals, this represented an improvement of 3.4 items (95%CI 2.6-4.3) compared with studies published in 2000, and of 1.7 items (95%CI 0.9-2.5) compared with studies published in 2004.
Conclusion: Reporting completeness of diagnostic accuracy studies improved in the 10 years after STARD's launch, but remains suboptimal.

[1]*The results presented in this abstract have been published in the following two articles:*
- Korevaar DA, van Enst WA, Spijker R, Bossuyt PM, Hooft L. Reporting quality of diagnostic accuracy studies: a systematic review and meta-analysis of investigations on adherence to STARD. Evidence-Based Medicine 2014;19(2):47-54.
- Korevaar DA, Wang J, van Enst WA, Leeflang MM, Hooft L, Smidt N, Bossuyt PM. Reporting diagnostic accuracy studies: small improvement after ten years of STARD. Radiology 2015;274(3):781-9.

Parallel session A2 – Methodology and scientometrics

A2.1 Increasing Transparency through a Multiverse Analysis

Sara Steegen[1], Francis Tuerlinckx[1], Andrew Gelman[2], Wolf Vanpaemel[1]

[1] University of Leuven

[2] Columbia University

Empirical research inevitably includes constructing a dataset by processing raw data into a form ready for statistical analysis. Data processing often involves choices among several reasonable options for excluding, transforming and coding data. We suggest that instead of performing only one analysis, researchers could perform a multiverse analysis, which involves performing all analyses across the whole set of alternatively processed data sets corresponding to a large set of reasonable scenarios. Using a worked example focusing on the effect of fertility on religiosity and political attitudes, we show that a single data set can be misleading and propose a multiverse analysis as an alternative practice. A multiverse analysis offers an idea of how much the conclusions change because of arbitrary choices in data construction, and gives pointers as to which choices are most consequential in the fragility of the result.

A2.2 Positive studies are cited twice as often as negative ones: a meta-analysis of citation bias

Bram Duyx[1,2], Miriam J.E. Urlings[1,2], Gerard Swaen[1,2], Lex M. Bouter[3,4], Maurice P. Zeegers[1,2]

[1] CAPHRI School for Public Health and Primary Care, Maastricht University, Maastricht

[2] NUTRIM School of Nutrition and Translational Research in Metabolism, Maastricht University, Maastricht

[3] Department of Epidemiology and Biostatistics, VU University Medical Center, Amsterdam

[4] Department of Philosophy, Faculty of Humanities, VU University, Amsterdam

Introduction: Citation bias, the selective citation of previous literature based on its outcome, can distort the evolution of knowledge and has been studied in different fields. In this systematic review we bring together all evidence and quantify the pooled impact for the first time.
Method: An extensive search strategy was developed and applied to the Web of Science Core Collection. Study outcome was operationalised in several ways. For each operationalisation a random effects meta-analysis was performed on those articles that contained sufficient information to pool count data.
Preliminary results: We identified 29 articles across disciplines, mostly biomedical. Random effects meta-analyses show that statistically significant studies are cited almost twice as often as non-signficant ones, and that studies supporting a specific hypothesis are cited more than two times as often as non-supportive ones.
Discussion: Positive studies are on average cited twice as often as negative studies. It seems likely that this imbalance threatens the valid evolution of knowledge.

A2.3 The integrity of peer review

Jelte M. Wicherts

Department of Methodology and Statistics, Tilburg School of Socials and Behavioral Sciences, Tilburg University

Peer review is a core method of quality control in the sciences, but several large-scale surveys and scandals have highlighted that it does not always function optimally. For instance, substandard peer review in some Open Access journals (as shown by John Bohannon submitting fake papers), reviewing stings in which researchers reviewed their own submissions, potential nepotism or bias in peer review, peer review's role in maintaining publication bias, and cases of scientific misconduct that were not uncovered by peer reviewers have sparked debate on how to improve the quality of peer review. Here I discuss ways in which transparency of the peer review system could help improve accountability and quality in peer review. I present recent work we have done to rate the transparency of the peer review system at academic journals and discuss ways to strengthen the system, including the use of automated tools to detect errors in the reporting of statistical results that appear to be particularly common in my own field of psychology.

A2.4 Researcher's Intuitions about Power in Psychological Research

Marjan Bakker, Chris H.J. Hartgerink, Jelte M. Wicherts, Han L.J. van der Maas
Tilburg University

Because of relatively small effect and sample sizes, many psychological studies are statistically underpowered. Yet most published studies are apparently not based on formal power analyses. Even if researchers understand the importance of well-powered research designs, their intuitions about power might be incorrect. In two studies, we surveyed a total of 505 psychological researchers concerning their power intuitions and found large discrepancies between the preferred amount of power and the power calculated based on their typical sample size, effect size, and alpha. Furthermore, 89% of the respondents overestimated the power when asked to estimate the power of specific research designs, and 95% underestimated the sample size to obtain 80% power for studying small effect sizes. Neither experience nor knowledge predicted bias in self-reported power intuitions. Because many respondents based sample sizes on rules of thumb or common practice in the field, we recommend the reporting of formal power analyses.

A2.5 Which factors drive citation? A Citation Network Analysis of trans fatty acid literature

Miriam J.E. Urlings[1,2], Bram Duyx[1,2], Gerard Swaen[1,2], Lex M. Bouter[3], Maurice P. Zeegers[1,2]
[1] CAPHRI School for Public Health and Primary Care, Maastricht University, Maastricht
[2] NUTRIM School of Nutrition and Translational Research in Metabolism, Maastricht University, Maastricht
[3] Department of Epidemiology and Biostatistics, VU University Medical Center, Amsterdam
[4] Department of Philosophy, Faculty of Humanities, VU University, Amsterdam

Introduction: Referring to different study outcomes in an unbalanced manner, can lead to skewed knowledge development and biased scientific consensus. This phenomenon, where selected parts of the study results are overrepresented in the list of references, is called citation bias. Currently, it is not known how authors select their citations. This study aims to identify which factors influence the likelihood that an article gets cited. As an example, the literature on trans fatty acids and cholesterol will be subject of this citation network analysis.
Methods: A citation network is identified, comprising all articles on the relationship between trans fatty acids and cholesterol, including observational studies, trials, reviews and opinion papers. Each article is scored on potential determinants of citation: study outcome, hedging, journal impact factor, number of affiliations, study design, sample size, funding source, authority of the authors, country, affiliation, gender, language, self-citation, title of the publication, and number of references. The unit of analysis is the potential citation relation in the network, which are either utilized or not. Random effect logistic regression will be used to relate these study characteristics to the chance of being cited two or more years from its publication date.
Results: Currently, analyses are being performed. Preliminary results will be presented at the conference.
Conclusion: To prevent citation bias and promote responsible research conduct, more insight in citation behaviour is required. This research will show the benefit of Citation Network Analysis as a proper tool to gain this insight.
Keywords: citation bias, questionable research practice, citation network analysis, trans fatty acids, meta-research

Parallel session A3 – Policy & data management

A3.1 Roles and responsibilities of a Medical Research Ethics Committee (MREC): A qualitative study investigating views and experiences of committee members

Benjamin Drukarch[1], Mariëlle Diepeveen[2], Jeroen J.G. Geurts[1], Wieke E. van der Borg[2]; Guy A.M. Widdershoven[2]
[1] Department of Anatomy and Neurosciences, VUmc, Amsterdam, The Netherlands
2 Department of Medical Humanities, VUmc, Amsterdam, The Netherlands

Background: In the Netherlands, medical research with human subjects is governed principally by the law on Medical-Scientific Research with Human Subjects (Dutch abbreviation: WMO). Within the WMO, protection of rights, safety and well-being of research participants is a primary concern which is entrusted to MREC's. MREC's in other countries perform similar tasks. Accordingly, MREC's are expected to judge on admissibility of research proposals by carefully weighing (potential) benefits against (potential) risks of participation and to follow-up on results.
The work of MREC's is relevant for research integrity. First and foremost, MREC's guard the ethical quality of research by balancing burdens and benefits. As part of this process, MREC's judge the methodological quality of the proposal to establish whether inadequate research would impose an unnecessary burden on research participants, and thus is unethical. Yet, besides protecting the interests of research subjects, evaluation by MREC's is also intended to increase research quality by "educating" investigators and foster their methodological and ethical expertise. These roles are discussed at length in literature. Surprisingly little, however, is known about the way in which MREC members themselves view and value their roles and responsibilities?
Aims: Aim of this pilot study, is to obtain insight into the views and experiences of individual MREC members concerning the role of an MREC as impartial judge of admissibility of medical research proposals and their own position in this process.
Method: Approximately 10 members of the MREC -VUmc will be interviewed individually using a prepared topic list. In a focus-group, for which 6-8 interested members of the MREC will be invited, the themes and issues identified by an iterative analysis of the outcome of the interviews will be validated and discussed in depth.

A3.2 Conflict of interests between publication embargo, patient safety and study progress: example case

Denhard J. de Smit[1,2], Martina C. Cornel[1,2]
[1] Community Genetics & Public Health Genomics, Department of Clinical Genetics, VU University Medical Center, Amsterdam
[2] EMGO Institure for Health and Care Research, VU University Medical Center, Amsterdam

During the execution phase of a clinical trial we encountered a situation in which we had restricted access to new research data that were relevant for us. A publication had been submitted on a potential safety risk. However, the publication was under embargo. Firstly this limited our abilities to maintain integrity in the evaluation of new research data concerning the safety of our trial. Secondly, it hampered the progress of our recruitment and threatened our financial resources.
The situation was as follows. In our multicenter study group one of the parties (X) was involved in a study on comparable relations between the exposure and outcome that was the object of our trial. The study results indicated a negative effect of the exposure. X informed the project leaders and the Medical Ethical Committee of this outcome, but did not provide the submitted publication or the data set. The main investigator was appointed to a non-participating research centre. This prohibited the researcher to communicate full details before acceptance of the publication.
We will present practical details of how we dealt with this situation and make recommendations how to deal with these potential conflicts in the future.

A3.3 Research Integrity Clinical Research Management System

Áine Honohan, J.G. van der Bom
Center for Clinical Transfusion Research, Sanquin Leiden

Background: In order to facilitate the conduct of research integrity within our clinical research department a research integrity clinical research management system (RICRMS) was set up under the guidance of the recently founded local scientific integrity committee (SIC). This system will automatically imposes the conduct of Good Clinical Practice and stimulate the principles of FAIR (Findable, accessible, interoperable re/usable) data within.

Method: After the appointment of a research integrity officer the RICSMS was launched on January 1st 2015. Research grants proposal, research protocols and articles are entered to the RICRMS generating a unique RICRMS number, which is used in all following documentation and correspondence. The scientific board (re)-views research at a number of designated stages defined within our RICRMS; prior to submission to grant funding boards and medical ethics boards, articles are (re)-viewed in proposal form before statistical analysis and prior to submission for peer review. After publication in a peer-reviewed journal the authors are required to complete a research integrity checklist and return datasets syntaxes, revisions and reviewers comments for archive. An annual audit is carried out in which a table or figure from a randomly selected article published in that year will be reproduced.
Results/Conclusion: Implementation of the system requires a change in research behavior and therefore takes time to implement. The assignment of a research integrity officer is essential for implementation success. The development of such a system forces self-reflection and exposes crevices in the current system. The RICRMS is being continually updated as unforeseen situations arise.

A3.4 Biomedical Research Integrity in China-A Systematic Review of Empirical Research

Nannan Yi[1], Benoit Nemery de Bellevaux[2], Kris Dierickx[1]

[1] University of Leuven, Department of Public Health and Primary Care, Centre for Biomedical Ethics and Law, Leuven, Belgium

[2] University of Leuven, Department of Public Health and Primary Care, Centre for Environment and Health, Leuven, Belgium

Background: With a growing number of scientific publications retracted and other scandals reported in recent years, Chinese biomedical research integrity is questioned. This presentation aims to: (1)provide more evidence to answer if Chinese biomedical scientific misconduct is so prevelant as thought, (2)explore the deeply related factors of biomedical scientific misconduct in China and (3)formulate suggestions to promote research integrity and responsible conduct of research in Chinese biomedical research.
Methods: A literature search of Chinese and international databases was conducted with keywords in Chinese or English. Studies were included if empirical, peer-reviewed and original. Based on these studies, we had a comprehensive overview of the status quo of research integrity and misconduct in Chinese biomedical research.
Results: We found that research misconduct, including fabrication and falsification of data, plagiarism, inappropriate authorship and duplicate submission could also be found in China. Furthermore, the reasons related to the scientific system, Chinese culture and other important aspects were explored in our work.
Conclusion: China is one of those countries facing the problem of research misconduct. In order to promote research integrity in Chinese biomedical research, improvement of training, evaluation metrics and other fields should be sought.

A3.5 Peer reporting of academic integrity violations

Gjalt de Graaf

From previous research in the public sector (e.g. De Graaf and Huberts, 2008), we know that in the direct surroundings of integrity violators, there are often signals about the wrong behavior; peers of corrupt officials often had suspicions – sometimes even evidence – of something wrong long before the investigation, but kept the information to themselves.
In my presentation at the first NRIN Research Conference I would like to report on recent research I conducted on reporting systems (De Graaf, 2015 (forthcoming)). The respondents from this study are from all kinds of public sector organizations, including Higher Education.
Whistleblowing and Whistleblowers have received a great deal of attention over the last decade in academic research. By now we know quite a bit about the whistle*blower*. Whistle*blowing,* and in particular internal reporting systems, has received less attention. Here the focus is on one possible part of an internal reporting system: the *confidential integrity advisor* (CIA). The research question is: *What is the most effective internal reporting system and what role should the confidential integrity advisor play?* The analysis is based on a

mixed-methods design. CIAs can play an important role in an effective reporting system and addressing the role of supervisors is important.
In my presentation I would like to present this study and apply it further to the academic context. How can we best lower thresholds in academia, in order for researchers to come forward with information about wrongdoing?

Parallel session B1 – Scientific (mis)behaviours

B1.1 The storybook image of the scientist

Coosje Veldkamp
Department of Methodology and Statistics, Tilburg School of Socials and Behavioral Sciences, Tilburg University

Do lay people and scientists themselves realize that scientists are human and therefore prone to human fallibilities such as error, bias, and even dishonesty? We sought to answer this question in a series of four studies (N = 3,752). In the first two studies, we found that both lay people and scientists attributed much more objectivity, rationality, open-mindedness, intelligence, integrity, and communality to scientists than to than other highly-educated people. Strikingly, scientists perceived a larger difference between scientists and other highly-educated people than lay people did. In two subsequent studies we examined whether these results may be related to in-group bias by studying whether scientists are prone to in-group favoritism in terms of professional level and gender. Here we found that established scientists and female scientists displayed signs of in-group bias. In addition, established scientists had a particularly negative view of early-career scientists: not only did they attribute much less objectivity, rationality, open-mindedness, integrity, and communality to early-career scientists than to established scientists, but they also attributed less of these features to early-career scientists than to PhD students. We discuss our results and their implications for scientists' willingness to adopt recently proposed research practices aimed at reducing human factors in science.

B1.2 Ranking importance of research misbehaviors

Lex Bouter[1,2], Nils Axelsen[3], Gerben ter Riet[4], Joeri Tijdink[5,6]
[1] Department of Epidemiology and Biostatistics, VU University Medical Center, Amsterdam, The Netherlands
[2] Department of Philosophy, Faculty of Humanities, VU University
[3] Statens Serum Institut, Copenhagen, Denmark
[4] Department General Practice, Academic Medical Center, Amsterdam, The Netherlands
[5] Department of Internal Medicine, VU University Medical Center, Amsterdam, The Netherlands
[6] Department of Psychiatry, Tergooi Hospital, Hilversum, The Netherlands

Background: Fabrication, falsification and plagiarism (FFP) are the officially recognized types of research misconduct and attract much attention. However, there is other research misbehavior and no agreement which is most detrimental.
Objective – Rank various research misbehaviors on their importance.
Methods: All attendees of the previous four World Conferences on Research Integrity were invited to complete a 60-item web-survey. Three subgroups of randomly selected attendees received 20 randomly selected misbehavior items, rating prevalence, effect on the validity of findings, effect on trust between scientists, and preventability (5-point scales). Impact was calculated as (prevalence × effect). We based rankings on means.
Results: 1345 emails were sent, 693 were opened. 227 (33%) persons responded. 46% worked in universities or hospitals, 58% in biomedicine. The top 5 misbehaviors in terms of impact on truth finding were insufficient supervision or mentorship of junior coworkers, insufficient reporting of study flaws and

limitations, inadequate note keeping of the research process, turning a blind eye to putative breaches of research integrity by others, and ignoring basic principles of quality assurance. Plagiarism ranked 1st on impact on trust. Complete fabrication, due to low estimates of occurrence, ranked 34th and 39th on impact on truth and trust, respectively, but 1st on priority.
Limitations: The survey had a low and possibly selective response and reflects subjective judgments. Some ranks hide small differences in mean values.
Conclusions: The top five items in terms of truth finding featured behaviors related to accuracy of procedures, accountability, supervision and honest reporting.

B1.3 A qualitative investigation of "honest" retractions by researchers in the Netherlands and other countries

Mohammad Hosseini[1], Medard Hilhorst[2], Inez de Beaufort[2], Daniele Fanelli[3]

[1] Master student at Utrecht University, The Netherlands

[2] Erasmus Medical Center, Rotterdam, The Netherlands

[3] Meta-Research Innovation Center at Stanford (METRICS), Stanford University, USA

Authors who retract their own papers following the discovery of an honest mistake show a highly proactive attitude with regards to their professional and moral responsibilities. In this qualitative study, we explored the circumstances, motivations and beliefs of scientists who retracted one of their own publications between 2010 and 2015. Scientists were selected primarily amongst Dutch researchers. Through semi-structured interviews, we sought to understand the experience of discovering an error and retracting one's own publication.
For all the participants, the process of retracting a publication had been one of the most stressful and pressing moments of their professional career. Nonetheless, in many cases having retracted a paper turned out to be a cause of praise for these scientists by their colleagues. We conclude that scientists' motivations to retract included a combination of prudential and moral considerations. However, the very act of communicating their mistake to the journals constitutes a realization of ethical and professional responsibilities, which should be praised and promoted.
This study offered preliminary insights, which may help the future promotion of scientific integrity in research and publication.

B1.4 The personality of fraudsters; a cross sectional survey among biomedical scientists

Joeri Tijdink[1], Lex Bouter[1], Coosje Veldkamp[2], Peter van de Ven[1], Jelte Wicherts[2], Yvo Smulders[1]

[1] VU University Medical Center

[2] Department of Methodology and Statistics, Tilburg School of Socials and Behavioral Sciences, Tilburg University

Background: Personality influences decision making and ethical considerations. Its influence on the occurrence of research misbehavior has never been studied. This study aims to determine the association between personality traits and self-reported questionable research practices and research misconduct. We hypothesized that narcissistic, Machiavellianistic and psychopathic traits, and self-esteem are associated with research misbehavior.
Methods: In a cross-sectional study design, we included 535 Dutch biomedical scientists (response rate 65%) across all hierarchical layers of 4 university medical centers in the Netherlands. We used validated personality questionnaires such as the Dark Triad (narcissism, psychopathy and Machiavellianism), Rosenberg's Self-Esteem Scale, the Publication Pressure Questionnaire (PPQ), and also demographic and job-specific characteristics to investigate the association with a composite research misbehavior severity score.
Findings: Machiavellianism was positively associated (beta 1.28, CI 1.06 - 1.53) with self-reported research misbehavior, while narcissism, psychopathy and self-esteem were not. Exploratory analysis revealed that, among persons in higher academic ranks (i.e., professors), narcissism and research misconduct were more severe ($p<0.01$ and $p<0.001$, respectively), and self-esteem scores and publication pressure were lower ($p<0.001$ and $p<0.01$, respectively) compared to postgraduate PhD fellows.

Conclusions: Machiavellianism may be a risk factor for research misbehaviour. Narcissism and research misbehaviour were more prevalent among biomedical scientists in higher academic positions. These results suggest that personality will have impact on research behavior and should be accounted for in fostering responsible conduct of research.

Parallel session B2 – Fostering RI

B2.1 Fostering a research integrity environment @ Ghent University

Stefanie van der Burght
Ghent University

2015 was the year of the big 'research integrity'-launch at Ghent University; with a policy plan the research integrity advisor started promoting our institutions vision on the topic. From the start it was clear we didn't want to focus on the negative story of an anti-fraud policy focused on a small group of researchers who misbehave and how we have to catch them. We firmly chose the positive story of a quality policy aimed at the large group of researchers doing their best every single day. The implementation of the plan turned out to be a lot more than putting a number of initiatives into practice. Our aim is to allow all members of staff and students to fully experience the core values of research integrity in their day-to-day activities, create a second nature. Changing the minds by touching the harts.
This presentation is a first 'look back' at our achievements with specific attention to the impact of initiatives. What have we accomplished, what have we learned, how to proceed. We focus on the experience from a researchers perspective since they have key roles in the implementation success of the plan.

B2.2 Promoting integrity in academia: experiences from an interactive policy process

Erwin van Rijswoud, Leo Huberts
Department for Political Science and Public Administration, Faculty of Social Science, VU University Amsterdam

In promoting integrity in science, attention has shifted from implementing checks and balances for the individual researcher towards promoting an organizational environment that fosters integrity in multiple ways. This shift co-occurred with a changed understanding of how integrity violations in academia come about (Anderson 2013) and should be prevented.
From this perspective and a research tradition of studying violations of integrity in other sectors, we have set up a policy oriented research project on academic integrity. This project addresses integrity in science, education and governance in an integral and interactive fashion, and is both executed at and focused on the Faculty of Social Science at the VU university Amsterdam.
In the presentation we will present our preliminary results of a survey on academic integrity, distributed in the Gamma domain of the university. We will address the following questions:

- Can integrity in research, education and governance of a faculty be studied and stimulated with similar tools and means, or is a more tailored approach mandated by the results?
- Are there notable differences between the three faculties in the survey, and between academic staff and support staff? How can these differences be interpreted?
- Are reported cases of misconduct comparable to cases of misconduct in other public sectors (Viz Huberts et al 2009)?

The presentation will conclude with our view on how, with these results available, incentives for fostering integrity can be developed.

B2.3 Using Moral Case Deliberation to foster responsible conduct of research

Laura A Hartman[1], Bert Molewijk[1,5], Fenneke Blom[2,3], Lex M Bouter[2,3,4], Guy Widdershoven[1,4]
[1] Department of Medical Humanities, VU University Medical Center

[2] Department of Epidemiology and Biostatistics, VU University Medical Center
[3] Netherlands Research Integrity Network
[4] Department of Philosophy, Faculty of Humanities, VU University
[5] Centre for Medical Ethics, Oslo

Background: Research integrity has received growing attention in the last decades. Until now, initiatives to foster responsible research practices have focused on the development of codes of conduct for researchers (i.e. providing clarity on what is expected) and on education on responsible conduct of research. However, several studies on improving integrity of professionals in other domains (health care ethics, military ethics, business ethics), show that the availability of rules and guidelines is usually not sufficient for increasing awareness and changing practice. Rules and guidelines do not support the dealing with moral dilemmas and questions regarding actual practices. Many of these dilemmas originate in conflicts of interest and rewards systems that stimulate other behavior than the rules and guidelines prescribe.
Moral Case Deliberation: Moral case deliberations (MCDs) are an established technique to stimulate moral learning of health care practitioners, making them more aware of the moral dilemmas they encounter and conscious of different action strategies. MCD is a reflective dialogue, in which, through a structured method, a concrete moral issue is analysed by a group of professionals in order to come to a shared moral perspective and a deepened insight as to which values and norms are underlying possible courses of action. This dialogue is moderated by a trained facilitator, who guides the joint reflection. Evaluation research shows that MCD is a fruitful method for clinical ethics support. The method is already used in a course on research integrity for PhD students of the VU University Medical Center.
The content of MCDs can be a rich data source for the actual moral dilemmas professionals experience with respect to research integrity in a particular research department. We therefore propose to analyze the content of series of MCDs within 5 departments and to develop for each department a tool that is context sensitive and which supports researchers in dealing with moral dilemmas concerning research integrity and responsible conduct of research (for instance a checklist for making certain decisions, a moral compass, change in consultation structure etc.). Which tool is needed is highly dependent on the specific conditions of that department.
Aim of the proposed study:

- To support researchers at research departments when dealing with moral questions regarding research integrity and to improve their moral competency;
- To stimulate a dialogue within 5 research departments on these moral dilemmas;
- To develop 5 context sensitive tools based on these MCDs related to one of the central moral issues supporting responsible conduct of research.

Methods: We will organise a series of eight MCD's (over one year) in 5 research departments with the involved researchers. The MCDs will be tape-recorded and transcribed. Through a thematic content analysis we will analyse the moral dilemmas, the dealing with these dilemmas, and the specific research culture that is typical for that research department. We will discuss the results of the overall analyses in a focus group with the departments and explore how and with what kind of tools the researchers would feel supported with the presented moral dilemmas. Besides this, we will organize one focus group with representatives of all 5 departments in order to come to a deepened insight in the differences and similarities between the departments (in total 6 focus groups). There is a great chance that the developed tools will be relevant for other departments, but this cannot be guaranteed beforehand (because of its intended context sensitivity).

B2.4 Are the R.I.Ch. getting richer? Lessons learnt from the in-house development of a research integrity workshop

Dieter de Bruyn, Nele Bracke, Katrien de Gelder, Stefanie van der Burght
Ghent University

One of Ghent University's strategies to further foster the responsible conduct of research is the organization of generic training for all those involved in research (PhD students, postdocs, professors, administrative and

technical staff, etc.) In 2014, four research coordination officers were trained by an external partner experienced in helping organizations to develop internal trainings ('train-the-trainer'). This project resulted in 'Fostering Responsible Conduct of Research', a one-day interactive workshop addressing researchers from all disciplines and in all stages of their research careers.
What characterizes the workshop is that it (1) is fully aligned with Ghent University's overall research policy and regulations; (2) encompasses the full scope of research integrity; (3) focuses on common examples and best practices rather than on exceptional cases of serious misconduct. At the end of the training the participants are expected to be more proficient in recognizing research integrity issues, more confident in responding to them, and fully ready to become research integrity champions (R.I.Ch.) within their research environment.
The proposed presentation will first briefly discuss the process of developing the one-day interactive workshop on research integrity and the training format that resulted from the train-the-trainer. The second part of the talk will focus on the lessons learnt from the first six sessions of the workshop. Issues to be dealt with here include measuring impact in the long term, customizing the format to better meet the needs of the research community, and increasing the sustainability of the workshop (e.g. by bringing it to the international academic market).

Parallel session B3 – The RI concept

B3.1 Deviance in science – Towards a Criminological Understanding

Marijke van Buggenhout, Jenneke Christiaens
Department of Criminology, Vrije Universiteit Brussels, Belgium

Scientific misconduct - either in the format of fraud or questionable research practices - forms a great threat to research integrity. High profile misconduct in the scientific enterprise has been neglected in criminological studies, not unlikely because "*deviance in one's own family is the last to be recognized or acknowledged*" (Bechtel & Pearson, 1990, p. 667). How can it be explained that highly intelligent, educated, honored and respected academics deliberately choose to gain prestige in a deviant way?
In this presentation I argue that a criminological understanding of deviance in science is necessary in order to explore the ecology of today's scientific practice as a knowledge economy, and the elements that are catalyst for misconduct.
Our starting point lies in analyzing concepts such as integrity and misconduct, taking into account definitional issues and different levels of conflicting organizational, institutional and cultural values in modern science. Further our aim is to contextualize scientific fraud or misconduct in relation to corporate, governmental, professional and/or white collar crime - continued by an exploration of the applicability of criminological insights (social control, rationalization, strain, …) to the understanding of deviance in science. Our interim findings based on a literature study point to the added value that criminological understanding, hitherto lacking, offers in broadening the perspective on scientific misconduct. Benefits of this approach will be to understand both the causes and motivations of scientific misconduct on an interdisciplinary level and can be applied for prevention strategies and policy in higher education.

B3.2 The Irreducible Plurality of the Value Basis of Research Integrity

Rik Peels, Jeroen de Ridder
Department of Philosophy, VU University Amsterdam

Guidelines for research integrity and good research practices, such as the European Code of Conduct for Research Integrity, contain all sorts of principles and norms regarding honesty and diligence, reliability, replicability, independence, responsibility for future generations, impartiality, open communication, avoidance of fabrication, plagiarism, and falsification, and so forth. This paper argues that such aspirational

research guidelines and concrete norms do *not* derive from a single value basis, but from an irreducible plurality of values.
This paper focuses on *epistemic* values, such as knowledge and understanding, rather than *societal* values, such as the wellbeing of future generations, or *moral* values, such as honesty and intellectual modesty. It is argued that even within the fairly narrow set of epistemic values, we find an irreducible plurality of values from which issue research responsibilities. This is illustrated by showing that such distinct epistemic values as knowledge, truth, rationality, reliability, understanding, and avoiding falsehood lead to somewhat different research principles and that, remarkably, these principles can even, at least potentially, be in conflict with each other. Finally, it is shown that unless we acknowledge and continue to explore this irreducible plurality of epistemic values that underpin research principles, we cannot do full justice to the variety of research responsibilities that both individual scientists and research groups nowadays bear.

B3.3 Value conflicts in academic teaching

Gjalt de Graaf

Just as in public service delivery and in other professions, there are conflicting values (and loyalties) in academic education on all levels, including the Graduate School. Integrity rules are not always clear and choices are not always between good and bad. Many professors will recognize the dilemma between the values of professionalism and of collegiality. Everyone will endorse these two values as important (intrinsic) in the academic context, yet professors sometimes find them conflicting.
The central research question here is threefold: which value conflicts do academic (VU) teachers perceive, what strategies are used to deal with these conflicts, and what are the advantages and disadvantages of these strategies?
Value conflict in itself is not a problem; perhaps value conflicts bring forth change for the better through innovation and alertness. Plus value conflict is unavoidable, it is a fact of academic life. Coping strategies (or coping mechanisms as they are also called in the literature) should prevent a state of paralysis for those who face intrinsic value conflicts.
In the proposed explorative study a survey will be administered in the month of April among all those who work in the gamma domain of the Vrije Universiteit Amsterdam, on the perceived dilemmas (/difficult decisions) and value conflicts, including in PhD supervision (e.g. authorship of publications). At the 1st NRIN Research Conference, I hope to present the first results.

B3.4 Understanding the various meanings of 'scientific integrity'

Serge Horbach

Integrity is widely considered to be an essential aspect of research, but there seems to be little consensus about the definition or exact meaning of 'scientific integrity'. The understanding of integrity ranges from the minimal (FFP) to the maximum, blending into science ethics, but underneath this obvious range are more subtle differences that are not immediately obvious. The absence of a clear and commonly held understanding of integrity in research is sometimes believed to hamper the promotion of scientific integrity and the prevention of misconduct.
Rather than performing a conceptual analysis through philosophical reasoning and discussion, we aimed at clarifying the discourse of 'scientific integrity' by studying its usage in daily practice. To this end, large numbers of scientific publications, policy documents and newspaper articles were analysed by means of scientometric- and content analysis techniques. We sampled articles referring to integrity in science from the Web of Science database, European policy documents and worldwide English newspapers. The texts were analysed on their usage of the term 'integrity' and of frequently co-occurring terms and concepts. A comparison was made between the usage in the various media, as well as between different temporal periods in which they were published.

From the analysis we conclude that there are clear differences between the discourse of 'integrity' in policy documents and scientific publications as well as between distinct temporal periods. Remarkably we see a clear difference between the tendency of 'promoting good science' in scientific publications and older policy documents, while there is the tendency of 'repressing misconduct' in more recent policy documents.