04/2013

UNCOVER is an FP7-funded project under Contract No 282574

AIT
DUK UNC

Edgar Schiebel
Maria-Elisabeth Züger

Bibliometric features of publication bias in clinical trials

UNCOVER project deliverable D2.2

This report should be cited as follows:
Schiebel, E., Züger, M.-E., Deliverable D2.2 of the UNCOVER FP7-funded project under contract number 282574: Bibliometric features of publication bias in clinical trials.

D2.2 is a deliverable of the work package (WP2) "Quantitative analysis to reveal characteristic bibliometric features of the publication bias" of the UNCOVER project. WP2 addresses bibliometric data and measurable differences of "bibliometric profiles" for publications based on registered vs. non-registered trials.

This deliverable was prepared for the UNCOVER project consortium:

AIT Austrian Institute of Technology
Department for Foresight & Policy Development
Coordinating partner

Vienna, Austria

DUK Danube University Krems
Department for Evidence-based Medicine and Clinical Epidemiology

Krems, Austria

UNC University of North Carolina at Chapel Hill
Gillings School of Global Public Health

Chapel Hill, North Carolina, United States

Contact
Dr. Manuela KIENEGGER

Technology Management
Department for Foresight & Policy Development
AIT Austrian Institute of Technology GmbH
Donau-City-Straße 1
A-1220 Vienna
Austria

T +43(0) 50550-4530
F +43(0) 50550-4599

manuela.kienegger@ait.ac.at

www.ait.ac.at

Bibliometric features of publication bias in clinical trials

This deliverable was prepared by:

Edgar Schiebel
Maria-Elisabeth Züger

Other contributions to the deliverable by:

Dirk Holste
Manuela Kienegger
Silvia Steinbrunner

Content

Table Index

Figure Index

1 Executive Summary

In this deliverable we report on the usage of bibliometric features to shed light on some aspects of publication bias in the context of clinical trials, meta-analysis and systematic reviews. We focus on two issues: (1) publication of results derived from registered or non-registered studies, and (2) getting a broader view on research issues in the context of clinical trials as well as finding more relevant literature for systematic reviews and meta-analyses. We formulated hypotheses and central questions for further analysis.

Based on a framework of registered or non-registered studies, published or not published trial results we formulated the following hypotheses:

- The registration ID for clinical trials is cited in publications.
- The number of publications with a registration ID is growing.
- If editors insist on publications referring to registered studies, the number of such publications grows.

Following the hypotheses based on bibliometric approaches we formulated:

- The citation rate of publications with a registration code deriving from study registers is higher than for publications on clinical trials without a registration number. Higher citation rates are considered an incentive for editors and publishing trialists.
- Journals with a publication policy to only accept publications on clinical trials that have a registration number have an influence on the growing number of publications related to registered studies.
- Science maps broaden the view on research issues in systematic reviews and meta-analyses and allow for a more comprehensive selection of relevant literature.

For our bibliometric analysis we used three medical cases:

- Case 1: Antidepressants in the Pharmacologic Treatment of Adult Depression (APTAD)
- Case 2: Target Immune Modulators (TIM)
- Case 3: Diseases of the Cardiovascular System (DCS)

We generated datasets from three databases (ClinicalTrials.gov, PubMed and Web of Science) with search strategies for the research topic and a reference to a clinical study whether registered or not.

In a first step we developed a method to analyse automatically whether a publication reports on a registered clinical trial (CT) by using a textual approach and information in specific database fields containing a reference to a registered study. In ClinicalTrials.gov we find information on publications indexed in PubMed by a list of PubMed identifiers. Registered trial IDs are cited in the abstract and/or in the files of PubMed indexed publications. Web of Science offers information on registered trial IDs only in the abstracts. The number of publications with information on registered clinical trial IDs is still low, but has been increasing steadily since 2005.

After the establishment of the registration database ClinicalTrials.gov in 2001 and other registries later on the number of publications referring to clinical trial IDs shows a considerable growth. For the three medical cases we revealed the following: For APTAD the number of publications with a registration number increased quickly in 2007, but has stayed stable for the following years. For TIM we could identify just one publication which included a reference number for a clinical trial, whereas for DCS the number has been continuously growing. Most of the publications with a reference to a CT registry are published in journals that follow the rules of ICMJE. Although the clinical trial register ClinicalTrials.gov is a US initiative, not only authors from the US provide a registration number (NCT number) from ClinicalTrials.gov in their publications but also authors from many European countries do so. It is also notable that funding bodies acknowledged in publications that included registration numbers are to a large extent representatives of the pharmaceutical industry. With respect to different medical specialities, publications tend to refer more often to CT registration numbers in general internal medicine, cardiovascular system cardiology, hematology and nutrition dietetics than in oncology, pharmacology/pharmacy, neurosciences/neurology and surgery.

The citation analysis revealed that publications derived from registered CTs are higher cited than publications from not registered studies, which is considered a powerful incentive for researchers and editors.

We applied the science mapping approach to identify author networks and visualize a more general view of the research landscape of medical topics of the three analysed medical cases.

Finally, we demonstrated that bibliographic coupling helps to identify more relevant publications for meta-analysis and systematic reviews.

2 Introduction

2.1 Background

The UNCOVER project is a direct contribution to overcome non-publication of clinical studies that have been designed and executed as randomized controlled trials (RCTs).

UNCOVER's aim is three-fold:

- to apply established and develop novel, solid, and useful methods for fact-finding and interventions into the socio-economic system defined by causes and sources of the publication bias;
- to engage with stakeholders and identify strategies, barriers, and facilitating factors associated with the publication bias and its consequences; and
- to synthesize lessons learned and recommend feasible measures to deal with the publication bias.

RCTs are currently best practice to avoid or minimize both systematic and random errors in clinical studies. They provide the best utility as input to systematic medicinal reviews, one cornerstone of evidence-based medicine (EbM) for improved safety and efficacy / effectiveness of patient outcomes, and their end-users.

That is guaranteed when RCTs are both correctly registered and published at least once. Because non-publication, as well as publication with time delay of RCTs, may decisively reduce the advantage of such systematic reviews of drugs, medical devices or procedures, it affects the knowledge base. Therefore, in a perspective way, this project contributes pro better allocation of funds to sponsor studies, and contra duplication of work and patients risk.

2.2 Objectives of WP2

The objectives of work package 2 of the UNCOVER project are:

- Framing of hypotheses, definition of suitable and measurable bibliometric features, and definition of statistical indicators (based on features)
- Extraction, pre-processing, and standardization of data from information sources

- Investigation of the influence of registered vs. non-registered studies on the obtained bibliometric profile in the case of both a systematic review and a comprehensive thematic compilation of medicinal research studies
- Interpretation of characteristic features and conclusions given the current measures against publication bias

3 Content of this deliverable

This deliverable pursues the following two objectives: “Investigation of the influence of registered vs. non-registered studies on the obtained bibliometric profile in the case of both a systematic review and a comprehensive thematic compilation of medicinal research studies” and “Interpretation of characteristic features and conclusions given by the current measures against publication bias. It reports on the results of Task 2.2 “Bibliometric analysis of characteristic features distinctive between registered v. non-registered studies and” and Task 2.3 “Characteristic bibliometric features of publication bias and conclusions." In this deliverable we stress two main issues: (1) publication of results derived from registered or non-registered clinical studies, and (2) getting a broader view on research issues in the context of clinical trials as well as finding more relevant literature for systematic reviews and meta-analyses. In the following we formulate hypotheses and central questions for our analyses.

Based on a framework of registered or non-registered studies, published or not published trial results we formulated the following hypotheses (see also Deliverable 2.1[1]):

- The registration ID for clinical trials is cited in publications.
- The number of publications with a registration ID is growing.
- If editors insist on publications referring to registered studies, the number of such publications grows.
- Incentives make it more likely for authors and editors to publish registered studies than not registered studies.

Following the hypotheses based on bibliometric approaches we formulated:

- The citation rate of publications with a registration code deriving from study registers is higher than for publications on clinical trials without a registration

[1] Schiebel, E., Palensky, B., Züger, M.-E. Deliverable D2.1 of the UNCOVER FP7-funded project under contract number 282574: Data sources for bibliometric analysis, 2013.

number. Higher citation rates are considered a powerful incentive for editors and publishing trialists.

- Journals with a publication policy to only accept publications on clinical trials with a registration number (from a registry) have an influence on the growing number of publications related to registered studies.
- Science maps broaden the view on research issues in systematic reviews and meta-analyses and allow for a more comprehensive selection of relevant literature.

The detection, delineation, and visualization of research issues and research findings have drawn growing attention in bibliometric and scientometric research. The visualization of such relational bibliometric approaches is referred to as science mapping. Networks of authors are drawn in a graph that consists of nodes (authors) and edges calculated by the similarity of authors measured by their co-occurrence of co-authors in the same publication. Content maps are drawn by bibliographic coupling and co-citation analysis.

Several publications report on these techniques. Price (1965)[2] introduced the concept of research fronts based on citations and Kessler (1965)[3] introduced bibliographic coupling of publications sharing references. In 1973, co-citation was introduced by Small (1973)[4] and Marshakova (1973)[5]. Chen & Morris (2003)[6] identified clusters of co-cited articles. In recent work Shibata *et al.* (2009)[7] analysed the performance co-citation, bibliographic coupling and direct citation in detecting research fronts. Boyack & Klavans (2010)[8] added a bibliographic coupling-based citation-text hybrid approach to the three mentioned bibliometric approaches and compared accuracies of cluster solutions for a very large set of articles. Both author groups gave a sophisticated overview of science mapping methods

[2] Price, D.D. (1965). Networks of scientific papers. Science, 149, 510-515.

[3] Kessler M. M. (1963). Bibliographic coupling between scientific papers. American Documentation, 14(1), pp. 10-25.

[4] Small H. G. (1973). Co-citation in the scientific literature: a new measure of the relationship between two documents. Journal of the American Society for Information Science, 24, 265-269.

[5] Marshakova I.V. (1973). System of document connections based on references. Nauchno- Tekhnicheskaja Informacya, ser.2, N6, 3-8 (in Russian).

[6] Chen, C., & Morris, S. (2003). Visualizing evolving networks: Minimum spanning trees versus pathfinder networks. Proceedings of IEEE Symposium on Information Visualization (pp 67-74), Seattle, WA: IEEE Computer Society Press.

[7] Shibata et al, (2009). Comparative Study on Methods of Detecting Research Fronts Using different Types of Citation. JASIST 60(3):571-580.

[8] Boyack & Klavans, (2010). Co-Citation Analysis, Bibliographic Coupling, and Direct Citation: Which Citation Approach represents the Research Front Most Accurately?, JASIST 61(12): 2389-2404.

to detect emerging research fronts. In a recent publication Schiebel (2011)[9] proposed a visualization technique based on areal density of bibliographically coupled publications or co-cited references projected by a spring model in a three dimensional coordinate system.

Research fronts are understood as a growing research activity reported by publications on a research topic. Scientists working in such an area use previously published knowledge shared by colleagues in a publication or in a presentation given at a conference. The trickier such a research topic, the more intensive the research work. Persson (1994)[10] distinguishes between the research front and the intellectual basis: “In bibliometric terms, the citing articles form a research front, and the cited articles constitute an intellectual base“. This means that a knowledge base of a research topic can be made visible by generating clusters of co-cited references and research fronts can be identified by agglomerations of papers based on common references. We use this concept to draw relational maps and calculate densities of spatial density of objects such as references, citing publications or links in between. References are taken as a vector for each citing publication and the scalar product as a similarity measure using the Jaccard index. Accordingly citing publications is taken as a vector for each cited reference. Large scale maps enable the inclusion of several thousand publications and references respectively. We are therefore able to obtain an overview of a whole field of research from two points of view: the intellectual bases on one hand and the research fronts on the other hand. Two dimensional maps show papers as points. Areas of high or low density are identified with high density areas denoting areas of high research activity on a delineated research topic.

The relative importance of a journal within its field is measured by the so-called Journal Impact Factor (JIF). It counts the average number of citations to actual articles published in the journal. Publications of journals with higher impact factors are more often used as a knowledge base for new publications; therefore, they have more impact in the scientific community. The impact factor was introduced by Eugene Garfield (1955)[11], the founder of the Institute for Scientific Information.

[9] Schiebel, E. (2011). Research Fronts and Areal Density of Bibliographically Coupled Publications, Proceedings of the ISSI 2011 Conference, 13th International Conference of the International Society for Scientometrics & Informetrics, Durban South Africa, 04-07 July

[10] Persson, O (1994). The Intellectual base and research fronts of JASSIS 1986-1990. JASSIS 45(1): 31-38.

[11] Garfield, Eugene (1955). "Citation indexes for science...". Science (AAAS) 122 (3159): 108–111. doi:10.1126/science.122.3159.108.

4 Identification of registration IDs for clinical trials in publications

In this chapter we analyse whether publications on clinical trials provide clear evidence on the registration of the trials in registries. We limited the analysis to the registration in two registries: ClinicalTrials.gov and ISRCTN Register.

"ClinicalTrials.gov is a registry and results database of publicly and privately supported clinical studies of human participants conducted around the world… A service of the U.S. National Institutes of Health… ClinicalTrials.gov currently lists 143,755 studies with locations in all 50 states and in 184 countries."[12]

"The ISRCTN is a simple numeric system for the unique identification of randomised controlled trials worldwide. The ISRCTN Register also accepts registration of other forms of studies designed to assess the efficacy of health-care interventions."[13]

Three different medical cases were analysed:

- Case 1: Antidepressants in the Pharmacologic Treatment of Adult Depression (APTAD)
- Case 2: Target Immune Modulators (TIM) and
- Case 3: Diseases of the Cardiovascular System (DCS)

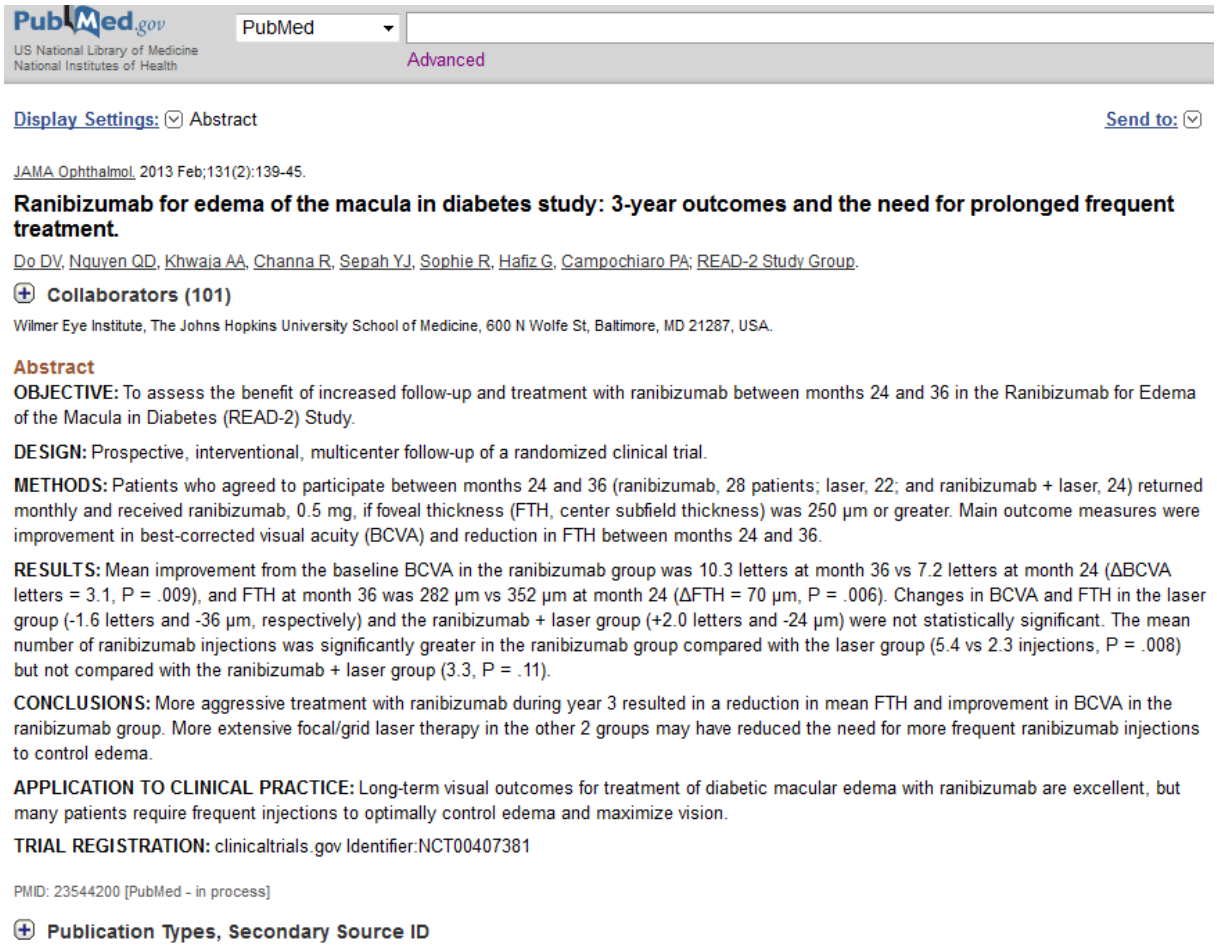
Publications were collected from the databases Web of Science (WoS) and PubMed. For details on the method data bases and search strategy see Deliverable 2.1[14].

Figure 4.1 provides an example for the display of a publication in the database PubMed. Both the trial registry and the registration number are listed at the end of the abstract, i.e. ClinicalTrials.gov and the registration number NCT00407381.

[12] ClinicalTrials.gov

[13] www.isrctn.org

[14] Schiebel, E., Palensky, B., Züger, M.-E. Deliverable D2.1 of the UNCOVER FP7-funded project under contract number 282574: Data sources for bibliometric analysis, 2013.



PubMed.gov
US National Library of Medicine
National Institutes of Health

PubMed

Advanced

Display Settings: Abstract

Send to:

JAMA Ophthalmol. 2013 Feb;131(2):139-45.

Ranibizumab for edema of the macula in diabetes study: 3-year outcomes and the need for prolonged frequent treatment.

Do DV, Nguyen QD, Khwaja AA, Channa R, Sepah YJ, Sophie R, Hafiz G, Campochiaro PA; READ-2 Study Group.

Collaborators (101)

Wilmer Eye Institute, The Johns Hopkins University School of Medicine, 600 N Wolfe St, Baltimore, MD 21287, USA.

Abstract

OBJECTIVE: To assess the benefit of increased follow-up and treatment with ranibizumab between months 24 and 36 in the Ranibizumab for Edema of the Macula in Diabetes (READ-2) Study.

DESIGN: Prospective, interventional, multicenter follow-up of a randomized clinical trial.

METHODS: Patients who agreed to participate between months 24 and 36 (ranibizumab, 28 patients; laser, 22; and ranibizumab + laser, 24) returned monthly and received ranibizumab, 0.5 mg, if foveal thickness (FTH, center subfield thickness) was 250 µm or greater. Main outcome measures were improvement in best-corrected visual acuity (BCVA) and reduction in FTH between months 24 and 36.

RESULTS: Mean improvement from the baseline BCVA in the ranibizumab group was 10.3 letters at month 36 vs 7.2 letters at month 24 (ΔBCVA letters = 3.1, P = .009), and FTH at month 36 was 282 µm vs 352 µm at month 24 (ΔFTH = 70 µm, P = .006). Changes in BCVA and FTH in the laser group (-1.6 letters and -36 µm, respectively) and the ranibizumab + laser group (+2.0 letters and -24 µm) were not statistically significant. The mean number of ranibizumab injections was significantly greater in the ranibizumab group compared with the laser group (5.4 vs 2.3 injections, P = .008) but not compared with the ranibizumab + laser group (3.3, P = .11).

CONCLUSIONS: More aggressive treatment with ranibizumab during year 3 resulted in a reduction in mean FTH and improvement in BCVA in the ranibizumab group. More extensive focal/grid laser therapy in the other 2 groups may have reduced the need for more frequent ranibizumab injections to control edema.

APPLICATION TO CLINICAL PRACTICE: Long-term visual outcomes for treatment of diabetic macular edema with ranibizumab are excellent, but many patients require frequent injections to optimally control edema and maximize vision.

TRIAL REGISTRATION: clinicaltrials.gov Identifier:NCT00407381

PMID: 23544200 [PubMed - in process]

Publication Types, Secondary Source ID

Figure 4.1: Structured abstract from the literature database PubMed providing a reference to a clinical trial registry (ClinicalTrials.gov) and a registration number.

In the PubMed database the database field “DB Accession number“ contains the registration identifier. Table 4.1 lists some examples for the case APTAD.

Table 4.1: Examples of PubMed entries with references (DB Accession number) to registered studies.

PM ID	Title	DBAccessionNumber	DBName
16540613	Maintenance treatment of major depression in old age.	NCT00178100	ClinicalTrials.gov
16554525	Bupropion-SR, sertraline, or venlafaxine-XR after failure of SSRIs for depression.	NCT00021528	ClinicalTrials.gov
16554526	Medication augmentation after the failure of SSRIs for depression.	NCT00021528	ClinicalTrials.gov
17548243	Effect of sertraline on symptoms and survival in patients with advanced cancer, but without major depression: a placebo-controlled double-blind randomised trial.	ISRCTN72466475	ISRCTN Register

The Web of Science database does not include a database field for registration numbers of clinical trials. A reference to a registry can only be found if the ID of the registered clinical trial is included in the abstract. Table 4.2 provides two examples of how trial identifiers are included in abstracts. The first example is a small meta-study referring to two clinical trials and the second one is a publication on a clinical trial. Both publications deal with results of registered clinical studies that can be identified by the search string "NCT0*" for a registered trial in ClinicalTrials.gov and the string "ISRCTN" for the other considered registry database.

Table 4.2: Two exemplary publications on APTAD in Web of Science citing an NCT or an ISRCTN registration ID.

WoS ID	Abstract
WOS:000301008200024	Objective. To determine response with duloxetine versus placebo in patients with osteoarthritis (OA) of the knee using the Outcome Measures in Rheumatoid Arthritis Clinical Trials-Osteoarthritis Research Society International (OMERACT-OARSI) responder index and other clinically relevant outcomes including minimal clinically important improvement (MCII) and patient acceptable symptom state (PASS) for pain and function. Methods. Data were pooled from two 13-week, randomized, double-blind, placebo-controlled trials comparing duloxetine 60 to 120 mg/day with placebo in patients with symptomatic OA of the knee. Treatment response was determined according to the OMERACT-OARSI responder index, >= 30% pain reduction, >= 50% pain reduction, and MCII and PASS for pain and function. (ClinicalTrials.gov identifiers **NCT00433290** and **NCT00408421**) Results. Duloxetine-treated patients were 33% more likely to experience an OMERACT-OARSI response than placebo-treated patients [p < 0.001, number needed to treat (NNT) = 6]. A significantly greater percentage of duloxetine-treated patients, compared with placebo-treated patients, reported 30% improvement in pain from baseline to endpoint (p < 0.001, NNT = 5) and 50% improvement in pain relative to baseline (p < 0.001, NNT = 7). The duloxetine-treated patients were also more likely to fulfill MOT criteria for pain (p < 0.001, NNT = 6) and function (p < 0.001, NNT = 7), and to achieve PASS for pain (p < 0.001, NNT = 6) and function (p = 0.009, NNT = 9). More duloxetine-treated patients compared with placebo-treated patients experienced 1 treatment-emergent adverse event (p = 0.003, number needed to harm = 8). Conclusion. Significantly more patients receiving duloxetine than placebo achieved an OMERACT-OARSI response, improvements in pain and function exceeding the level accepted as MCII, and reached PASS. Results support the clinical relevance of outcomes of prior duloxetine studies in symptomatic OA of the knee. (First Release Dec 1 2011; J Rheumatol 2012;39:352-8; doi:10.3899/jrheum.110307)
WOS:000301188800023	Background: Previous studies suggest that electroacupuncture possesses therapeutic benefits for depressive disorders. The purpose of this study was to determine whether dense cranial electroacupuncture stimulation (DCEAS) could enhance the antidepressant efficacy in the early phase of selective serotonin reuptake inhibitor (SSRI) treatment of major depressive disorder (MDD). Methods: In this single-blind, randomized, controlled study, patients with MDD were randomly assigned to 9-session DCEAS or noninvasive electroacupuncture (n-EA) control procedure in combination with fluoxetine (FLX) for 3 weeks. Clinical outcomes were measured using the 17-item Hamilton Depression Rating Scale (HAMD-17), Clinical Global Impression-severity (CGI-S), and Self-rating Depression

WoS ID	Abstract
	Scale (SDS) as well as the response and remission rates. Results: Seventy-three patients were randomly assigned to n-EA (n = 35) and DCEAS (n = 38), of whom 34 in n-EA and 36 in DCEAS group were analyzed. DCEAS-treated patients displayed a significantly greater reduction from baseline in HAMD-17 scores at Day 3 through Day 21 and in SDS scores at Day 3 and Day 21 compared to patients receiving n-EA. DCEAS intervention also produced a higher rate of clinically significant response compared to n-EA procedure (19.4% (7/36) vs. 8.8% (3/34)). The incidence of adverse events was similar in the two groups. Conclusions: DCEAS is a safe and effective intervention that augments the antidepressant efficacy. It can be considered as an additional therapy in the early phase of SSRI treatment of depressed patients. Trial Registration: Controlled-Trials.com **ISRCTN88008690**

We analysed the number of publications citing IDs of registered studies for the case APTAD and restricted to the occurrence of "NCT0…" in abstracts of WoS and PubMed data or the DB Accession Number of PubMed. The results are summarized in Table 4.3. Concerning publications on clinical trials referred to in the systematic review on APTAD, we found only a few (less than 8%) references to registered study IDs. Further below in this report, we will show that more recent publications include more often references to trial registries. As was expected, the numbers do not differ for the two databases. However, the topic search in WoS (data set WoS II) resulted in 18 additional publications with a reference to registered studies in comparison to the PubMed search. This number derives from the difference between the data sets of WoS II and WoS I. The WoS I dataset of publications is based on a title-wise match with the PubMed data, therefore the set of 3,227 has 18 more publications referring to registered clinical studies.

Table 4.3: Number of publications citing an NCT registered trial ID for the case APTAD indexed in the databases PubMed and WoS. Time span: 1983 to 2011.

Databases	Number of publications	String "NCT0*" in publication abstract	NCT ID in field "DB Accession Number" (for PubMed)**
PubMed	785	35	59
WoS I (data set match with PubMed by titles)	742	17	n.r.
WoS II (data set retrieved by topic search)	3227	35	n.r.

**n.r.: not relevant

It should be noted that the analysis in Table 4.3 is restricted to NCT IDs, which refer only to ClinicalTrials.gov and not to other clinical trials registries. However, this approach seems a good approximation for registration in general, as ClinicalTrials.gov is currently the most commonly used registry. Wikipedia gives the following numbers: "…top five registries (as

of August 2012): ClinicalTrials.gov (130,756 trials), EU register (18,660 trials), ISRCTN (10,853 trials), Japan registries trwork (JPRN, 10,511 trials), Australia and New Zealand (ANZCTR, 6,916 trials)"[15].

We found that the hypothesis "The registration ID for clinical trials is cited in publications" is valid for the case APTAD and the databases PubMed and WoS. We restricted the analysis to the occurrence of the string "NCT0*". This registration ID refers to registered studies in ClinicalTrials.gov. Yet, the reference to registered studies is sparse.

5 Characteristics of publications that cite a registration ID

5.1 Number of registered studies in ClinicalTrials.gov

The citation of a reference number to a clinical trial registry depends on the availability of registered studies. We started our analysis with the number of registered clinical trials in ClinicalTrials.gov, the registry with the highest number of registered clinical trials currently. Figure 5.1 depicts the cumulated number of registered studies. While this registry was established in 2000, a remarkable acceleration of registrations has been observed since 2006.

5.2 Increase of number of publications with a reference to registered clinical trials

The development of the number of publications that cite a registration ID over time was examined by an analysis in the Web of Science database for publications where the strings "NCT0*" and ISRCTN occur as a topic (Table 5.1).

[15] http://en.wikipedia.org/wiki/Clinical_trials_registry, accessed on 4 April 2013.

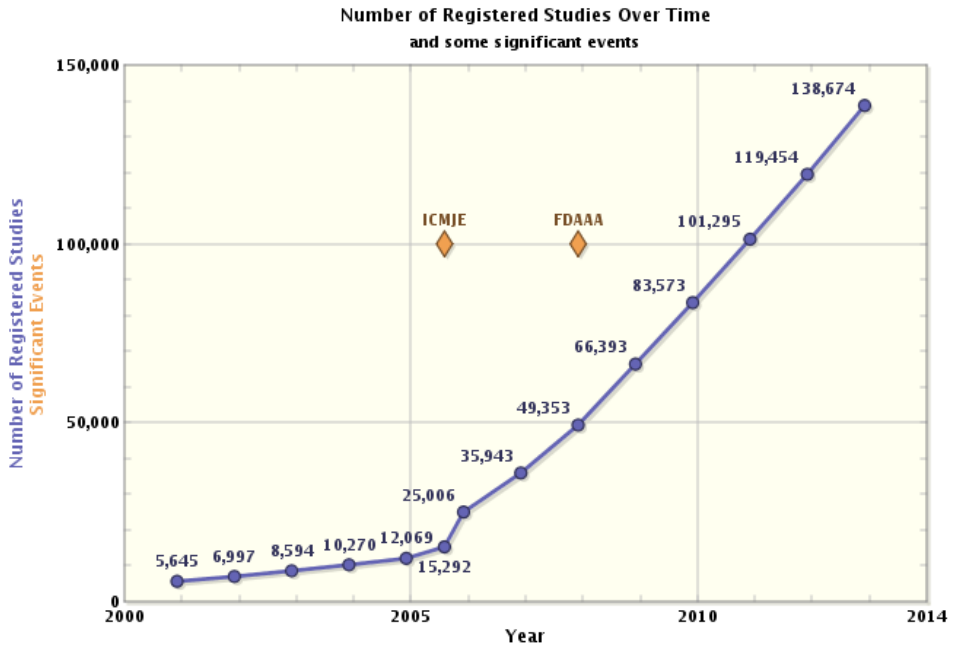


Figure 5.1: Number of registered studies in ClinicalTrials.gov over time. ICMJE: Indicates when the International Committee of Medical Journal Editors started to require trial registration as a precondition for publication under the Uniform Requirements for Manuscripts Submitted to Biomedical Journals (URM) (September 2005). FDAAA: Indicates when the expanded registration requirements of FDAAA began and were implemented on ClinicalTrials.gov (December 2007). Source: http://clinicaltrials.gov/ct2/resources/trends, accessed on 4 April 2013.

Table 5.1: Search Strategy for publications with a registration ID, Source: Web of Science.

Search step	Search strings and Boolean operations	Hits
1	NCT0*	6,563
2	ISRCTN*	1,642
3	#1 OR #2	8,164

Figure 5.2 shows the number of publications per year. It can be seen that the first document citing a registration ID was published in 2001. Following the increase of numbers of registered studies in ClinicalTrials.gov the number of publications rose considerably in 2008.

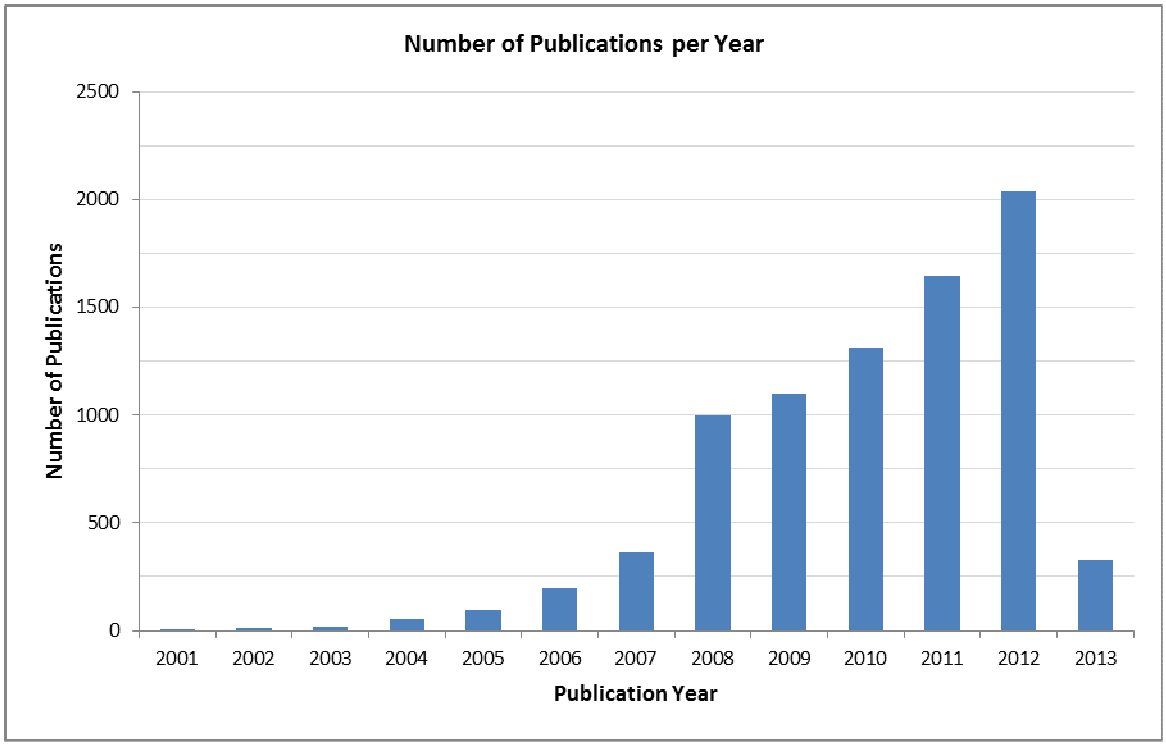


Figure 5.2: Number of publications per year citing a registration number for clinical trials. All publications cite a registration number (NCT or ISRCTN). Total number of publications: 8,164; date of search: 3 2013; data source: Web of Science.

The above analysis was performed on the basis of publications within the database Web of Science citing a registry number. In a further step we looked at the publications of clinical trials for the three medical cases “Antidepressants in the Pharmacologic Treatment of Adult Depression (APTAD)”, “Target Immune Modulators (TIM)” and “Diseases of the Cardiovascular System (DCS)”.

For the case APTAD we collected 742 publications on clinical trials by common files in PubMed and WoS. The first publication that refers to a registered study ID was published in 2006 (Figure 5.3). The number increases quickly to 14 publications with an NCT number and two with an ISRCTN number in the year 2007, but afterwards we see a stagnation with around ten publications, which are about a quarter of all publications on clinical trials of a year. This relative high number of publications citing a registry ID could have been promoted by the Journal of Clinical Psychiatry that follows the rules of ICMJE.

For the case “Target Immune Modulators (TIM)” we identified a smaller share of publications on clinical trials referring to a registration ID (Figure 5.4). There is one publication with an ISRCTN ID in each of the years 2008, 2009 and 2010; and 18, 16 and 1 publications with an NCT number from the years 2009, 2010 and 2011, respectively.

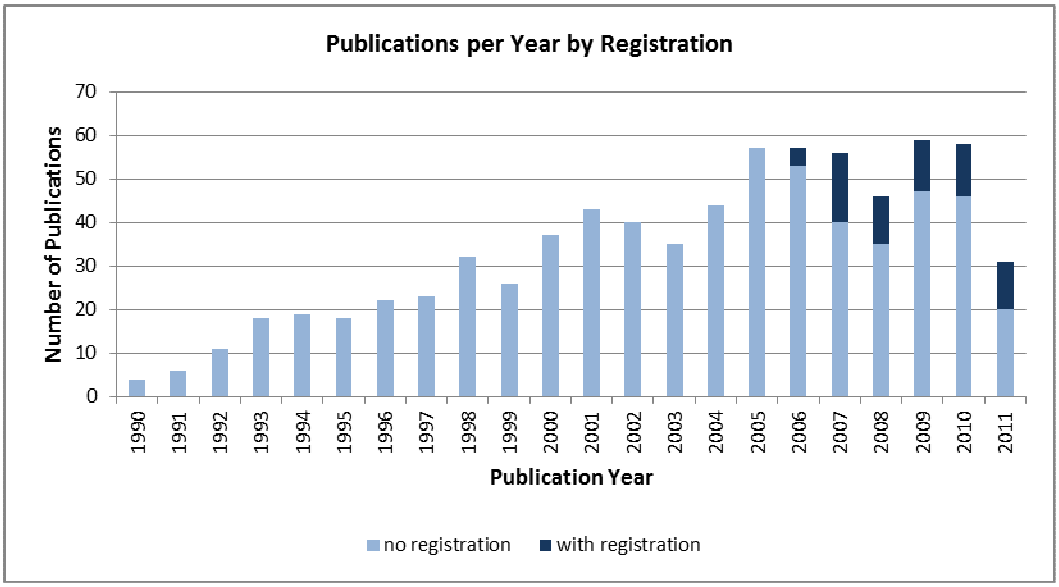


Figure 5.3: Number of publications per year for the medical case "Antidepressants in the Pharmacologic Treatment of Adult Depression (APTAD)". Publications which cite a registration number (NCT or ISRCTN) are coloured dark blue; total number of publications: 742; date of search: 4 2012; data source: Web of Science.

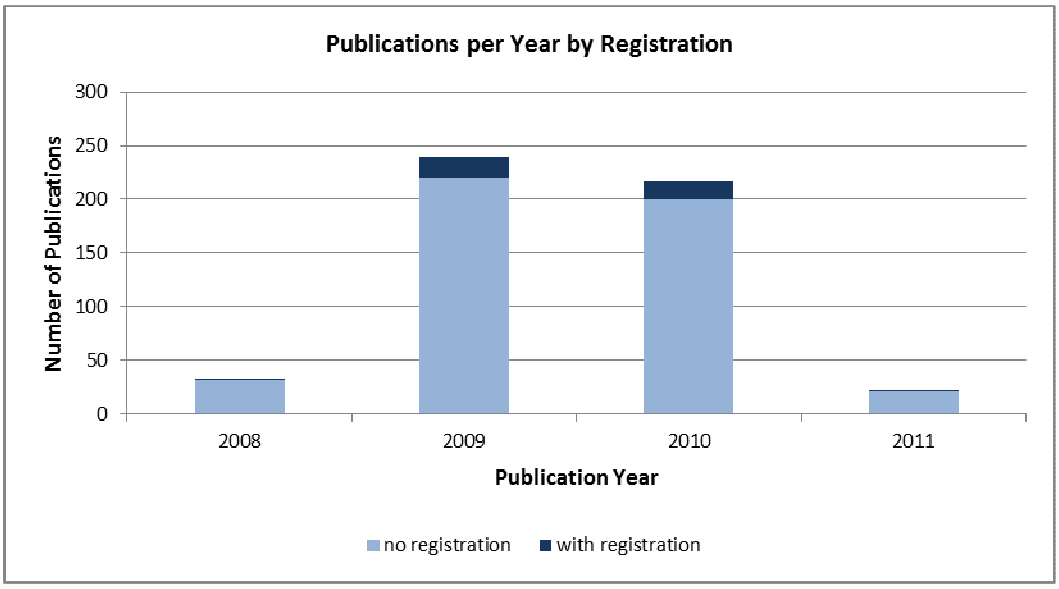


Figure 5.4: Number of publications per year for the medical case "Target Immune Modulators (TIM)". Publications which give a registration number (NCT or ISRCTN) are coloured dark blue. Total number of publications: 511; date of search: 4 2012; data source: Web of Science.

The number of publications referring to a registration ID is growing from 1 ISRCTN registration ID in the year 2005 to 81 (76 NCT and 5 ISRCTN) for the case "Diseases of the Cardiovascular System (DCS)".

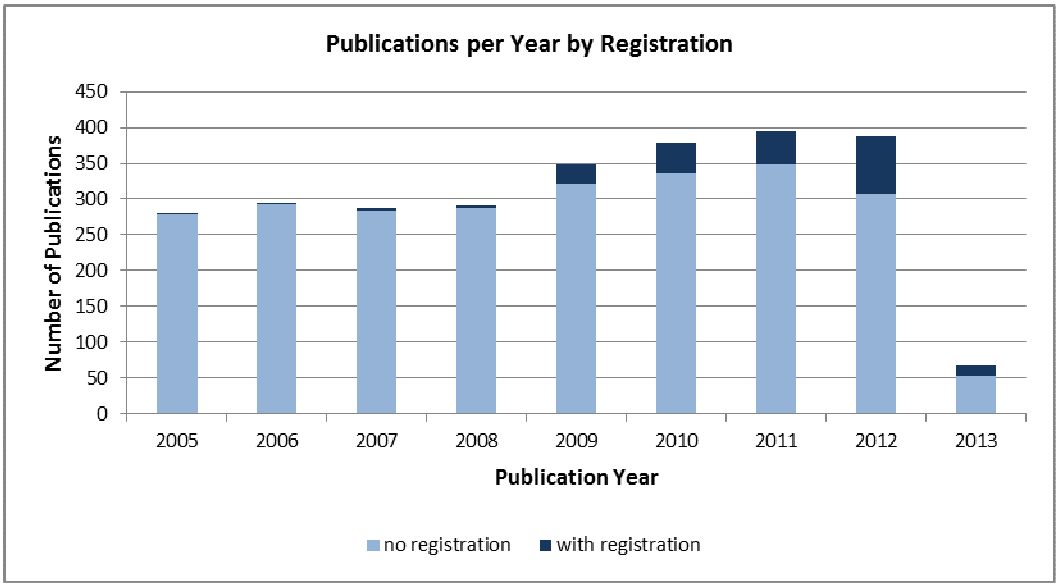


Figure 5.5: Number of publications per year for the medical case "Diseases of the Cardiovascular System (DCS)". Publications which give a registration number (NCT or ISRCTN) are coloured dark blue. Total number of publications: 2,727; date of search: 3 2013; data source: Web of Science.

5.3 Publication sources of publications citing a registration ID

Table 5.2 provides an overview of the top 30 scientific journals (publication sources) that have references to the trial registry IDs, the number of publications, the share of all publications and whether the journal follows the ICMJE requirements. These journals contribute remarkably to the growing number of publications that cite one of our examined registration IDs. In addition, the median of the Journal Impact Factors (JIF) for the listed journals that follow the ICMJE uniform requirements (6.7) differs slightly from that of the journals that do not follow the requirements (4.3).

Table 5.2: Source titles (titles of journals) with the number of publications (Top 30). All publications provide a registration number (NCT or ISRCTN). The ICMJE column indicates whether the journal officially follows the uniform requirements for manuscript by the ICMJE; JIF: Journal Impact Factor. Total number of publications: 8,164; date of search: 3 2013; data source: Web of Science.

Source titles (journals)	Number of publications	Share of total (% of 8,164)	ICMJE uniform req.	JIF
NEW ENGLAND JOURNAL OF MEDICINE	573	7.0	Yes	53.289
LANCET	564	6.9	Yes	38.278
BLOOD	434	5.3	No	10.558
AMERICAN JOURNAL OF CLINICAL NUTRITION	382	4.7	Yes	6.7
JOURNAL OF THE AMERICAN COLLEGE OF CARDIOLOGY	303	3.7	Yes	14.292

Source titles (journals)	Number of publications	Share of total (% of 8,164)	ICMJE uniform req.	JIF
LANCET ONCOLOGY	231	2.8	No	22.59
CURRENT MEDICAL RESEARCH AND OPINION	142	1.7	Yes	2.38
JAMA JOURNAL OF THE AMERICAN MEDICAL ASSOCIATION	134	1.6	Yes	30.026
INVESTIGATIVE OPHTHALMOLOGY VISUAL SCIENCE	132	1.6	Yes	3.466
CLINICAL THERAPEUTICS	123	1.5	No	2.321
TRIALS	123	1.5	No	2.5
JOURNAL OF CLINICAL PSYCHIATRY	113	1.4	Yes	5.8
LANCET NEUROLOGY	111	1.4	No	23.462
PLOS ONE	103	1.3	No	2
JACC CARDIOVASCULAR INTERVENTIONS	98	1.2	No	6.8
BRITISH JOURNAL OF SURGERY	88	1.1	Yes	4.606
CHEST	83	1.0	Yes	6.225
ANNALS OF INTERNAL MEDICINE	82	1.0	Yes	16.733
ANNALS OF SURGERY	80	1.0	Yes	7.492
DIABETOLOGIA	80	1.0	Yes	6.814
HAEMATOLOGICA THE HEMATOLOGY JOURNAL	78	1.0	Yes	6.424
BMC PUBLIC HEALTH	73	0.9	Yes	2
JOURNAL OF INFECTIOUS DISEASES	73	0.9	Yes	6.41
BMC MUSCULOSKELETAL DISORDERS	70	0.9	No	1.58
BRITISH MEDICAL JOURNAL	67	0.8	Yes	14.093
BMC HEALTH SERVICES RESEARCH	62	0.8	no	1.66
GASTROINTESTINAL ENDOSCOPY	61	0.7	Yes	4.878
HUMAN REPRODUCTION	60	0.7	No	4.475
VACCINE	58	0.7	No	4.251
CIRCULATION	56	0.7	Yes	14.816

5.4 Geographical aspects

As the two analysed IDs are from registries sited in USA and Canada, we also examined the countries of the authors' affiliations of the publications (Figure 5.6). The analysis revealed that the USA is the leading country concerning publications citing an NCT or an ISRCTN registration ID. Canada is on the fourth position. In addition, authors from European organizations in England, Germany, the Netherlands, France, Italy, Belgium and others play also an important part in registering clinical studies in these two registries, or performing meta-analysis or systematic reviews that refer to registered clinical trials. It can be deduced that the initiative of ICMJE together with the implementation of ClinicalTrials.gov and other registries have an important impact on the growing transparency of the publication of clinical trials.

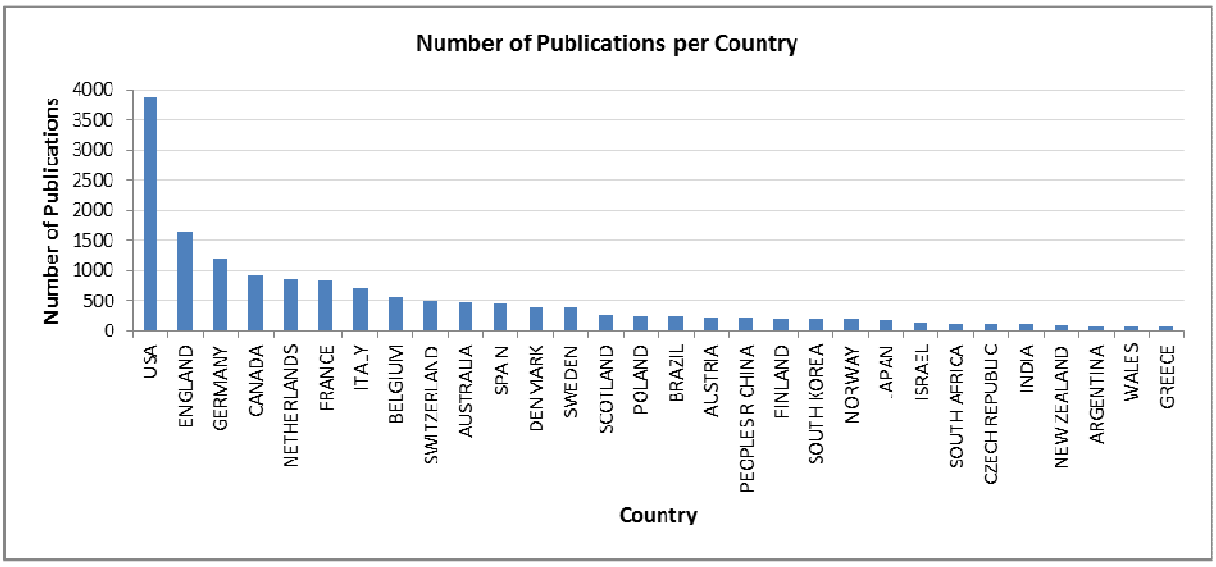


Figure 5.6: Number of publications per country (TOP 30). All publications provide a registration number (NCT or ISRCTN). Total number of publications: 8,164; date of search: 3 2013; data source: Web of Science.

5.5 Funding bodies

As the major sponsor of clinical trials is the pharmaceutical industry as a, we were also interested in the question whether the funding for the published research referring to registered clinical trials comes from private or public organisations. Table 5.3 lists the top 30 funding bodies for the published clinical trials by the number of publications and the share of the number of publications of the whole set of 8,164 publications. The share of publications referring to private sponsors out of the 30 is 58.9 %. This indicates that the major part of publications with a reference to clinical trial IDs is sponsored by this sector.

Table 5.3: Funding bodies and number of publications (Top 30). All publications cite a registration ID (NCT or ISRCTN). Total number of publications: 8,164; date of search: 3 2013; data source: Web of Science.

Funding bodies	Number of publications	Share of total (% of 8,164)
PFIZER	466	5.7
NOVARTIS	463	5.7
GLAXOSMITHKLINE	448	5.5
ASTRAZENECA	413	5.1
BRISTOL MYERS SQUIBB	350	4.3
SANOFI AVENTIS	308	3.8
MERCK	305	3.7
NATIONAL INSTITUTES OF HEALTH	278	3.4
ROCHE	251	3.1

Funding bodies	Number of publications	Share of total (% of 8,164)
BOEHRINGER INGELHEIM	243	3.0
ELI LILLY	201	2.5
ABBOTT	196	2.4
MEDTRONIC	174	2.1
AMGEN	142	1.7
BOSTON SCIENTIFIC	142	1.7
NIH	127	1.6
SCHERING PLOUGH	126	1.5
BAYER	125	1.5
NATIONAL HEART LUNG AND BLOOD INSTITUTE	117	1.4
NATIONAL CANCER INSTITUTE	100	1.2
PFIZER INC	100	1.2
TAKEDA	94	1.2
WYETH	94	1.2
DAIICHI SANKYO	86	1.1
GENENTECH	86	1.1
SERVIER	86	1.1
ABBOTT VASCULAR	84	1.0
JOHNSON JOHNSON	82	1.0
NOVO NORDISK	78	1.0
ASTELLAS	70	0.9

5.6 Medical specialities

We examined whether the number of publications referring to a registration number differs with respect to medical specialties. Web of Science provides different medical specialties in the database field “research areas”. As Table 5.4 shows ”General internal medicine” is the dominating category with publications that refer to registered clinical trials followed by cardiovascular system cardiology, hematology, oncology and nutrition dietetics.

We compared the results with the relative number of publications in the same specialty retrieved by the keyword “clinical trial” in the topic field of Web of Science. We suppose that such publications refer to clinical trials in general whether they are registered or not. In this case WoS lists a more than ten times higher number of publications: 92,702 vs. 8,164 that cite a registered clinical trial ID. In “general internal medicine” a much higher relative number (25.1%) of publications cite a registration number than publications with the keyword “clinical trial” (9.2%). We have a similar picture for cardiovascular system cardiology, hematology and nutrition dietetics. On the other hand, publications referring to registered studies show a lower relative number in the specialties oncology, pharmacology pharmacy, neurosciences neurology and surgery.

Table 5.4: Research areas (WoS categories) with the number of publications (Top 30). (a) All publications cite a registration number (NCT or ISRCTN), (b) publications use the keyword "clinical trial", date of search: 3 2013; data source: Web of Science**.**

Research Areas	Number of publications	(a) % of 8164	(b) % of 92702
GENERAL INTERNAL MEDICINE	2053	25.1	9.206
CARDIOVASCULAR SYSTEM CARDIOLOGY	968	11.9	6.267
HEMATOLOGY	625	7.7	2.727
ONCOLOGY	624	7.6	11.383
NUTRITION DIETETICS	461	5.6	1.734
RESEARCH EXPERIMENTAL MEDICINE	430	5.3	5.242
PHARMACOLOGY PHARMACY	361	4.4	10.271
NEUROSCIENCES NEUROLOGY	360	4.4	7.787
PSYCHIATRY	345	4.2	4.47
SURGERY	328	4.0	8.165
IMMUNOLOGY	314	3.8	3.651
INFECTIOUS DISEASES	273	3.3	2.19
RESPIRATORY SYSTEM	258	3.2	2.212
OBSTETRICS GYNECOLOGY	212	2.6	3.23
GASTROENTEROLOGY HEPATOLOGY	208	2.5	3.49
PSYCHOLOGY	194	2.4	2.695
RHEUMATOLOGY	189	2.3	2.026
ENDOCRINOLOGY METABOLISM	179	2.2	2.563
OPHTHALMOLOGY	173	2.1	3.826
PUBLIC ENVIRONMENTAL OCCUPATIONAL HEALTH	168	2.1	3.383
PEDIATRICS	164	2.0	2.709
MICROBIOLOGY	158	1.9	1.1
HEALTH CARE SCIENCES SERVICES	143	1.8	2.384
ORTHOPEDICS	119	1.5	2.132
SCIENCE TECHNOLOGY OTHER TOPICS	103	1.3	0.756
REPRODUCTIVE BIOLOGY	94	1.2	0.734
ANESTHESIOLOGY	85	1.0	1.484
UROLOGY NEPHROLOGY	81	1.0	2.86
RADIOLOGY NUCLEAR MEDICINE MEDICAL IMAGING	76	0.9	2.111
ALLERGY	61	0.7	0.521

5.7 Citation analysis

In this chapter we examine the following hypotheses:

1. The citation rate of publications with a registration ID from study registers is higher than for publications on clinical trials without a registration number.
2. Higher citation rates are considered an incentive for editors and publishing trialists for publication of registered clinical trials.
3. Incentives make it more likely for authors and editors to publish registered studies than not registered studies.

As the registration of clinical trials is an important contribution to reduce publication bias it is interesting to examine if publications on registered trials are higher cited or not. The citation rate of published work is of growing interest for individual authors and editors. For authors the times cited of their publications is an important indicator for the visibility of their work and is commonly measured by the Hirsch-Index that ranks the publications of one researcher by the descending number of times cited.

The attractiveness of a journal is measured by the Journal Impact Factor that is also based on the citation of articles in these journals. Editors are interested in journals with a high Journal Impact Factor.

If the citation rate of publications referring to a registration ID is higher than the citation rate of not registered studies then it can be a powerful incentive to register studies from the point of view of editors and authors.

Figure 5.7 shows the times cited of each publication per year for the medical case “Antidepressants in the Pharmacologic Treatment of Adult Depression (APTAD)”. In the years 2006, 2007 and 2010 the publications with the highest number of citations refer to registered trial IDs. In the years following 2005, the medians are higher or at least equal for publications with a registration ID in comparison to those without registration.

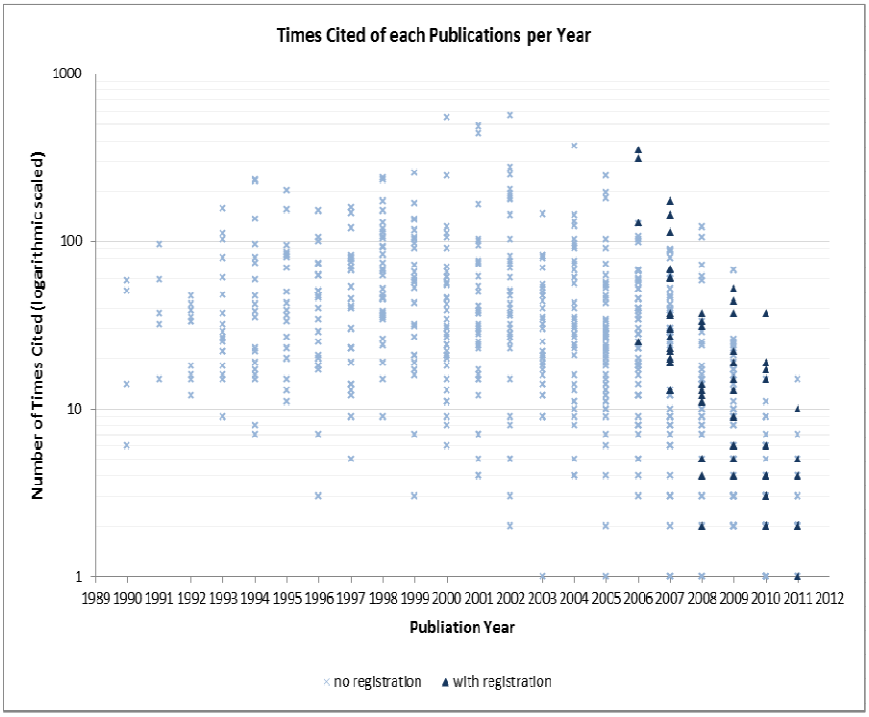


Figure 5.7: Times cited of each publication per year for the medical case "Antidepressants in the Pharmacologic Treatment of Adult Depression (APTAD)". Times cited is logarithmic scaled (base 10). Publications which cite a registration ID (NCT or ISRCTN) are coloured dark blue; total number of publications: 742; date of search: 4 2012; data source: Web of Science.

Table 5.5: Number of publications and median of times cited per year for the medical case "Antidepressants in the Pharmacologic Treatment of Adult Depression (APTAD)". Total number of publications: 742; date of search: 4 2012; data source: Web of Science.

Publication year	Number of publications			Median of times cited		
	Non-registered	NCT	ISRCTN	Non-registered	NCT	ISRCTN
1990	4			32,5		
1991	6			35		
1992	11			33		
1993	18			28		
1994	19			38		
1995	18			41		
1996	22			43		
1997	23			41		
1998	32			58		
1999	26			58		
2000	37			31		
2001	43			31		
2002	40			41		
2003	35			25		
2004	44			31		

Publication year	Number of publications			Median of times cited		
	Non-registered	NCT	ISRCTN	Non-registered	NCT	ISRCTN
2005	57			23		
2006	53	4		28	220	
2007	40	14	2	25	29	105
2008	35	11		14	12	
2009	47	11	1	8	15	6
2010	46	11	1	2	4	37
2011	20	10	1	2	2	10

For the case “Target Immune Modulators (TIM)” the highest cited publications in the years 2009, 2010 and 2011 have a reference to a registered clinical trial. Only in the year 2008 we see a highest cited publication that does not refer to a registered clinical trial (Figure 5.8). The median values for “times cited” are listed in Table 5.6, where we find higher citation rates for publications with a registration ID.

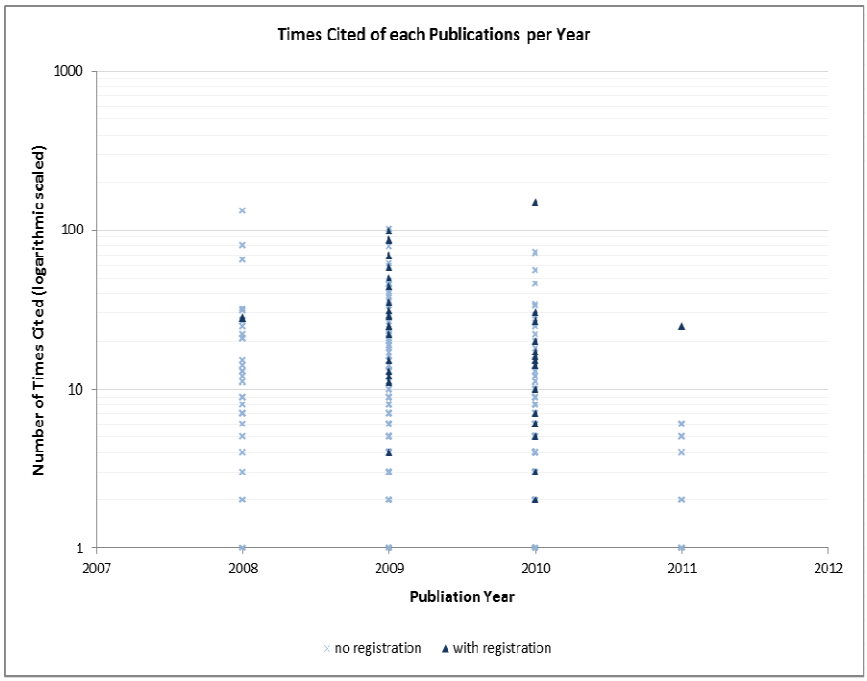


Figure 5.8: Times cited of each publication per year for the medical case “Target Immune Modulators (TIM)”. Times cited is logarithmic scaled (base 10). Publications which cite a registration ID (NCT or ISRCTN) are coloured dark blue; total number of publications: 511; date of search: 4 2012; data source: Web of Science.

Table 5.6: Number of publications and median of times cited per year for the medical case "Target Immune Modulators (TIM)". Total number of publications: 511; date of search: 4 2012; data source: Web of Science.

Publication year	Number of publications			Median of times cited		
	Non-registered	NCT	ISRCTN	Non-registered	NCT	ISRCTN
2008	32		1	10		28
2009	220	18	1	7.5	30	12
2010	200	16	1	3	12	6
2011	21	1		1	25	

For the medical case "Diseases of the Cardiovascular System (DCS)" we also have clear evidence for a higher citation rate of publications with a reference to a registered clinical trial ID. Since 2007 all medians have been lower for the category of publications without a reference to a registered clinical trial (Table 5.7). Figure 5.9 shows that in the years 2009 and 2012 the highest cited publications refer to a registered clinical trial.

Table 5.7: Number of publications and median of times cited per year for the medical case "Diseases of the Cardiovascular System (DCS)". Total number of publications: 2,727; date of search: 3 2013; data source: Web of Science.

Publication year	Number of publications			Median of times cited		
	Non-registered	NCT	ISRCTN	Non-registered	NCT	ISRCTN
2005	278		1	11		10
2006	292	1	1	13	2	2
2007	284	3	1	9	25	158
2008	287	4		6	35	
2009	322	24	2	6	30	1
2010	336	40	3	4	16	38
2011	348	46	1	2	8	8
2012	307	76	5	0	2	4
2013	54	13	2	0	1	1

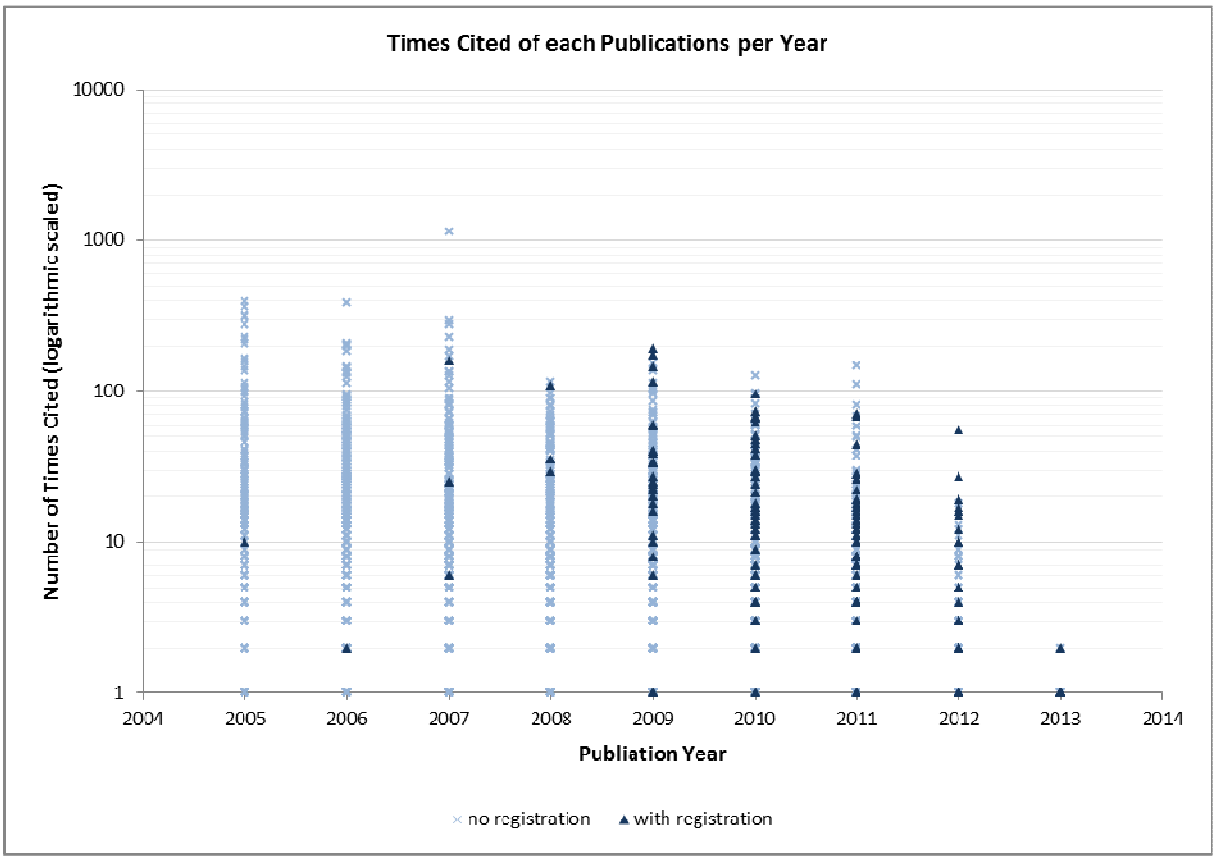


Figure 5.9: Times cited of each publication per year for the medical case "Diseases of the Cardiovascular System (DCS)". Times cited is logarithmic scaled (base 10). Publications which cite a registration ID (NCT or ISRCTN) are coloured dark blue; total number of publications: 2,727; date of search: 3 2013; data source: Web of Science.

The citation analysis shows that publications referring to a registered clinical trial ID have a higher citation rate than other publications. There could be several reasons for this positive effect. Firstly, the clinical trials of a medical topic can be easily found in the registration database. Additionally, the PubMed number of published work is given in a data field of the registration database. Trialists who publish the work of their own trial can easily find other trials and publications and cite them. Researchers who work on a meta-analysis or a systematic review are also able to easily identify relevant previously published publications or data about clinical trials. Therefore, the probability of citation of a journal article is much higher if it refers to a registered clinical trial. It can be expected that the citation rate of published work of registered clinical trials will have a higher citation rate in the future due to a better visibility and reliability.

The higher citation rate is a powerful incentive for trialists and editors to publish results of clinical trials that have been registered.

6 Science maps broaden the view on research issues in systematic reviews and meta-analyses and allow for a more comprehensive selection of relevant literature

In this chapter we discuss two issues: Firstly, we will demonstrate that relational bibliometric techniques, namely the co-authorships, the co-citation and bibliographic coupling opens the view on research issues of single clinical trials.

We analyse the thematic research activities from two perspectives. The first perspective is the so called knowledge base. Knowledge bases are identified by a co-citation analysis of backward citations. All cited references represent the intellectual base of written publications. The co-citation landscape visualizes the agglomerations of similar references by their common occurrence in publications. The agglomerations are named by keywords and titles of the citing documents.

Secondly, we will show that bibliographic coupling is a sound method to identify publications that are not found by a database search based on keywords as it is usually done for systematic reviews.

6.1 Case “Antidepressants in the Pharmacologic Treatment of Adult Depression (APTAD)”

For the case “Antidepressants in the Pharmacologic Treatment of Adult Depression (APTAD)” we used two data sets. The on-topic search as it is described in Deliverable 2.1 resulted in 3,227 publications in the Web of Science database (set I). The search for publications on clinical trials for this case resulted in 742 publications that where identical in the title of the two databases PubMed and Web of Science (set II).

6.1.1 Map of Co-Authorships

Table 6.1 presents a list of the most active authors. Interestingly, the most active authors in data set II are identical with the most active authors in data set I. In set I we found 213 publications of the author Fava, M with 18 publications that refer to a registered clinical trial. The specific search for clinical trials in data set II identifies 69 publications in total and 12 publications that refer to a NCT registration ID of the same author. For the following authors the on-topic search identifies one or more publications with a reference to an NCT registration number. It is remarkable that the most active authors have publications that

refer to a clinical study. There are only a few exceptions such as Kasper,S, Serretti, A, Rosenbaum, JF and others.

We did not examine the relevance of a single publication but just demonstrated that the on-topic search included more publications on clinical trials by an author-based list that could be used for the identification of relevant literature for a systematic review or a meta-analysis. That could be a contribution to reduce publication bias by a more complete set of relevant literature.

Table 6.1: Authors (top 30 with highest number of publications) with total number of publications in the data set I and data set II for the case "Antidepressants in the Pharmacologic Treatment of Adult Depression (APTAD)". *On-Topic Search (data set I)*: Total number of publications: 3,227; total number of publications with registration number: 86. *WoS matched with PubMed search (data set II):* total number of publications: 742; total number of publications with registration number: 66; date of search: 4 2012; data source: Web of Science.

Author	On-topic search (WoS data) (Set I)				PubMed matched with WoS data (Set II)			
	Number of publications	% of 3227	Number of publ. with reg. ID	% of 86	Number of publications	% of 742	Number of publ. with reg. ID	% of 66
Fava, M	213	6.6	18 NCT	20.9	69	9.3	12 NCT	18.2
Thase, ME	113	3.5	12 NCT	14.0	57	7.7	11 NCT	16.7
Rush, AJ	106	3.3	11 NCT	12.8	45	6.1	8 NCT	12.1
Trivedi, MH	91	2.8	14 NCT	16.3	43	5.8	13 NCT	19.7
Nierenberg, AA	88	2.7	11 NCT	12.8	32	4.3	8 NCT	12.1
Papakostas, GI	62	1.9	1 NCT	1.2	7	0.9		
Wisniewski, SR	57	1.8	12 NCT	14.0	25	3.4	9 NCT	13.6
Alpert, JE	47	1.5	1 NCT	1.2	12	1.6		
Detke, MJ	45	1.4	4 NCT	4.7	14	1.9	2 NCT	3.0
Kasper, S	43	1.3			9	1.2		
Wohlreich. MM	43	1.3	6 NCT	7.0	18	2.4	5 NCT	7.6
Mallinckrodt, CH	41	1.3	6 NCT	7.0	17	2.3	5 NCT	7.6
Kornstein, SG	40	1.2	4 NCT	4.7	30	4.0	6 NCT	9.1
Kennedy, SH	35	1.1	2 NCT	2.3	2	0.3		
Raskin, J	35	1.1	7 NCT	8.1	13	1.8	5 NCT	7.6
Serretti, A	35	1.1			2	0.3		
Rosenbaum, JF	34	1.1			13	1.8		
Warden, D	31	1.0	8 NCT	9.3	17	2.3	6 NCT	9.1
Dunner, DL	29	0.9	3 NCT	3.5	14	1.9	5 NCT	7.6
Emslie, GJ	27	0.8			-	-		
Mischoulon, D	27	0.8	1 NCT	1.2	8	1.1		
Keller, MB	26	0.8	1 NCT	1.2	18	2.4	2 NCT	3.0

Author	On-topic search (WoS data) (Set I)				PubMed matched with WoS data (Set II)			
	Number of publications	% of 3227	Number of publ. with reg. ID	% of 86	Number of publications	% of 742	Number of publ. with reg. ID	% of 66
Perlis, RH	24	0.7			7	0.9		
Watkin, JG	24	0.7	1 NCT	1.2	11	1.5	1 NCT	1.5
McGrath, PJ	23	0.7	5 NCT	5.8	14	1.9	3 NCT	4.5
Prakash, A	23	0.7	4 NCT	4.7	6	0.8	3 NCT	4.5
Ninan, PT	22	0.7	1 NCT	1.2	8	1.1	3 NCT	4.5
Smeraldi, E	22	0.7			8	1.1		
Andersen, HF	22	0.7			7	0.9		
Montgomery, SA	22	0.7	1 NCT	1.2	14	1.9	1 NCT	1.5

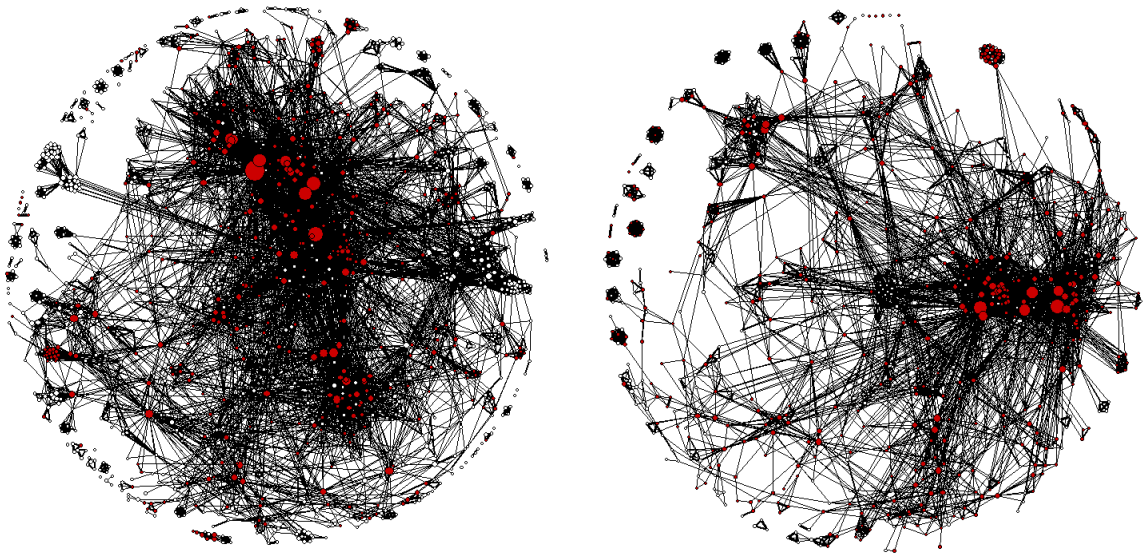


Figure 6.1: Map of Authors (Co-Publications) for the case "Antidepressants in the Pharmacologic Treatment of Adult Depression (APTAD)". Coloured authors (red) are identified in both maps. Grouping of authors on their common appearance in publications; circle: author, the size corresponds to the number of publications; edge: Jaccard index of co-frequencies; date of search: 4 2012; data source: Web of Science. *On-Topic Search, set I (left)*: Total Number of publications: 3,227; each author occurs at least in 3 publications; number of nodes: 1,264; number of edges: 7,858. *PubMed Search set II (right)*: Total number of publications: 742; each author occurs in at least two publications; number of nodes: 716; number of edges: 4,671.

The networks of co-publishing authors are drawn in Figure 6.1. The left graph shows the interrelationships of authors of set I and the right graph shows the authors of set II.

The most active authors with a high centrality (relative number of edges) in the network are identical in both datasets. We have a high interrelationship in the community of researchers that are active in the research field “Antidepressants in the Pharmacologic Treatment of Adult Depression (APTAD)”. The giant component of the graph (largest sub-network of interconnected authors) is very huge and only a few small components can be found in the periphery.

6.1.2 Knowledge bases (Map of cited References)

Figure 6.2 shows the landscape of references in the case “Antidepressants in the Pharmacologic Treatment of Adult Depression (APTAD)”. We find huge agglomerations for the three subthemes: (1) fluoxetine; out-patients; (2) fluoxetine; obsessive-compulsive disorder and (3) venlafaxine (XR); bupropion. The knowledge base fibromyalgia and pain syndromes is in the periphery but consists or a relative high number of 109 different references. We find diagnostic technologies such as (quantitative) electroencephalography (QEEG) – cordance and positron emission tomography and other themes such as children or adolescent depression, pharmacogenetics polymorphism, treatment-resistant/refractory depression, escitalopram citalopram, report QIDS-SR (Quick Inventory of Depressive Symptomatology – self report) and STAR*D (Sequential Treatment Alternative to Relieve Depression), escitalopram citalopram 2, duloxetin and placebo.

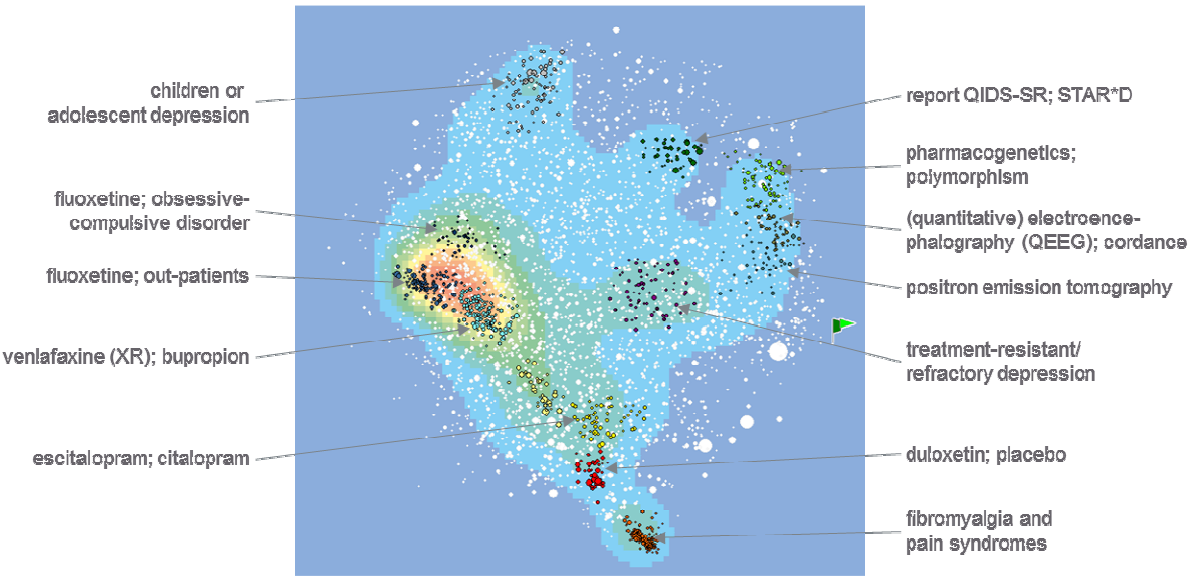


Figure 6.2: Map of knowledge bases (co-citation analysis) for the case “Antidepressants in the Pharmacologic Treatment of Adult Depression (APTAD)”. Nodes are coloured by topics. Grouping of cited references by their common appearance in publications; circle: reference, the size corresponds to the number of citing publications; edges: Jaccard index of co-frequencies in citing

publications; background: density map of the local number of references weighted by the Jaccard index of all links, cos-weighted moving average filter, total number of publications: 3,227; date of search: 4 2012; each reference occurs in at least 6 publications; number of references: 3,784; number of edges: 484,417 (not plotted); data source: Web of Science.

Table 6.2 lists the knowledge bases and some indicators such as the number of references that form the agglomeration, the number and share of references that are cited by publications referring to a registration ID, the number of the citing publications in total and citing a registration ID.

The knowledge bases with the highest number of citing publications are “venlafaxine (XR); bupropion” (667), “escitalopram; citalopram 2” (458), “fluoxetine; out-patients” (419) and “treatment-resistant/refractory depression” (410). Publications with a reference to a registration ID cite a large share of references in the following knowledge bases: report QIDS-SR and STAR*D (65.9% of all references in this agglomeration), duloxetin; placebo (61.8%), escitalopram; citalopram 2 (52.6%) and treatment-resistant/refractory depression (51.5%). Such knowledge bases contribute more than 50% to publications on registered clinical trials by their number of references and have the highest number of citing publications.

Table 6.2: Knowledge bases on “Antidepressants in the Pharmacologic Treatment of Adult Depression (APTAD)” with the share of publications with a registration ID. Total number of publications: 3,227; date of search: 4 2012; data source: Web of Science**.**

	Topics (1)	Number of references (2)	Number of ref. cited by publ. with registration ID (3)	Share of ref. cited by (6) in % (4)	Number of citing publ. (5)	Number of citing publ. with a registration ID (6)	Share of publ. with registration in topic (6)/(5) in % (7)
1.	fluoxetine; out-patients	135	4	3.0	419	5	1.2
2.	venlafaxine (XR); bupropion	134	27	20.1	667	24	3.6
3.	fibromyalgia and pain syndromes	109	36	33.0	194	13	6.7
4.	children or adolescent depression	82	15	18.3	228	4	1.8
5.	pharmacogenetics; polymorphism	71	18	25.4	228	8	3.5
6.	treatment-resistant/refractory depression	70	36	51.4	410	24	5.9
7.	escitalopram; citalopram	55	21	38.2	309	14	4.5

	Topics (1)	Number of references (2)	Number of ref. cited by publ. with registration ID (3)	Share of ref. cited by (6) in % (4)	Number of citing publ. (5)	Number of citing publ. with a registration ID (6)	Share of publ. with registration in topic (6)/(5) in % (7)
8.	(quantitative) electroencephalography (QEEG); cordance	53	0	0.0	125	0	0.0
9.	fluoxetine; obsessive-compulsive disorder	52	2	3.8	290	2	0.7
10.	report QIDS-SR; STAR*D	44	29	65.9	354	35	9.9
11.	positron emission tomography	42	8	19.0	181	3	1.7
12.	escitalopram; citalopram 2	38	20	52.6	458	17	3.7
13.	duloxetin; placebo	34	21	61.8	351	28	8.0

6.1.3 Research fronts – bibliographically coupled publications

The map of research fronts represents agglomerations of similar publications. The similarity is measured by the relative number (Jaccard Index) of common references and positioned by a spring algorithm.

Figure 6.3 shows the map of research fronts. Published research from the medical case APTAD forms a strong agglomeration in the centre of the map. The reason for this is that in general most of the knowledge bases form the intellectual base of the publications in a statistical sense: Sets of publications refer to most of the findings in different knowledge bases in the case of APTAD. We have separated research topics in the issues “children or adolescent depression” and “duloxetine: fibromyalgia”. The publications spread over the landscape do not have similar reference lists to a bigger agglomeration of a research front. The reasons are that we have single publications that do not refer to an established knowledge base or just use one or more keywords of the search strategy in a different context than publications of the “main stream”.

The on-topic analysis in Web of Science indicated some dominance in research in “main stream” issues of the topics listed in Table 6.3. The Research Fronts “duloxetine; placebo” (10.2) publications) and “STAR*D” (18) have a higher number of publications that refer to an ID of a registered study (Table 6.3).

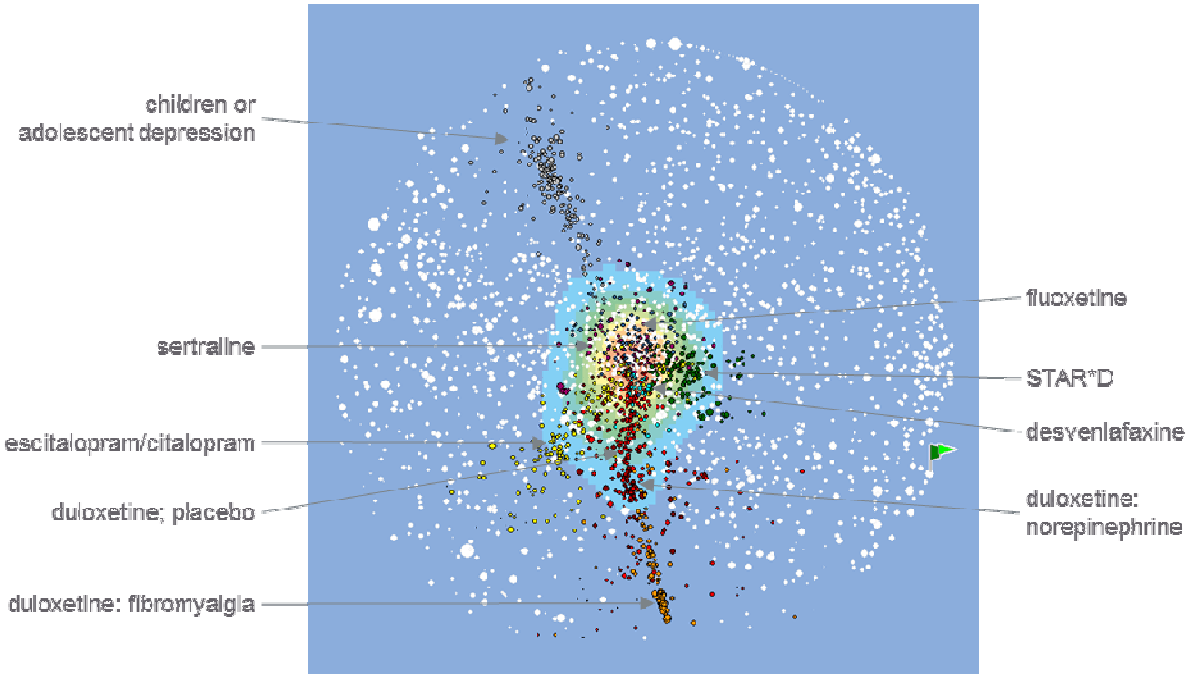


Figure 6.3: Map of research fronts (bibliographically coupled publications) for the case "Antidepressants in the Pharmacologic Treatment of Adult Depression (APTAD)". Nodes coloured by topics. Grouping of publications by their common references; circle: publication, the size corresponds to the number of references; edges: Jaccard index of co-references; background: density map of bibliographically coupled publications, cos-weighted moving average filter; total number of publications: 3,227; date of search: 4 2012; number of nodes: 3,015; number of edges: 766,986 (none shown); data source: Web of Science.

Table 6.3: Research fronts for the case "Antidepressants in the Pharmacologic Treatment of Adult Depression (APTAD)" with share of publications with a registration ID. Total number of publications: 3,227; date of search: 4 2012; data source: Web of Science.

	Research Fronts (RF)	Number of publications	Number of pub. with registration ID	Share of pub. with registration ID in%
1.	duloxetine; placebo	206	21	10.2
2.	escitalopram/citalopram	179	4	2.2
3.	children or adolescent depression	171	4	2.3
4.	STAR*D	120	18	15.0
5.	duloxetine: fibromyalgia	111	4	3.6
6.	duloxetine: norepinephrine	103	4	3.9
7.	fluoxetine	100	2	2.0
8.	sertraline	83	1	1.2
9.	desvenlafaxine	28	4	14.3

6.2 Case "Target Immune Modulators (TIM)"

Gartlehner *et al*. published a systematic review on this topic, which was used for a bibliometric analysis. In their publication they provide the following definition of TIM: *"Targeted immune modulators (TIMs) – commonly referred to as biological response modifiers or simply biologics – are a relatively new category of medication used in the treatment of certain types of immunologic and inflammatory diseases, including rheumatoid arthritis (RA), juvenile rheumatoid arthritis (JRA), ankylosing spondylitis (AS), psoriatic arthritis (PsA), plaque psoriasis, Crohn's disease, and ulcerative colitis (UC). "*[16]

For the case "Target Immune Modulators (TIM)" we used two data sets. The on-topic search as it is described in Deliverable 2.1[17] identified 10,688 publications in the Web of Science database (set I). The search for publications on clinical trials in this topic resulted in 511 publications that where identical in the title for the two databases PubMed and Web of Science (set II).

6.2.1 Map of co-authorships

Table 6.4 gives a list of the most active authors. Similar to the APTAD case the most active authors in data set II are identical with the most active authors in data set I. In set I we found 303 publications of the most active author Emery, M with only 3 publications that refer to a registered clinical trial. The specific search for clinical trials in data set II identified 21 publications in total and two publications including an NCT registration ID by the same author. For most of the following authors the on-topic search identified one or more publications with a reference to an NCT registration number. Similar to the case APTAD the most active authors have publications that refer to a clinical study but there are far fewer.

[16] Gartlehner et al., Drug Class Review on Targeted Immune Modulators. Final Report 2007, RTI-UNC Evidence-based Practice Center, Cecil G. Sheps Center for Health Services Research, University of North Carolina at Chapel Hill (http://www.donau-uni.ac.at/imperia/md/content/department/evidenzbasierte_medizin/abstracts_publikationen_gerald/drug_class_review_on_targeted_immune_modulators.pdf)

[17] Schiebel, E., Palensky, B., Züger, M.-E. Deliverable D2.1 of the UNCOVER FP7-funded project under contract number 282574: Data sources for bibliometric analysis, 2013.

Table 6.4: Authors (Top 30) and number of publications and registrations on "Target Immune Modulators (TIM)". *On-Topic Search*: total number of publications: 10,688; total number of publications with registration number: 85. *PubMed Search*: total number of publications: 511; total number of publications with registration number: 38. Date of search: 4 2012; data source: Web of Science.

Author	On-topic search (WoS data) (Set I)				PubMed matched with WoS Data (Set II)			
	Number of publ.	% of 10688	Number of publ. with reg. ID	% of 85	Number of publ.	% of 511	Number of publ. with reg. ID	% of 38
Emery, P	303	2.8	3 NCT	3.5	21	4.1	2 NCT	5.3
Kavanaugh, A	151	1.4	3 NCT	3.5	7	1.4	1 NCT	2.6
Dougados, M	134	1.3	2 NCT	2.4	8	1.6	1 NCT	2.6
Keystone, E	128	1.2	1 NCT	1.2	7	1.4	1 NCT	2.6
Westhovens, R	123	1.2	1 NCT	1.2	2	0.4		
van der Heijde, D	122	1.1	1 NCT	1.2	10	2.0	2 NCT	5.3
Keystone, EC	117	1.1	5 NCT	5.9	6	1.2	1 NCT	2.6
Tak, PP	116	1.1	3 NCT/ISRCTN	3.5	11	2.2	3 NCT/ISRCTN	7.9
Rutgeerts, P	111	1.0	3 NCT	3.5	15	2.9	5 NCT	13.2
Dijkmans, BAC	110	1.0	2 NCT/ISRCTN	2.4	18	3.5	2 NCT/ISRCTN	5.3
Braun, J	99	0.9			2	0.4		
Kupper, H	97	0.9	5 NCT	5.9	10	2.0	5 NCT	13.2
Breedveld, FC	92	0.9	2 NCT	2.4	6	1.2	1 NCT	2.6
Mease, PJ	89	0.8	6 NCT	7.1	7	1.4	3 NCT	7.9
Sandborn, WJ	87	0.8	6 NCT	7.1	20	3.9	5 NCT	13.2
Smolen, JS	85	0.8	4 NCT	4.7	7	1.4	2 NCT	5.3
Li, T	83	0.8	3 NCT	3.5	4	0.8	2 NCT	5.3
Moreland, LW	79	0.7	1 NCT	1.2	1	0.2		
Burmester, GR	77	0.7			8	1.6		
Klareskog, L	77	0.7			2	0.4		
Aranda, R	73	0.7	1 NCT	1.2				
Fleischmann, RM	73	0.7	1 NCT	1.2				
Schiff, M	73	0.7	1 NCT	1.2	3	0.6	1 NCT	2.6
Sieper, J	73	0.7	4 NCT	4.7	6	1.2	4 NCT	10.5
Smolen, J	73	0.7			2	0.4		
Becker, JC	72	0.7	1 NCT	1.2	2	0.4		
Vermeire, S	72	0.7	2 NCT	2.4	10	2.0	2 NCT	5.3
Kalden, JR	70	0.7			1	0.2		
van Vollenhoven, RF	69	0.6	2 NCT	2.4	8	1.6	2 NCT	5.3
Van Assche, G	67	0.6	3 NCT	3.5	9	1.8	1 NCT	2.6

The networks of co-publishing authors are depicted in Figure 6.4. The left graph shows the interrelationships of authors of set I and the right graph those of authors of set II. The most active authors with a high centrality (relative number of edges) in the network are identical in both datasets as it was for the case APTAD. Again, we identified a high interrelationship among the authors in the community of researchers that are active in the research field "Target Immune Modulators (TIM)". The major component of the graph (biggest sub-network of interconnected authors) is again very huge in both graphs and only a few small components can be found in the periphery. However, we see some differences to the APTAD case. The left graph of the on-topic search shows a higher density (number of links) over all authors than the PubMed search for clinical trials. The giant component of set II shows two communities (agglomeration of authors with a high density of links) with relative weak linkages between the two communities. There are many small communities in the periphery of set II. The whole research community in this research field (set I) is much stronger interconnected than the scientists who co-publish work about clinical trials.

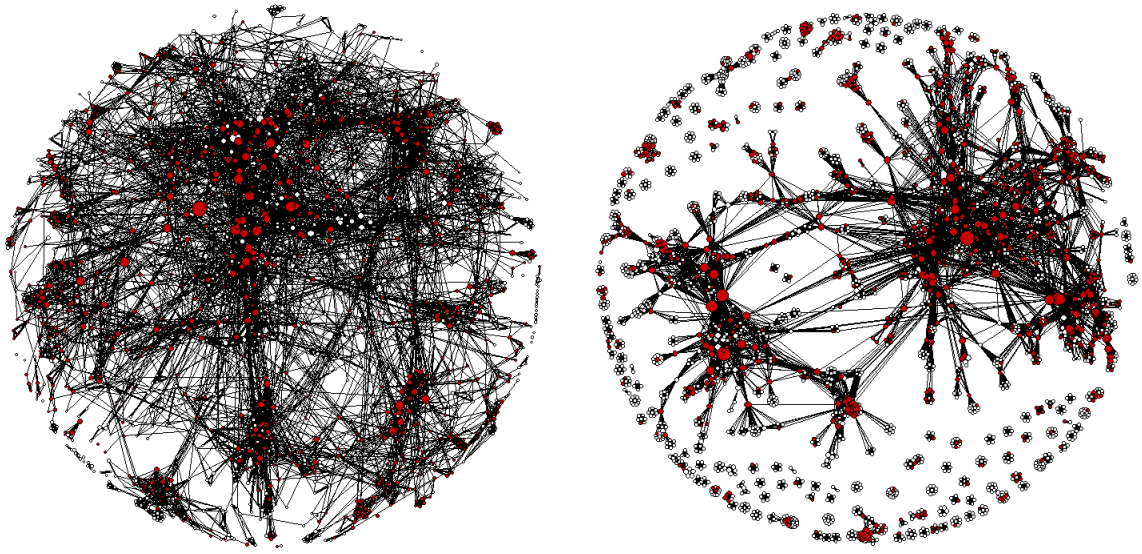


Figure 6.4: Map of authors (co-publications) for the case "Target Immune Modulators (TIM)". Coloured authors (red) are identified in both maps. Grouping of authors on their common appearance in publications; circle: author, the size corresponds to the number of publications; edge: Jaccard index of co-frequencies; date of search: 4 2012; data source: Web of Science *On-Topic Search, set I (left)*: total number of Publications: 10,688; each author occurs in at least 6 publications; number of nodes: 1,687; number of edges: 18,716 (only 8,617 shown) *PubMed Search set II (right)*: total number of publications: 511; number of nodes: 2,764; number of edges: 16,702.

6.2.2 Knowledge bases (map of cited references)

Figure 6.5 shows the clearly structured landscape of co-cited references in the case “Target Immune Modulators (TIM).” We identified two huge agglomerations: ”B-cell (lymphocyte) depletion; rituximab” with 321 references and “inflammatory bowel disease; ulcerative colitis” (300) together with “inflammatory bowel disease; crohn's disease;” (145) and “inflammatory bowel disease; anti-tumor necrosis factor” (112). The topic “inflammatory bowel disease; crohn's disease; virus infection; hepatitis B/C” (158) is separated in the 6 o’clock position.

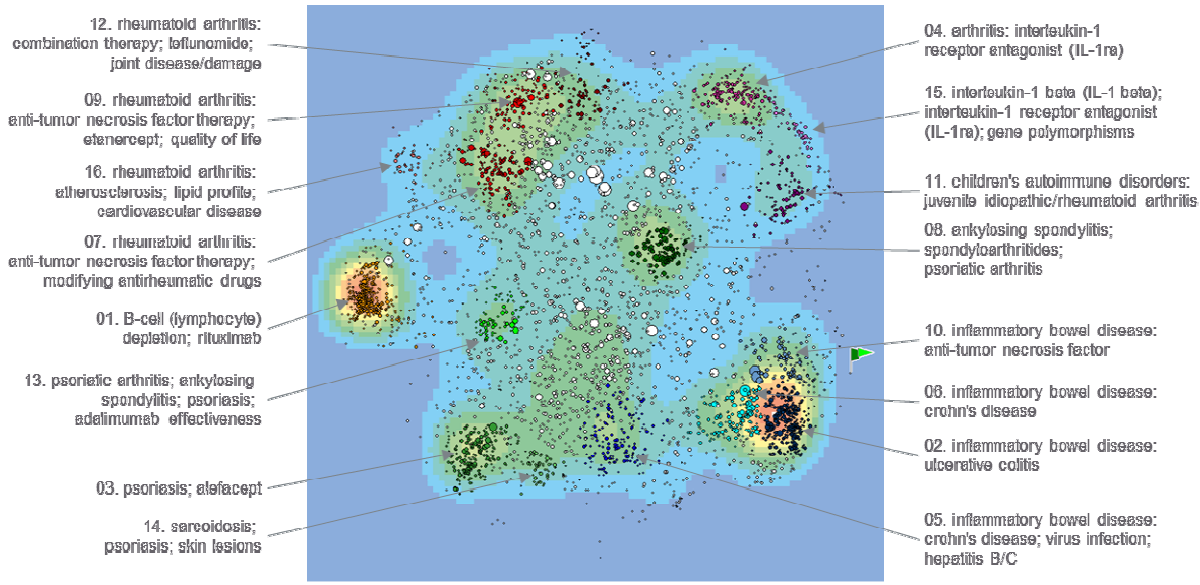


Figure 6.5: Map of knowledge bases (co-citation analysis) for the case “Target Immune Modulators (TIM)”. Nodes are coloured by topics. Grouping of references on their common appearance in publications; circle: reference, the size corresponds to the number of publications; edges: Jaccard index of co-frequencies in citing publications; background: density map of the local number of references weighted by the Jaccard index of all links of the node, cos-weighted moving average filter; total number of publications: 10,688; date of search: 4 2012; each reference occurs in at least 11 publications; number of nodes: 4,101; number of edges: 785,000 (none shown); data source: Web of Science.

Table 6.5 lists the knowledge bases and some indicators such as the number of references which form the agglomeration, the number and share of references which are cited by publications referring to a registration ID, the number of the citing publications in total and citing a registration ID.

The knowledge bases with the highest number of citing publications are “rheumatoid arthritis: anti-tumor necrosis factor therapy; modifying antirheumatic drugs” (1,454), “inflammatory bowel disease; crohn's disease” (1,386) and “rheumatoid arthritis: anti-

tumor necrosis factor therapy; etanercept; quality of life" (1,386). Publications with a reference to a registration ID cite a high proportion of references in the following knowledge bases: "psoriatic arthritis; ankylosing spondylitis; psoriasis; adalimumab effectiveness" (52.1% of all references in this agglomeration)," rheumatoid arthritis: anti-tumor necrosis factor therapy; modifying anti-rheumatic drugs" (40.7%). Such knowledge bases contribute more than 40% of their number of references to publications on registered clinical trials and have the highest number of citing publications. The two knowledge bases dealing with "rheumatic arthritis" are cited by the highest number of publications that refer to a registration ID. But all in all there are just a few publications on registered clinical trials.

Table 6.5: Knowledge bases for the case "Target Immune Modulators (TIM)" with the share of publications with a registration ID. Total number of publications: 10,688; date of search: 4 2012; data source: Web of Science.

	Topics (1)	Number of references (2)	Number of ref. cited by publ. with registration ID (3)	Share of ref. cited by (6) in % (4)	Number of citing pub. (5)	Number of citing pub. with registration ID (6)	Share of publ. with registration in topic (6)/(5) in % (7)
1.	B-cell (lymphocyte) depletion; rituximab	321	70	21.8	954	13	1.4
2.	inflammatory bowel disease; ulcerative colitis	300	84	28.0	1.236	13	1.1
3.	psoriasis; alefacept	180	56	31.1	932	14	1.5
4.	arthritis: interleukin-1 receptor antagonist (IL-1ra)	160	7	4.4	758	4	0.5
5.	inflammatory bowel disease; crohn's disease; virus infection; hepatitis B/C	158	3	1.9	1.027	3	0.3
6.	inflammatory bowel disease; crohn's disease;	145	23	15.9	1.390	12	0.9
7.	rheumatoid arthritis: anti-tumor necrosis factor therapy; modifying antirheumatic drugs	123	50	40.7	1.454	32	2.2
8.	ankylosing spondylitis; spondyloarthritides; psoriatic arthritis	118	35	29.7	1.110	10	0.9
9.	rheumatoid arthritis:	114	39	34.2	1.386	25	1.8

	Topics (1)	Number of references (2)	Number of ref. cited by publ. with registration ID (3)	Share of ref. cited by (6) in % (4)	Number of citing pub. (5)	Number of citing pub. with registration ID (6)	Share of publ. with registration in topic (6)/(5) in % (7)
	anti-tumor necrosis factor therapy; etanercept; quality of life						
10.	inflammatory bowel disease; anti-tumor necrosis factor	112	14	12.5	1.212	9	0.7
11.	children's autoimmune disorders: juvenile idiopathic/rheumatoid arthritis	98	39	39.8	674	10	1.5
12.	rheumatoid arthritis: combination therapy; leflunomide; joint disease/damage	89	17	19.1	974	11	1.1
13.	psoriatic arthritis; ankylosing spondylitis; psoriasis; adalimumab effectiveness	71	37	52.1	594	13	2.2
14.	sarcoidosis; psoriasis; skin lesions	66	1	1.5	463	1	0.2
15.	interleukin-1 beta (IL-1 beta); interteukin-1 receptor antagonist (IL-1ra); gene polymorphisms	49	0	0.0	269	0	0.0
16.	rheumatoid arthritis: atherosclerosis; lipid profile; cardiovascular disease	34	11	32.4	206	3	1.5

6.2.3 Research fronts – bibliographically coupled publications

The map of research fronts draws agglomerations of similar publications. The similarity is measured by the relative number (Jaccard Index) of common references and positioned by a spring algorithm.

Figure 6.6 shows the map of research fronts on “Target Immune Modulators (TIM)”. Published research for the case TIM resulted in a more disperse map of research fronts than for the case APTAD. The reason for this might be that in this case more diseases are affected.

We have corresponding knowledge bases and research fronts “B-cell (lymphocyte) depletion”, “inflammatory bowel disease; crohn's disease and ulcerative colitis”, ”psoriasis, alefacept”, “rheumatoid arthritis”, “inflammatory bowel disease”, “ankylosing spondylitis”; ” children's autoimmune disorders: juvenile idiopathic/rheumatoid arthritis”, “anti-tumor necrosis factor therapy and infections: tuberculosis” and “interleukin-1 receptor antagonist (IL-1ra); (rheumatoid) arthritis”.

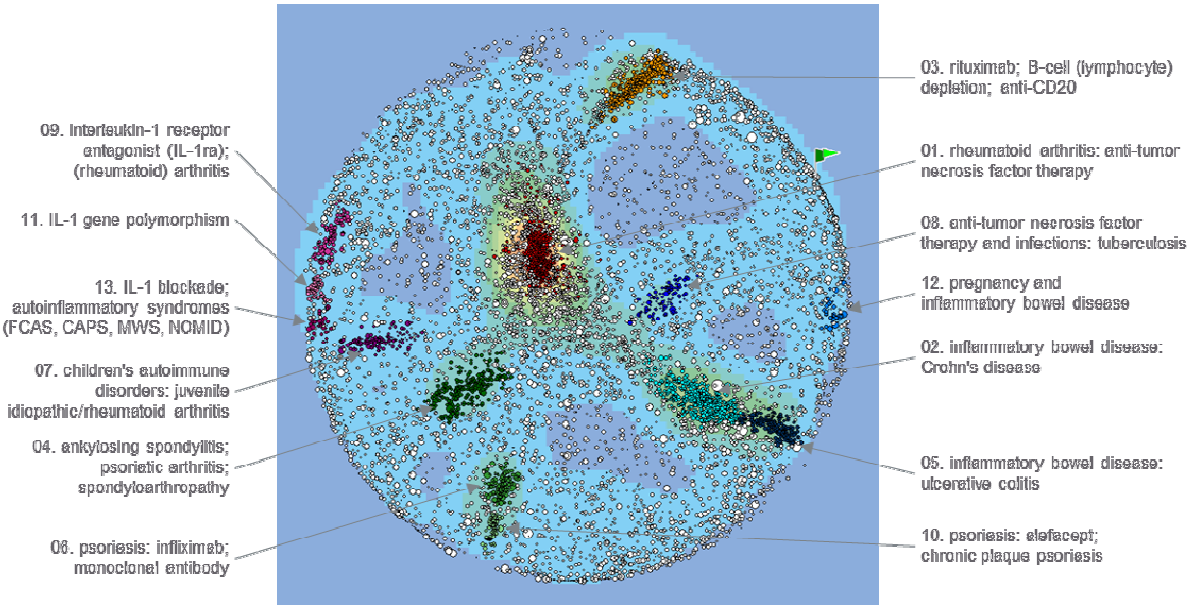


Figure 6.6: Map of research fronts (bibliographically coupled publications) on “Target Immune Modulators (TIM)”. Nodes are coloured by topics. Grouping of publication on their common references; circle: publication, the size corresponds to the number of references; edges: Jaccard index of co-frequencies; background: density map of bibliographically coupled publications, cos-weighted moving average; total number of publications: 10,688; date of search: 4 2012; number of nodes: 8,712; number of edges: 3,844,561 (none shown); data source: Web of Science.

The research fronts “psoriasis: infliximab; monoclonal antibody”, “ankylosing spondylitis; psoriatic arthritis; spondyloarthropathy” and “IL-1 blockade; autoinflammatory syndromes (FCAS, CAPS, MWS, NOMID)” have a share of more than three per cent of publications that refer to a registration ID of a clinical trial. This is smaller compared to APTAD. There is a clear need to catch up in the registration of clinical studies.

The publications spread over the landscape do not have similar reference lists to a bigger agglomeration of a research front. The reasons are that there are single publications that do not refer to an established knowledge base or just use one or more keywords of the search strategy in a different context than publications of the “main stream” research fronts.

Table 6.6: Research fronts for the case "Target Immune Modulators (TIM)" and number of publications with a registration ID. Total number of publications: 10,688; date of search: 4 2012; data source: Web of Science.

	Research fronts	Number of publi-cations	Number of publ. with registra-tion ID	Share of publ. with registra-tion ID in%
1.	rheumatoid arthritis: anti-tumor necrosis factor therapy	590	10	1.7
2.	inflammatory bowel disease: Crohn's disease	410	6	1.5
3.	rituximab; B-cell (lymphocyte) depletion; anti-CD20	293	5	1.7
4.	ankylosing spondylitis; psoriatic arthritis; spondyloarthropathy	239	9	3.8
5.	inflammatory bowel disease: ulcerative colitis	217	3	1.4
6.	psoriasis: infliximab; monoclonal antibody	130	6	4.6
7.	children's autoimmune disorders: juvenile idiopathic/rheumatoid arthritis	123	3	2.4
8.	anti-tumor necrosis factor therapy and infections: tuberculosis	110	1	0.9
9.	interleukin-1 receptor antagonist (IL-1ra); (rheumatoid) arthritis	91	0	0.0
10.	psoriasis: alefacept; chronic plaque psoriasis	88	0	0.0
11.	IL-1 gene polymorphism	58	0	0.0
12.	pregnancy and inflammatory bowel disease	39	0	0.0
13.	IL-1 blockade; autoinflammatory syndromes (FCAS, CAPS, MWS, NOMID)	32	1	3.1

6.3 Case "Diseases of the Cardiovascular System (DCS)"

We analysed a dataset for the case "Diseases of the Cardiovascular System (DCS)". The on-topic search, as it is described in Deliverable 2.1[18] resulted in 2,727 publications in the Web of Science database.

6.3.1 Knowledge bases (map of cited references)

Figure 6.7 shows the landscape of references for the medical case "Diseases of the Cardiovascular System (DCS)". We identified 18 agglomerations that represent issues of research in the case of DCS. The largest knowledge bases are formed by more than 150 references within the whole set of 2,727 publications: stem cells; myocardial infarction

[18] Schiebel, E., Palensky, B., Züger, M.-E. Deliverable D2.1 of the UNCOVER FP7-funded project under contract number 282574: Data sources for bibliometric analysis, 2013.

(238 references), atrial fibrillation; ablation (radiofrequency, catheter, pulmonary-vein, ...) (173), diabetes mellitus (type 2); pioglitazone (153) and drug-eluting stents; bare-metal stents (154).

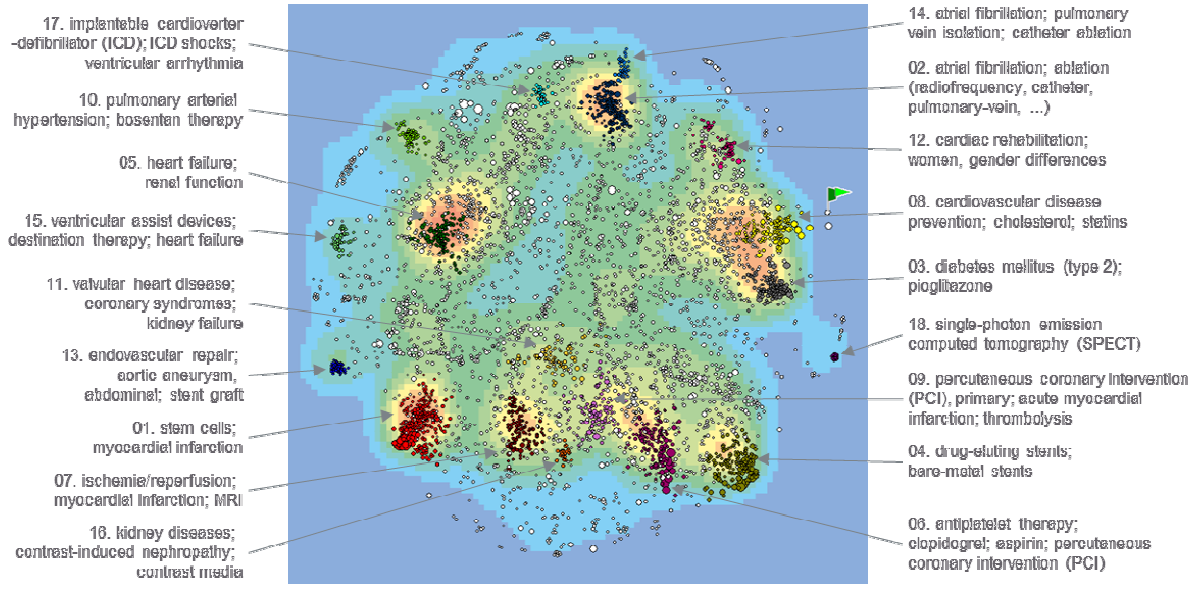


Figure 6.7: Map of knowledge bases (co-citation analysis) for the case "Diseases of the Cardiovascular System (CSD)". Nodes are coloured by topics. Grouping of cited references by their common appearance in publications; circle: reference, the size corresponds to the number of publications; edges: Jaccard index of co-frequencies in citing publications. Background: density map of the local number of references weighted by the Jaccard index of all links of the node, cos-weighted moving average filter; total number of publications: 2,727; date of search: 3 2013; each reference occurs in at least 3 publications; number of nodes: 4,860; number of edges: 213,200 (none shown); data source: Web of Science.

Table 6.7 lists the knowledge bases and some indicators such as the number of references that form the agglomeration, the number and share of references that are cited by publications referring to a registration ID, the number of the citing publications in total and citing a registration ID.

Publications with a reference to a registration ID cite a high share of references of the following knowledge bases: "drug-eluting stents; bare-metal stents" (82.5% of all references in this knowledge base), "drug-eluting stents; bare-metal stents" (54.5%), atrial fibrillation; ablation (radiofrequency, catheter, pulmonary-vein, ...) (45.1%) and ischemia/reperfusion; myocardial infarction; MRI (43.7%). Such knowledge bases contributed by more than 40% to publications on registered clinical trials by their number of references. The three knowledge bases "drug-eluting stents; bare-metal stents" (18.7 % share of publications with a registration ID), "heart failure; renal function" (13.5%) and "cardiovascular disease prevention; cholesterol; statins" (13.3%) are cited by the highest

number of publications that refer to a registration ID. This case has the highest shares of the three examined ones.

Table 6.7: Knowledge bases for the case "Diseases of the Cardiovascular System (CSD)" and number of publications with a registration ID. Total number of publications: 2,727; date of search: 3 2013; data source: Web of Science.

	Topics (1)	Number of references (2)	Number of ref. cited by publ. with registration ID (3)	Share of ref. cited by (6) in % (4)	Number of citing publ. (5)	Number of citing publ. with registration ID (6)	Share of publ. with registration in topic (6)/(5) in % (7)
1.	stem cells; myocardial infarction	238	7	2.9	148	2	1.4
2.	atrial fibrillation; ablation (radiofrequency, catheter, pulmonary-vein, ...)	173	78	45.1	141	15	10.6
3.	diabetes mellitus (type 2); pioglitazone	153	21	13.7	139	11	7.9
4.	drug-eluting stents; bare-metal stents	154	127	82.5	198	37	18.7
5.	heart failure; renal function	144	46	31.9	141	19	13.5
6.	antiplatelet therapy; clopidogrel; aspirin; percutaneous coronary intervention (PCI)	136	32	23.5	195	15	7.7
7.	ischemia/reperfusion; myocardial infarction; MRI	135	59	43.7	114	14	12.3
8.	cardiovascular disease prevention; cholesterol; statins	121	66	54.5	180	24	13.3
9.	percutaneous coronary intervention (PCI), primary; acute myocardial infarction; thrombolysis	85	15	17.6	138	9	6.5
10.	pulmonary arterial hypertension; bosentan therapy	77	20	26.0	40	2	5.0
11.	valvular heart disease; coronary syndromes; kidney failure	73	13	17.8	167	16	9.6
12.	cardiac rehabilitation; women, gender	69	6	8.7	70	6	8.6

	Topics (1)	Number of references (2)	Number of ref. cited by publ. with registration ID (3)	Share of ref. cited by (6) in % (4)	Number of citing publ. (5)	Number of citing publ. with registration ID (6)	Share of publ. with registration in topic (6)/(5) in % (7)
	differences						
13.	endovascular repair; aortic aneurysm, abdominal; stent graft	55	0	0.0	27	0	0.0
14.	atrial fibrillation; pulmonary vein isolation; catheter ablation	44	15	34.1	33	2	6.1
15.	ventricular assist devices; destination therapy; heart failure	40	7	17.5	41	1	2.4
16.	kidney diseases; contrast-induced nephropathy; contrast media	36	0	0.0	24	0	0.0
17.	implantable cardioverter-defibrillator (ICD); ICD shocks; ventricular arrhythmia	34	6	17.6	39	3	7.7
18.	Single-photon emission computed tomography (SPECT)	19	0	0.0	5	0	0.0

6.3.2 Research fronts – bibliographically coupled publications

The map of research fronts draws agglomerations of similar publications. The similarity is measured by the relative number (Jaccard Index) of common references and positioned by a spring algorithm.

Figure 6.8 shows the map of research fronts. Published research findings in the case DCS form several research fronts distributed over the landscape

We separated dominating research activities in the themes: “drug-eluting stents; bare-metal stents; restenosis; thrombosis” (122 publications), “stem cells; myocardial infarction” (105), “percutaneous coronary intervention (PCI); antiplatelet therapy; drug-eluting stents” (99), “cardiovascular disease prevention: diabetes, cholosterol; coronary-heart-disease” (86) and “implantable cardioverter-defibrillator (ICD); cardiac resynchronization therapy” (72).

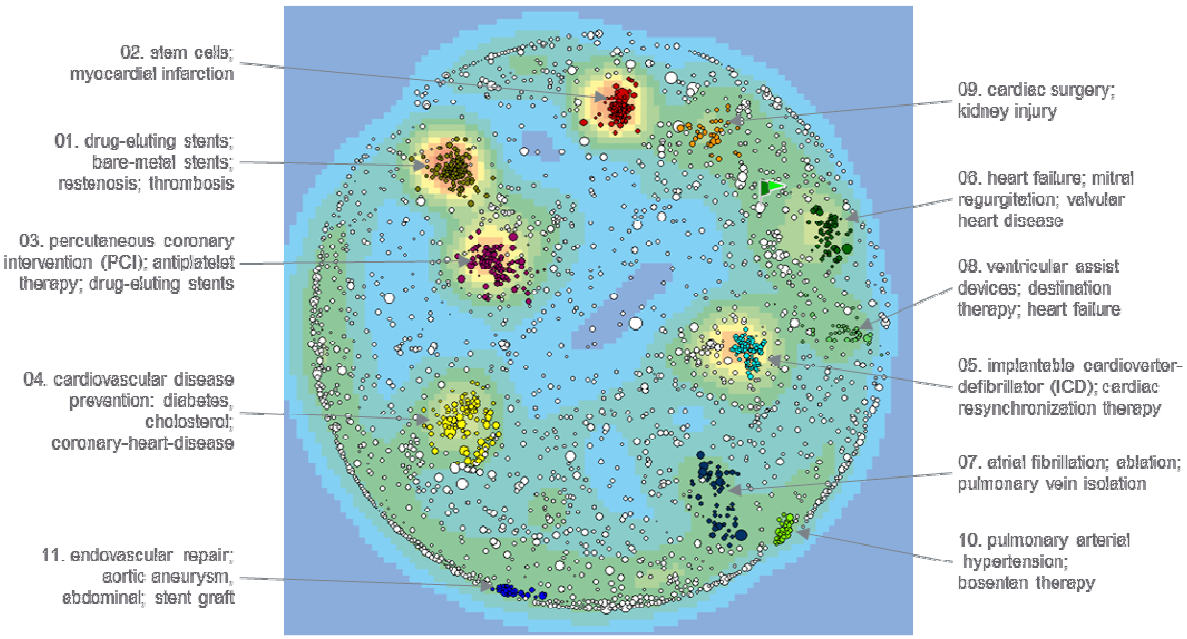


Figure 6.8: Map of research fronts (bibliographically coupled publications) for the case "Diseases of the Cardiovascular System (CSD)". Nodes are coloured by topics. Grouping of publication on their common references; circle: publication, the size corresponds to the number of references; edges: Jaccard index of co-frequencies; background: density map of bibliographically coupled publications, cos-weighted moving average filter; total number of publications: 2,727; date of search: 3 2013; number of nodes: 2,468; number of edges: 63,906 (none shown); data source: Web of Science.

Table 6.8 lists the research fronts together with the number of publications with a registration ID. The following research fronts have a relative high number of publications on registered clinical trials: "drug-eluting stents; bare-metal stents; restenosis; thrombosis" (22.1%) "cardiac surgery; kidney injury" (12.9%) and "cardiovascular disease prevention: diabetes, cholesterol; coronary-heart-disease" (10.5%).

Table 6.8: Research fronts of the case "Diseases of the Cardiovascular System (CSD)" with the share of publications with a registration ID. Total number of publications: 2,727. Date of search: 3 2013; data source: Web of Science**.**

	Topics	Number of publications	Number of publ. with registration	Share of publ. with registration in topic in %
1.	drug-eluting stents; bare-metal stents; restenosis; thrombosis	122	27	22.1
2.	stem cells; myocardial infarction	105	1	1.0
3.	percutaneous coronary intervention (PCI); antiplatelet therapy; drug-eluting stents	99	6	6.1
4.	cardiovascular disease prevention: diabetes, cholesterol; coronary-heart-disease	86	9	10.5

	Topics	Number of publications	Number of publ. with registration	Share of publ. with registration in topic in %
5.	implantable cardioverter-defibrillator (ICD); cardiac resynchronization therapy	72	5	6.9
6.	heart failure; mitral regurgitation; valvular heart disease	57	4	7.0
7.	atrial fibrillation; ablation; pulmonary vein isolation	52	5	9.6
8.	ventricular assist devices; destination therapy; heart failure	36	1	2.8
9.	cardiac surgery; kidney injury	31	4	12.9
10.	pulmonary arterial hypertension; bosentan therapy	28	1	3.6
11.	endovascular repair; aortic aneurysm, abdominal; stent graft	26	0	0.0

6.4 Identification of additional publications on clinical trials by bibliographic coupling

One of the challenges in systematic reviews and meta-analyses is to find all relevant publications. Keyword-oriented searches can result in too many not relevant publications on the one hand or exclude relevant publications on the other hand if the publication text does not use the expected keywords. In this chapter we propose to use bibliometric coupling as an approach to identify “more” relevant publications on clinical trials with the following procedure:

1. Generate a set I of publications with an on-topic search by keywords in WoS, collect a big set of publications without restrictions to keywords that identify clinical studies;
2. Use the PubMed fields for an on-topic search with a restriction to clinical studies to collect a set II of publications;
3. Match the publications of set I with the publications of set II by the title of the publication to produce a set III of common publications on clinical trials in the topic;
4. Select all publications of set III in the WoS research fronts map of bibliographically coupled publications of set I;
5. Select all publications that are similar to set III as set IV;
6. Identify publications in set IV that cite a registration ID in the text to make sure that additional publications on a registered clinical trial are found.

We applied this procedure for the cases APTAD and TIM. The results are given in Table 6.9. For APTAD we identified 579 publications with a Jaccard index above a threshold of 0.1. These are publications that share at least 10% of common references with publications on clinical trials. For TIM we identified 2,027 publications above the same threshold. To make sure that we find additional publications on clinical trials we selected publications that cite a registration ID of ClinicalTrials.gov or ISRCTN or publish in an ICMJE listed journal. For APTAD we identified 16 and 159 publications, respectively, and for TIM 25 publications.

Table 6.9: Number of publications for the cases APTAD and TIM above a similarity threshold for the Jaccard Index being >0.1.

Data set	Number of publications above threshold	Number of publications with registration (NCT/ISRCTN)	Number of publications in ICMJE listed journals
APTAD	579	16	159
TIM	2,037	25	unknown

Table 6.10 presents a list of the most similar publications measured by the Jaccard index for the case APTAD. The publication with the WoS ID WOS:A1991GG98000021 has at least one common reference to any of 33 other publications. The sum of all Jaccard index values of the 33 publications is 2.312. One of the 33 publications has a Jaccard index value of 0.811. Publications citing a registration ID are not identified in the most similar 20 publications, but some publish in a journal that is listed in ICMJE.

Table 6.10: Publications similar to any publication on a clinical trial ranked by the maximum of Jaccard index for the case APTAD. Frequency is the number of publications on clinical trials with at least one common reference to the listed publication; maximum of Jaccard is the highest value of similarity of one of the publications listed under “Frequency”; sum of Jaccard is the sum of all values of the publication number listed by “Frequency”. Registration gives the registration ID in a clinical trial registry and ICMJE if the publication is an ICMJE listed journal.

WoS ID	Frequency	Sum of Jaccard	Maximum of Jaccard	Registration	ICMJE
WOS:A1991GG98000021	33	2.312	0.811		ICMJE listed Journal
WOS:000280312800009	23	1.974	0.515		
WOS:000250506100058	28	4.864	0.5		
WOS:A1991FN81700004	16	2.605	0.476		
WOS:000178966000011	40	1.198	0.473		ICMJE listed Journal
WOS:000248727200013	33	8.458	0.471		ICMJE listed Journal
WOS:000267738000005	58	7.586	0.449		
WOS:000269810700002	28	9.566	0.404		
WOS:000168347800008	28	8.012	0.364		

WoS ID	Frequency	Sum of Jaccard	Maximum of Jaccard	Registration	ICMJE
WOS:000225239400006	35	2.261	0.348		
WOS:000239476100012	34	3.299	0.347		
WOS:000225508200016	38	7.012	0.344		
WOS:000169026600011	32	7.453	0.324		
WOS:000240205100011	49	7.89	0.319		ICMJE listed Journal
WOS:000234367100010	71	6.259	0.313		ICMJE listed Journal
WOS:000174613400010	21	6.998	0.31		ICMJE listed Journal
WOS:000177314900009	17	2.607	0.31		
WOS:000267738000004	44	8.68	0.304		
WOS:000223799200031	46	4.328	0.3		
WOS:000244845100007	34	1.859	0.3		

The following two publications are examples of identified additional publications on registered clinical trials in the case APTAD.

Painful physical symptoms and treatment outcome in major depressive disorder: a STAR*D (Sequenced Treatment Alternatives to Relieve Depression) report *Leuchter, AF; Husain, MM; Cook, IA; Trivedi, MH; Wisniewski, SR; Gilmer, WS; Luther, JF; Fava, M; Rush, AJ*

Publication Year: 2010 **Registration**: NCT **Abstract**: Background. Painful physical symptoms (PPS) are both common and reduce the likelihood of remission in major depressive disorder (MDD), based upon results of clinical trials in selected populations. Whether PI'S significantly contribute to poorer treatment outcome overall in primary or specialty psychiatric care settings remains Unclear. Method. Out-patients (n = 2876) with MDD were treated in the first step of the Sequenced Treatment Alternatives to Relieve Depression (STAR*D) trial with citalopram up to 60 mg/day for LIP to 14 weeks. Presence of painful symptoms, as well as severity of depression, physical illness, and demographic and treatment factors were examined. Time to and overall rates of remission were analysed in relation to the presence of PPS. Results. Of the participants, 80% complained of PPS. These patients, both in primary and specialty psychiatric settings, had significantly, lower remission rates and took longer to remit. Increasing severity of PIS was associated with greater physical illness burden, lower socio-economic status, absence of private insurance and being female, African-American or Hispanic. After adjustment for these factors, patients with PPS no longer had significantly poorer treatment outcomes. Conclusions. Presence and severity of PPS is an indicator of MDD that may have poorer treatment outcome with in initial selective serotonin reuptake inhibitor. These poorer treatment outcomes are multifactorial, however, and are not explained by the presence and severity of pain per se.

Times Cited: 11 **ID**: WOS:000274422600007 **Author Keywords**: Antidepressant medication; major depression; pain; treatment response **Keywords Plus®**: DIAGNOSTIC SCREENING QUESTIONNAIRE; REPORT QIDS-SR; QUICK INVENTORY; RATING-SCALE; PRIMARY-CARE; PSYCHOMETRIC EVALUATION; BACK-PAIN; SYMPTOMATOLOGY; TRIAL; ANTIDEPRESSANTS

Transcranial magnetic stimulation in the treatment of major depressive disorder: A comprehensive summary of safety experience from acute exposure, extended exposure, and during reintroduction treatment *Janicak, PG; O'Reardon, JP; Sampson, SM; Husain, MM; Lisanby, SH; Rado, JT; Heart, KL; Demitrack, MA*

Publication Year: 2008 **Registration**: NCT and ICMJE listed Journal **Abstract**: Background: Transcranial magnetic stimulation (TMS) has demonstrated efficacy in the treatment of major depressive disorder; however, prior studies have provided only partial safety information. We examined the acute efficacy of TMS in a randomized sham-controlled trial, under openlabel conditions, and its durability of benefit. Method: Aggregate safety data were obtained from a comprehensive clinical development program examining the use of TMS in the treatment of major depressive disorder. There were 3 separate clinical protocols, including 325 patients from 23 clinical sites in the United States, Australia, and

Canada. Active enrollment occurred between January 2004 and August 2005. Adverse events were assessed at each study visit by review of spontaneous reports with separate reporting of serious adverse events. Safety assessments were also completed for cognitive function and auditory threshold. Assessment of disease-specific risk included the potential for worsening of depressive symptoms. Finally, the time course and accommodation to the most commonly appearing adverse events were considered. Results: TMS was administered in over 10,000 cumulative treatment sessions in the study program. There were no deaths or seizures. Most adverse events were mild to moderate in intensity. Transient headaches and scalp discomfort were the most common adverse events. Auditory threshold and cognitive function did not change. There was a low discontinuation rate (4.5%) due to adverse events during acute treatment. Conclusions: TMS was associated with a low incidence of adverse events that were mild to moderate in intensity and demonstrated a largely predictable time course of resolution. TMS may offer clinicians a novel, well-tolerated alternative for the treatment of major depressive disorder that can be safely administered in an outpatient setting. Trial Registration: clinicaltrials.gov Identifier: NCT00104611.

Times Cited: 47 **ID**: WOS:000253506300008 **Keywords Plus®**: STAR-ASTERISK-D; FAILED MEDICATION TREATMENTS; CONTROLLED-TRIAL; SINGLE-PULSE; EPILEPSY; SEIZURE; AUGMENTATION; TOLERABILITY; MIRTAZAPINE; PARAMETERS

7 Summary and conclusions

This deliverable offers some answers to central questions about the publication of registered clinical trials that have been formulated as hypotheses.

The first hypothesis formulated was that *the registration ID for clinical trials is cited in publications*. We have shown that for Web of Science the trial registration ID is part of the abstract, provided usually at the end of a structured abstract or mentioned somewhere else in the text. Additionally, PubMed offers the database field “DB Accession Number” where the trial registry number is given. Yet, as we have shown, the information provided is still very sparse.

The hypothesis *the number of publications that refer to a registration ID is growing* could also be confirmed. One of the preconditions is that clinical trials are registered in a publically available database. Currently, the most comprehensive registry is ClinicalTrials.gov (based in the USA) with a remarkable growing number of registrations and with about 140 thousand registered trials in 2013. Accordingly, we have shown that the number of publications that refer to a registration ID has also grown remarkably to about 2,000 publications in Web of Science in 2012. The authors of the publications are spread mostly over the USA, Canada and even Europe, although the above mentioned registry is based in the USA. Furthermore, it is also remarkable that the funding of reported research findings is dominated by private pharmaceutical companies such as Pfizer, Novartis, GlaxoSmithKline, Astra Zeneca, Bristol Myers Squibb, Sanovi Aventis, Merck, Roche, Böhringer Ingelheim Eli Lilly, rather than by public funding bodies such as the National Institute of Health. It also seems that the tendency to cite registration IDs differs between medical specialities. For example, publications in the area of general internal

medicine or cardiovascular system cardiology tend to cite registration IDs more frequently than publications on topics concerning oncology pharmacology, neurosciences or surgery.

We established the Hypothesis: *If editors insist on publications referring to registered studies, the number of such publications growths*. To accelerate the publication of registered studies the ICMJE initiative forced trialists to register studies by only accepting publications on clinical trials that have been registered before. By analysing the publication source, i.e. the journals, of publications on registered studies we revealed that 19 out of 30 journals with the highest number of publications referring to a registration ID follow ICMJE requirements. They play an important role concerning the registration of clinical trials. It seems that the ICMJE initiative has a favourable influence on the growing number of publications that refer to a registration ID.

We found that *publications that refer to a registered trial ID have higher citation rates than others*. The citation rate of published research is of growing interest for individual authors and editors. For authors the times cited of their publication is an important indicator for the visibility of their work and is commonly measured by the Hirsch Index that ranks the publications of the individual researcher by the descending number of times cited. The attractiveness of a journal is measured by the Journal Impact Factor that is also based on the citation of articles in these journals. Editors are interested in journals with a high Journal Impact Factor. We conclude that higher citation rates are powerful incentives which make it more likely for authors and editors to publish clinical studies that are registered than those that are not registered.

We applied relational bibliometric techniques, especially co-authorships, co-citation analysis and bibliographic coupling, to enlarge the view on medical research topics for systematic reviews and meta-analyses. The analysis of the networks of co-authorships showed that an on-topic search without a restriction on clinical trials identifies authors who are engaged in the research topics and that publications referring to clinical trials are only a part of their work. More comprehensive systematic reviews should take into account such work. Knowledge bases and research fronts offer a broad view of worldwide publications in a medical issue. Some research is done by clinical studies but clinical studies are not the only approach. Additional aspects should also be reflected in evidence-based medicine.

The analysis of bibliometric research landscapes demonstrated that it is possible to broaden the view on research issues in systematic reviews and meta-analysis. We analysed the thematic research activities from two perspectives. The first perspective is the so called intellectual knowledge base. It is identified by a co-citation analysis of backward citations. All cited references represent the intellectual base of written publications. The co-citation landscape visualizes the agglomerations of similar references by their common occurrence

in publications. The agglomerations are named by keywords and titles of the citing documents.