EJM
Erasmus Journal of Medicine
No.2
Oktober 2013 • vol 3 - no 2
Erasmus Journal of Medicine: independent scientific journal

Treatment of idiopathic thrombocytopenic purpura

Systematic review

Perioperative chemotherapy

Systematic review

Editorial comment

Ethics of placebo treatment

Opinion

Knee replacement in the morbidly obese

Colophon

The Erasmus Journal of Medicine (EJM) is a scientific magazine by and for medical students, especially students of Erasmus MC University Medical Center Rotterdam. It was initiated by the MFVR (the medical students' organization of Erasmus MC). The journal appears twice a year. It is published on paper (1000 copies) and on the EJM website (www.erasmusjournalofmedicine.nl).

The main purpose of the EJM is to encourage young medical students to conduct research (empirical studies or systematic reviews) and report on this research, while becoming acquainted with a professional publishing process as authors, reviewers or editors. A secondary purpose is to make the results of excellent student-driven research known to others.

The journal contains articles describing original research, systematic reviews, extended abstracts (summaries of recently conducted studies), calls from research projects for students to participate, opinion papers written by students, editorial comments and letters to the editor.

The Erasmus Journal of Medicine is funded by Erasmus MC, through an unrestricted grant. Papers may be submitted by everyone, but the first author has to be a medical student.

Publisher
Erasmus MC University Medical Center Rotterdam.

Formatting
Ditems Media, Monnickendam

Printing
Verhagen Communicatie BV, Hooglanderveen

Correspondence
Erasmus Journal of Medicine
Petra Erkens, editorial assistant
Room Gk-658
Postbus 2040
3000 CA Rotterdam
E-mail: ejm@erasmusmc.nl

Website
www.erasmusjournalofmedicine.nl. Like us on www.facebook.com/erasmusjournalofmedicine.

Foreword

EJM: young yet established

Erasmus Journal of Medicine begins its third year of publication. It is a young scientific journal that has matured very fast. This became clear during In Praise of Medicine, the annual public lecture organized by Erasmus MC on biomedical science and people, held on 4 October at Congress Center De Doelen in Rotterdam. Following lectures on pain research by pediatricians and researchers Dick Tibboel (Erasmus MC) and Sunny Anand (University of Tennessee), the Erasmus MC Fellowships were awarded to promising young researchers who, despite their impressive achievements, are still at the start of their careers.

The fellowships are intended for researchers to develop their own line of research and spread their wings to the international community of scientific research. Arfan Ikram, Carolien van Deurzen, Bert Mik, Hester Lingsma and Peter de Keizer are this year's Erasmus MC Fellows and are all an inspiring example for our students. Especially when you realize that the Fellows themselves only recently graduated and received their doctorates.

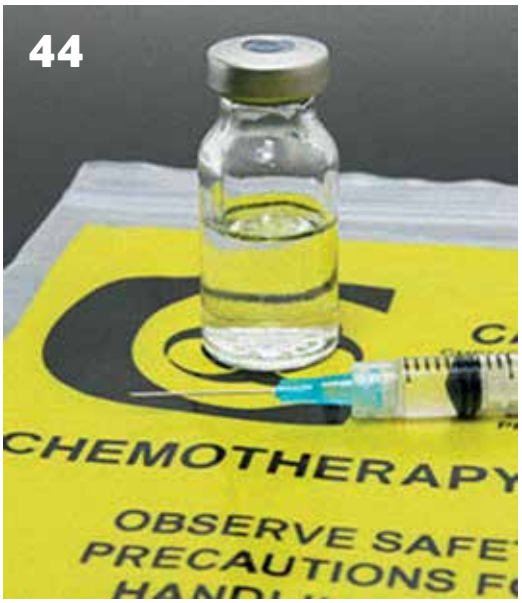
It was therefore particularly nice that I also had the honor of presenting the first ever prize for the best article in Erasmus Journal of Medicine, in addition to presenting the Fellowships during In Praise of Medicine. This creates a continuous line between the medical students who conducted their first real scientific research; the Erasmus MC Fellows who have only just passed this stage and are now emerging as a new generation of excellent researchers; and the established researchers who are often the face of their research institute and even their field of research.

It was a great honor to present the EJM award for the best article to Marlene Mende for an article that she co-authored with Annemarie van Leeuwen, Marleen Bakker and Coenraad Koegelenberg. According to the selection committee, their article 'Liquid based vs. conventional cytology for evaluation of fine needle aspiration biopsies obtained by pulmonary physicians. A pilot study' gives an accurate and realistic description of scientific research in a clinical setting. The selection committee, consisting of Axel Themmen, Aart Jan van der Lelij and Bas Hullegie, reviewed the article on originality of the research question, execution of the research and interpretation of the research results in the article.

The introduction of the EJM award for the best article is a wonderful achievement for our own journal. Given the age of its authors and readers, the journal will always stay young, but at the same time it is working at becoming an established institute.

Prof. Jaap Verweij, Dean

Erasmus Journal of Medicine



Half a dozen of EJMs: we are getting somewhere!

Erasmus Journal of Medicine has grown. The journal started out as a small paper run by a small group of staff members and medical students who formed the editorial board and performed all the reviewing and other work.

Now, three years later, we have a regularly (twice yearly) appearing journal with a well-organized basis of professionals and students. All positions are manned by staff members ánd students, because our aim is to educate and train our students in every aspect of writing, editing and reviewing. We have a large board of reviewers (staff members and students), whose concise recommendations assist us in making adequate decisions. At the moment, we have assured ourselves of the help of dedicated and experienced medical language editors: Charles Frink and Ed Hull. We are assisted by a professional editorial manager: Petra Erkens. Our new web-based editorial management system helps us to keep track of all submitted papers and respond promptly. Annually, we report to our supervisory board on finance, logistics and our editorial policy.

Almost all of the received papers are unsolicited, and about 20% of the papers need to be rejected. We distinguish ourselves from other journals not primarily aiming at medical students in providing more concrete help and constructive feedback during the entire process of rewriting and resubmitting. When a paper is accepted for publication, an interactive process of language editing starts. In our view, submitting a paper to the Erasmus Journal of Medicine offers a unique and valuable learning experience for those who are interested in a scientific career in medicine.

We are proud of our journal and what it has grown into. Much has yet to be done, and new issues and challenges are looming ahead, that require fresh insights and new energy. It is therefore time for me to step back. Ajda Rowshani will take over the chair of the editorial board. We are very happy to welcome her in this position. I wish the journal a long and healthy life.

Diederik Dippel, MD, PhD
neurologist (chairman of the editorial board)

Ajda Rowshani (future chairman of EJM Editoral Board) and Marlene Mende (Winner EJM Award 2013)

A lot happened

Herewith, the students of the editorial board present you the sixth issue of the Erasmus Journal of Medicine. We are proud that EJM is still an evolving project, giving medical students a unique chance for scientific experience.

A lot happened since launching our last edition. First, one of the establishers of our journal, prof. Huibert Pols, has left his functions as dean of the medical faculty and honorary editor in chief of our journal. The editorial board is very thankful for all his efforts in EJM. We wish him all the best in his new function of Rector Magnificus of the Erasmus University Rotterdam.

We would like to present you our new honorary editor in chief, prof. Jaap Verweij, the new dean of the Erasmus Medical Center. We are looking forward to the years of cooperation with him and we are glad with his support for our journal.

His support was clearly stated at the last edition of In Praise of Medicine. Here he solemnly handed out the first Erasmus Journal of Medicine Award to the author of the best article of our last volume. We would like to congratulate Marlene Mende and co-authors with this marvelous achievement. We hope that the establishment of this Award will be another stimulus for students to get the best out of their scientific ambitions.

But above all, a lot of wonderful research is done by medical students since launching or last issue. In this edition of the Erasmus Journal of Medicine, an overview of the best papers is presented.

Placebo treatment: is it ethical?

In 2006 in Boston 82 healthy paid volunteers were recruited by means of an online advertisement. [1] Each participant was informed by brochure about a new opioid analgesic, but it was actually a placebo pill. After randomization, half of the participants were informed that the drug had a regular price of $2.50 per pill and half that the price had been discounted to $0.10 per pill (no reason for the discount was mentioned). All participants received identical placebo containing pills.

Electrical shocks to the wrist were calibrated to each participant's pain tolerance. Visual analog scale ratings for pain were converted to a 100-point scale, the post-pill score for each voltage was subtracted from the pre-pill score, and the mean of these differences was calculated for each participant. Considering all voltages tested, pain reduction was greater for the regular-priced pill. The study shows that patients' expectations can influence treatment outcome. In this experiment the price of the drug influenced expectations. There are numerous examples of physician related factors influencing outcome. I recommend you do take a look at the You Tube movie "The Strange Powers of the Placebo effect".[2]

In daily practice I sometimes see patients complaining of headache, dizziness or abdominal pain, for which no clear cause can be identified. I never prescribe placebo treatment, but I do occasionally initiate treatment with drugs that in my view have a rather low intrinsic therapeutic effect. What I typically explain the patients is that it is my intention to try out the treatment for a defined period of time, that I have seen stunning successes in previously treated patients with similar symptoms and that I am eager to see the effects. I am convinced that the expectations that I raise positively do influence outcome. In complementary medicine a large proportion of the treatment effect is based on the patient-doctor interaction, and traditional medicine can learn a lot from complementary medicine on how to use this determinant of outcome to the advantage of the patient.

In this issue of EJM Nico Jansen discusses the use of placebos in the management of medically unexplained symptoms. [3] While use of placebos in (double-blind) clinical trials is generally accepted, for patient care such placebos are more controversial, and for patients with medically unexplained symptoms this is also the case. Patients may interpret prescription of a placebo as a fake treatment, and as a sign that their doctor thinks their symptoms are deliberately faked as well. In the past in patients with medically unexplained symptoms trials have been performed with tricyclic antidepressants (TCAs) and selective serotonin reuptake inhibitors (SSRIs). These drugs with substantial side-effects at best were effective in small proportions of patients only. Assuming a better adverse event profile for placebos it is tempting to investigate the added value of placebos in the treatment of medically unexplained symptoms, but in such trials ethical and psychological considerations should not be forgotten.

[1]. Waber RL, Shiv B, Carmon Z, Ariely D. Commercial features of placebo and therapeutic efficacy. JAMA 2008;299:1016-7.
[2]. http://www.youtube.com/watch?v=yfRVCaA5o18
[3]. Jansen N. Placebos in clinical practice: management of medically unexplained symptoms. EJM 2013; 3:57-59.

Teun van Gelder, MD, PhD
Professor of Clinical Pharmacology,
Erasmus MC Medical Center Rotterdam, the Netherlands
E-mail: t.vangelder@erasmusmc.nl

In the supermarket are the sirens of 'sugar, salt and saturated fat' enchanting: 'take me home, eat me'

In this issue of Erasmus Journal of Medicine, Roth and Bessems discuss the scientific and ethical grounds of unwillingness of orthopedic surgeons to operate on morbid obese patients. We all have heard about the global catastrophe of the 'obesity epidemic'. This is certainly not hot news. We see obese people every day, everywhere. In Europe, according to the WHO, obesity is one of the greatest public health challenges of the 21st century. Its prevalence has tripled in many European countries since the 1980s, and the numbers of those affected continue to rise, particularly among children. But who is to blame? The patients? The big food companies? The governments? McDonalds? The gnomes?

Stephen Sanger, CEO of General Mills, one of America's largest food industries says in 1999: 'Don't talk to me about nutrition. Talk to me about taste, and if this stuff tastes better, don't run around trying to sell stuff that doesn't taste good'. Sanger is cited in 'Salt, sugar, Fat: How the Food Giants Hooked us' by Michael Moss. Virtually everything you can buy in a supermarket that is not an outer-aisle pure food has been fiddled with sugar, salt or saturated fat. Almost everything! Why is that? Because we like it. And because we like it, the big food companies add it to our food. In substantial quantities. We choose our food on taste, not on ingredients. That sells! Food scientists use cutting-edge technology to calculate the 'bliss point' of sugary beverages or enhance the 'mouthfeel' of fat by manipulating its chemical structure. The food companies' marketing campaigns are designed to redirect concerns about the health risks of the processed products (just like the tobacco companies did, and still do). Consumption of food that is rich in added high fructose corn syrup (HFCS) results in increased visceral adiposity, lipid dysregulation and decreased insulin sensitivity. This results in the metabolic syndrome, increased risk for cardiovascular disease and type-2 diabetes. And every day almost everyone is eating food with added HFCS, mostly without knowing it. All those added sweeteners pose serious dangers to our health.

I think the food companies are not to blame. Without sugar, salt and saturated fat, the companies cease to exist. Their millions of users are addicted to the taste of their products, silently making them obese and sick. In the supermarket are the Sirens of sugar, salt and saturated fat enchanting: 'take me home, eat me'.

When the food companies are not to blame, who then is responsible for obesity and its devastating consequences? The individuals buying the food? Yes certainly, as it is a personal choice where you stick your fork in. But I would also put responsibility on the shoulders of the governments who allows the food companies to add sugar, salt and fat in sick making quantities. Governments should act in the same way as they (should) do in accordance with tobacco and alcohol. Telling that HFCS will kill you and regulate it.

What should health care providers do in this epidemic? First prevention. But only telling that eating all that processed sweet food makes us thick and sick does not work as we can obviously see. Physicians provide treatment. Mostly symptomatic treatment of the effects of obesity; treating high blood sugars, insulin resistance, hypertension, coronary heart disease and osteoarthritis with pharmaceuticals, stents, bariatric surgery and…orthopedic interventions.

We do not refuse a second coronary stent or third CABG as symptomatic treatment. We treat patients with multiple comorbidities. But, let us be honest, this is not curative medicine, this is palliative medicine. The palliative treatment of diseases resulting from overeating is core business of modern medicine. And as long as governments allow the food industries to add sick making large amounts of sugar, salt and fat to our food, the patients are not primarily to blamed and health care providers should provide palliative measures for the symptoms. Because that is what we do everyday. Maybe we should see knee replacement in obese patients only as a palliative measure and don't be idealistic about changing the lifestyle of our patients. Or we should give more sound arguments why we are resistant to provide palliative treatment in general to the enormous numbers of overweight people addicted to sweet taste.

1. Roth KC, Bessems, JHJM. Sorry, but you will have to lose weight before receiving your knee replacement. Erasmus J Med 2013;6:54-57.

Erwin Kompanje
Clinical ethicist, Department of Intensive care medicine. Erasmus MC University Medical Center Rotterdam, the Netherlands.
E-mail: erwinkompanje@me.com

Sleep problems in childhood predict later anxiety symptoms

A systematic review

Marloe van Berkel[a], Lauren Dautzenberg[a], Floor V.A. van Oort[b]
[a] Medical student, Erasmus MC University Medical Center Rotterdam, the Netherlands
[b] supervisor Dept. of Child and Adolescent Psychiatry, Erasmus University Medical Center Rotterdam, the Netherlands
Correspondence: L. Dautzenberg, email: 340978ld@student.eur.nl

Objective: **Prevention of anxiety and sleep problems is not effective. Sleep problems co-occur with anxiety in childhood. The relationship is unclear. We tested the following hypotheses in children: (1) sleep problems predict later anxiety symptoms and (2) anxiety symptoms predict later sleep problems.**
Methods: **We systematically searched the PubMed database for articles reporting the predictive relationship of sleep problems and anxiety in children, first measurement of anxiety or sleep problems starting at the age of 0-18 years.**
Results: **Seven studies met our inclusion criteria. Six out of seven studies reported a positive relationship between both general and subtypes of sleep problems from the age of 6 months until 19 years and later anxiety symptoms. Three studies investigated hypothesis 2 and only the presence of anxiety at the age of 18 months showed a significant association with nightmares 6 months later.**
Conclusions: **Both general and subtypes of sleep problems in childhood predict later anxiety symptoms. Subsequently, sleep problems should be treated to prevent later anxiety symptoms. Perhaps, an intervention more early in the process of developing sleep problems and anxiety problems is possible. We suggest further research into this. There was little support for an association between anxiety symptoms and later nightmares or other sleep problems. We suggest more investigations in childhood into underlying risk factors for both sleep problems and anxiety symptoms.**

Introduction

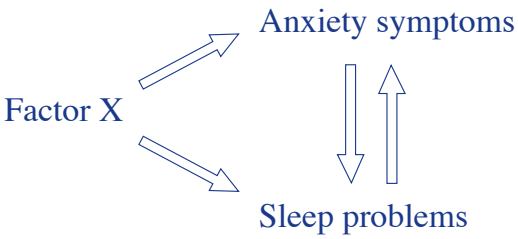
Anxiety symptoms and sleep problems are the most common psychiatric symptoms in children and adolescents, with prevalence rates in lifetime between 2-10% for anxiety symptoms [1,2] and 20-25% for sleep problems.[3,4]

Anxiety symptoms and sleep problems are linked in an unbalanced way. The diagnosis of anxiety is possible in the presence of sleep disturbances. In contrast, the diagnosis of a primary sleep problem is only possible after exclusion of an anxiety disorder.[5] According to the DSM criteria, the diagnosis Generalized Anxiety Disorder is associated with three or more of the following six symptoms: "restlessness of feeling keyed up or on edge; being easily fatigued; difficulty concentrating or mind going blank; irritability; muscle tension and sleep disturbance".[5] For this diagnosis in children, only one of these six symptoms is required. However, the DSM criteria do not address a potential long term, predictive relationship between anxiety and sleep problems. The existence of sleep problems in childhood might predict anxiety later on and vice versa. Hypothetically, in children, anxiety symptoms result directly in sleep problems; or sleep problems result directly in anxiety problems; or anxiety problems and sleep problems have a similar underlying risk factor X, for example environment, genetic variation and child characteristics(Figure 1).[6] Longitudinal research is essential to understand the complex relation between anxiety and sleep problems in children and adolescents.[3,4]

Prevention of sleep problems and anxiety symptoms could be more effective if the predictive relationship between these two problems is further clarified. Our goal was to determine the predictive relationship between sleep problems and anxiety symptoms in children and adolescents. Specifically, we did not search for factor X, but we tested the following hypotheses: in children (1) sleep problems predict later anxiety symptoms and (2) anxiety symptoms predict later sleep problems.

In a systematic review, we compared the supporting evidence for these 2 hypotheses. Specifically we addressed the following research question. Do sleep problems in childhood predict later anxiety symptoms, or do anxiety symptoms in childhood predict later sleep problems?

Figure 1-
Relationship between anxiety symptoms, sleep problems and factor X



Methods

Search strategy

On January 12th 2012, we searched the PubMed electronic database for English-language articles using the following Medical Subject Headings (MeSH): "Sleep Disorders"[Mesh] AND ("Anxiety Disorders"[Mesh] OR "Child Behaviour Disorders"[Mesh]) AND (longitudinal OR cohort OR follow-up). We limited the search to articles about All Children, 0-18 (years).

Selection criteria and quality assessment
We read the title and abstract of the articles to determine whether an article could be used in this review. To be eligible, the studies had to include all of the following items: 1) anxiety or sleep problems measured in childhood; 2) longitudinal studies; 3) report on a predictive association between sleep problems and anxiety. A study was excluded when: 1) the type of publication was a review or case study; 2) the studied populations suffered from other primary diagnoses than sleep disorders or anxiety symptoms, for example asthma, sleep disordered breathing or autism. If the articles seemed eligible, we read the full text. The full text also had to meet the inclusion and exclusion criteria.

Analysis
The primary outcome measure was the association between sleep problems and anxiety symptoms. First we produced an overview of all the included articles with author, publication date, study population, study subject, time-points of measurement, used tests, the measure of outcome of the studies and the results that were relevant to our review objective (Table 1, 2). We reported results that were adjusted for confounding factors (Table 2), in order to minimize false positive associations.

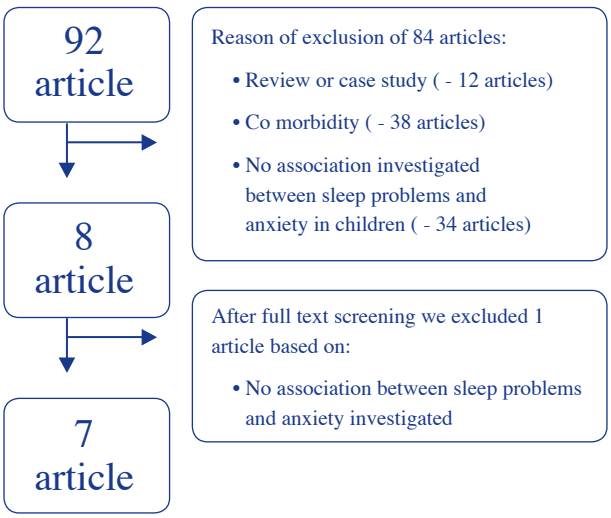
Results

Description of studies
Our Pubmed search resulted in 92 publications, of which 7 articles were included (Figure 2). The included studies are described in Table 1.

Subjects
The studies were all conducted in the Western World during peace-time, the follow-up period varied from 3 to 21 years, and six of these studies were population cohorts.[6-11] Approximately 50% of all the participants were male, most of the children were white, ranging from 81.7% to 96.6% and in most studies, the full range of social economic status was represented. In only one study, the results were not adjusted for confounding factors(Table 2).[12]

The ages at which sleep problems were assessed varied from the age of 2 months [6] to 19 years.[7] The ages at which anxiety was assessed varied from the age of 18 months [6] to 32 years.[7]

Figure 2 -
Flow chart representing the selection of studies during the literature search



Sleep problems
One study did not mention the frequency of sleep problems.[9] One study found that 21% of the children aged three to four years old had general sleep problems.[10] Three studies described sleep problems more in detail.[6, 7,11] One study focused on lower amounts of childhood sleep and found 24.5% of the children, aged 6-12 years, slept less than 7.5 hours a day.[11] One study found, at the age of 12 months, a frequency of 6% for > 2 awakenings at night, 30% slept < 12.5 hours, 22.2% had a presence of nightmares and 7.7% had an unstable sleep pattern.[6] One study found that 13% of the participants aged 4 to 19 years old were reported to sleep less than others, 4% were reported to be overtired and 9% were reported to have trouble sleeping.[7] Two studies focused on persistent sleep problems and found that 12.3% of the children with sleep problems at the age of 8-10 months still had sleep problems 3-4 years later [12] and 12.4% of the children with sleep problems, aged 5-7 years had sleep problems 2-4 years later.[8]

Anxiety problems
Two studies did mention the frequency of anxiety.[8,11] In a population cohort, 15.3% of the children aged 6-12 years old were anxious/depressed.[11] In children with persistent sleep problems, aged 5-9 years old, the frequency of anxiety in adulthood, 16-21 years later, was 46%.

Hypothesis 1
Early sleep problems predict later anxiety symptoms.
Both general and subtypes of sleep problems predict later anxiety symptoms, supported by all the included studies, except Lam et al [12], after controlling for several confounding factors (Table 2). The association between sleep problems and anxiety symptoms increased in mid-adolescence age.[9]

Hypothesis 2
Early anxiety symptoms predict later sleep problems.
Three studies investigated this hypothesis [6, 9, 12] and only the symptom presence of anxiety at the age of 18 months showed an association with nightmares 6 months later.[6]

Table 1 - Details of included studies

Author	Population	Country	Follow- up + Time-points of measurements	Used tests
Gregory et al., 2002 (9)	Patientcontrolled adoptive cohort	USA, Colorado	Follow-up: 11 years	
	N=490		Time-points of measurement of sleep problems: age 4,7,9,10,11,12,13,14 and 15 years	CBCL* (sleep problem items)
	54% male			
			Time-points of measurement of behaviour and emotional problems: age 4,7,9,10,11,12,13,14 and 15 years	CBCL (anxious/depressed scale)
Gregory et al., 2004 (10)	Twin pairs Population	England, Wales	Follow-up: 3-4 years	
	Cohort		Time-points of measurement of sleep problems: age 3 and 4 years	Questionnaire
	N=6491			
	49% male		Time-points of measurement of anxiety: age 3,4 and 7 years	Strenghts and Difficulties Questionnaire
Gregory et al., 2005 (8)	Population cohort	New Zealand, Dunedin	Follow-up: 16-21 years	
	N=943		Time-points of measurement of sleep problems: age 5,7 and 9 years	Questionnaire
	52% male			
			Time-points of measurement of anxiety: age 21 and 26 years	Standardized interview
Gregory et al., 2008 (7)	Population cohort	The Netherlands, Zuid-Holland	Follow-up: 14 years	
	N=2076		Time-points of measurement of sleep problems: age 4-16, 6-17, 8-19, 9-18 and 12-19 years	CBCL (sleep problem items)
	49% male			
			Time-points of measurement of emotional and behavioural symptoms: age 18-32 years	YASR**
Jansen et al., 2011 (6)	Population cohort	The Netherlands, Rotterdam	Follow-up: 3 years	
	N=4782		Time-points of measurement of sleep problems: age 2 months and 24 months	Parental questionnaire
	56% male			
			Time-points of measurement of anxiety/depression: age 18 months and 3 years	CBCL
Lam et al., 2003 (12)	Infants with sleep problems	Australia, Melbourne	Follow-up: 3 years	
	at 6-12 months		Time-points of measurement of sleep problems:	Standardized maternal
	N=114		age 6-12 months, 3-4 years	questionnaire
	57% male		Time-points of measurement of anxiety/depression: age 3-4 years	CBCL
Silva et al., 2011 (11)	Population cohort	USA, Arizona	Follow-up: 4-7 years	
	N=304		Time-points of measurement of sleep problems: age 6-12, 11-17 years	Polysomnogram***
	51% male			
			Time-points of measurement of anxiety: age 6-12, 10-18 years	CBCL

*CBCL: Child Behaviour Check List: a parent reported questionnaire (13)
** YASR: Young adult self-report, self-report of anxiety symptoms (14)
*** Polysomnogram: Continious recording of specific physiologic variables during sleep.

Table 2 - Summery of the results

Author	Hypothesis 1 Sleep → Anxiety	Hypothesis 2 Anxiety → Sleep	Covariates
Gregory et al.,	ß = 0.13 (0.04)	No significant association	Child sex, adoptive status and stability
2002 (9)	(p < .01)		of behavioural/emotional problems
sleep problems	This association between sleep problems		
	and anxiety increased with age		
Gregory et al.,	ß = 0.12	Δ	Anxiety level at age 3-4 years
2004 (10)	(p < 0.001)		
sleep problems			
Gregory et al.,	OR 1.60 (1.05-2.45)	Δ	Childhood internalizing problems,
2005 (8)			sex and socioeconomic status
persistent sleep problems			
Gregory et al.,	16-24% of the children aged 4 to 19 years old	Δ	Sex, age, socioeconomic status,
2008 (7)	with different types of sleep problems have		parentrated scores through development
	anxiety symptoms at the age of 18 to 32 years		for the difficulty being assessed
- sleeping less	OR 1.43 (1.07-1.90)		
than most kids			
- overtiredness	OR 1.37 (1.02-1.84)		
- trouble sleeping	OR 1.39 (1.02-1.89)		
Jansen et al.,			Child age, ethnicity, gender and
2011 (6)			the CBCL
- > 2 awakenings at night	OR 1.61 (1.19-2.17)	OR 1.05 (0.63-1.74)	
- sleep < 12.5 hrs	OR 1.32 (1.07-1.62)	OR 0.98 (0.73-1.32)	
- presence of nightmares	OR 1.34 (1.13-1.61)	OR 1.28 (1.00-1.65)	
- unstable sleep pattern	OR 1.23 (0.95-1.60)	OR 1.09 (0.79-1.51)	
Lam et al.,	Persistent sleep problems	No significant association	
2003 (12)	CBCL* scores:		
- no sleep problems	53(4)		
- sleep problems	53(5) p>.05		
Silva et al.,		Δ	BMI (kg/m^2), athnicity, sleep disordered
2011 (11)			breathing, age, caffeine use and baseline
			values
- sleep > 7.5 - 9 hrs	OR 2.4 (0.6-9.44)		
- sleep < 7.5 hrs	OR 3.3 (0.83-13.5)		

OR ( ): odds ratio with 95% confidence interval, indicates the risk of developing anxiety symptoms as a result of sleep problems (hypothesis 1) or the risk of developing sleep problems as a result of anxiety symptoms (hypothesis 2)

ß: standardized coefficients, Δ: not reported

*CBCL: Child Behaviour Check List: a parent reported questionnaire (13)

Discussion

Our study supports the hypothesis that sleep problems in childhood predict later anxiety symptoms. We consider our results reliable, because six of the seven studies were in large population cohort studies.

Hypothesis 1

Sleep problems from the age of 6 months until 19 years did show a positive relationship with later anxiety symptoms. This association increased during life until young-adults. Persistent sleep problems after the age of 4 years might have a greater impact. Persistent sleep problems before the age of four years did not show any association with anxiety.[12]

It is suggested in literature that dysregulation of the hypothalamic-pituitary-adrenal (HPA) axis is the link between sleep problems and anxiety symptoms in children.[15,16] Possibly, only after the age of 4 years, this mechanism occurs. More investigation is needed to establish the role of stress hormones in the association with sleep problems and anxiety symptoms.

Hypothesis 2

In only one study the presence of anxiety at the age of 18 months predicted one item of sleep problems, nightmares, 6 months later. Although the support for anxiety predicting nightmares is still preliminary, findings of other studies also support this association. [17,18] The reason for this weak association is unclear. However, because of the fact that only a few studies were available, this might be the result of investigation bias. Another reason for this weak association between anxiety and sleep problems might be the fact that anxiety is a difficult diagnosis in childhood.

Strengths and limitations

A strength of this review is the hypothetical character. We studied both plausible directions in the association between sleep problems and anxiety symptoms.

A limitation is that, as we searched in PubMed only, we may have missed relevant articles in other databases, for example PsychInfo. Although, four of the seven included studies were performed by the group of Alice Gregory, which may give the impression that she studies the same population several times, the four studies were conducted in different countries. The small number of research groups merely shows how under-represented this research area is. A reason for this under-representation might be the difficulty of classification of the primary disorders. It is possible that many patients classified as having a generalized anxiety disorder are actually only suffering from sleep problems according to the DSM criteria.

The included studies also had limitations. In all of the included studies except one [11], the parents reported the sleep problems of their children, which might have resulted in reporter bias. However, because in all the studies the validated parental questionnaire CBCL was used, this bias exists in all the included studies, and therefore, the results are comparable. However, if the CBCL is less reliable and therefore leads to bias, this bias will exist in all the studies and therefore the relationship between sleep problems and anxiety symptoms is overestimated.

For future studies, the reporting of sleep problems by an objective actigraphy (a method of monitoring human rest-activity cycles) might be an option. However, research shows that data of sleep reported by parents (CBCL) is useful.[19] Because different definitions of sleep problems were used in the studies, it was impossible to compare the results directly. Research in the future may work on developing definitions of sleep problems.

Conclusion

Findings from this review support the hypothesis that sleep problems in childhood are associated with later anxiety symptoms. With this knowledge, sleep problems in childhood should be taken serious and should be treated well to prevent later anxiety symptoms. Perhaps, an intervention more early in the process of developing sleep problems and anxiety problems is possible, like the prevention of sleep problems. We suggest further research into this. There was little support for an association between anxiety symptoms and later nightmares.

In this review, we did not search for factor X. We suggest future investigations into underlying risk factors for both sleep problems and anxiety symptoms. We don't think the use of an objective actigraphy in future investigations is necessary, because it is quite invasive and an objective actigraphy seems to be not more exactly than sleep data reported by parents.

References

1. *Costello E.J. et al., Prevalence and development of psychiatric disorders in childhood and adolescence. Arch Gen Psychiatry. 2003; 60: 837-844.*
2. *Verhulst F.C. et al., The prevalence of DSM-III-R diagnoses in a national sample of Dutch adolescents. Arch Gen Psychiatry. 1997; 54: 329-336.*
3. *van Litsenburg R.R. et al., Sleep habits and sleep disturbances in Dutch children: a population-based study. Eur J Pediatr. 2010; 169: 1009-1015.*
4. *Roberts R.E. et al., Restricted sleep among adolescents: prevalence, incidence, persistence, and associated factors. Behav Sleep Med. 2011; 9: 18-30.*
5. *American Psychiatric Association, Washington DC en Londen UK: DSM-IV-TR. United States: 2006.*
6. *Jansen P.W. et al., Does disturbed sleeping precede symptoms of anxiety or depression in toddlers? The generation R study. Psychosom Med. 2011; 73: 242-249.*
7. *Gregory A.M. et al., Parent-reported sleep problems during development and self-reported anxiety/depression, attention problems, and aggressive behavior later in life. Arch Pediatr Adolesc Med. 2008; 162: 330-335.*
8. *Gregory A.M. et al., Prospective longitudinal associations between persistent sleep problems in childhood and anxiety and depression disorders in adulthood. J Abnorm Child Psychol. 2005; 33: 157-163.*
9. *Gregory A.M. et al., Sleep problems in childhood: a longitudinal study of developmental change and association with behavioral problems. J Am Acad Child Adolesc Psychiatry. 2002; 41: 964-971.*
10. *Gregory A.M. et al., Etiologies of associations between childhood and behavioral problems in a large twinsample. J Am Acad Child Adolesc Psychiatry. 2004; 43: 744-751.*
11. *Silva G.E. et al., Longitudinal Association between Short Sleep, Body Weight, and Emotional and Learning Problems in Hispanic and Caucasian Children. Sleep. 2011; 34: 1197-1205.*
12. *Lam P. et al., Outcomes of infant sleep problems: a longitudinal study of sleep, behavior, and maternal well-being. Pediatrics. 2003; 111: 203-207.*
13. *Achenbach TM. Manual for the Child Behavior Checklist and 1991 Profile. Burlington: Department of Psychiatry, University of Vermont; 1991.*
14. *Achenbach TM. Manual for the Young Adult Self-report and Young Adult Behavior Checklist. Burlington: Department of Psychiatry, University of Vermont; 1997.*
15. *Buckley T.M. et al., On the interactions of the hypothalamic-pituitary-adrenal (HPA) axis and sleep: normal HPA axis activity and circadian rhythm, exemplary sleep disorders. J Clin Endocrinol Metab. 2005; 90: 3106-3114.*
16. *Feder A. et al., Twenty-four-hour cortisol secretion patterns in prepubertal children with anxiety or depressive disorders. Biol Psychiatry. 2004; 56: 198-204.*
17. *Spoormaker V.I. et al., Nightmares: from anxiety symptom to sleep disorder. Sleep Med Rev. 2006; 10: 19-31.*
18. *Smedje, H. et al., Short-term prospective study of sleep disturbances in 5-8-year-old children. Acta Paediatr. 2001; 90: 1456-1463.*
19. *Gregory A.M. et al., Sleep items in the child behavior checklist: a comparison with sleep diaries, actigraphy, and polysomnography. J Am Acad Child Adolesc Psychiatry. 2011; 50: 499-507.*

Efficacy and safety of rituximab, splenectomy and dexamethasone in patients with Idiopathic thrombocytopenic purpura (ITP)

A systematic review

Fahid Chohan[a], Jenny Giang[a], Peter te Boekhorst[b]
[a] Medical student, Erasmus MC University Medical Center Rotterdam, the Netherlands
[b] Supervisor, Department of Hematology, Erasmus MC University Medical Center Rotterdam, the Netherlands
Correspondence: Fahid Chohan, e-mail: 329472mc@student.eur.nl

Objective: **The purpose of this systematic review was to evaluate the efficacy and safety of rituximab, splenectomy and high dose (HD) dexamethasone for potential use as first-line treatment for idiopathic thrombocytopenic purpura (ITP).**
***Background*: Prednisone is the current first-line treatment for ITP. Treated patients have a 70-90% response rate, but only 20-40% of patients have a stable reponse either complete or partial.**
***Methods*: We searched the literature from December 2011 until January 2012. Inclusion criteria: a randomized design aimed at evaluating the efficacy and safety of the various treatments for ITP, specifically concerning rituximab, splenectomy and HD dexamethasone, with the primary outcomes consisting of platelet response and toxicity.**
***Results*: Our PubMED search yielded 95 publications, of which 5 were used in our systematic review. Rituximab had a complete response (CR) of 79%. Splenectomy showed high short- and long term CR rates of 79.5% and 67.7% respectively. HD dexamethasone as first-line treatment has better long term effect compared to second-line HD dexamethasone (CR; 39% vs 17%). HD Dexamethasone plus rituximab yielded greater results than HD dexamethasone as monotherapy in previously untreated ITP patients (63% vs 36%). HD dexamethasone has less severe side-effects (grade 5) compared to rituximab and splenectomy.**
***Conclusion*: Splenectomy, rituximab and HD dexamethasone are effective treatment modes in patients with ITP. However, HD dexamethasone can be designated as most effective and safe when compared to the other treatments reviewed in this article. Therefore, we suggest to consider it as a first-line treatment in patient with idiopathic thrombocytopenic purpura.**

Introduction

Idiopatic Trombocytopenic Purpura (ITP) is a common hematological autoimmune disorder, without a clear underlying cause of thrombocytopenia. ITP is characterized by a low platelet count due to auto-antibodies formation and possibly a low production of thrombocytes in the bone marrow.[1] The current first-line treatment for ITP is prednisone, to which 70-90% of patients respond. However, only 20-40% of the patients will have a stable response, complete or partial.[2]
Many studies have shown that other treatment modalities also show beneficial effects in treating ITP. The second-line treatments that we compared in this review are splenectomy, rituximab and high dose (HD) dexamethasone. Rituximab is a monoclonal antibody directed against CD20 and targets B-lymphocytes and causes them to go into apoptosis. Rituximab has been used for the treatment of B-cell lymphoma in combination with chemotherapy, resulting in higher remission rates compared to classic chemotherapy. Recently, rituximab has shown its usefulness in treating ITP.[4]
Splenectomy results in an increased platelet count in 2/3 of the patients.[2] A previous study by Cheng et al has shown that after one course of HD dexamethasone, long-term remission is obtained in more than 40% of the patients.[4] Moreover, HD dexamethasone has shown fewer side-effects than conventional prednisone treatment. Conventional prednisone treatment also resulted in much earlier treatment failure when compared with HD dexamethasone.
The promising results of rituximab, splenectomy and HD dexamethasone suggest that, one of these treatments could be used as a first-line treatment in newly diagnosed patients with ITP. Therefore, we systematically reviewed the literature to answer the following question. Could rituximab, splenectomy or HD dexamethasone be used as a first-line treatment with better efficacy and safety in terms of platelet response and side-effects than conventional prednisone treatment?

Methods

Search strategy
Our search took place from December 2011 until January 2012. Pubmed was used for the search with Medical Subject Heading (MeSH) terms. The following search was performed: ((((“Purpura, Thrombocytopenic, Idiopathic/drug therapy”[Majr]) AND”Purpura, Thrombocytopenic,Idiopathic/blood”[Mesh]) NOT”Thrombopoietin”[Mesh])NOT “Immunoglobulins,Intravenous ”[Mesh]) OR (((“Purpura, Thrombocytopenic, Idiopathic”[Majr]) AND”Splenectomy”[Mesh]) AND “Treatment Outcome”[Mesh]). We searched for articles published between 2005 and 2012. We specifically excluded review articles and studies in a language other than English. We also searched the references in the articles manually for additional studies that matched our search criteria. Articles were selected by abstract analysis, title and date of publication.

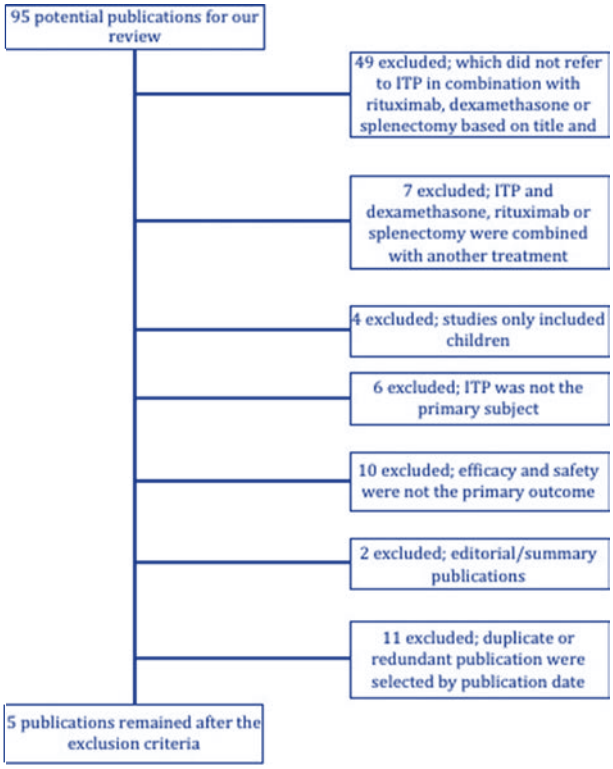
Study selection
Inclusion criteria were; a randomized design aimed at evaluating the efficacy and safety of the various treatments for ITP as the main subject, specifically aimed at rituximab, splenectomy and HD dexamethasone. Studies were eligible if they treated at least 10 patients diagnosed with ITP. We did not differentiate between chronic and severe ITP. Exclusion criteria were secondary causes of thrombocytopenia such as viral infections, splenomegaly, pregnancy and medication. We also excluded childhood studies and editorials/summaries.

Endpoints
To determine the efficacy of the treatment the following response criteria were used: complete response (CR), partial response (PR), sustained response (SR), minor response (MR) and no respond (NR). CR was defined as a rise in the platelet count of >100 x 109/L, PR as a rise in the platelet count of >50 x 109/L, MR as a rise in the platelet count, but not above 50 x 109/L, and NR was defined as no increase in the platelet count.

Results

Our Pubmed search produced 95 publications. Based on title and abstract of the articles, we excluded 49 articles which did not refer to Idiopathic Trombocytopenic Purpura in combination with rituximab, dexamethasone or splenectomy. Of the remaining 46 publications 7 studies were combined with another treatment other than rituximab, HD dexamethasone and splenectomy. Another 23 were excluded because the publications consisted of childhood studies and did not describe ITP as the primary topic. Editorials/summaries and duplicate articles were excluded by publication date (Figure 1).
We included the remaining 5 articles in our systematic review. Two of these were randomized controlled trials with a follow up duration of 6 and 54 months, and remaining three articles were cohort-studies. The total population consisted of 340 patients, almost all of whom had received prednisone or other medication prior to receiving second-line treatment.
However, HD dexamethasone was given as first-line treatment in Braendstrup et al and Zaja et al only enrolled untreated patients in their study.[6,9] Patient age range was 5 to 86 years. Of the 340 patients, 122 were men and 218 women (Table 1).
In two studies, rituximab was given once per week as an infusion of 375mg/m2 for 4 weeks. [7,9] In one study, patients were randomized 1:1 to receive dexamethasone with or without rituximab. A daily dose of 40 mg HD dexamethasone was given to both arms for 4 consecutive days. The experimental group received 375mg/m2 intravenous rituximab.[6]

Figure 1. Article selection.



In one study, 40mg HD dexamethasone was given during an 8-day course.[8] In one study, laparoscopic splenectomy (LS) was used as a treatment; a three-port technique, performed under general anesthesia, was applied.[5]

Platelet count response
According to our defined criteria as defined in the Methods section, rituximab treatment resulted in 11 CR (79%), 2 PR (14%) and the overall response was 93 %. Alasfoor et al reported a median duration of remission of 12.5 months (2-19).[7] Braendstrup et al showed that 17 patients had a CR (18%) to rituximab, 6 had a PR (15%), with an overall response rate of 44%.[9] The median duration of remission in this study was 18.3 months (2-31.5).

However, Zheng et al reported 101 CR from splenectomy (79.5%) and 12 PR (9.5%).[5] The initial response at two months after laparoscopic splenectomy was 89%; the long-term response was 80.3% with a mean follow up of 43.6 months.

Table 1 - Characteristics of patients with ITP (n=340)

Study, year (reference)	Patients. n	Country	Median age (patient age range)	Duration of ITP (range months)	Platelet count before second-line treatment x 10^9 cells/L	Mean follow up (follow up range in months)
Zheng et al., 2011(5)	154	China	30.4 (5-78)	2-240	78-103	43.6 (9-114)
Zaja et al., 2010 (6)	101	Italy	47	0-103	0-22	20 (4-40)
Alasfoor et al., 2009 (7)	14	Kuwait	35 (12-72)	1-133	NRD*	18 (2-35)
Borst et al., 2004 (8)	36	The Netherlands	52.4 (16-86)	0-516	3-56	12.5 (1-54)
Braendstrup et al., 2005 (9)	35	Denmark	53 (17-82)	49 (1-288)	14 (1-49)	NRD*

*NRD: not reported

In the study done by Borst et al HD dexamethasone was given as first-line and second-line treatment, this led to an acute response (rise of platelet count above 50 x 109/L) in 83% of the patients.[8] In the same study, HD dexamethasone was given as first-line treatment, 59% of the patients were in remission after 31 months. As second-line treatment it showed that 50% of the patients were in remission after 5 months, declining to 25% after 54 months.

The study of dexamethasone in combination with rituximab reported that a SR at 6 months after initiation (a platelet count greater than 50 x 109/L) was higher in patients who were treated with both dexamethasone and rituximab compared to dexamethasone treatment only (63% vs 36%, P = 0.004, 95% CI; 0.079-0.455).[6]

Toxicities

The side-effects related to rituximab were moderate symptoms such as dizziness, tachycardia, cramps or hypotension in 9 patients (18%) reported by Alasfoor et al and Braendstrup et al.[7,9] Due to side-effects, 2 patients (4%) discontinued treatment: one had a severe anaphylactic reaction after rituximab treatment and the other showed severe muscular pains and swelling of the legs. Five patients (10%) from a total of 49 experienced severe or life threatening complications.[7,8] Braendstrup et al, reported two patients (4%) who died during treatment caused by respiratory insufficiency.[9] A 71 year-old female with severe chronic lung disease died of respiratory insufficiency 6 days after start of treatment. A 73 year-old man who also had severe chronic obstructive lung disease died of pneumonia 13 weeks after the last rituximab treatment.[9]
Zheng et al showed severe side-effects in 20 patients (13%). Including pancreatitis, pneumonia, hematocoelia and wound hematomas.[5] Out of the 127 patients with long-term follow-up, 2 patients suffered from cerebral infarction at 3 and 5 months after LS and 5 developed pneumonia 3-35 months after LS. During the course of this study, 2 patients (1%) died: 1 patient died from spontaneous intracranial bleeding 10 months after LS and 1 male patient died from subphrenic abcess and sepsis 25 days after LS treatment.
From a total of 88 HD dexamethasone treatments, 5 patients (6%) had side-effects leading to discontinuation of this treatment.[6,8] One of these patients had developed herpes zoster of the skin. Most frequent side-effects were weight gain, loss of appetite, epigastric discomfort, tiredness and dizziness.
However, HD dexamethasone in combination with rituximab treatment caused severe side-effects in 4 patients (8%), including seizure, supraventricular tachycardia and hospitalization due to low platelet count and pneumonia.[6]
No deaths were reported from either HD dexamethasone alone or HD dexamethasone plus rituximab treatment.

Discussion

This systematic review shows that HD dexamethasone can be used as first-line treatment, because it has the same long-term platelet count response as prednisone and results in fewer side-effects when compared to rituximab, splenectomy and conventional prednisone treatment. Rituximab resulted in a CR in 79% of the patients according to Alasfoor et al but only 18% of the patients in Braendstrup et al.[7,9] However, splenectomy showed high short-term and long-term CR rates of 79.5% and 67.7%, respectively.[5] When HD dexamethasone was given as first-line treatment, it had better long-term effects compared to second-line HD dexamethasone (CR; 39% vs. 17%). In contrast, the short-term effects of second-line HD dexamethasone were better than first-line treatment (CR;71% vs. 56%).[8]
HD dexamethasone and rituximab yielded better results than HD dexamethasone as monotherapy in previously untreated ITP patients (63% vs. 36%).[6]
Based on our results, we believe that HD dexamethasone should be considered as a treatment for newly diagnosed patients with ITP. HD dexamethasone has fewer severe side-effects (grade 5) compared to rituximab or splenectomy (Table 3). Splenectomy and rituximab resulted in 2 deaths and showed more life threatening side-effects. HD dexamethasone does not have the same efficacy as splenectomy and rituximab. Nevertheless, the combination of good response rates (long and short term) and the relatively moderate side-effects make HD dexamethasone a better candidate.
In our review we observed that HD dexamethasone has the same long-term response rate as prednisone. Prednisone has been the standard first-line medication for the treatment of ITP.[2] However, HD dexamethasone can also achieve the same results. The alternatives are splenectomy and rituximab. Splenectomy is a treatment option that should be considered as second or third-line treatment, due to its side-effects. Although splenectomy is a laparoscopic procedure it still can cause major morbidity. Patients who have undergone splenectomy have a compromised immune system, which requires systematic monitoring.
Splenectomy patients also have to undergo various vaccinations prior and after LS. Altogether, this makes laparoscopic splenectomy an expensive treatment.
Rituximab has good response rates in patients who have already received corticosteroid treatment.

Table 2 - Complete response (CR), Partial response (PR), Minor response (MR), and No platelet count response (NR) after second-line treatment in patients with ITP, as reported in 5 studies.

Study	Characteristics of study	CR	PR	MR	NR
Splenectomy					
Zheng et al., 2011(5)	Short-term responses	101 (79.5%)	12 (9.5%)	NRD*	14 (11%)
	Long-term responses	86 (67.7%)	16 (12.6%)	NRD*	25 (19.7%)
Rituximab					
Alasfoor et al., 2009 (7)	Responses	11 (79%)	2 (14%)	NRD*	1 (7%)
Braendstrup et al., 2005 (9)	Responses	17 (18%)	6 (15%)	4 (10%)	NRD*
HD dexamethasone					
Borst et al., 2004 (8)	Short-term responses (first-line)	10 (56%)	6 (33%)	2 (11%)	NRD*
	Long-term responses (first-line)	7 (39%)	4 (22%)	6 (33%)	1 (6%)
	Short-term responses (second-line)	13 (71%)	1 (6%)	3 (17%)	1 (6%)
	Long-term responses (second-line)	3 (17%)	0	8 (44%)	7 (39%)

*NRD: not reported

Table 3 - Toxicities observed after second-line treatment in patients with ITP as reported in 5 studies.

Study	Patients	Grade 1-2	Grade 3-4	Grade 5
Splenectomy				
Zheng et al., 2011(5)	154	NRD*	Focal cerebral infarction (2)	Spontaneous intracranial
			Pneumonia (7)	bleeding (1)
			Hematocoelia (2)	Subprhenic abcess and
			Pancreatitis (3)	sepsis (1)
			Adrenal crisis (1)	
Rituximab			Hematomas (2)	
Alasfoor et al., 2009 (7)	14	Infusion toxicity (2)	Tonic clonic seizure (1)	NRD*
			Trombocytosis VTE (1)	Chronic lung disease (1)
			Sepsis (1)	Chronic obstructive lung
Braendstrup et al., 2005 (9)	35	Restlessness (1)	Anaphylactoid reaction (1)	disease (1)
		Swelling of the fingers and feet (1)	Muscular pains and swelling of	
Dexamethasone		Unspecified exanthema (1)	the legs (1)	
Zaja et al., 2010 (6)	101	NRD*	Supraventricular tachycardia (1)	NRD*
			seizure (1)	
Borst et al., 2004 (8)	36	Weight gain (5) loss of appetite (6)	Severe to discontinue treatment:	NRD*
		tiredness (6)	Herpes Zoster (1) and other (4)	
		Dysphoria (4)		
		Hyperglycaemia (3)		

Toxicity grades are National Cancer Institute Common Terminology Criteria for Adverse Events, version 4.0 [10]: Grade 1-2 is mild to moderate; Grade 3-4 is serious to life-threatening; Grade 5 is fatal. NRD means not reported.

This shows its great potential in reducing B-lymphocyte count and increasing platelet levels, even in patients who did not respond to corticosteroids.

Limitations

The studies that described the efficacy and safety of rituximab had small patient sample sizes. One study included only 14 patients and the other only 35 patients.[7,9] Braendstrup et al only showed 27 results of a total of 35 patients who were enrolled in the study.[9] A second limitation is that the patient samples in each article were not the same. Consequently, there is a possibility of confounding factors in the smaller studies that could have influenced our conclusions. However, a homogenous patient population would have made our study selection even smaller. Even with a heterogeneous population, we included only 5 out of 95 articles. A third limitation is that Borst et al compared first-line HD dexamethasone with second-line HD dexamethasone, without a test for significance.[8] This study did not report any p-values or confidence interval.
A more general limitation is that we were unable to find 5 studies with newly diagnosed patients. As a result, the duration of ITP in months and platelet count before second-line treatment was not the same. This could be one of the reasons why rituximab showed such variable results. However, HD dexamethasone and splenectomy were also affected by this same limitation.
Finally, we did not include other second-line treatments, such as intravenous immune globulin (IVIG) and thrombopoietin receptor agonist. Instead, we only compared rituximab, splenectomy and HD dexamethasone. The addition of these treatments to our systematic review would have enabled us to draw a more general conclusion about the efficacy and safety of second-line treatments for ITP.

Conclusion

The current first-line prednisone treatment does not have a better long-term platelet count response than HD dexamethasone. However, it does result in more side-effects than HD dexamethasone. By treating ITP patients with HD dexamethasone, unnecessary side-effects can be avoided. Splenectomy, rituximab and HD dexamethasone are all effective treatments in patients with ITP.

For future research, we suggest larger studies in which rituximab is given as a first-line treatment in patients with ITP. In addition, in a randomized controlled trial HD dexamethasone could be compared with prednisone to directly asses their effectiveness and safety.

References

1. *McMillan R. Chronic idiopathic thrombocytopenic purpura. N Engl J Med 1981; 304:1135-1147.*
2. *Szold A, Sagi B, Merhav A et al. Optimizing laparoscopic splenectomy. Technical details and experince in 59 patients. Surg Endosc 1998; 12:1078-1081.*
3. *Coiffer B, Lepage E, Briere J et al. CHOP chemotherapy plus Rituximab compared with CHOP alone in elderly patients with diffuse large B-cell lymphoma. N Eng J Med 2002; 346:235-242.*
4. *Cheng Y, Wong RS, Soo YO et al. Initial treatment of immune thrombocytopenic purpura with high-dose dexamethasone. N Engl J Med. 2003; 349:831-836.*
5. *Zheng C, Zheng D, Chen L et al. Laparoscopic splenectomy for immune thrombocytopenic purpura at a teaching institution. Chin Med J 2011; 124: 1175-1180*
6. *Zaja F, Baccarani M, Mazza P et al. Dexamethasone plus rituximab yields higher sustained response rates than dexamethasone monotherapy in adults with primary immune thrombocytopenia. Blood 2010; 115:2755-2765.*
7. *Alasfoor K, Alrasheed M, Alsayegh F et al. Rituximab in the treatment of idiopathic thrombocytopenic purpura (ITP). Ann Hematol 2009; 88: 239-243.*
8. *Borst F, Keuning JJ, van Hulsteijn H et al. High-dose dexamethasone as a first- and second-line treatment of idiopathic thrombocytopenic purpura in adults. Ann Hematol 2004; 83:764-768.*
9. *Braendstrup P, Bjerrum OW, Nielsen OJ et al. Rituximab Chimeric Anti-CD20 Monoclonal Antibody Treatment for Adult Refractory Idiopathic Thrombocytopenic Purpura. Am J Hematol. 2005; 78:275-280.*
10. *National Institute of Health. National Cancer Institute. Common Terminology Criteria for Adverse Events, Version 4.0. Accessed at http://evs.nci.nih.gov/ftp1/CTCAE/About.html on Jan 2012.*

Gene therapy for Pompe disease in a murine model

A systematic review

Nicole Sonneveld[a], Aart Cornelis van der Vlist[a], Merel Stok[b], Gerard Wagemaker[b]
[a] Medical student, Erasmus MC University Medical Center Rotterdam, the Netherlands
[b] Department of Hematology, Erasmus MC University Medical Center Rotterdam, the Netherlands
Correspondence: Nicole Sonneveld, e-mail: n.sonneveld@erasmusmc.nl

STRUCTURED ABSTRACT

Objective: **Pompe disease is a lysosomal storage disorder for which current treatment is not curative. Gene therapy is a promising strategy and therefore thoroughly investigated. A systematic review of preclinical evaluation in murine models was performed to gain a comprehensive overview of the efficacy of different viral vectors.**

Methods**: The MEDLINE database was searched from January first until January 17th 2012. Acid α-glucosidase activity levels in murine tissues were analysed from the studies included.**

Results**: A total of six publications were included. Transgene expression levels were reported in all studies. Viral vectors used were adenoviral vectors, adeno-associated virus vectors, lentiviral vectors and an adenoviral and adeno-associated virus hybrid vector. The age of the mice, the time of treatment and the delivery methods varied. Wild type acid α-glucosidase activity levels were 6.43 nmol/h/mg protein in cardiac muscle, 6.43 nmol/h/mg protein in the diaphragm and 14.29 nmol/h/mg protein in skeletal muscle. Transgene expression varied between 0.4 and 355.4 nmol/h/mg protein for the cardiac muscle, 16.1 and 190 nmol/h/mg protein for the diaphragm and 5.2 and 406 nmol/h/mg protein in skeletal muscle.**

Conclusions**: First attemps of gene therapy for Pompe disease where performed using an adenoviral vector. Over the years, adeno-associated viral vectors and later lentiviral vectors were introduced. Efficacy of the vectors varied. Due to variations in methods used, the results of the studies were difficult to compare. Two main problems were observed: immune response to the vector and recombinant genetic material and the inability to reach all target cells. We recommend more investigation in the method using hematopoietic stem cells for gene delivery to reduce immune response.**

Introduction

Pompe disease, or glycogen storage disease type II (GSDII), is an autosomal recessive inherited disease. The incidence is estimated at 1/40000 in the Netherlands.[1]

The cause of GSDII is a deficiency of acid α-glucosidase (GAA). This enzyme is responsible for the degradation of glycogen into glucose in lysosomes. In normal cells, newly synthesized GAA is transferred to the rough endoplasmic reticulum where glycosylation occurs. Subsequently, this glycosylated GAA is transferred to the Golgi-apparatus, where mannose-6-phosphate (M6P) is added to the N-terminus of GAA to label proteins destined for lysosomes.[2] This results in a GAA precursor (110 kDa) of which the majority is transported in vesicles to the lysosomes. A small amount of the precursor GAA is secreted outside the cells. The secreted precursor can bind to a M6P receptor of other cells and be transferred to the lysosomes by endocytosis.[3] It should be noted that it is suggested that there also is a M6P-receptor independent pathway, which probably uses asialoglycoprotein and mannose receptors.[4] In the lysosomes, proteolytic intra-lysosomal enzymes change the precursor GAA to its mature isoforms, respectively 67 and 76 kDa. These mature isoforms degrade glycogen to glucose.
In GSDII patients, glycogen is not degraded and accumulates in the lysosomes. Eventually the normal functioning of the cell is disrupted, which leads to organ disfunction.[5]

Among the 300 GAA mutations described, not all lead to GSDII.[6] The mutations that lead to GSDII cause either limited synthesis of GAA, or the precursor GAA cannot be transformed into the mature form and therefore catalytic activity is limited or absent.[7]

Based on the clinical features, two major phenotypes are distinguished.

When there is less than 1% of normal GAA activity newborns will present with progressive muscle weakness, respiratory dysfunction and cardiac failure. This infantile form is the most life threatening phenotype of GSDII; death occurs within the first year of life when untreated.[8]
The late-onset phenotype varies in severity of the symptoms, manifest as progressive myopathy. Limb-girdle weakness is in general the first symptom and can present at any age from early childhood. Most patients will eventually become wheelchair-dependent and in need of ventilation support.[9]

The first and only approved treatment for GSDII is enzyme replacement therapy (ERT), registered in 2006.[10] The therapy involves recombinant human GAA (Myozyme®) injected intravenously in the patient. Although the treatment extends life, improves the quality of life and leads to significant reduction of glycogen in lysosomes, there are some downsides. First of all, patients need to be treated every other week for a few hours. This is demanding, especially for children. Secondly, some of the treated patients present an immune response to the ERT. This response results in reduced efficacy of the enzyme replacement treatment and may cause serious adverse events.[11] Also, the costs of ERT are high.

To overcome the immune response and the high costs of the ERT, new therapies need to be explored. A relatively new approach is gene therapy, in which a gene coding for human GAA (hGAA) is placed into the patient's genome. This can be achieved with a wide variation of vectors, both viral and non-viral.[12] In an ideal situation, the patients receive gene therapy with hGAA only once and then produce the missing enzyme themselves. In this review we will give an overview from vectors that have been explored for GSDII and their results in GAA-knock-out mice.

Methods

Literature search and study selection

In order to select studies we searched the MEDLINE database using the following MeSH terms in PubMed: ("Glycogen Storage Disease Type II/therapy"[Majr]) AND "Gene Therapy"[Majr]. We limited our search to English language. Our last search was on the 17th of January 2012.

Studies were included based on the following criteria:

I. Primary publication
II. The use of a non-specific promoter (i.e. CMV, CB, MND) only
III. Use of the GAA-transgene
IV. The use of the GAA-KO mouse as a model for GSDII only
V. Measurements of GAA-activity in cardiac muscle, diaphragm or skeletal muscle.

Both authors screened all abstracts and full-text articles.

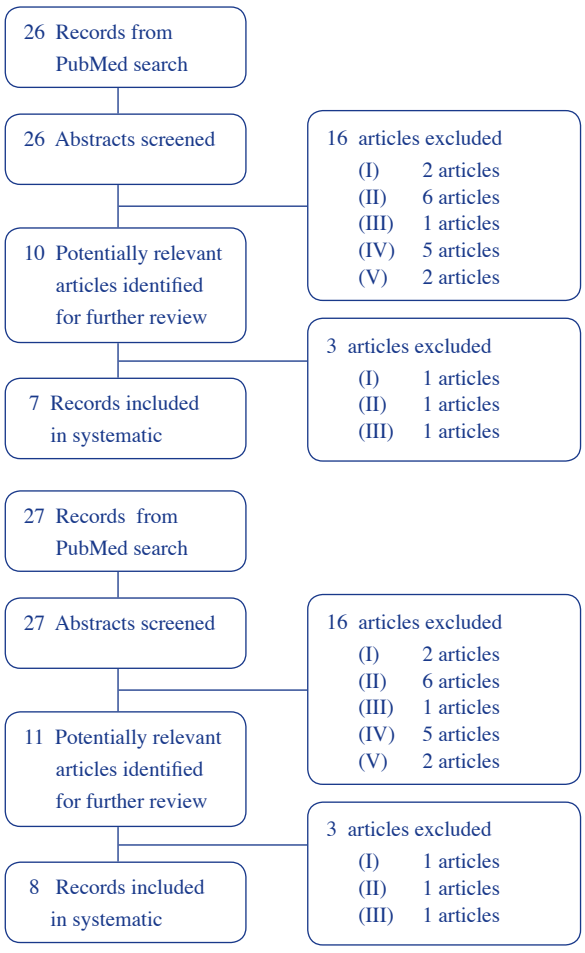
Because of the Poenaru et al. review published in July 2000 we excluded all publications before this date (VI).[12]

Data analysis

In order to compare the included studies, we extracted the most essential features of each article. This included the viral vector used, the promoter, the age of the mice at the time of treatment and the delivery method.

The specific outcomes in each article were GAA-activity in cardiac muscle, GAA-activity in the diaphragm and GAA-activity in skeletal muscle. Not all articles contained all specific end-point parameters.

Figure 1. Flow chart of study selection.



When outcomes were only given in a segmented column chart or histogram and not in exact data, the specific outcomes were analysed by both authors independently and the median was calculated if the deviation was less than 5%.

If the included studies reported multiple measurements of hGAA expression, the median expression level over time was calculated to improve intercomparison of the included studies.

Results

Literature flow and study characteristics

Using the described search strategy, we found 26 articles of which six met our inclusion criteria.[7,13-17] A flow chart of the study selection process is shown in Figure 1.

Vectors used in these studies were lentiviral vectors, adenoviral vectors, an adeno-associated virus (AAV) vector and an adenoviral-AAV (AdAAV) hybrid vector (although the latter was not injected into mice). These vectors were delivered into the mice with a variety of different methods; three studies delivered the vector intravenous, one intramuscular, one gel-mediated and one by ex vivo transduction of hematopoietic stem cells. Furthermore, the age of mice used differed in all studies varying from two days old to 21 months old. Immune responses, being a serious disadvantage, as it reduces the amount of functional hGAA, were investigated in three studies.[13,14] The essential features of the studies are summarized in Table 1.

Table 1 - Essential features of the included studies.

References	Vector	Promoter	Age of GAA-KO mice at time of treatment	Method of delivery
Van Til et al. .[17]	Lentiviral	SFFV	8-12 weeks	Bone marrow transplantation with transduced HSCs
Mah et al. .[14]	AAV1	CMV	3, 9 and 21-month-old	Gel-mediated into diaphragm
Kyosen et al. .[13]	Lentiviral	CMV	≤ 2 days	Intravenous, v. temporalis superficialis
Xu et al. .[7]	Adenoviral	CMV	12-14 months and 17-19 months	Intravenous, retro-orbital sinus
Sun et al. .[16]	AAV2 and AAV6	CMV	3 months	Intravenous, v. portae
Martin-Touaux .[15]	Adenoviral	CMV	4 days	Intramuscular, m. gastrocnemius

AAV = adeno-associated virus. SFFV = spleen focus-forming virus; CMV = cytomegalovirus; HSC = hematopoietic stem cell

GAA-activity in cardiac muscle after vector delivery of hGAA

As shown in Table 2, all studies reported different GAA-activity levels in cardiac muscle. We have included two studies using a lentiviral vector. The Van Til et al. study, using a spleen focus-forming virus (SFFV) promoter and hematopoietic stem cells as delivery-vehicle in eight to twelve weeks old mice, showed a GAA-activity of 17 ± 2 nmol/h/mg protein measured two weeks after injection.[17] The Kyosen et al. study delivered the cytomegalovirus (CMV) promoter driven vector intravenously into 48 hours old mice. This resulted in an average GAA-activity of 355.4 ± 167 nmol/h/mg protein after an average of 20 weeks after injection.[13]

Table 2 - GAA-activity in cardiac and skeletal muscle and diaphragm.

References	GAA-activity (nmol/h/mg protein)		
	Cardiac muscle	Diaphragm	Skeletal muscle
Van Til et al. .[17]	17 ± 2	37 ± 12	16 ± 3
Mah et al. .[14]	0.4	-	-
Kyosen et al. .[13]	355.4 ± 167*	16.1 ± 9.3*	5.2 ± 1.7*
Xu .[7]	185 ± 150	190 ± 100	65 ± 30
Sun AAV2	6	83	9
.[16] AAV6	1.1	50	20
Martin-Touaux .[15]	0.42 ± 0.3*	-	406 ± 38.7*

* = This value is calculated as an average of results from multiple measurements post-injection.

The Sun et al. and Mah et al. studies both used an AAV-vector. Mah et al. focussed on the diaphragm, but did also report the GAA-activity in cardiac muscle. They used a CMV-promoter driven AAV1-vector and delivered it gel-mediated to the diaphragm. This resulted in GAA-activity levels of 0.4 nmol/h/mg protein 3 months post-treatment.[14] Sun et al. investigated the potential of an AdAAV-hybrid. However, in mice they used both an AAV2- and an AAV6-vector. In cardiac muscle, the GAA-activity of the AAV2-vector was higher then the AAV6-vector, with GAA-activity levels of 6 nmol/h/mg protein and 1.1 nmol/h/mg protein respectively.[16]
The remaining two studies used an adenoviral vector. Xu et al. used old mice in their study (12-14 and 17-19 months old) with intravenous injection of the CMV-driven adenoviral vector. This resulted in a GAA-activity of 185 ± 150 nmol/h/mg protein at 17 days post-injection.[7] Martin-Touaux et al. injected the vector intramuscular into the gastrocnemius of four days old mice. Levels of 0.42 ± 0.3 nmol/h/mg protein GAA-activity were detected on in average 9.5 weeks.[15]

GAA-activity in diaphragm after vector delivery of hGAA
Although Mah et al. focussed on the diaphragm, they only reported GAA-activity as a percentage of normal GAA levels. Unfortunately, they did not mention what they used as a reference and so we could not determine the GAA-activity in nmol/h/mg protein.[14] Also, the Martin-Touaux et al. study did not report any data on GAA-activity levels in the diaphragm.[15]

Four studies show data about the GAA-activity in the diaphragm. Xu et al. reported the highest levels of GAA-activity with 190 ± 100 nmol/h/mg protein.[7] The AAV2 and AAV6-vectors in the Sun et al. study resulted in levels of 83 and 50 nmol/h/mg protein respectively.[16] Van Til et al. reported 37 ± 12 nmol/h/mg protein and lowest levels were reported by Kyosen et al. with 16.1 ± 9.3 nmol/h/mg protein.[13,17]

GAA-activity in skeletal muscle after vector delivery of hGAA
Skeletal muscle was the main subject in the Martin-Touaux et al. study. With intramuscular injection of the adenoviral vector, they reached GAA-activity levels of 406 ± 38.7 nmol/h/mg protein. This was the highest GAA-activity level reported.[15] GAA-activity levels using the AAV2-vector in the Sun et al. study were 9 nmol/h/mg protein. Whereas in cardiac muscle and diaphragm AAV2-mediated delivery reported better results, in skeletal muscle the AAV6-vector showed higher levels, with 20 nmol/h/mg protein.[16]
Other results were 16 ± 3 nmol/h/mg protein in the Van Til et al. study, 5.2 ± 1.7 nmol/h/mg protein GAA-activity levels in the Kyosen et al. study and 65 ± 30 nmol/h/mg protein levels in the Xu et al. study.[7,13,17]

GAA-activity levels in wild type (WT) mice
To describe the effect of the different methods used in the selected studies on the GAA-activity levels, normal levels are essential to identify. According to Van Til et al. GAA-activity levels in WT-mice are 6.43 nmol/h/mg protein in the cardiac muscle, 6.43 nmol/h/mg protein in the diaphragm and 14.29 nmol/h/mg protein in skeletal muscle.[17]

Discussion

Levels of GAA-activity differed in all studies due to the different methods used in the studies.

First of all, the age of the mice at the time of treatment differed. When treatment starts at a young age, immune tolerance is assumed to be better.[18] As anti-GAA antibodies can deplete the hGAA, GAA-activity levels may decrease. Interestingly, whereas GAA-activity levels in the cardiac muscle are highest in the Kyosen et al. study using neonatal mice, it is not in the other tissues.[13] The Xu et al. study, using old mice, also reported relatively high levels of GAA-activity, although they report no significant glycogen depletion compared to a previous study in younger mice.[7] High GAA-levels were explained by the assumption of a progressive reduction in immune response when mice get old.[19]

Apparently, immune responses to the recombinant hGAA need to be avoided. Van Til et al. aimed for immune tolerance using hematopoietic stem cells.[17] The idea that inserting the transgene into these stem cells would lead to immune tolerance was first introduced by Douillard-Guilloux et al. in 2008. They used this method to induce immune tolerance for ERT.[20] Van Til et al. aimed for a complete phenotypic correction. Levels of GAA-activity were significantly higher in cardiac muscle and diaphragm than in WT-mice. Although this seems a good result, skeletal muscle was not fully corrected yet. Moreover, the SFFV promoter used in this study is very strong and therefore not preferred in clinical application because, as Van Til et al. state, this promoter is more likely to hit and activate proto-oncogenes than cellular promotors.[17]

All other studies used a CMV-promoter.[7,13-16] This promoter is less strong and has already been used in clinical trials.[21]

The idea of correcting the GAA enzymatic activity in Pompe disease was raised in the late 1990's. The first publication of Pauly et al. in 1998 described the use of an adenoviral vector to correct the GAA-activity in neonatal rats.[22] Hereafter, a more preferable GAA-KO mouse model for GSDII was used. Although adenoviral vectors were still being studied, in 2002 the first study investigating the adeno-associated virus in an animal model was published.[23] The use of an AAV-vector instead of the adenoviral vector is the assumption that there would be a lower immune response. Furthermore, there are multiple serotypes with different immune responses due to different exposure rates of humans to these serotypes.[24]

In 2008 Richard et al. published the first study where the lentiviral vector was used.[25] This vector can integrate into the human genome. Although the transgene will stay in the host cell persistently, there is a chance of affecting proto-oncogenes or tumor-suppressor genes. Additionally, the lentiviral vector can stably transduce quiescent stem cells, whereas adenoviral and AAV-vectors do not. This also enhances the persistence of the transgene in the GSDII patients.
The traditional approach for gene therapy is to inject the viral vector intravenous or intramuscular. A relatively new method is to transduce human stem cells ex vivo.[17,20] As already discussed, this would be a potential benefit with respect to immune tolerance induction.

The first clinical trial for gene therapy in Pompe disease was started in 2011. This trail uses an AAV-vector delivered to the diaphragm in GSDII patients with respiratory failure despite ERT. During this phase I/II study, patients will still receive enzyme replacement therapy for their cardiac and skeletal muscles.[24] In view of the reported results, there is reason to assume that the efficacy of the vectors and the methods differ. However, we could not compare the vectors due to the large number of other variables in the included studies (e.g. age of the mice at time of injection, amount of vector particles injected (data not shown), time of measurement after injection (data not shown) and method of delivery).

In order to compare the included studies, we used the reported data of GAA-activity levels. Some studies reported expression of GAA at multiple time points after treatment. Because other studies reported the activity only at one time point after injection, we calculated the median expression level. This enabled us to compare the studies, but could mask an up- or downward trend of expression levels after treatment.

Furthermore, the age of the mice at the time of treatment differed as well as the method of delivery. This made comparison of the vectors used in the included studies difficult. Different expression levels may either be due to the used vectors or to one or more of the other variables.

We excluded all studies in which a tissue-specific promoter was used. This made the comparison of the studies less complicated. However, a potential benefit of these promoters is therefore not noted in this review.

The Poenaru review already showed two main problems regarding vector-based gene therapy.[12] One is the inability to reach all target cells with a single injection. The second problem is the immune response to the used vector and the recombinant genetic material. These problems also occurred in the first phase I/II trial in children.[26] In this study, the objective was to improve vertilatory function and to monitor safety, while using an AAV1 CMV-driven vector. While also using inspiratory muscle conditioning exercises, a non-significant improvement of spontaneous ventilatory endurance of 425% was reported, as well a significant improvement of 28.8% of the best tidal volume. However, the GAA-activity was limited to the diaphragm, in which the vector was injected, and there was a clear immune response.

We conclude that more research in animal models is required to define the optimal treatment resulting in cure of GSDII patients. We recommend more investigation in the method using gene therapy in hematopoietic stem cells as described in the study of Van Til et al.[17] We hope that with a reduced immune response the efficacy improves. On top of that, more research needs to be done in order to reach all target cells.

References

1. *Ausems M.G., Verbiest J., Hermans M.P., et al. Frequency of glycogen storage disease type II in The Netherlands: implications for diagnosis and genetic counselling. Eur J Hum Genet. 1999; 7: 713-716.*
2. *Chen Y.T., Amalfitano A. Towards a molecular therapy for glycogen storage disease type II (Pompe disease). Mol Med Today. 2000; 6: 245-251.*
3. *Pauly D.F., Fraites T.J., Toma C., et al. Intercellular transfer of the virally derived precursor form of acid alpha-glucosidase corrects the enzyme deficiency in inherited cardioskeletal myopathy Pompe disease. Hum Gene Ther. 2001; 12: 527-538.*
4. *Dittmer F., Ulbrich E.J., Hafner A., et al. Alternative mechanisms for trafficking of lysosomal enzymes in mannose 6-phosphate receptor-deficient mice are cell type-specific. J Cell Sci. 1999; 112 (Pt 10): 1591-1597.*
5. *van der Ploeg A.T., Reuser A.J. Pompe's disease. Lancet. 2008; 372: 1342-1353.*
6. *Kroos M., Pomponio R.J., van Vliet L., et al. Update of the Pompe disease mutation database with 107 sequence variants and a format for severity rating. Hum Mutat. 2008; 29: E13-26.*
7. *Xu F., Ding E., Migone F., et al. Glycogen storage in multiple muscles of old GSD-II mice can be rapidly cleared after a single intravenous injection with a modified adenoviral vector expressing hGAA. J Gene Med. 2005; 7: 171-178.*
8. *van den Hout H.M., Hop W., van Diggelen O.P., et al. The natural course of infantile Pompe's disease: 20 original cases compared with 133 cases from the literature. Pediatrics. 2003; 112: 332-340.*
9. *Hagemans M.L., Winkel L.P., Van Doorn P.A., et al. Clinical manifestation and natural course of late-onset Pompe's disease in 54 Dutch patients. Brain. 2005; 128: 671-677.*
10. *Schoser B., Hill V., Raben N. Therapeutic approaches in glycogen storage disease type II/Pompe Disease. Neurotherapeutics. 2008; 5: 569-578.*
11. *Kishnani P.S., Goldenberg P.C., DeArmey S.L., et al. Cross-reactive immunologic material status affects treatment outcomes in Pompe disease infants. Mol Genet Metab. 2010; 99: 26-33.*
12. *Poenaru L. Approach to gene therapy of glycogenosis type II (Pompe disease). Mol Genet Metab. 2000; 70: 163-169.*
13. *Kyosen S.O., Iizuka S., Kobayashi H., et al. Neonatal gene transfer using lentiviral vector for murine Pompe disease: long-term expression and glycogen reduction. Gene Ther. 2010; 17: 521-530.*
14. *Mah C.S., Falk D.J., Germain S.A., et al. Gel-mediated delivery of AAV1 vectors corrects ventilatory function in Pompe mice with established disease. Mol Ther. 2010; 18: 502-510.*
15. *Martin-Touaux E., Puech J.P., Chateau D., et al. Muscle as a putative producer of acid alpha-glucosidase for glycogenosis type II gene therapy. Hum Mol Genet. 2002; 11: 1637-1645.*
16. *Sun B., Chen Y.T., Bird A., et al. Packaging of an AAV vector encoding human acid alpha-glucosidase for gene therapy in glycogen storage disease type II with a modified hybrid adenovirus-AAV vector. Mol Ther. 2003; 7: 467-477.*
17. *van Til N.P., Stok M., Aerts Kaya F.S., et al. Lentiviral gene therapy of murine hematopoietic stem cells ameliorates the Pompe disease phenotype. Blood. 2010; 115: 5329-5337.*
18. *Mah C., Pacak C.A., Cresawn K.O., et al. Physiological correction of Pompe disease by systemic delivery of adeno-associated virus serotype 1 vectors. Mol Ther. 2007; 15: 501-507.*
19. *Ginaldi L., De Martinis M., D'Ostilio A., et al. The immune system in the elderly: II. Specific cellular immunity. Immunol Res. 1999; 20: 109-115.*
20. *Douillard-Guilloux G., Richard E., Batista L., et al. Partial phenotypic correction and immune tolerance induction to enzyme replacement therapy after hematopoietic stem cell gene transfer of alpha-glucosidase in Pompe disease. J Gene Med. 2009; 11: 279-287.*
21. *Fabb S.A., Dickson J.G. Technology evaluation: AAV factor IX gene therapy, Avigen Inc. Curr Opin Mol Ther. 2000; 2: 601-606.*
22. *Pauly D.F., Johns D.C., Matelis L.A., et al. Complete correction of acid alpha-glucosidase deficiency in Pompe disease fibroblasts in vitro, and lysosomally targeted expression in neonatal rat cardiac and skeletal muscle. Gene Ther. 1998; 5: 473-480.*
23. *Lin C.Y., Ho C.H., Hsieh Y.H., et al. Adeno-associated virus-mediated transfer of human acid maltase gene results in a transient reduction of glycogen accumulation in muscle of Japanese quail with acid maltase deficiency. Gene Ther. 2002; 9: 554-563.*
24. *Byrne B.J., Falk D.J., Pacak C.A., et al. Pompe disease gene therapy. Hum Mol Genet. 2011; 20: R61-68.*
25. *Richard E., Douillard-Guilloux G., Batista L., et al. Correction of glycogenosis type 2 by muscle-specific lentiviral vector. In Vitro Cell Dev Biol Anim. 2008; 44: 397-406.*

Systematic review

Efficacy and safety of catheter ablation for supraventricular tachycardia in the pediatric age group *A systematic review*

Fahid Chohan[a], Herbert Kroon[a], Maarten Witsenburg[b]
[a] Medical student, Erasmus MC University Medical Center Rotterdam, the Netherlands
[b] Supervisor, Department of Cardiology, Erasmus MC University Medical Center Rotterdam, the Netherlands
Correspondence: Herbert Kroon, e-mail: 334858hk@student.eur.nl

Objective: **The purpose of this systematic review was to evaluate the outcomes, efficacy and safety of catheter ablation for supraventricular tachycardia (SVT) in the pediatric age group.**
***Background*: In adults, the success rates of catheter ablation for tachyarrhythmias have been established. However, in the pediatric population less is known about the short-term success rates, recurrences during follow-up and procedure-related complications.**
***Methods*: We searched the database PubMed for articles published before the 25th of September 2012. We included articles with a design aimed at studying the outcomes of catheter ablation for SVT in the pediatric age group. Our primary endpoints were short-term success rate, recurrence rate during follow-up and procedure-related complications. We subdivided the included studies in Group I and Group II studies. Group I studies gave their results per patient, while Group II studies published results per accessory bundle (substrate).**
***Results*: The results show an overall short-term success rate of 93.1% and 93% for Group I and Group II studies. The recurrence rate was 8.8% in Group I studies and 9.2% in Group II studies. Complications were reported, with a major complication rate of 1.6% and a minor complication rate of 2.5%.**
***Conclusion*: In conclusion, we found high success rates for both short- and long-term in the pediatric population with reasonable complication rates. The complications are mostly minor, but in rare cases can be major.**

Introduction

Supraventricular tachycardia (SVT) is the most common tachyarrhythmia in the pediatric population.[1] In children, atrioventricular reciprocating tachycardia (AVRT) and atrioventricular nodal re-entry tachycardia (AVNRT) are the most common forms of supraventricular tachycardia.[2] In atrioventricular reciprocating tachycardia and atrioventricular nodal re-entry tachycardia, the AV node plays a crucial role in the re-entry circuit. Other SVTs are (ectopic) atrial tachycardia ((E)AT), atrial flutter (AFL), atrial fibrillation (AF), intra-atrial re-entry tachycardia (IART) and junctional ectopic tachycardia (JET). The incidence of supraventricular tachycardia has been estimated one in 250 to 1000 children. Approximately, 95% of the tachycardia in children is a supraventricular tachycardia.[1,3]

The mechanisms of supraventricular tachycardia have been described previously.[4] In short, supraventricular tachycardia is often marked by a narrow QRS-complex tachycardia.[1] Furthermore, heart frequencies are generally higher than 220 and 180 beats per minute in infants and children.[5] Mostly, infants tolerate SVT hemodynamically on the short-term, but on the long-term, infants can present with congestive heart failure or shock.[6] In comparison, older children will present themselves with complaints of palpitations, chest pain, light-headedness or dizziness.[1] In general, most forms of supraventricular tachycardia are not life threatening. Nevertheless, atrial fibrillation could progress into ventricular fibrillation which could cause sudden cardiac death. This mechanism is also described in patients with Wolff-Parkinson-White syndrome.[2]

Since the onset of radiofrequency catheter ablation in the 1980s in the adult population, it has become a successful tool in the treatment of SVT.[2] This is because catheter ablation of supraventricular tachycardia in adults is associated with high procedural success rates and clinical improvements during follow-up.[7] Radiofrequency ablation (burning) is perfectly suited to place a focal lesion.[8] But, harm could be done to surrounding cardiac tissue. An alternative is cryoablation, freezing of the arrhythmogenic substrate.[8] Cryoablation offers advantages compared to radiofrequency ablation, especially with regard to safety near the AV-node.[8] With cryomapping, doctors have the ability to evaluate the acute effects of cryoablation for the substrates to treat, before this will be a permanent lesion.[8] With radiofrequency ablation, each lesion made is a permanent one. Also, the formation of ice at the catheter electrode tip causes adhesion of the catheter tip to the endocardium: this avoids dislodgement during cryomapping and –ablation which provides a more precise lesion to be possible.[8] However, lower acute success rates and higher recurrence rates are reported compared to radiofrequency catheter ablation.[9]

Furthermore, recent technological progress has made it possible for cardiologists to perform catheter interventions in small children.[2] Increased experience with catheter ablation has caused this technique to be considered as a first line therapy for arrhythmias in children.[2,10] However, concerns regarding the safety and long-term outcomes of this procedure in children with a developing heart are persistent. Damage to several structures, like valves and coronary arteries, is reported as well as ventricular aneurysms. Also, the development of new arrhythmias and sudden cardiac death has already been reported.[10] Lastly, in one large pediatric series a high rate of late recurrences were observed after an initially successful ablation of the supraventricular tachycardia.[11]

Therefore it is important to evaluate the outcomes of catheter ablations in children with supraventricular tachycardia.

In this systematic review we study the results of catheter ablation in children with supraventricular tachycardia. We specifically aimed our review to examine the short-term and long-term outcomes and procedure-related complications of catheter ablation.

Methods

Search strategy

We searched the MEDLINE electronic database PubMed for articles published before the 25th of September 2012. The search was as follows: (((("Tachycardia, Supraventricular/surgery"[Mesh]) AND "Catheter Ablation"[Mesh])) AND ((((("Infant"[Mesh]) OR "Child, Preschool"[Mesh]) OR "Child"[Mesh]) OR "Adolescent"[Mesh])) NOT (((((("Adult"[Mesh]) OR "Young Adult"[Mesh]) OR "Middle Aged"[Mesh]) OR "Aged"[Mesh]) OR "Aged, 80 and over"[Mesh]).

We included English-written articles, published after January 1st 1997 of which the full text was available in the library of the Erasmus Medical Center. Furthermore, the references of the included literature were checked for articles that could match the search criteria. We also examined the latest review about this subject for references.

Our search consisted of two kinds of studies: studies which gave their results per patients and studies which gave their results per accessory bundle (substrate). We named them Group I and Group II studies, respectively. Group II studies gave their results per amount of accessory bundles rather than per amount of patients. For example, two patients could have atrioventricular reciprocating tachycardia, one of them having two accessory bundles. If both were to be treated successfully, this would be a 100% success rate in both Group I and II. However the numbers differ: there are 2 out of 2 patients successfully treated, with 3 out of 3 accessory bundles ablated.

Study selection

We specifically did not include review articles and case reports. Based on title and abstract, we excluded articles which did not refer to catheter ablation, the pediatric age group or supraventricular tachycardia. Any disputes were resolved by agreement of the two reviewers. We searched for articles with a design aimed at studying the outcomes of catheter ablation of supraventricular tachycardia in the pediatric age group.

Endpoints

We defined short-term success as non-inducibility of the supraventricular tachycardia via the standard stimulation protocol after catheter ablation by means of radiofrequency ablation or cryoablation. Long-term success was defined as a recurrence free follow-up.

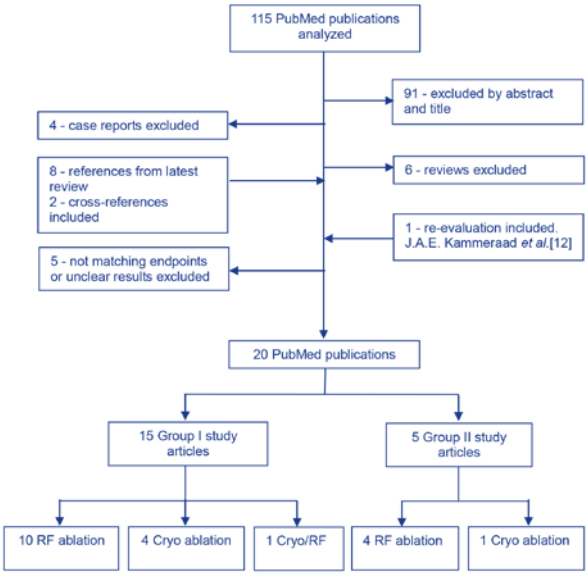
We did not mark adverse events that were apparent during the procedure and disappeared by the end of the procedure. Complications apparent after the ablation were defined as procedure-related complications, except complications which were transient and disappeared within 24 hours after the procedure. We divided complications into minor and major. Major complications were pericardial effusion, valvular damage or high grade AV block. Complications such as a low grade AV block or bruises were classified as minor complications.

Results

Our PubMed search produced 115 publications. After applying the exclusion criteria 14 articles remained. Cross-references and references of the latest review yielded 2 and 8 publications, respectively, making a total of 24 publications. Hereafter, we excluded two articles because of unclear diagnoses of patients at baseline and including patients with non-inducible tachycardia during the procedure. One article, "Catheter ablation of tachyarrhythmia substrates in children" by J.A.E. Kammeraad et al, had been added after reevaluation.[12] Another three articles were excluded because they did not investigate the endpoints we studied. Eventually, a total of 20 publications were included in our research.

We originally found 3 articles by cross-references, but "Pediatric radiofrequency catheter ablation registry success, fluoroscopy time, and complication rate for supraventricular tachycardia: comparison of early and recent eras" from Kugler JD et al was excluded, because Van Hare et al did a more recent and extensive research with comparable endpoints and patients from the same Radiofrequency Catheter Ablation Registry as stated in that article.[10]

Figure 1. Flowchart article selection.



Short-term success

High success rates of the first ablation for all the investigated supraventricular tachycardias were reported by Kammeraad JAE et al, Lee S et al, Collins KK et al and Lee P et al in Group I.[12-15] These showed success rates ranging from 86% to 100% (Table 1). There were several small sample size studies that also showed high success rates of 88% to 100%.[16-21] However, low success rates were reported by Collins KK et al, Kriebel T et al and Miyazaki A et al, the lowest being 63%.[22-24] Group II studies which reported a high success rate were Mandapati R et al and Van Hare GF et al (Table 1).[10,25] Chiu SN et al also showed high success rates, but they had a smaller sample size.[26] Low success rates were reported by Kirsh JA et al.[27]

The short-term success rates for Group I were 90% (100%), 96%, 93% and 100% for atrioventricular reciprocating tachycardia (Wolff-Parkinson-White), atrioventricular nodal re-entry tachycardia, atrial tachycardia and junctional ectopic tachycardia. For Group II studies, overall success rates were 93% (100%), 94%, 69% and 60% for atrioventricular reciprocating tachycardia (WPW), atrioventricular nodal re-entry tachycardia, atrial tachycardia and junctional ectopic tachycardia/other. In total, the short-term success rates were 93.1% and 93% for Group I and II, respectively.

Long-term success

The same studies reporting high short-term success rates in Group I, also reported low recurrence rates for radiofrequency ablation (Table 2).[12-15] These recurrence rates ranged from 1.7% to 4.7%. However, one of these studies reported a relatively high recurrence rate of 7.4% for the patients who underwent cryoablation.[14]

Table 1 - Short-term success rates for Group I and II studies.

Group I studies		Successful first ablations (SFA) / patients				
Study	Ablation Technique	AVRT (%) - WPW (%)		AVNRT (%)	AT (%)	JET (%)
Benito F, et al	RF	4/4	3/3	1/1		
1997 [16]		(100)	(100)	(100)		
Kammeraad JAE, et al	RF	36/36		16/16*	6/7	
2000 [12]		(100)		(100)	(86)	
Lee S, et al	RF	39/42		13/13	22/22	
2000 [13]		(93)		(100)	(100)	
Collins KK, et al	RF	10/15		7/8		1/1
2002 [22]		(67)		(88)		(100)
"Kammeraad JAE, et al	RF			26/26**		
2004 [17]*"*				(100)"		
Drago F, et al	Cryo	11/12	10/10	13/14		
2005 [18]		(92)	(100)	(93)		
Kriebel T, et al	Cryo	13/19		11/13		
2005 [23]		(68)		(85)		
Miyazaki A, et al	Cryo	5/8		21/22		
2005 [24]		(63)		(95)		
Collins KK, et al	Cryo			54/57		
2006 [14]				(95)		
	RFCA			60/60		
				(100)		
Drago F, et al	Cryo	21/22	21/21	28/29	1/1	1/1
2006 [19]		(95)	(100)	(97)	(100)	(100)
Vida VL, et al	RF	15/17		9/11		
2006 [28]		(88)		(82)		
Lee P, et al	RF	129/140		64/66	19/22	
2007 [15]		(92)		(97)	(86)	
Cummings RM, et al	RF				22/25	
2008 [20]					(88)	
Moltedo JM, et al	RF				4/5	
2009 [29]					(80)	
Toyohara K, et al	RF				35/35	
2011 [21]					(100)	
Total		283/315	34/34	323/336	109/117	2/2
(%)		(90)	(100)		(93)	(100)
Group II studies		**Successful first ablations (SFA) / substrates**				
Study	**Ablation Technique**	**AVRT (%) - WPW (%)**		**AVNRT (%)**	**AT (%)**	**JET /Other (%)**
Mandapati R, et al	RF	127/136				
2003 [25]		(93)				
Van Hare GF, et al	RF	364/379		150/156	2/3	Other 1/2 (50)
2004 [10]		(96)		(96)	(67)	
Kirsh JA, et al	Cryo	18/30 ***		25/30	0/2	2/3
2005 [27]		(60)		(83)	(0)	(67)
Chiu SN, et al	RF	15/16		7/7	7/8	
2009 [26]		(94)	(100)	(100)	(88)	
Total		524/561	2/2	182/193	9/13	3/5
(%)		(93)	(100)	(94)	(69)	(60)

The category AT consists of EAT, AT, AFL, AF and IART. EAT = Ectopic Atrial Tachycardia.
AT = Atrial Tachycardia. AFL = Atrial Flutter. AF = Atrial Fibrillation. IART = Intra-Atrial Re-entrant Tachycardia.
* Dual AV node physiology was still apparent in 1/16 patients.
** Dual AV node physiology was still apparent in 2/26 patients.
*** Kirsh JA et al reported 31 cases of AVRT, but treated 30 cases.

Benito F et al and Vida VL reported a rate of 0% for 5 and 28 patients.[16,28] Drago F et al, Cummings RM et al, Collins KK et al and Moltedo JM et al found high recurrence rates, ranging 25-40.9%.[18,20,22,29] After correction for these articles, the overall recurrence rate for Group I studies is 4.9%. The majority of Group I studies did not mention second or third ablations after recurrences.

In Group II studies, relatively high recurrence rates ranging from 6.7% to 10.3% were reported (Table 2). Those with a large sample size had rates below 10%.[25,27,30]

We found an overall recurrence rate of 8.8% in Group I studies. There were 23 (9), 15 and 18 recurrences for atrioventricular reciprocating tachycardia (Wolff-Parkinson-White), atrioventricular nodal re-entry tachycardia and atrial tachycardia. We found no recurrences of junctional ectopic tachycardia. One study reported 7 recurrences, but did not specify from which arrhythmia these were.[22] These were counted when calculating the overall recurrence rate for Group I studies (8.8%). Patients who had recurrences often got a redo ablation. The success rates for redo ablations were 3/3 (100%), 4/5 (80%) and 12/13 (92.3%) for AVRT, AVNRT and AT in Group I studies.

Table 2 - Long-term results after successful first ablation (SFA) in Group I and II studies.

Group I studies Study	Recurrences after SFA (per patient) AVRT (WPW)	AVNRT	AT	Recurrence rate % of SFA (number of SFAs)	Time of recurrence	Mean ± SD Range (x) M = month(s), Y = year(s)	
Benito F, et al				0% (5)		18.4 m	
1997 [16]					NR	(13 - 30 m)	
Kammeraad JAE, et al		1	1	3.4% (58)		AVRT Median	AVNRT Median
2000 [12]						1.9 y (0.2-3.8)	2.1 y (0.1-4.0)
						EAT Median	IART Median
						0.8 & 1.4 y	1.3 y (0.5-3.3)
Lee S, et al	1		2	4.1% (74)	< 5 m	23.5 ± 9.4 m	
2000 [13]						(8.2 - 39.7 m)	
Collins KK, et al	NR*	NR*		38.9% (18)	< 9 y	Median 9.9 y	
2002 [22]						(8.6 – 11.1 y)	
Kammeraad JAE, et al		3		11.5% (26)	< 3 m	Median 25 m	
2004 [17]						(3 - 61 m)	
Drago F, et al		3		29% (24)	< 1 y	(1 - 22 m)	
2005 [18]	4 (4)						
Kriebel T, et al				8.3% (24)	< 2 m	8.9 m	
2005 [23]	2					(1 - 15 m)	
Miyazaki A, et al		1		11.5% (26)	< 4 w	8.2 m	
2005 [24]	2					(0.8-14.4 m)	
Collins KK, et al *Cryo*		4		7.4% (54)	< 4 m	Median 1y	
						(5 m - 1 y)	
2006 [14] *RFCA*		1		1.7% (60)	< 1 m	Median 1 y	
						(6 m - 1 y)	
Drago F, et al		2		13.7% (51)		13 m	
2006 [19]	5 (5)				< 4 m	(5 - 30 m)	
Vida VL, et al				0% (24)		13.69 ± 7.16 m	
2006 [26]							
Lee P, et al			1	4.7% (212)	< 1 m, n=8	86 ± 38 m	
2007 [15]	9					(0.5 – 185 m)	
Cummings RM, et al			9 ***	40.9% (22)	< 2 w	(6 m - 7 y), N = 14	
2008 [20]							
Moltedo JM, et al			1	25% (4)	< 1 m	14 ± 4.6 m	
2009 [29]							
Toyohara K, et al			4	11.4% (35)	< 1 m	36 ± 5 m	
2011 [21]						(2 - 8 y)	
Total	23 (9)	15	18	8.8 % (717)	-		
Group II studies **Study**							
Mandapati R, et al	9			7% (127)	NR		
2003 [25]							
Van Hare GF, et al	43	8		9.9% (517)	(0 – 12 m)	(0 - 12 m)	
2004 [30]							
Kirsh JA, et al	1	2		6.7% (45)	NR	3 m, N = 28	
2005 [27]							
Chiu SN, et al	2		1	10.3% (29)	< 6 y	5.4 y	
2009 [26]						(1 - 13 y)	
Total	55	10	1	9.2% (718)			

The category AT consists of EAT, AT, AFL, AF and IART. EAT = Ectopic Atrial Tachycardia.
AT = Atrial Tachycardia. AFL = Atrial Flutter. AF = Atrial Fibrillation. IART = Intra-Atrial Re-entrant Tachycardia.
NR Not reported.
* A total of 7 recurrences were reported, but not specified from which type of arrhythmia.
** Third ablation abolished the recurrence in the last patient.
*** 2 Patients were lost to follow-up.

For Group II studies, we found an overall recurrence rate of 9.2%. In numbers, these were 55, 10 and 1 for atrioventricular reciprocating tachycardia, atrioventricular nodal re-entry tachycardia and atrial tachycardia, respectively. None of these studies mentioned any redo ablations for the recurrences.

Complications

Major complications were 2nd and 3rd degree (complete) AV block, valve regurgitation, reversible brachial plexus injury, intracardiac thrombosis and pericardial effusion. Minor complications were 1st degree AV block, right bundle branch block, hematomas or bruises.

Table 3 - Procedure-related complications in Group I and II studies.

Study	2nd / 3rd degree AV block	Valve regurgi-tation	Reversible brachial plexus injury	Pericardial effusion	Intracardiac thrombosis	1st degree AV block	RBBB	Local he-matomas/ bruises	Other
Benito F, et al 1997 [16]				2/5 40%					
Kammeraad JAE, et al 2000 [12]	1/59 1.7%								
Lee S, et al 2000 [13]		1/77 1.3%		1/77 1.3%					
Mandapati R, et al 2003 [25]	4/118 3.4%						5/118 4.2%		
Van Hare GF, et al 2004 [10]		5/481 1.0%			1/481 0.2%	6/481 1.2% *			8/481 1.9% **
Kammeraad JAE, et al 2004 [17]	1/26 3.8%								
Collins KK, et al 2006 [14] *Cryo* / *RFCA*				1/57 1.75%					
Vida VL, et al 2006 [28]									
Lee P, et al 2007 [15]	2/228 0.9%		2 / 228 0.9%					10 / 228 4.4%	
Chiu SN, et al 2009 [26]	1/27 3.7%								
Total	9/1413 0.6%	6/1413 0.4%	2/1413 0.1%	4/1413 0.3%	1/1413 0.07%	13/1413 0.9%	5/1413 0.4%	10/1413 0.7%	8/1413 0.6%

* The type of AV block was not defined.
** Other complications consisted of chest pain (1x), skin burn (1x), thrombosis (1x) and 'Other' complications (5x).
*** The total population was calculated without the study populations of Cummings et al and Toyohara et al, because these studies did not mention complications.

First degree AV block was the most common complication affecting 0.9% (13/1413) of the total study population (Table 3).[10,15,29] Nine cases (0.6%) of higher degree AV block were reported after catheter ablation.[12,15,17,25,26] Valve regurgitation and pericardial effusion was reported in 6 and 4 patients, respectively.[10,13,14,16] Remarkably, we found reversible brachial plexus injury in 2 patients. This was reported by one study.[15] Another major complication was intracardiac thrombosis, reported by Van Hare et al in one patient, categorized under 'other complications'.[10] Other complications in this category were chest pain, skin burn, and non-intracardiac thrombosis. Van Hare et al stated another 5 complications as 'other complications', which were not defined.[10]

If we focus on the larger studies, 53 complications were reported.[10,12-15,25] In total, 6 cases of valve regurgitation were seen.[10,13] We found two studies which reported pericardial effusion in two patients.[13,14] Second and third degree AV block occurred in 7 patients.[12,15,25] Right bundle branch block was only reported by Mandapati R et al, accounting for 5 patients.[25] A total of 13 cases of 1st degree AV block were found.[10,15,28] One study reported reversible brachial plexus injury and local hematomas or bruises in 2 and 10 patients.[15] Lastly, Van Hare GF et al listed 9 other complications, as explained earlier.[10]

Overall, the major complication rate is 1.6% (22/1413), while the rate for minor complications is 2.5% (36/1413).

Discussion

We systematically reviewed the efficacy and safety of catheter ablation for supraventricular arrhythmias in the pediatric population. The results show overall short-term success rates of 93.1% and 93% for Group I and II studies. Recurrence rates are 8.8% and 9.2% for Group I and Group II studies. The major complication rate is 1.6%, while the rate for minor complications is 2.5%.

We think that short-term success rates of 93.1% and 93% are reasonably high; especially in infants and children it's a more complex procedure because of the smaller anatomy compared to adults. Recurrence rates of 8.8% and 9.1% are on the higher range. This could be due to the fact that safety of the procedure is an important aspect. Catheter ablation must be done with caution, because normal heart tissue surrounding the bundle needs to be preserved. Four articles had extreme values of recurrences, ranging from 25% to 40.9%.[18,20,22,29] These studies had small populations ranging from 5-26 patients, one less or extra recurrence has a high impact on the recurrence rate. When we correct for these, we find an overall recurrence rate of 4.9% for Group I studies. A minor complication rate of 2.5% is high. However, 10 out of 36 were local hematomas or bruises. The major complication rate of 1.6% is high, also. This means that one out of 63 patients has severe morbidity after catheter ablation.

Relatively low short-term success rates for atrial tachycardia and junctional ectopic tachycardia/others were reported in Group II studies.[10,27] These were 69% and 60%, respectively. Kirsh JA et al had 2 out of 2 failed ablations which contributed significantly to the short-term success rate for atrial tachycardia.[27] The low success rate for junctional ectopic tachycardia/others was primarily caused by only having 5 patients in that group of which only 3 had a successful ablation.[10,27]

However, high short-term success rates were seen for atrioventricular reciprocating tachycardia and atrioventricular nodal re-entry tachycardia (93% and 94%). These results are comparable with those of Group I studies. The overall recurrence rate is 9.2%.

Limitations

This systematic review has its limitations. Firstly, this review was not controlled, without any randomized comparison between RF ablation and cryoablation. Secondly, we did not differentiate between patients with structurally normal hearts and those with cardiomyopathy or congenital heart disease. Also, we did not evaluate the type of mapping the studies used.

We did not correct for the varieties in follow-up among the studies. There was a lot of variation between studies in range of follow-up and number of patients for which the follow-up was available. Some studies have a few months of follow-up for a limited number of patients. One study reported 10 years of follow-up. The consequence of this variation in follow-up is there's risk of underreporting the recurrences. With this in mind, our results might have given a more optimistic view regarding long-term success. Also, some studies did not mention the definition of a recurrence or a failed first ablation.

Moreover, the age ranges of the studies varied from 2.5 months to 20 years. Benito et al for example only reported patients ranging from 2.5 to 8 months old. Moltedo et al had a different age group with a range of 7 to 18 years. An 18 year old patient for example has an easier procedure if we look at the anatomical problem compared to a 2.5 months old patient. It is possible that a negative effect of the smaller anatomy on the short- and long-term success is of less importance in older patients.

Several articles did report the short-term results, but were unclear about long-term success and procedure-related complications. That made it impossible, in some cases, to determine the recurrences and complications.

Conclusion

In adults, catheter ablation is already an accepted and successful therapy of supraventricular tachycardia. In the pediatric age group this was unclear. This systematic review shows optimistic results from the current literature. We found reasonably high success rates for the short-term. Overall, catheter ablation is an effective method for treating supraventricular tachycardia. However, the safety remains an issue because the rates of recurrences and complications are considerably high.

For future research, larger studies need to be conducted where radiofrequency catheter ablation and cryoablation are compared with each other in treating supraventricular tachycardia. It is important that these studies focus on a complete follow-up after catheter ablation.

References

1. *Samson RA, Atkins DL. Tachyarrhythmias and defibrillation. Pediatr Clin North Am. 2008; 55: 887-907.*
2. *Chun TUH, Van Hare GF. Advances in the Approach to Treatment of Supraventricular Tachycardia in the Pediatric Population. Current Cardiology Reports 2004, 6: 322–326.*
3. *Ludomirisky A, Garson A. Supraventricular tachycardia. In: Gillette PC, Garson A, editors. Pediatric arrhythmias: electrophysiology and pacing. Philadelphia: WB Saunders 1990; 380–426.*
4. *Shakespeare CF, Anderson M, Camm AJ. Pathophysiology of supraventricular tachycardia. European Heart Journal 1993; 14: 2-8.*
5. *Fisher DJ, Gross DM, Garson A Jr. Rapid sinus tachycardia: differentiation from supraventricular tachycardia. Am J Dis Child 1983; 137: 164–166.*
6. *Gikonyo BM, Dunnigan A, Benson DW Jr. Cardiovascular collapse in infants: association with paroxysmal atrial tachycardia. Pediatrics 1985; 76: 922–926.*
7. *Yee R, Connolly S, Noorani H. Clinical review of radiofrequency catheter ablation for cardiac arrhythmias. Canadian Journal of Cardiology 2003; 19: 1273-1284.*
8. *De Santis A, Fazio G, Silvetti MS, et al. Transcatheter ablation of supraventricular tachycardias in pediatric patients. Current Pharmaceutical Design 2008; 14: 788-793.*
9. *Buddhe S, Singh H, Du W, et al. Radiofrequency and cryoablation therapies for supraventricular arrhythmias in the young: five-year review of efficacies. PACE 2012; 35: 711-717.*
10. *Van Hare GF, Javitz H, Carmelli D, et al. Prospective assessment after pediatric Cardiac Ablation: Demographics, Medical Profiles and Initial outcomes. J Cardiovasc Electrophysiol 2004; 15: 759-770.*
11. *Kugler JD, Danford DA, Houston K, et al. Radiofrequency catheter ablation for parox-ysmal supraventricular tachycardia in children and adolescents without structural heart disease. Pediatric EP Society, Radiofrequency Catheter Ablation Registry. Am J Cardiol 1997; 80: 1438-1443.*
12. *Kammeraad JAE, Hoorntje T, Ramanna H, et al. Catheter ablation of tachyarrhythmia substrates in children. Cardiologie 2000; 7: 412-416.*
13. *Lee S, Schueller WC. Tachycardias in infants, children and adolescents: safety and effectiveness of radiofrequency catheter ablation. Cardiology 2000; 94: 44-51.*
14. *Collins KK, Dubin AM, Chiesa NA, et al. Cryoablation versus radiofrequency ablation for treatment of pediatric atrioventricular nodal re-entrant tachycardia: initial experience with 4mm cryocatheter. Heart Rhythm 2006; 3: 564-570.*
15. *Lee P, Hwang B, Chen S, et al. The results of radiofrequency catheter ablation of supraventricular tachycardia in children. PACE 2007; 30: 655-661.*
16. *Benito F, Sánchez C. Radiofrequency catheter ablation of accessory pathways in infants. Heart 1997; 78: 160-162.*
17. *Kammeraad JAE, Ten Cate FU, Simmers T, et al. Radiofrequency catheter ablation of atrioventricular nodal reentrant tachycardia in children aided by the LocaLisa mapping system. Europace 2004; 6: 209-214.*
18. *Drago F, De Santis A, Grutter G, et al. Transvenous cryothermal catheter ablation of reentry circuit located near the atrioventricular junction in pediatric patients. JACC 2005; 45: 1096-1103.*
19. *Drago F, Silvetti MS, De Santis A, et al. Lengthier cryoablation and a bonus cryoablation is associated with improved efficacy for cryothermal catheter ablation of supraventricular tachycardias in children. J Interv Card Electrophysiol 2006; 16: 191-198.*
20. *Cummings RM, Mahle WT, Strieper MJ, et al. Outcomes following electroanatomic mapping and ablation for the treatment of ectopic atrial tachycardia in the pediatric population. Pediatr Cardiol 2008; 29: 393-397.*
21. *Toyohara K, Fukuhara H, Yoshimoto J, et al. Electrophysiologic studies and radiofrequency catheter ablation of ectopic atrial tachycardia in children. Pediatr Cardiol 2011; 32: 40-46.*
22. *Collins KK, Chiesa NA, Dubin AM, et al. Clinical outcomes of children with normal cardiac anatomy having radiofrequency ablation ≥ 10 years earlier. Am J Cardiol 2002: 89: 471-475.*
23. *Kriebel T, Broistedt C, Kroll M, et al. Efficacy and safety of cryoenergy in the ablation of atrioventricular reentrant tachycardia substrates in children and adolescents. J Cardiovasc Electrophysiol 2005; 16: 960-966.*
24. *Miyazaki A, Blaufox AD, Fairbrother DL, et al. Cryoablation for septal tachycardia substrates in pediatric patients: mid-term results. JACC 2005; 45: 581-588.*
25. *Mandapati R, Berul CL, Triedman JK, et al. Radiofrequency catheter ablation of septal accessory pathways in the pediatric age group. Am J Cardiol 2003; 92: 947-950.*
26. *Chiu SN, Lu CW, Chang CW, et al. Radiofrequency ablation of supraventricular tachycardia in infants and toddlers. Circ J 2009; 73: 1717-1721.*
27. *Kirsh JA, Gross GJ, O'Conner S, et al. Transcatheter cryoablation of tachyarrhythmias in children. Initial experience from an international registry. JACC 2005; 45: 133-136.*
28. *Vida VL, Calvimontes GS, Macs MO, et al. Radiofrequency catheter ablation of supraventricular tachycardia in children and adolescents. Pediatr Cardiol 2006; 27: 434-439.*
29. *Moltedo JM, Cannon BC, Fenrich AL, et al. Radiofrequency ablation of nonautomatic focal atrial tachycardia in children with structurally normal hearts. J Interv Card Elec-trophysiol 2009; 26: 225-229.*
30. *Van Hare GF, Javitz H, Carmelli D, et al. Prospective assessment after pediatric cardiac ablation: recurrence at 1 year after initially successful ablation of supraventricular tachycardia. Heart Rhythm 2004; 1: 188-196.*

The new oral anticoagulants on the rise; prevention of stroke in atrial fibrillation

A systematic review

Hrvoje Kovacic[a], Michel Veenman[a], Dimitri A. Kool[a] and Frank W.G Leebeek[b]
[a] Medical Student, Erasmus MC University Medical Center Rotterdam, The Netherlands
[b] Supervisor,Department of Hematology, Erasmus MC University Medical Center Rotterdam, The Netherlands
Correspondence: Hrvoje Kovacic, e-mail: 330124hk@student.eur.nl

ABSTRACT

Objective: **To assess whether new oral anticoagulants are better than vitamin K antagonist warfarin in the prevention of stroke in patients with atrial fibrillation.**

Methods**: We researched which new oral anticoagulants already were tested in phase-III trials. After that we did a search in the Medline database based on several MeSH-terms.**

Results**: Our initial search resulted in eighteen possible articles. After applying the exclusion criteria only three articles remained. Rivaroxaban was found to be noninferior to warfarin in the prevention of stroke and systemic embolism (HR: 0.88; 95% CI 0.75 – 1.03; p < 0.001). Dabigatran was tested in two doses namely 110 mg and 150 which were noninferior (HR: 0.91; 95% CI 0.74 – 1.11; p < 0.001) and superior (HR: 0.66; 95% 0.53 – 0.82; p = 0.003) respectively in the prevention of stroke and systemic embolism. Apixaban is superior to warfarin in the prevention of stroke and systemic embolism (HR: 0.79; 95% CI 0.66 – 0.95; p = 0.01).**

Conclusions**: The new oral anticoagulants have been proven to be noninferior or even superior in preventing stroke in patients with atrial fibrillantion. Apixaban showed the most promising results, but further research is needed.**

Introduction

Atrial fibrillation

Atrial fibrillation is the most common cardiac arrhythmia. This results in the loss of coördinated contractions of the atria followed by blood stasis and thrombus formation and therefore increase the risk of stroke. [2] Studies have shown that the incidence of atrial fibrillation increases with age, occurring in 1 percent of patients <60 years and 8 percent of patients >80 years. [2] It also increases the risk of an ischemic stroke fivefold. [3]

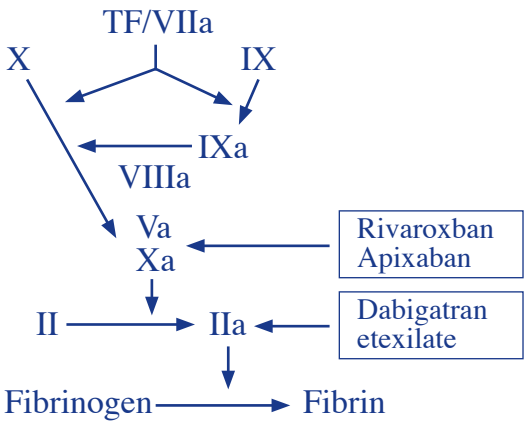
Warfarin

The current therapy for the prevention of a stroke in patients with atrial fibrillation is a vitamin K antagonist. In most countries warfarin is used, but in the Netherlands other coumarinderivates such as acenocoumarol and fenprocoumon are used. Multiple times it has been proven that warfarin is efficient in reducing the risk of having a stroke. [4] It reduces the risk of stroke with 62 percent. [5] However, warfarin has several side effects. There is a great variation in response on warfarin between en within individuals; therefore the extent of anticoagulation has to be monitored frequently with the INR. Studies have shown that patients with atrial fibrillation taking warfarin are outside the target range for 50 percent of the time. Also, warfarin increases the risk of bleeding and has multiple food and medication interactions. This all makes it difficult to monitor patients. [6]

New oral anticoagulants

Despite the fact that warfarin successfully reduces the risk of stroke in patients with atrial fibrillation, its side effects made investigators search for alternatives. They developed a few Xa inhibitors and IIa inhibitors. At this moment most of the studies researched the effects of rivaroxaban, apixaban and dabigatran. (Figure 1) Their advantage over warfarin is that the patients do not need to be monitored regularly. It is also easier in use and there are fewer interactions with food and other medication, which makes the pharmacokinetics more predictable. However, a disadvantage of the new oral anticoagulants is the lack of an antidote to oppose the anticoagulant effects. [7] In this review we discuss whether new oral anticoagulants are better than vitamin K antagonist warfarin in the prevention of stroke in patients with atrial fibrillation.

Figure 1- Target areas of new oral anticoagulants(1).



Methods

Literature search

We searched the PubMED database on the 12th of January, 2012 for articles that researched if the use of new oral anticoagulants in the prevention of stroke in patients with atrial fibrillation was better than warfarin. The following search terms were included: ("Factor Xa/antagonists and inhibitors"[Mesh] OR "Thrombin/antagonists and inhibitors"[Mesh] OR "apixaban" [Supplementary Concept] OR "dabigatran etexilate" [Supplementary Concept] OR "rivaroxaban" [Supplementary Concept]) AND "Warfarin"[Mesh] AND "Stroke"[Mesh] AND "Atrial Fibrillation"[Mesh]. We restricted the search to randomized controlled trials that were published in English with no date limitation.

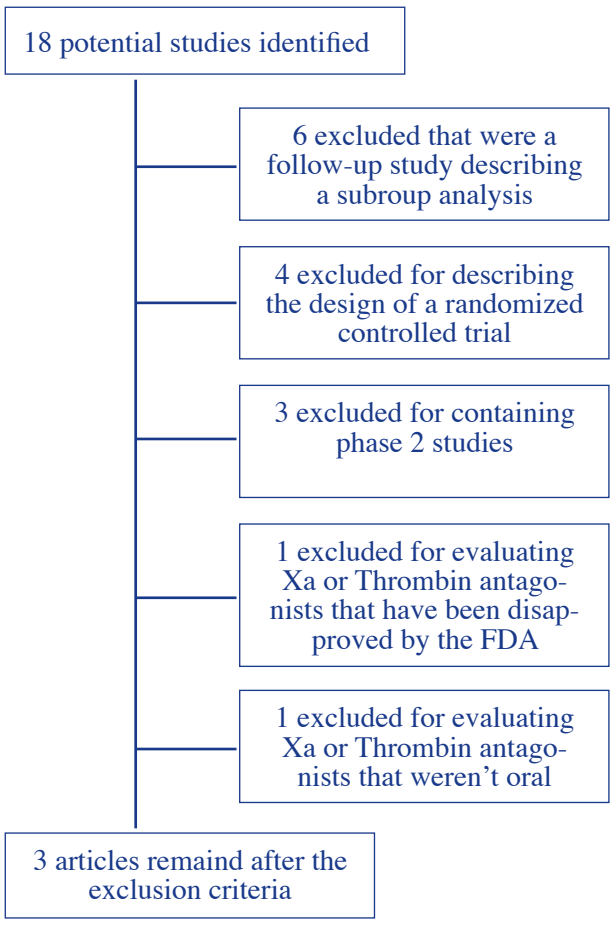
Exclusion criteria

We excluded articles based on the following criteria: Original articles that were a follow-up study describing a subgroup analysis, studies describing the design of a randomized controlled trial, articles containing phase 2 studies, studies evaluating Xa or thrombin antagonists that have been disapproved by the FDA, evaluating Xa or thrombin antagonists that weren't oral medication. (Figure 2)

Study Quality Assessment

All three authors independently assessed study quality according to randomization methods, event rate differences in the warfarin treatment group, adherence and definitions of stroke and major bleeding.

Figure 2- Flowchart for selection of studies



Results

Eighteen publications were found with our initial PubMED search on the 12th of January. After exclusion, three publications remained. (Figure 2) In the three trials warfarin was compared with either, Rivaroxaban, Apixaban or Dabigatran in the prevention of stroke in patients with atrial fibrillation.

Study population

Population at risk was defined in the CHADS2 score, this is a measure of the risk of stroke from ranges 1-6, with higher scores indicating an increased risk. The CHADS2 groups differed in the three studies. Among the studies the ARISTOTLE had a large population at risk, from CHADS2 score 1 to 6. The RELY study had a population with a moderate to mild risk (32% of patients had a CHADS2 score between 3-6) and the used population in the ROCKET-AF study had a moderate to severe risk (87% had CHADS between 3-6). [8-10]

Furthermore, the RELY and ARISTOTLE studies excluded patients with a severe renal impairment, 30 ml/min and 25ml/min respectively.

New oral anticoagulants compared with Warfarin

The ROCKET AF (Rivaroxaban Once Daily Oral Direct Factor Xa Inhibition Compared with Vitamin K Antagonist for Prevention of Stroke and Embolism Trial in Atrial Fibrillantion) study compared Rivaroxaban with Warfarin in a randomized controlled, double blind trial. In this trial 14264 patients with nonvalvular atrial fibrillation who had an increased risk for stroke were randomly assigned to receive either Rivaroxaban at a daily-dose of 20mg or dose-adjusted warfarin. The median duration of the follow-up was approximately 1.9 years.The trial found that rivaroxaban was noninferior to warfarin in the prevention of stroke or systemic embolism (HR 0.88; 95%CI, 0.75-1.03; $p < 0.001$ for noninferioty) (Table 1). Furthermore, the study found that rates of major bleeding were similar in rivaroxaban and warfarin (HR 1.04; 95%CI 0.90-1.20); $p = 0.58$ for superiority) (Table 2). [8]

The RE-LY (Randomised Evaluation of Long-Term Anticoagulation Therapy) trial compared Dabigatran in two doses, specifically 110mg twice a day and 150mg a day with dose-adjusted warfarin in a randomized controlled, trial. Dabigatran was administered in a blinded fashion and warfarin was administered unblinded. In this trial 18,113 patients with atrial fibrillation and a risk of stroke were randomly assigned and the median follow-up was 2.0 years. The trial showed that a dose of 110mg of dabigatran was noninferior to dose-adjusted warfarin (HR 0.91; 95%CI 0.74-1.11; $p <0.001$ for noninferioty, $p=0.34$ for superiority). In the group of patients that received 150mg dabigatran the study found that this dose was superior to warfarin (HR: 0.66; 95%CI 0.53-0.82; $p < 0.001$ for superiority) (Table 1). Major bleeding occurred less in the group of patients receiving 110mg of dabigatran (HR 0.80; 95%CI 0.69-0.93; $p = 0.003$ for superiorty). Major bleeding in the group of patients receiving 150mg of dabigatran had a similar rate compared to the group receiving warfarin (HR: 0.93; 95%CI 0.81-1.07; $p = 0.31$ for superiority) (Table 2). [9]

The ARISTOTLE (Apixaban for Reduction in Stroke and Other Thromboembolic Events in Atrial Fibrillation) trial compared apixaban at a dose of 5mg twice daily with warfarin. This study used a double-blind, double-dummy design. The 18,201 patients with atrial fibrillation and at least one additional factor for stroke were randomly assigned to treatment with apixaban or dose-adjusted warfarin. The median duration of follow-up was 1.8 years. The trial showed that apixaban was noninferior to warfarin in prevention to stroke (HR: 0.79; 95%CI 0.66-0.95; $p = <0.001$; $p = 0.01$ for superiority) (Table 1). The group receiving apixaban showed less major bleedings than the group receiving warfarin (HR: 0.69; 95%CI 0.60-0.80; $p < 0.001$ for superiority) (Table 2). [10]

Table 1 - Primary Efficacy Outcome of Stroke or Systemic Embolism

	Treatment	N (number of patients)	Event Rate (Treatment)	Event Rate (Warfarin)	Hazard Ratio (95% CI)	P Value	
						Non-inferiority	Superiority
ROCKET AF	Rivaroxaban	14171	2.1	2.4	0.88(0.75-1.03)	<0.001	0.12
ARISTOTLE	Apixaban	18201	1.27	1.6	0.79(0.66-0.95)	<0.001	0.01
RELY*	Dabigatran	12037	1.53	1.69	0.91(0.74-1.11)	<0.001	0.34
							<0.001
RELY	Dabigatran	12098	1.11	1.69	0.66(0.53-0.82)	<0.001	0.4%

* Dabigatran dose of 110 mg, N(treatment) =6015, N(Warfarin)= 6022
** Dabigatran dose 150 mg,N(treatment)= 6076, N(Warfarin)= 6022

Table 2 - Safety outcomes, Rates of Major Bleeding Events

	Treatment	N (Patients with Event)	Event Rate %/yr (Treatment)	Event Rate %/yr (Warfarin)	Hazard Ratio (95% CI)	P Value	
						Non-inferiority	Superiority
ROCKET AF	Rivaroxaban	781	3.6	3.4	1.04(0.9-1.20)		0.58
ARISTOTLE	Apixaban	789	2.13	3.09	0.69(0.60-0.80)		<0.001
RELY*	Dabigatran	719	2.71	3.36	0.80(0.69-0.93)		0.003
RELY	Dabigatran	772	3.11	3.36	0.93(0.81-1.07)		0.31

* Dabigatran dose of 110 mg, N(treatment) =322, N(Warfarin)= 397
** Dabigatran dose 150 mg, N(treatment) =375, N(Warfarin)= 397

Discussion

The results showed that apixaban was superior in the prevention of stroke and systemic embolism (p = 0.01), rivaroxaban was noninferior in the prevention of stroke and systemic embolism (p = <0.001), dabigatran 110mg was noninferior in the prevention of stroke and systemic embolism (p = <0.001), dabigatran 150mg was superior in the prevention of stroke and systemic embolism (p = <0.001). [8-10] Therefore, the new oral anticoagulants seem to be comparible or similar with regard to outcome.

The main advantage of these new oral anticoagulants is that there is no more need for the patients to be monitored regularly in comparison to warfarin where the anticoagulation effect is difficult to control. [6] However, there are regions where the anticoagulation effect of warfarin is under good control. The question is if the new oral anticoagulants are cost-effective in those regions. [11] However, a few questions are raised.

We noticed that all three trials had a high discontinuation rate for which the reason was largely unknown. In the RELY study the rates of discontinuation for 110 mg, 150 mg and warfarin were respectively 20.7%, 21.2% and 16.6%. In the ROCKET AF study a discontinuation of 23.7% was seen with rivaroxaban and 22.2% with warfarin. The rates in the ARISTOTLE study were 25.3% for apixaban and 27.5% for warfarin. The discotinuation rate is approximately the same in the three studies, therefore it had no effect on the outcome. [8-10]

Side effects

The new oral anticoagulants were not without side effects. First if all, in the RELY study in patients who were using dabigatran dyspepsia was more common. In 11.8% of the patients using dabigatran dose of 110 mg and in 11.3% of the patients using a dabigatran dose of 150mg, compared with 5.8% in the warfarin group. [9]

Gastrointestinal bleedings were more frequent in patients using dabigatran and rivaroxaban compared with the warfarin groups. In 1.0% of 110mg dabigatran group, 1.3% of 150mg dabigatran group compared with 0.9% in the warfarin group. In the rivaroxaban group gastrointestinal bleedings were significantly higher occuring in 3.2% of the patients compared with 2.2%, P<0,001, in the warfarin group. [8, 9]

Furthermore, myocardial infarction was more frequent in the 110-mg and 150-mg dabigatran groups than in the warfarin group. This might be because of the possible protecting effect of warfarin on myocardial infarction. [12, 13]

There were also several differences between the new oral anticoagulants. First of all there is a difference between the studies in the used population at risk. As seen in the results the studies we reviewed used different populations at risk defined with the CHADS2 score.

Furthermore, as described in the results, the RELY and ARISTOTLE studies excluded patients with a severe renal impairment. With an increase age, a decrease in renal function is seen. A normal renal clearance for a 75 year old person is 65 ml/min. However, in the used patientgroup there is a risk for further renal impairment because of cardiovascular diseases. Exclusion of the high-risk groups results in a bias. Therefore, the tested medication may seem overrated in those high-risk groups.

Definition of stroke and major bleeding

The definition of stroke was approximately the same in the studies that we included for our review. However, the definitions of major bleeding were different in the three studies. The ROCKET AF study defined major bleedings as events involving the central nervous system that met the definition of stroke were adjudicated as hemorrhagic strokes and include in both the primary efficacy and safety end points. [8]

In the RELY study major bleeding was defined as a reduction in the haemoglobin level of at least 20g per liter, transfusion of at least 2 units of blood, or symptomatic bleeding in a critical area or organ. [9] Finally, the ARISTOTLE study described a major bleeding, according to the ISTH criteria, as clinically overt bleeding accompanied by a decrease in the haemoglobin level of at least 2 g per decilitre or transfusion of at least 2 units of packed red cells, occurring at a critical site, or resulting in death. [10] In the ROCKET AF study a major bleeding needed to meet the definition of a stroke and therefore it might be that fewer bleedings were defined as major bleedings compared with the ARTISTOLE and RELY studies. However, the warfarin rate of bleeding was similar in ROCKET AF(3.4%), ARISTOTLE(3.09%), RELY(3.36%). Therefore, the different types of definitions result in similar outcomes.

Conclusion

In conclusion, the new oral anticoagualants have proven to be non-inferior or superior to warfarin in the prevention of a stroke or systemic embolism. Furthermore, there was a reduction of major bleeding events for low dose dabigatran and apixaban versus warfarin. However, there are several things still to be investigated. For example, the new medication has a few adverse effects that need to be more investigated because of the complications that can occur. The cost efficiency is different in several regions. Some regions, for instance Scandinavia and the Netherlands, already managed to keep the INR inside of the therapeutic range for more than 70% the time. [14] The question is whether the new anticoagulants would be beneficial in those regions. [15]

Still, the most important disadvantage of the new oral anticoagulants is the lack of an antidote. In critical situations, like traumatic events, an antidote is important to counteract the anticoagulation effect and stop the bleeding.

When comparing the studies we found that the new oral anticoagulants are not inferior to warfarin, some were even superior. Of the three anticoagulants, it seems that apixaban is the most promising. However, we think that there are still some fields that need further investigation to be safe enough for administration.

There is stil a lot of research needed to determine whether or not the new oral anticoagulants will be the standard therapy in the prevention of stroke in patients with atrial fibrillation. For example, research will need to be conducted in finding an antidote. Also, at this moment no test is available to measure the compliance, maybe this will not be the case in the future. There are several other trials with new oral anticoagulants, like edoxaban. [16] However, these are phase-II trials and therefore in development. Finally, the cost- effectiveness of the new oral anticoagulants should be investigated. The results imply that the new oral anticoagualants could be effective. However, if it is not cost-effective insurance companies may not be interested and would rather invest in new ways to improve the complaince of vitamin K antagonists.

References

1. *Eikelboom JW, Weitz JI. A replacement for warfarin: the search continues. Circulation. 2007; 2: 131-133.*
2. *Gutierrez C, Blanchard DG. Atrial fibrillation: diagnosis and treatment. Am Fam Physician. 2011; 1: 61-68.*
3. *Wolf PA, Abbott RD, Kannel WB. Atrial fibrillation as an independent risk factor for stroke: the Framingham Study. Stroke. 1991; 8: 983-988.*
4. *Hart RG et al., Antithrombotic therapy to prevent stroke in patients with atrial fibrillation: a meta-analysis. Ann Intern Med. 1999; 7: 492-501.*
5. *Kakar P, Lip GY. Atrial fibrillation and stroke prevention. Expert Rev Neurother. 2006; 10: 1523-1530.*
6. *Pirmohamed M. Warfarin: almost 60 years old and still causing problems. Br J Clin Pharmacol. 2006; 5: 509-511.*
7. *Garcia D, Libby E, Crowther MA. The new oral anticoagulants. Blood. 2010; 1: 15-20.*
8. *Patel MR et al., Rivaroxaban versus warfarin in nonvalvular atrial fibrillation. N Engl J Med. 2011; 10: 883-891.*
9. *Connolly SJ et al., Dabigatran versus warfarin in patients with atrial fibrillation. N Engl J Med. 2009; 12: 1139-1151.*
10. *Granger CB et al., Apixaban versus warfarin in patients with atrial fibrillation. N Engl J Med. 2011; 11: 981-992.*
11. *Pink J et al., Dabigatran etexilate versus warfarin in management of non-valvular atrial fibrillation in UK context: quantitative benefit-harm and economic analyses. Bmj. 2011; 343: d6333.*
12. *Uchino K, Hernandez AV. Dabigatran Association With Higher Risk of Acute Coronary Events: Meta-analysis of Noninferiority Randomized Controlled Trials. Arch Intern Med. 2012; 5: 397-402.*
13. *Hurlen M et al., Warfarin, aspirin, or both after myocardial infarction. N Engl J Med. 2002; 13: 969-974.*
14. *Wallentin L. et al., Efficacy and safety of dabigatran compared with warfarin at different levels of international normalised ratio control for stroke prevention in atrial fibrillation: an analysis of the RE-LY trial. Lancet. 2010; 9745: 975-983.*
15. *Hankey GJ. Replacing aspirin and warfarin for secondary stroke prevention: is it worth the costs? Curr Opin Neurol. 2010; 1: 65-72.*
16. *Weitz JI et al., Randomised, parallel-group, multicentre, multinational phase 2 study comparing edoxaban, an oral factor Xa inhibitor, with warfarin for stroke prevention in patients with atrial fibrillation. Thromb Haemost. 2010; 3: 633-641.*

Systematic review

Diagnostic value of FDG-PET/CT in fever of unknown origin

A systematic review

Emma L. van der Ende[a], Rani V.G. Punwasi[a], Paul L.A. van Daele[b], Karin A.M. Schurink[c]
[a] Medical students, Erasmus MC University Medical Center Rotterdam, the Netherlands
[b] Supervisor, Department of Clinical Immunology, Erasmus MC University Medical Center Rotterdam, the Netherlands
[c] Supervisor, Department of Medical Microbiology and Infectious Diseases, Erasmus MC University Medical Center Rotterdam, the Netherlands
Correspondence: Rani Punwasi, e-mail: 345751rp@student.eur.nl

ABSTRACT

Background purpose: **The diagnosis of patients with fever of unknown origin (FUO) remains a challenging problem for internal medicine. Several studies have suggested that Fludeoxyglucose with PET/CT (FDG-PET/CT) could provide an outcome. Our aim was to determine the diagnostic value of FDG-PET/CT in classic FUO patients and to suggest the correct stage for using FDG-PET/CT.**
Methods: **We systematically searched the online database Pubmed for articles published between 26 September 2007 and 26 September 2012 which met our predetermined criteria: studies had to include adults with classic FUO that underwent FDG-PET/CT and provide sufficient statistical data. Reviews and case-reports were excluded.**
Results: **Eight studies (six retrospective and two prospective) met our inclusion criteria. Collectively 285 patients were analyzed; between 29% and 41.7% of them remained undiagnosed. Final diagnosis involved infections (19-37%), non-infectious inflammatory diseases (8-37.5%), malignancy (0-25%). Diagnostic value showed a range of 41.7-66.7%.**
Conclusion: **Due to the heterogeneity of the studies, it remains difficult to draw a definitive conclusion. However, FDG-PET/CT provided a significant, or malignancy contribution to a final diagnosis, for example in comparison with CT scan (38.5 vs. 58.3 % ; 95% CI 0.0524-0.3327). This suggests that FDG-PET/CT should be used at an earlier stage and could be cost-effective. Moreover some articles suggest a correlation between FDG-PET/CT and CRP. More research is required to verify our suggested algorithm for FUO work-up.**

Introduction

Identifying the source of fever of unknown origin (FUO) remains a major medical challenge. Approximately 1.5-3% of all hospitalized patients are admitted for FUO[1], and 10-51% of FUO patients remain undiagnosed.[2] The main causes of FUO are infectious disease, followed by non-infectious inflammatory disease and neoplasm. Early identification of the cause of FUO is essential in guiding further diagnostic procedures and for early initiation of treatment.

FUO was defined in 1961 by Petersdorf and Beeson as recurrent fever of 38.3° or higher lasting for at least three weeks, and without a diagnosis after 1 week of hospital evaluation.[3] Nowadays, FUO is generally interpreted as no diagnosis after appropriate inpatient or outpatient evaluation.[4] FUO can be divided into four main groups: classic FUO, nosocomial FUO, immunodeficient FUO and HIV-associated FUO.

Currently, identification of the cause of FUO relies on extensive laboratory testing and conventional anatomic imaging modalities such as ultrasound, computed tomography (CT) and magnetic resonance imaging (MRI). However, these imaging modalities can only detect lesions that produce substantial anatomical change, thus some lesions cannot be detected at an early stage.[5] Also, whole-body scanning is rarely used, although this could be useful in FUO. The use of a positron emission tomography (PET) scan can overcome these limitations.

PET using 18F-FDG (a glucose analogue, abbreviated as FDG) as a radiotracer is a well-established clinical tool for assessment of a wide range of malignancies.[5] To visualize uptake of FDG throughout the body, FDG is coupled to a radioactive Fluorine molecule. Uptake of FDG indicates a high intracellular glucose metabolism, such as in malignant cells. Increased uptake is also seen in inflammatory cells, so it can also be used to identify infectious processes and non-infectious inflammatory processes. FDG-PET may therefore also be useful in the investigation of FUO. PET, however, has a low spatial resolution; thus the exact anatomic location of increased FDG uptake would be difficult to determine.[5] In 2001, a hybrid PET and CT using FDG as a tracer (FDG-PET/CT) was developed, allowing anatomic location and PET uptake to be combined in one image.

As FDG-PET/CT is relatively new, few studies have been done to assess its value in FUO investigation. Also, it is unclear at which stage an FDG-PET/CT should be considered. Studies imply that FDG-PET/CT may be cost-effective when applied at an earlier stage than currently.[1] In this systematic review, we addressed the following research questions:

1. What is the diagnostic value, sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV) of FDG-PET/CT in patients with classic FUO? We defined diagnostic value as the ratio of true positive FDG-PET/CTs to total FDG-PET/CTs.
2. What is the added value of FDG-PET/CT compared to conventional diagnostic methods? Specifically, is FDG-PET/CT significantly more effective than the CT scan alone for investigating FUO?
3. What are the most important advantages and disadvantages of FDG-PET/CT?
4. Is FDG-PET/CT cost-effective as a method of early diagnosis?
5. At which stage of FUO work-up should FDG-PET/CT be used?

Methods

Data Sources and Searches

We searched in the electronic database PubMed for studies evaluating the value of 18F-FDG-PET/CT as a diagnostic tool in FUO patients. Our final search was: "fluorodeoxyglucose 18F" OR "18f FDG" AND "PET" AND "diagnosis" AND "FUO" OR "fever of unknown origin" OR "febris eci." We limited the search to studies published between 26 September 2007 and 26 September 2012 and to articles written in English. In addition, we limited the search to "humans." We screened the references of the obtained studies for useful articles.

Study Selection

Studies were selected for inclusion when they met the following criteria: Patients in the studies had to meet the criteria for classic FUO: recurrent fever of 38.3°or higher, lasting for at least three weeks, and without a diagnosis after appropriate inpatient or outpatient examinations.[4]

Studies had to exclude patients who had recently undergone a surgical procedure, patients with an immune-compromised status, or patients with nosocomial fever.

Studies that included a particular age category e.g. only children were excluded, as this would increase the heterogeneity of the articles.

Data in the studies must have been sufficient to determine the number of true positives (TP), false positives (FP), true negatives (TN) and false negatives (FN) and to allow us to calculate specificity, sensitivity, PPV and NPV.

A case was defined as TP when the FDG-PET/CT result led to the correct diagnosis, either directly or after recommended tests (such as biopsy). FP was defined as an abnormality on the FDG-PET/CT which could not be identified as the cause of fever.

FN was defined as a negative FDG-PET/CT, while a disease causing FUO was identified by another method. TN was defined as no foci on the FDG-PET/CT and no evidence of disease after clinical follow-up.

Reviews and case reports were excluded. Studies that used FDG-PET instead of FDG-PET/CT were excluded.

The articles were screened by title and abstract by two independent reviewers (EE and RP). Articles were rejected if they clearly did not meet the drawn criteria.

Data extraction and analyses

For each included study, we extracted the following information: their definition of classic FUO, which pre-exams were applied before FDG-PET/CT was conducted, patient characteristics (number of patients, mean age), number of undiagnosed patients, final diagnosis and mean duration of follow-up. We used these data to calculate diagnostic value, sensitivity, specificity, PPV and NPV.

Results

Our initial literature search yielded 68 results. Of these studies 8 met our inclusion criteria after reading the abstract and full paper and were included in our review; 6 studies were retrospective and 2 studies were prospective. All studies used the same preparation for the FDG-PET/CT, such as administration of FDG and fasting times. Sample sizes ranged from 10 to 68 patients. All studies except one used follow-up to determine the outcome of patients that had a negative FDG-PET/CT scan. Follow-up duration varied from 4 to 43.2 months. Table 1 shows the characteristics of each study in detail. Much variation was seen in the pre-exams carried out per study. All studies categorized causes of FUO into four main groups: infectious, non-infectious inflammatory, neoplastic or miscellaneous. Table 2 presents the exact distribution per study.

The most common causes of infectious disease were pneumonia, tuberculosis, salmonella, abcesses, Q fever, CMV, endocarditis, pyelonefritis, osteomyelitis, Dengue fever and hepatitis. The most common inflammatory diseases were vasculitis, adult-onset Still's disease, sarcoidosis, thyroiditis, Crohn's disease, Sjögren's syndrome, SLE, Wegener's syndrome, reactive arthritis, polyartritis, giant cell arteritis, and pancreatitis. Neoplastic causes include non-Hodgkin's lymphoma's, myelodysplastic syndrome, various carcinoma's, chronic lymphocytic leukemia and in one case prolactinoma. Miscellaneous causes were drug fevers or transplant rejection.

We found a sensitivity of 64.7-88.9%, and a specificity of 33.3-100%. Positive predictive value was 71.4-100% and negative predictive value was 45.5-77.8%. Diagnostic value ranged from 45.0-66.7% (table 3).

In patients with a false positive FDG-PET/CT, either the fever subsided spontaneously, or another cause of FUO was found unrelated to the focus of increased uptake. Several studies found increased FDG uptake in lymph nodes. After biopsy, it was concluded that this was not the cause of FUO.[13]

Figure 1 - Flow chart of systematic review process

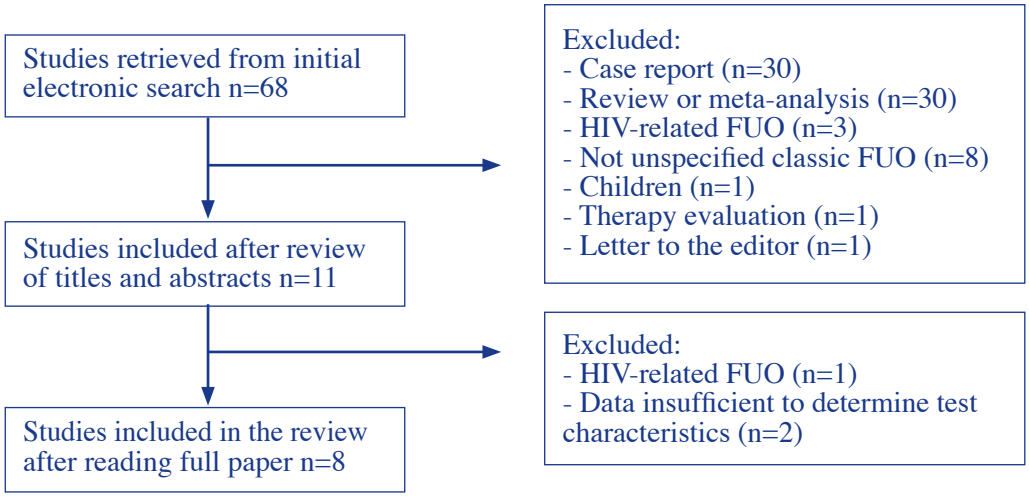


Table 1 - Study characteristics

Study	Population (n)	Mean age (years)	Definition FUO	Pre-exams	Duration of follow-up (months)	Study design
Kei et al.[6]	12	13-75	Classic FUO with 3 days of inpatient or 2 of weeks outpatient investigation	Blood routine investigation and conventional imaging modalities	N/A	Retrospective
Seshadri et al.[7]	23	33-83	Classic FUO with ≥ 1 week of inpatient investigation	N/A	≥6	Prospective
Balink et al.[8]	68	23-91	Classic FUO with appropriate inpatient or outpatient investigation	N/A	4-24	Retrospective
Sheng et al.[9]	48	24-82	Classic FUO with > 1 week of inpatient investigation	First*- , second**- & third***- line	≥6	Retrospective
Federici et al.[10]	10	25-74	Classic FUO with ≥ 1 week of inpatient investigation	C-reactive protein, cellular blood count, electrolytes, creatinine, protein electrophoresis, alanine amino-transferase, alkaline phosphatase, lactate dehydrogenase, antinuclear antibodies, urinalysis, blood culture, urine culture, tuberculosis tests (tuberculin skin test, sputum or urine analysis), chest radiography and abdominal ultrasonography	-	Retrospective
Keidar et al.[11]	48	24-82	Classic FUO with > 1 week of inpatient investigation	Routine laboratory tests, urinanalysis, blood and urine cultures, chest x-ray and abdominal US or CT	12-36	Prospective
Pedersen et al.[12]	52	32-64	Classic FUO (only admitted patients were included)	N/A	16.8-43.2	Retrospective
Pelosi et al.[13]	24	14-81	Classic FUO after appropriate inpatient or outpatient investigation	X-ray chest or CT-scan, abdominal US, routine blood chemistry, urinanalysis and in-depth physical examination	≥6	Retrospective

* First-line Complete blood count with leucocyte differentiation stool and urine routine examination ESR, CRP, TP, ALT, AST, ALP, LDH, CK, GGT, albumin–globulin ratio, level of blood glucose and fat, electrolytes, renal function tests, cultivation for blood, urine, throat and sputum,(when needed) rheumatoid factor, antistreptolysin O titer, Purified Proteine Derivatives, chest radiography, ECG, abdominal ultrasonography
** Second-line Serologic antibodies tests for cytomegalovirus, Epstein–Barr virus, rubeola, toxoplasma, hepatitis viruses, and HIV serology C3, C4, protein electrophoresis ANA, AMA, ANCA, hydrothorax and seroperitoneum test for tumor cell, thyreoid function test, Widal and Wright agglutination tests, biopsy for bone marrow, lymph node, and skin, (when diagnostic clues recommend), bone marrow culture, CT and/or MRI of abdomen, chest or cerebrum, echocardiography, colonoscopy, ECT scanning (e.g., bone, parotid, thyroidea), PET/CT scanning
***Third-line Liver biopsy, surgery (e.g., splenectomy and pancreectomy)

Table 2 - Distribution of final diagnoses per study

Study	Diagnosis				
	Infection	Inflammatory	Neoplasm	No diagnosis	Misc.
Kei et al.[6]	33.35%	8.3%	16.7%	41.7%	
Seshadri et al.[7]	26%	35%	4%	35%	
Balink et al.[8]	37%	21%	3%	25%	4%
Sheng et al.[9]	31%	19%	25%	25%	
Federici et al.[10]	40%	30%	0%	30%	
Keidar et al.[11]	18.8%	20.8%	6.3%	39.6%	14.6%
Pedersen et al.[12]	19.2%	32.7%	7.7%	40.3%	
Pelosi et al.[13]	25%	33%	12.5%	29.5%	

Another common FP result was inhomogeneous uptake in the spine. An FDG-PET/CT scan shows physiologically homogenous uptake in the spine, thus inhomogeneous uptake is an indication for bone marrow or skeletal biopsy. Incorrect interpretation could lead to homogenous uptake being categorized as a positive FDG-PET/CT. Increased uptake was proven to be FP through biopsy, urine culture or ultrasound. In Kei et al.[6], one of the false-positive cases had increased uptake around the kidney suggesting pyelonefritis, but was categorized as FP due to sterile urine culture. However, broad-spectrum antibiotics had been given at the time of urine sampling, thus it is uncertain whether or not this is an FP result.

In Seshadri et al.[7] causes of FN were adult-onset Still's disease and polymyalgia rheumatica. In Balink et al.[8] FN results were drug fevers (2), chronic lymphatic leukemia (1), Sweet syndrome (1), arteritis temporalis (1), and Churg-Strauss (1). In Sheng et al.[9] FN results were caused by Adult-onset Still's disease (1), viral hepatitis (1), upper respiratory infection (1), and pneumonia (1).

In Federici et al.[10] FN results were later diagnosed with Adult-onset Still's disease. In Keidar et al.[11] FN was caused by urinary tract infections (2), typhoid fever (1), Q-fever (1), CMV (1) and drug fever (1). In Pedersen et al.[12] causes of FN results were Adult-onset Still's disease (1), necrotizing vasculitis (1), rheumatological (1), and two cases died with fever without a cause found (but death presumably related to fever). In Pelosi et al.[13], causes of FN were auto-immune diseases (1) (not otherwise specified), lower-limb vasculitis (1), prolonged viruses (1), and biliary microlithiasis (1), which was seen as uptake on FDG-PET/CT, but interpreted incorrectly.

Discussion

Regarding the first research question, we found a diagnostic value of 41.7-66.7%, a sensitivity of 64.7-88.9% and a specificity of 33.3-100%. Positive predictive value was 71.4-100% and negative predictive value was 45.5-77.8%. False-negative results were largely due to systemic disease without a focus. For example, adult-onset Still's disease was a common FN result. FDG-PET/CT can only detect focal disease or systemic disease with an active focus. Test characteristics for focal disease are excellent, reaching up to 100% NPV and sensitivity.[6,7,10] Accordingly, it follows that a negative FDG-PET/CT scan is highly valuable to exclude many diseases. If a patient continues to have fever, then non-focal systemic disease, such as adult onset Still's disease, should be considered. Dong et al.[21] conducted a meta-analysis on the use of FDG-PET/CT in FUO, analyzing 174 pooled patients. They found a pooled sensitivity of 0.982 (95% CI 0.936-0.998) and specificity of 0.859 (95% CI 0.750-0.934). Pooled diagnostic value was 62.1%. The sensitivity (66.7-88.9%), specificity (33.3-100%) and diagnostic value found in our review are slightly lower, probably due to correction in Dong et al.[21] for heterogeneity.

Regarding the second question, our findings suggest that FDG-PET/CT is significantly more valuable than the CT scan alone (19%; 95% CI 0.1803- 0.44) [15], liver biopsy (14-17%; 95% CI 0.2018- 0.4579) [16,17] or bone marrow cultures (0-0.2% ; 95% CI 0.3941- 0.6067).[18,19] Of our evaluated studies, only Federici et al. [10] reported the superiority of FDG-PET/CT above CT chest and abdomen scan (58.3 % versus 38.5% 95 %CI (0.0524-0.3327). Crouzet et al. [20], reported that FDG-PET/CT had a diagnostic value of 73.8% compared to 62.3% (95% CI -0.0212 - 0.2457) in CT chest and abdomen scans.
FDG-PET/CT was considered essential in 24.6% of patients, because no other investigation was helpful. Crouzet et al. also compared baseline features of patients with and without contributory FDG-PET/CT. By multivariate analysis, presence of adenopathy (OR: 9.25 (1.84-46.52); p=0.01), CRP >30mg/L (OR 6.44 (1.65-25.11); p=0.01) and anemia (OR 5.03 (1.31-19.33); p=0.02) were significantly associated with contributory FDG-PET/CTs. Patients who previously used empiric antibiotic therapy for FUO more often had non-contributory FDG-PET/CT results (OR 0.19 (0.05-0.72); p=0.02).
Using these baseline predictors could therefore provide an indication in which patients an FDG-PET/CT is particularly valuable. However, it is uncertain whether patients with these positive baseline predictors also have more contributory CT scans.

Regarding the third research question, advantages of the FDG-PET/CT include good test characteristics, whole body imaging (head to mid-thigh), and high patient convenience. Disadvantages include high radiation doses, relatively high cost, limited availability and a high rate of false-positive results. Another disadvantage of the FDG-PET/CT is its inability to detect systemic, non-focal disease.[5] It cannot detect disease in certain organs due to high glucose metabolism (brain, heart) or due to excretion of FDG (gastro-intestinal and urinary tract).[5] The latter explains the large proportion of FN patients with urinary tract infection. It is important to carry out sufficient urine testing and culture to exclude urinary tract infections. Patients who are receiving corticosteroids to treat their fever have a much higher risk of a false negative FDG-PET/CT because inflammation is suppressed.[5] FDG-PET/CT may be less reliable in diabetic patients due to altered glucose metabolism.[5] Furthermore, distinguishing between malignancy, infection or inflammation is not possible.
Although our reviewed articles show FDG-PET/CT was contributive in 50.8% (45-66.7%) of the cases in identifying the cause of fever, clinicians should not base their final diagnosis on FDG-PET/CT alone.[6-13] For instance, although FDG-PET/CT has proven to be highly sensitive and specific for large cell vasculitis, an increased uptake of the tracer in artherosclerotic plaque could be mistaken for vasculitis.[12] In Pelosi et al.[13], the FDG-PET/CT showed an increased uptake in the abdominal lymph nodes, although after biopsy it was shown not to be the cause of the fever. Unfortunately, the impact or consequence of these specific, invasive investigations were not described in those articles. We assume this was due to either the retrospectively acquired information or because this was not their primary aim. However, the negative effect of unnecessary lymph node biopsies and other investigations due to false-positive FDG-PET/CT must not be underestimated.

Regarding the fourth research question, despite its high cost FDG-PET/CT could be still be cost-effective if it is performed early in a FUO investigation. Early diagnosis limits the number of other noncontributory and invasive tests and reduces time to diagnosis, which in turn reduces the duration of hospitalization. Beccara Nakayo et al.[1] analyzed the cost-effectiveness of FDG-PET/CT and suggested a place in FUO workup. They found a sensitivity of 78.57%, a specificity of 83.33%, PPV of 91.67% and NPV of 62.50%, which is comparable to results of other studies. According to their calculations, performing FDG-PET/CT early (at the end of the second diagnostic week) would save €5471.00 per patient.[1]

Table 3 - Results and test characteristics of FDG-PET/CT per study

Study	Patients with diagnosis, n (%)	Abnormal FDG-PET/CT, n (%)	False positive results, n	False negative results, n	Diagnostic value	Sensitivity	Specificity	PPV
Kei et al.[6]	7 (58.3%)	7 (58.3%)	2/12	2/12	0.417	0.714	0.600	0.714
Seshadri et al.[7]	15 (65.2%)	14 (60.1%)	2/23	2/23	0.522	0.857	0.778	0.857
Balink et al.[8]	44 (64.7%)	41 (60.3%)	3/68	6/68	0.559	0.863	0.875	0.927
Sheng et al.[9]	36 (75.0%)	40 (83.3%)	8/48	4/48	0.667	0.889	0.333	0.800
Federici et al.[10]	7 (70.0%)	5 (50.0%)	0/10	2/10	0.500	0.714	1.000	1.000
Keidar et al.[11]	29 (60.4%)	27 (56.3%)	5/48	6/48	0.458	0.759	0.737	0.815
Pedersen et al.[12]	15 (68.2%)	12 (54.5%)	2/22	5/22	0.450	0.667	0.714	0.833
Pelosi et al.[13]	17 (70.8%)	13 (54.2%)	2/24	6/24	0.458	0.647	0.714	0.846

In Figure 2 we suggest an algorithm for the management of FUO patients. Presumably, this algorithm will modify the term classic FUO as we know or reduce the number of patients with this diagnosis.

Initially, all patients presenting with fever of unknown origin should have an extensive history taken – including family, intoxication and travel history. This should be followed by full physical examination.[12-13] Particular attention should be given to exposure to animals, work environment and recent contact with persons exhibiting similar symptoms.[7,8,10,11] If any diagnostic clues are found, appropriate tests should be conducted. We recommend extensive blood and urine investigation, as shown in Figure 2, including ANA/ANCA ; Bleeker-Rovers et al. [2] reported that these titers were valuable in several FUO patients for the detection of SLE. Every patient should have a HIV, EBV and CMV test, plus extra serology based on local epidemiology.[11,12] We also recommend a tuberculin skin test, as tuberculosis was a common cause of FUO in the studies we reviewed.[7,9,10] It seems advisable to conduct an ECG, chest radiography and/or abdominal ultrasound. Furthermore, clinicians should consider a Doppler-ultrasonography in patients with a high risk of deep vein thrombosis. [14] Insufficient research has been conducted to determine exactly which tests and imaging are necessary at this stage. However, the recommended tests are simple, noninvasive and inexpensive, and can provide many diagnostic clues. If no diagnostic clues are found despite these tests, drug fever, habitual hyperthermia and factitious fever must be excluded.[6,8,11] Duke criteria must be applied, as FDG-PET/CT has a low sensitivity for endocarditis and endocarditis is a frequent cause of prolonged fever. [5,14] In patients with no diagnosis at this stage, we advise an FDG-PET/CT. Patients with a negative FDG-PET/CT should be followed up. Temporal artery biopsy should be considered in patients over 50 years, due to severe complications, aspecific presentation and frequent false-negative FDG-PET/CT[8,10]. In patients with a deteriorating condition, a therapeutic trial is worth considering. With the exception of temporal artery biopsy, it is not advisable to conduct specific biopsies such as liver biopsy or bone marrow biopsy or culture [14] unless there are diagnostic clues.

Our algorithm is mostly in agreement with those suggested by Bleeker-Rovers et al.[2] and Beccara-Nakayo et al.[1]. Beccara-Nakayo et al. also suggested the use of tumor markers, but the value of these is unknown. They also suggest testing for rare pathogens if PET/CT is negative but fever persists. However, the sensitivity of FDG-PET/CT for infections is high, and we found no evidence for further testing of rare infections in our review. Bleeker-Rovers et al.[2] suggests bone marrow biopsy if FDG-PET/CT is negative.

Limitations

Despite the promising results, drawing a definitive conclusion about the diagnostic value of FDG-PET/CT remains difficult due to the heterogeneity of the articles. This heterogeneity is caused by the variety of definitions for FUO, small populations with a wide range of ages and duration of follow-up (ranging from 0 to 43.2 months). Moreover, there was much diversity in pre-exams carried out before FDG-PET/CT, and some articles were not transparent in the pre-exams they applied before conducting FDG-PET/CT.[6,7,11] We think this is due to retrospectively achieved information and because centers do not adhere to a standard FUO protocol.

Conclusion

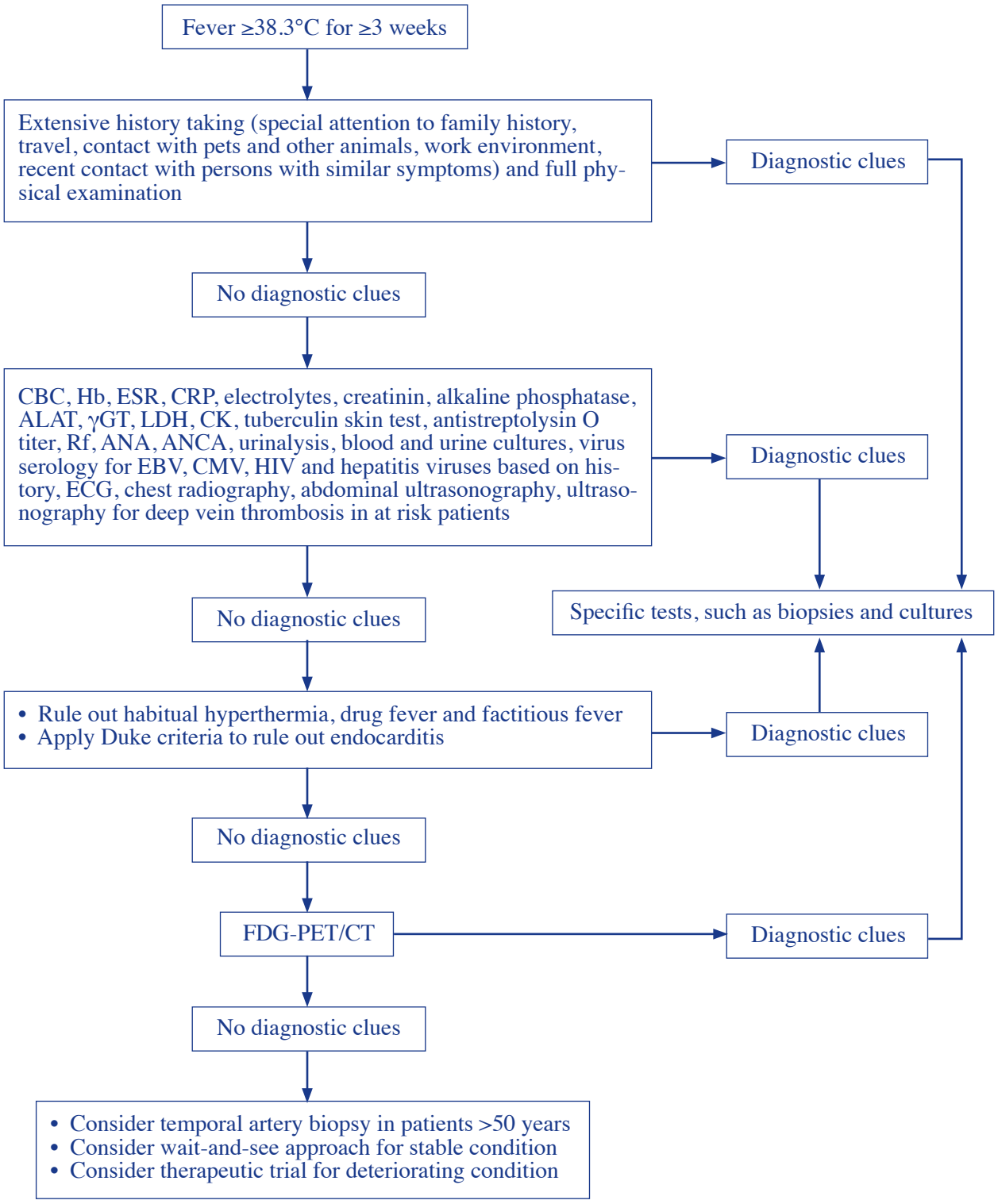
In conclusion , the gold standard for FUO diagnosis remains difficult to identify because of the wide range of causes. However, FDG-PET/CT could be used at an early stage in FUO work-up. In the case of FUO, it seems advisable to perform an FDG-PET/CT instead of a CT scan alone. The radiation dose is marginally higher [5] but the diagnostic value is much greater. Also, in patients with conditions such as lymphoma, an FDG-PET/CT scan can also be used for staging the disease[6]. The heterogeneity of the available studies limits the diagnostic value of FDG-PET/CT. Also, the correct FUO protocol is unclear.

We recommend a prospective randomized controlled trial with a larger sample, with one group managed according to our algorithm and another control group without FDG-PET/CT in the first or second line investigation. If this study confirms that FDG-PET/CT is suitable for early detection and is cost-effective, then integration of FDG-PET/CT in FUO work-up will just be a matter of time.

References

1. *Becerra Nakayo EM et al. Analysis of cost effectiveness in the diagnosis of fever of unknown origin and the role of (18)F-FDG PET-CT: a proposal of diagnostic algorithm. Rev Esp Med Nucl Imagen Mol. 2012; 31: 178-186.*
2. *Bleeker-Rovers CP et al. A prospective multicenter study on fever of unknown origin: the yield of a structured diagnostic protocol. Medicine(Baltimore). 2007; 86: 26-38.*
3. *Petersdorf RG, Beeson PB. Fever of unexplained origin: report on 100 cases. Medicine (Baltimore). 1961; 40:1-30.*
4. *Petersdorf RG. Fever of unknown origin. An old friend revisited. Arch Intern Med. 1992 ; 152: 21-22.*
5. *Gotthardt M et al. Imaging of inflammation by PET, conventional scintigraphy, and other imaging techniques. J Nucl Med. 2010; 51: 1937-1949.*
6. *Kei PL et al. [18F] FDG PET/CT in patients with fever of unknown origin: a local experience. Nucl Med Commun. 2010; 31: 788-792.*
7. *Seshadri N et al. Superiority of 18F-FDG PET compared to 111ln-labelled leucocyte scintigraphy in the evaluation of fever of unknown origin. J Infect. 2012; 65 : 71-79.*
8. *Balink H, Collins J, Bruyn GA, Gemmel F. F-18 FDG PET/CT in the diagnosis of fever of unknown origin. Clin Nucl Med. 2009; 34: 862-868.*
9. *Sheng JF et al. Diagnostic value of fluorine-18 fluorodeoxyglucose positron emission tomography/computed tomography in patients with fever of unknown origin. Eur J Interen Med. 2011; 22 :112-116.*
10. *Federici L et al. Value of (18)F-FDG-PET/CT in patients with feverof unknown origin and unexplained pro-longed inflammatory syndrome: a single centre analysis experience. Int J Clin Pract. 2010; 64: 55-60.*
11. *Keidar Z et al. Fever of unknown origin: the role of 18F-FDG PET/CT. J Nucl Med. 2008; 49: 1980-1986.*
12. *Pedersen TI et al. Fever of unknown origin: a retrospective study of 52 cases with evaluation of the diagnostic utility of FDG-PET/CT. Scand J Infect Dis. 2012; 44:1 8-23.*
13. *Pelosi E et al. Role of integrated PET/CT with [^{18}F]-FDG in the management of patients with fever of unknown origin: a single-centre experience. Radiol Med. 2011; 116: 809-820.*
14. *Mourad O, Palda V, Detsky A.S. A comprehensive evidence-based approach to fever of unkown origin. Arch Intern Med. 2003; 163: 545-551.*
15. *Quinn MJ et al. RR Computed tomography of the abdomen in evaluation of patients with fever of unknown origin. Radiology. 1980;136 : 407- 411.*
16. *Mitchell DPHanes TEHoyumpa AMSchenker S Fever of unknown origin: assessment of the value of percutaneous liver biopsy. Arch Intern Med. 1977;137 : 1001- 1004.*
17. *Holtz TMoseley RHScheiman JM Liver biopsy in fever of unknown origin: a reappraisal. J Clin Gastroenterol. 1993; 1729- 1732.*
18. *Volk EMiller MLKirskley BAWashington JA The diagnostic usefulness of bone marrow cultures in patients with fever of unknown origin. Am J Clin Pathol. 1998;110150- 110153.*
19. *Riley UCrawford SBarrett SPAbdalla SH Detection of mycobacteria in bone marrow biopsy specimens taken to investigate pyrexia of unknown origin. J Clin Pathol.1995; 48706- 48709.*
20. *Crouzet J et al. Place of (18)F-FDG-PET with computed tomography in the diagnostic algorithm of patients with fever of unknown origin. Eur J Clin Microbiol Infect Dis. 2012 ;31: 1727-1733*
21. *Dong MJ et al. A meta-analysis of the value of fluorodeoxyglucose- PET-CT in the evaluation of fever of unknown origin. Eur J Radiol. 2011 ;80: 834-844.*
22. *Roth AR, Basello GM. Approach to the adult patient with fever of unknown origin. Am Fam Physician. 2003; 68: 2223-2229*

Figure 2 - Algorithm for management of patients with FUO



Midline surface area of the corpus callosum in children with attention-deficit/hyperactivity disorder *A systematic review*

Reinier van Linschoten[a], Dean Clydesdale[a], Bram Dierckx[b], Sabine E. Mous[b,c]
[a]Medical Student, Erasmus MC University Medical Center Rotterdam, the Netherlands
[b]Child and Adolescent Psychiatry/Psychology, Erasmus MC University Medical Center Rotterdam, the Netherlands
[c]Supervisor, Generation R Study Group, Erasmus MC University Medical Center Rotterdam, the Netherlands
Correspondence: Reinier van Linschoten, e-mail: 353574rl@eur.nl

Summary

Objective: **To address the question of whether the midline surface area of the corpus callosum differs between children with Attention-Deficit/Hyperactivity Disorder (ADHD) and healthy controls.**
Method: **We systematically reviewed the existing literature that compared the midline surface area of the corpus callosum of children with ADHD with non-ADHD children.**
Results: **Seventeen articles were found after the initial search. Eight remained after the first selection round and eventually six were included to be used in this systematic review. Four of these articles stated that some areas of the corpus callosum were significantly smaller in the children with ADHD. There was no area of the corpus callosum that was found to be significantly smaller in more than 3 studies.**
Conclusions: **The studies reviewed in this article suggest that an association between the presence of ADHD and a smaller midline surface area of the corpus callosum exists. Further research is necessary to determine whether specific areas of the corpus callosum are involved more prominently or whether it involves the entire corpus callosum.**

Introduction

Attention-Deficit/Hyperactivity Disorder (ADHD) is one of the most common disorders of childhood. The prevalence of ADHD is 3-7%.[1] The defining symptoms of ADHD include high levels of activity, impulsivity and inattention. More boys are diagnosed with ADHD than girls.[2]

Abnormalities in the brain may result in ADHD symptomatology.[3] One of these brains areas that could be associated with the presence of ADHD is the corpus callosum. The corpus callosum is a bundle of nerve fibers that connects the left and right hemispheres of the brain. Previous research has shown that patients with a partly or entirely missing corpus callosum experience attention problems.[4] Since inattention is one of the core symptoms of ADHD, this leads to the hypothesis that there is an association between corpus callosum morphology and the presence of ADHD.
Research[5] has also demonstrated a positive correlation between the corpus callosum area and the number of fibers crossing the corpus callosum. This means that the callosal area can be used to estimate the total number of small nerve fibers connecting the two hemispheres and thereby provide information about connectivity. If this connectivity appears to be abnormal, this could explain certain impairments that are present in ADHD.
By linking ADHD with a difference in the corpus callosum, a better understanding of the disease could be achieved. If a significant difference exists in the surface area of the corpus callosum between ADHD patients and those without then it may be used in diagnosis of the disease.

No systematic review has yet been performed on the relationship between corpus callosum area and the presence of ADHD in children, so no consensus about this association has been reached. We therefore decided to conduct a systematic review to address the following research question: Is there an association between the presence of ADHD in children and the midline surface area of the corpus callosum?

Methods

Search methods:
Electronic Search:
We systematically searched Pubmed on January 9th 2013. The following Mesh terms were used:
"Corpus Callosum"[Majr]
AND "Attention Deficit Disorder with Hyperactivity"[Mesh]
AND ("Child"[Mesh] OR "Infant"[Mesh] OR "Adolescent"[Mesh])
AND "Magnetic Resonance Imaging"[Mesh]
AND English[lang]

No further limits were used.

The inclusion and exclusion criteria of this review were based on four aspects of the studies: types of studies, types of participants, types of assessment and outcome measures.

Types of studies:
Only studies that reported the midline surface area of the corpus callosum of an ADHD patient group were included. No limit was placed on the date of publication.

Types of participants:
Studies that included patients with severe co-morbidities were excluded, with exception of Oppositional Defiant Disorder (ODD) or Conduct Disorder (CD). The patients had to have a clinical diagnosis of ADHD. This could be any of the three types (Inattentive, Hyperactive, Combined). No restrictions were made on sex.

Types of assessment:
All studies had to involve structural Magnetic Resonance Imaging (MRI) to determine the midline surface area of the corpus callosum. Diffusion-tensor imaging (DTI) and other imaging techniques were excluded in order to maximize the comparability between the results of the studies in this review.

Types of outcome measures:
The primary outcome measure was the midline surface area of the corpus callosum.

Data collection:
A further selection was made based on inclusion and exclusion criteria. In the first round both authors read each abstract and in case of doubt the abstract would be discussed until agreement was reached. In the second round the same criteria were used to include studies based on their full text.

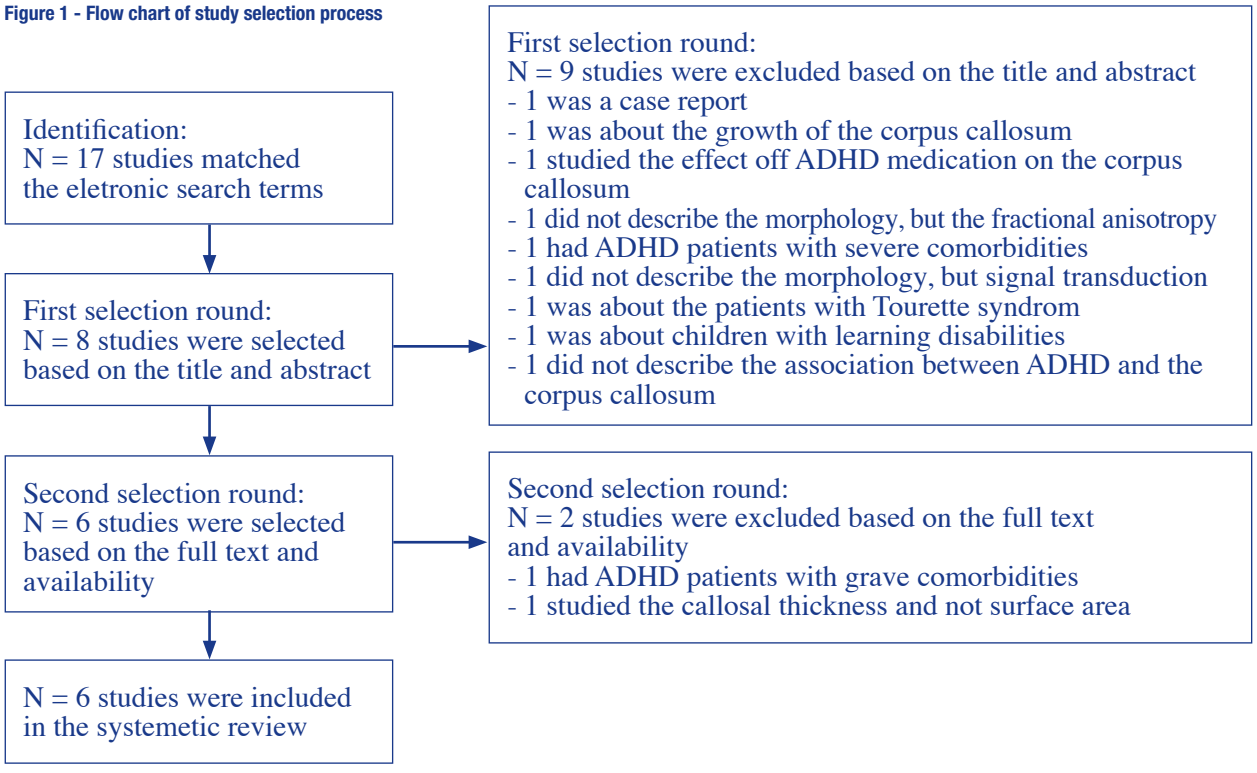
Data extraction:
From each study we extracted the following data:
1. Description, authors, publication year.
2. Number of participants and their characteristics, including type of ADHD, sex and presence of ODD/CD.
3. Results of MRI analyses and p-values corresponding to the association between the midline surface area of the corpus callosum from a midsagittal view and the clinical diagnosis of ADHD.

Results

Description of included studies:
The PubMed search yielded 17 articles. After the first selection round 8 of those were potentially relevant. After the second round 6 articles remained to be used in the analysis of this review (Figure 1). The 6 studies included 216 patients with 187 controls, who were matched at least for age and sex. All studies used their own control group. All patients had an IQ score above 80. Only two studies included girls in their patient group.[6,7] Some studies[6,8-10] included only right-handed patients and controls, because of the difference in brain differentiation between right- and left-handed people.[11] The control group of Overmeyer et al.[12] consisted of siblings of children with ADHD and of siblings of children with symptoms of ADHD.

The procedures used to diagnose ADHD differed between the studies. In the study of Lyoo et al.[5] two different procedures were used to determine the diagnosis of ADHD. The procedure with which the first group of patients was diagnosed used the Diagnostic Interview Schedule for Children (DISC) assessment and the second group was diagnosed by using chart reviews. Both procedures are used to clinically diagnose ADHD, so both patient groups were included. The other studies used rating scales, parent interviews, psychiatric interviews and the Clinical Diagnostic Interviewing Scales. Some studies[6,9,10] used the Diagnostic and Statistic Manual of Mental Disorders-III-revised (DSM-III-r) while others[7,8,12] used DSM-IV to diagnose ADHD (Table 1). Concordance between ADHD diagnoses based on DSM-III-r versus DSM-IV is very high [13], so we included studies that used either DSM-III-r or DSM-IV.

Figure 1 - Flow chart of study selection process



Risk of bias assessment

Potential confounders

Some of the patients used stimulant medication, whereas controls did not. This could potentially confound the relationship between corpus callosum area and ADHD. However, Castellanos et al.[14] showed that ADHD medication had no effect on the structure of the brain. Age and sex are possible confounders as well, since they are both related to the size of the brain and the prevalence of ADHD. However, all studies matched patients and controls for age and sex.

Publication bias

There were not enough studies or large enough studies to rule out publication bias. It is still possible that non-significant studies were not published.

Selection bias

The study by Overmeyer et al.[12] did not use a healthy control group. Instead they used siblings of children with ADHD, which might have influenced their results. In addition, most of the studies included patients with ODD[6-9,12] and CD.[6,8,9,12] It is not clear whether ODD and CD are related to the size of the corpus callosum, but if they are, their inclusion may have biased the results.

MRI data collection

All but one of the studies used a 1.5 Tesla MRI scanner from Siemens, General Electronics or Philips. Cao et al.[8] is the only study that used a 3 Tesla scanner. Because of the higher field strength, the image had an higher resolution. However, every study used a control group which was scanned with the same scanner, so the results were still comparable.

MRI data processing

Not all studies reported how they processed image data. Overmeyer et al.[12] used ANALYZETM software to process the data. McNally et al.[7] processed their data using a Medical Image Processing, Analysis, and Visualization program. These where the only authors that specified their imaging processing steps.

Primary outcomes

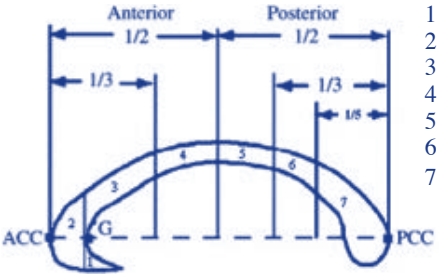
Every study used the midline surface area of the corpus callosum, or one of its anatomical areas, as its primary outcome measure. Five out of six studies used the areas as specified in Witelson et al.[11] (Figure 2) while only McNally et al.[7] choose to divide the corpus callosum as specified by Peterson et al.[15] (Figure 3). All studies calculated the area of the corpus callosum at the midsagittal plane.

McNally et al.[7] find no significant difference in the midline surface area in any of the anatomical areas of the corpus callosum. Neither did they find a significant difference in the midline surface area of the whole corpus callosum. Cao et al.[8] and Giedd et al.[9] found that the total midline surface area of the corpus callosum differed significantly between the groups. Both Lyoo et al.[6] and Semrud-Clikeman et al.[10] reported a significantly smaller splenium in the ADHD patients. Lyoo et al.[6] reported a significant smaller isthmus, but this was only present in the ADHD patients diagnosed according to the DISC. Both the anterior mid body and the isthmus were reported to be significantly smaller in the ADHD group of Cao et al.[8] The study of Giedd et al.[9] was the only study to report a significantly smaller rostrum and rostral body.

Discussion

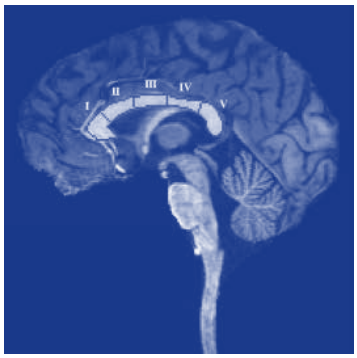
The studies in this review suggest an association between the presence of ADHD and a smaller midline surface area of the corpus callosum in children.

Figure 2 - Anatomical areas, adapted from Witelson et al.



1 Rostrum
2 Genu
3 Rostral Body
4 Anterior midbody
5 Posterior midbody
6 Isthmus
7 Splenium

Figure 3 - Anatomical areas, adapted from Peterson et al.



I - Genu
II - Rostral Body
III - Mid-body
IV - Isthmus
V = Splenium

Four [6,8-10] of the six articles in this review stated that some areas of the corpus callosum were significantly smaller in the children with ADHD. However, the areas that were smaller differed in each article. Two out of the four studies with a significant outcome showed that the splenium had a larger surface area in the ADHD patient group.[6, 10] In both patient groups used in the study by Lyoo et al.[6] the splenium was reported to be significantly smaller.

Two studies reported non-significant results. The study of Overmeyer et al.[12] did not use a healthy control group. Instead, the control group was made of siblings of children with ADHD. It has been reported that genetics explain two-thirds of the increased occurrence of ADHD in families[16], so it could be possible that this control group of siblings shows subclinical ADHD symptoms, which could explain why they did not find significant differences between the patient group and control group.

McNally et al.[7] also reported non-significance. An explanation for this could be that they used the strictest method to diagnose ADHD and the only comorbidity they allowed to be present was ODD.

The discrepancy in findings between the different studies was probably due to differences in methodology between the studies, such as the method used to diagnose ADHD. For example, the patient group of Semrud-Clikeman et al.[10] and Lyoo et al.[6] were not personally interviewed by the researchers. Instead they reviewed the data (e.g. summary notes, consultant reports) from the patients. They did not describe exactly how these patients were diagnosed. In all other studies the patients were interviewed according to a rating scale and/or an interview questionnaire. Overmeyer et al.[12] and McNally et al.[7] were the only ones who also interviewed the parents of the patients. By doing so their diagnosis might be more accurate.

Table 1 - Characteristics of included studies

Study	Diagnosis Method	Sample size/ Controls	DSM version	Significant for	P values	Scale
McNally et al (2009)	Rating scales and Parent interview	64/64	DSM-IV	None	$p > 0.05$	Peterson et al.
Cao et al. (2009)	CDIS*	28/27	DSM-IV	Anterior Mid Body	$p < 0.004$	Witelson et al
				Isthmus	$p < 0.001$	
				Corpus Callosum	$p < 0.008$	
Overmeyer et al. (2000)	Parent and psychiatric interview	15/15	DSM-IV	None	$p > 0.05$	Witelson et al
Lyoo et al. (1996)	DISC***	51/28	DSM-III-R	Splenium	$p = 0.041$	Wttelson et al
	Chart reviews****	25/20	DSM-III-R	Ishtmus	$p = 0.039$	
				Splenium	$p = 0.028$	
Giedd et al. (1994)	Rating scales	18/18	DSM-III-R	Rostrum	$p = 0.007$	Witelson et al
				Rostral Body	$p < 0.05$	
				Corpus Callosum	$p < 0.05$	
Semrud-Clikeman et al. (1994)	Data reviews**	15/15	DSM-III-R	Splenium	$p = 0.0265$	Witelson et al

*Clinical Diagnostic Interviewing Scales
**Historical, Behavioral and Psychometric data
***Diagnostic Interview Schedule for Children
****Diagnosis with help of summary notes, progress reports, consultant reports, and laboratory data

An alternative explanation for the difference in the area of the corpus callosum between children with ADHD and controls could be a difference in total brain volume between both groups. If the children in the ADHD group had smaller brains overall, their corpus callosum could have been smaller as well. However, most studies[7,9,12] used total brain volume as a covariate. Cao et al.[8] did not find a difference between the groups for this variable, and the other studies[6, 10] did not mention total brain size in their article. According to Johnson et al.[19], no correlation between total brain size and corpus callosum midline surface area was found in healthy controls. However, this could be different in children with ADHD. Therefore, we consider total brain volume an important potential confounding factor.

Also, normal variation in the size of the brain and the corpus callosum [17,18] does not explain the difference found between the patient and the control groups. Normal variation should be present in both groups equally and can therefore not explain the significant results that were found. Since the difference in callosal area between the groups was found to be significant, the chance that this is caused by the normal variation is very small.

Image quality is another important aspect to consider. Because of movement due to hyperactivity, it is possible that the quality of the MRI scans of the ADHD group was not as good as the control group. To avoid this problem, one study[9] sedated the patients when necessary. Another study [12] limited head movement by using a restraining band. In this study children could also watch a video and the parents were allowed to stay in the room with the child, which might have minimized distress and movement. The other studies[6-8,10] did not mention how they ensured good image quality.

As mentioned before, only 2 studies[6,7] included girls in their patient group. This was probably done because of the difference in corpus callosum area between boys and girls[11] and because ADHD is a predominant male disorder.[2]

Another potential problem is the presence of comorbidities. In this case the presence of ODD or CD, because all other comorbidities were excluded in this review. It is still unclear whether the presence of these disorders could be a possible confounder. A study that assessed the association between the thickness of the corpus callosum and the presence of ADHD (Luders et al.[20]) reported an interesting difference. When they excluded the children with ODD they found that some areas of the corpus callosum where significantly thinner, compared to when these children were included. However, Luders et al.[20] studied the thickness of the corpus callosum instead of the midline surface area. Consequently, we cannot fully compare Luders et al.[20] with the other studies in this review. However, we should not discard their findings.

Finally, because all studies had small sample sizes, a lack of power could be a problem in finding reliable significant differences between the two groups.

Conclusion

The studies reviewed in this article suggest an association between the presence of ADHD and a smaller midline surface area of the corpus callosum. Several studies have shown areas of the corpus callosum that are significantly smaller in children with ADHD. However, there was no general consensus which area of the corpus callosum is smaller in children with ADHD. Further research is necessary to determine whether specific areas of the corpus callosum are involved more prominently or whether the difference in morphology by ADHD involves the entire corpus callosum.

Since the association between ADHD and the size of the corpus callosum is not yet very clear, it cannot be used as an indicator to diagnose ADHD. However, understanding the neurobiology of the disorder does help in understanding why certain deficits or difficulties are present in children with ADHD. This might also serve as a starting point for studying potential treatments of the disorder.

Previous research[5] has shown a positive correlation between corpus callosum area and the number of fibers crossing the corpus callosum. This means that callosal area can be used to estimate the total number of small nerve fibers connecting the two hemispheres. Smaller midline surface area would therefore imply fewer connecting fibers, which would suggest an abnormality in cortical connectivity. This abnormal connectivity could in turn lead to certain impairments that are present in ADHD.

To credibly answer the question if there is an association between ADHD and corpus callosum size, future research should exclude ADHD patients with comorbid ODD and CD. However, this might be difficult, since these disorders are highly comorbid with ADHD.

Additionally, studying the different subtypes of ADHD (Inattentive, Hyperactive and Combined) separately could be very informative.

Apart from interviewing the patient group, the control group should also be interviewed to be completely sure that they are healthy and don't have any ADHD-like symptoms.

Previous research [21, 22] has shown that a multiple-informant approach in assessing childhood psychopathology is a preferred method, since different informants (such as parents and teachers) can provide different information, as well as complementing information from different situations. Information from multiple informants might be complementary in creating a more extensive overview of the child's mental health.

It is possible that different areas of the corpus callosum are involved in the various subtypes of the disorder. Because of this, studies should not only focus on the corpus callosum as a whole, but also study the different parts it is made up off. In addition, bigger sample sizes should be used, which would make the studies more powerful in detecting differences in corpus callosum size between ADHD patients and controls. Finally, studies should pay more attention to how they diagnose children with ADHD. Not only should the children themselves be interviewed, but also their parents and teachers.

Acknowledgments

This study is financially supported through the NWO Brain & Cognition (Hersenen & Cognitie) project number 433-09-228.

References

1. *American A.P.A. Diagnostic and Statistical Manual of Mental Disorders. In: Psychiatric Press: Washington D ed.; 1994;*
2. *Lahey B.B., Applegate B., McBurnett K., et al. DSM-IV field trials for attention deficit hyperactivity disorder in children and adolescents. The American journal of psychiatry. 1994; 151: 1673-1685.*
3. *Seidman L.J., Valera E.M., Makris N. Structural brain imaging of attention-deficit/hyperactivity disorder. Biological psychiatry. 2005; 57: 1263-1272.*
4. *Badaruddin D.H., Andrews G.L., Bolte S., et al. Social and behavioral problems of children with agenesis of the corpus callosum. Child psychiatry and human development. 2007; 38: 287-302.*
5. *Aboitiz F., Scheibel A.B., Fisher R. S., et al. Fiber composition of the human corpus callosum. Brain Research, 1992; 598: 143-153*
6. *Lyoo I.K., Noam G.G., Lee C.K., et al. The corpus callosum and lateral ventricles in children with attention-deficit hyperactivity disorder: a brain magnetic resonance imaging study. Biological psychiatry. 1996; 40: 1060-1063.*
7. *McNally M.A., Crocetti D., Mahone E.M., et al. Corpus callosum segment circumference is associated with response control in children with attention-deficit hyperactivity disorder (ADHD). Journal of child neurology. 2010; 25: 453-462.*
8. *Cao Q., Sun L., Gong G., et al. The macrostructural and microstructural abnormalities of corpus callosum in children with attention deficit/hyperactivity disorder: a combined morphometric and diffusion tensor MRI study. Brain research. 2010; 1310: 172-180.*
9. *Giedd J.N., Castellanos F.X., Casey B.J., et al. Quantitative morphology of the corpus callosum in attention deficit hyperactivity disorder. The American journal of psychiatry. 1994; 151: 665-669.*
10. *Semrud-Clikeman M., Filipek P.A., Biederman J., et al. Attention-deficit hyperactivity disorder: magnetic resonance imaging morphometric analysis of the corpus callosum. Journal of the American Academy of Child and Adolescent Psychiatry. 1994; 33: 875-881.*
11. *Witelson S.F. Hand and sex differences in the isthmus and genu of the human corpus callosum. A postmortem morphological study. Brain : a journal of neurology. 1989; 112 (Pt 3): 799-835.*
12. *Overmeyer S., Simmons A., Santosh J., et al. Corpus callosum may be similar in children with ADHD and siblings of children with ADHD. Developmental medicine and child neurology. 2000; 42: 8-13.*
13. *Biederman J., Faraone S.V., Weber W., et al. Correspondence between DSM-III-R and DSM-IV attention-deficit/hyperactivity disorder. Journal of the American Academy of Child and Adolescent Psychiatry. 1997; 36: 1682-1687.*
14. *Castellanos F.X., Lee P.P., Sharp W., et al. Developmental trajectories of brain volume abnormalities in children and adolescents with attention-deficit/hyperactivity disorder. JAMA : the journal of the American Medical Association. 2002; 288: 1740-1748.*
15. *Peterson B.S., Leckman J.F., Duncan J.S., et al. Corpus callosum morphology from magnetic resonance images in Tourette's syndrome. Psychiatry research. 1994; 55: 85-99.*
16. *Elia J., Sackett J., Turner T., et al. Attention-deficit/hyperactivity disorder genomics: update for clinicians. Current psychiatry reports. 2012; 14: 579-589.*
17. *Allen J.S., Damasio H., Grabowski T.J., Normal Neuroanatomical Variation in the Human Brain: An MRI-Volumetric Study. American Journal of Physical Anthropology. 2002; 118: 341-358.*
18. *Yani Yazgan M., Wexler B.E., Kinsbourne M., Peterson B, et al. Functional Significance of Individual Variations in Callosal Area. Neuropsychologia. 1995; 6: 769-779.*
19. *Johnson S.C., Farnworth T., Pinkston J.B., Bigler E.D., et al. Corpus Callosum Surface Area Across the Human Adult Life Span: Effect of Age and Gender. Brain Research Bulletin. 1994; 35: 373-377.*
20. *Luders E., Narr K.L., Hamilton L.S., et al. Decreased callosal thickness in attention-deficit/hyperactivity disorder. Biological psychiatry. 2009; 65: 84-88.*
21. *Achenbach T.M., As Others See Us: Clinical and Research Implications of Cross-Informant Correlations for Psychopathology. Current Directions in Psychological Science. 2006; 15: 94-99.*
22. *De Los Reyes A. Introduction to the Special Section: More Than Measurement Error: Discovering Meaning Behind Informant Discrepancies in Clinical Assessments of Children and Adolescents. Journal of Clinical Child & Adolescent Psychology. 2011; 40: 1-9.*

Does (neo)adjuvant or perioperative chemotherapy improve disease-free survival and overall survival in patients with resectable colorectal liver metastases?

Koen Hussaarts[a], Kimberly van der Willik[a], Kees Verhoef[b]
[a] Medical student, Erasmus MC University Medical Center Rotterdam, the Netherlands
[b] Supervisor, Department of Surgical Oncology, Erasmus University MC, Daniel den Hoed Cancer Center, Rotterdam, The Netherlands
Correspondence: Kimberly van der Willik, e-mail: 340175kw@student.eur.nl

Abstract

Objective: **Colorectal cancer is the second most common cause of cancer related death in Europe. More than 50% of the patients will develop liver metastases. We wanted to determine if recent chemotherapy regimens (since 2006) in adjuvant, neoadjuvant or both settings increases disease-free survival (DFS) and overall survival (OS) in patients with resectable colorectal liver metastasis compared to surgery only.**

Methods: **We used Pubmed to find articles for this review. We only included studies after 2006, which used 5-FU, leucoverin and oxaliplatin (FOLFOX) or 5-FU, leucoverin and irinotecan (FOLFIRI) as chemotherapy. Liver metastases had to be surgically removed. Studies reported DFS and OS as primary or secondary endpoint.**

Results: **We found seven articles on neoadjuvant and adjuvant chemotherapy and two articles on perioperative chemotherapy. Neoadjuvant chemotherapy did not significantly improve 3-year DFS and OS. Adjuvant chemotherapy showed higher 3-year DFS and OS compared to surgery alone. One study showed that perioperative chemotherapy improves 3-year DFS and 5-year OS, but not significantly more than adjuvant chemotherapy.**

Conclusions: **Some of our studies showed an improved 3-year DFS and OS in the adjuvant chemotherapy setting compared to surgery alone. However, these were all retrospective studies, which may have caused a bias in our results. Therefore we do not recommend using (neo)adjuvant or perioperative chemotherapy in the treatment of colorectal liver metastases (CRLM).**

Introduction

Colorectal cancer is the second most common cause of cancer-related death in Europe.[1] In 2010 nearly 60,000 cases of colorectal cancer were reported in the Netherlands, and more than 12,000 people died from this disease.[2] Because of the ageing of the European population, the incidence of colorectal cancer is likely to increase.[3]

Patients with colorectal cancer mostly die due to metastases. More than 50% of the patients develop liver metastases.[4] Without treatment, the prognosis of patients with colorectal liver metastases (CRLM) is poor.[5;6] Importantly, the presence of distant metastases from colorectal cancer does not exclude curative treatment. In case of CRLM only, resection of CRLM is the best therapeutic option.[2;7-9] However, only 10-30% of the patients with CRLM have resectable disease at diagnosis.[10]

The 5-year overall survival ranges from 23% to 48% after resection with curative intent.[11] After complete resection without adjuvant therapy, 65-70% of the patients develop disease relapse.[12] Colorectal intrahepatic recurrences can be resected by a repeated hepatectomy, with similar overall survivals which were obtained after the first resection.[13;14]

Since 2000, effective chemotherapies such as oxaliplatin and irinotecan, have been developed for colorectal cancer.

Frequently used chemotherapy combinations are FOLFOX (5-FU, leucoverin and oxaliplatin) or FOLFIRI (5-FU, leucoverin and irinotecan). However, it is still unknown which therapeutic strategy is the best for metastatic colorectal cancer.

Because of the high percentage of disease relapse and accompanied healtconsequences, reducing this is an important health issue. Neoadjuvant, adjuvant and/or perioperative chemotherapy might help to reduce the high rate of relapse. in the Netherlands, however, these are not standard treatments after resection of CRLM. Various studies have produced conflicting results about the effect of chemotherapy in addition to the resection of CRLM and had low numbers of patients.[15]

In our literature review, we therefore addressed the following research question. Does neoadjuvant, adjuvant and/or perioperative chemotherapy in combination with resection of CLRM improve disease free survival (DFS) and/or overall survival (OS)?

Methods

We used Pubmed search for articles for this review. Due to the fairly recent development of effective chemotherapy for colorectal cancer (since around 2000), we included only studies published after 2006, that used FOLFOX or FOLFIRI as neoadjuvant, adjuvant or perioperative chemotherapy.

Our other inclusion criteria were the following:
- CRLM had to be surgically removed.
- Studies measured DFS and OS as primary or secondary endpoint.
- Studies must have measured DFS and OS for at least three years.
- Published in English, with full text articles available at the Erasmus MC

We used the MeSH terms 'disease-free survival', 'survival rate' or 'survival', 'colorectal neoplasms' 'neoplasm metastasis', 'surgical procedures, operative' and 'drug therapy'.
We excluded studies that used chemoradiation, hepatic arterial infusion or hyperthermic isolated hepatic infusion as therapeutic strategy.

Results

In total we included 14 different studies. Table 1 presents an overview of all included study characteristics. It gives an overview of the median age, number of metastases. These characteristics about the population that recieved these therapeutic regimens may have influenced therapy outcomes.

Neoadjuvant chemotherapy

In the last few years neoadjuvant strategies are used more common.[16] There are two reasons for using preoperative chemotherapy. Induction chemotherapy is given when the CRLM are initially considered unresectable. This type of chemotherapy could downsize the CRLM, after which resection could be considered.[17] This results in 10-30% more potentially resectable patients, who were initially unresectable.[18] Second, neoadjuvant chemotherapy is preoperatively given to patients with initially resectable disease.[19] This could increase the percentage of R0 resections, limit the extent of the hepatectomy and eliminate micro metastases.[20] Because of the latter advantage of neoadjuvant chemotherapy, we also wanted to examine whether this type of chemotherapy improves the DFS and OS. However, not all included studies distinguished the use of induction and neoadjuvant therapy.

Study results

The main results of the individual studies are presented in Table 2. All seven studies measured the DFS and OS. Most studies used FOLFOX or FOLFORI as neoadjuvant chemotherapy. Five studies investigated retrospectively the effect of neoadjuvant chemotherapy on DFS and OS.[21-25] The other two studies used a prospective design.[26;27] The total patient number of all included studies was n = 1693.

Disease-free survival

The median DFS after neoadjuvant chemotherapy and CRLM resection varied from 6.9 to 24.7 months. Of all studies that measured the DFS in percentage, the lowest 3-year DFS percentage was 27% and, the highest was 50%.

Neoadjuvant chemotherapy plus resection versus resection only
Three studies investigated the difference in DFS and OS after neoadjuvant chemotherapy plus CRLM resection compared to resection alone.[21-23] Reddy et al. found a higher 1-year DFS in the neoadjuvant group, but after three years, the DFS in the no chemotherapy group was higher compared to the neoadjuvant chemotherapy group.[21] The other two studies found also a higher 3-year DFS in the patient group without neoadjuvant chemotherapy.[22;23] This finding was only significant in the study of Pinto et al.

Table 1 - Study characteristics

	Mean/ median age	Number of SCRLM	Gender (male)	Size of largest SCRLM (cm)	Pre-hepatectomy CEA (ng/ml)	Major hepatic resection (>3 segments)	Node-positive primary cancer	T3/T4 primary cancer
Reddy et al. (2009)	57 (49-66)	2 (1-3)	58.9%	3 (2-5)	6.8 (2.5–27.3)	51.9%	61.5%	84.2%
Pinto et al. (2011)	60.31 (22-89)	2.23 (1-15)	65.7%	3.3 (0.3–14.5)	45.4 (0.3–1800)	40.9%	54.8%	78.9%
Lubezky et al. (2009)	66	1.47 (1-3)	85.7%	3.4			75%	
Falcone et al. (2007)	64 (21-75)		69%		>100 (31%)			
Gruenberger et al. (2008)	62 (36-77)		68%		>20 (30%)		68%	
Ayez et al. (2011)	63 (30-86)	2 (1-10)	62%	3.5 (0.5-18)			58%	83%
Small et al. (2009)		> 2 (72%)	50%				67%	83%
Chan et al. (2011)	63 (29-88)	>1 (44.4%)	66.9%	>5 (18.7%)	12.4 (0.8-4280)	24.5%		
Ychou et al. (2009)	63 (27-75)	> 2 (36%)	58.8%			28.8%		
Liu et al. (2010)	<70 (52%)	>1 (42%)	68%	> 3 (40%)	>50 (22%)			
Kim et al. (2009)	<70 (83.9%)	>1 (49.4%)	65.4%		>5 (57.7%)		79.4%	92.5%
Kim et al. (2011)	55 (31-73)	>1 (38.3%)	61%	>3 (26.7%)	>5 (55%)			96.7%
Nordlinger et al. (2008)	63 (25-79)	> 1 (48%)	66%		>5 (63%)		58%	83%

Table 2 - Results of individual studies on the effect of neoadjuvant chemotherapy

Study	Number of patients	Treatment arms	3-year disease free survival (percentage)	Median disease free survival (months)	3- or 5-year overall survival (percentage))	Median overall survival (months)
Reddy et al.	412	Neoadjuvant v	27% v 37% v		67% v 77% v 80%	53 (95% CI 46–60) v
(2009)		adjuvant v peri-	30% v 35%**		v 47%	76 (95% CI 47–104) v
		operative v CRLM	p = 0.85		5-year OS:	67 (95% CI 54–80) v
		resection only			45% v 60% v 55%	36 (95% CI 24–48)
					v 35%	
					p < 0.01	
Pinto et al.	676	Oxaliplatin or	20% v 38%		59% v 71%	
(2011)		irinotecan based	5-year DFS: 13%		5-year OS:	
		chemotherapy v	v 26%		43% v 55%	
		CRLM resection	p < 0.0001		p = 0.009	
		only				
Ayez et al.	251	Oxaliplatin based	36% v 40%	13 (95% CI 9-17) v	53% v	65 (95% CI 44-86) v
(2011)[b]		chemotherapy v		16 (95% CI 12-20)	45%***	48 (95% CI 33-63)
		CRLM resection				
		only				
Lubezky et al.	56	Neoadjuvant v	50% v 49%	14.8 v	70% v 84%	
(2009)		adjuvant FOLFOX	p = ?	14.4[a]	p = ?	
		or FOLFIRI		p = ?		
Falcone et al.	244	Neoadjuvant		9.8 v 6.9***		22.6 v 16.7
(2007)		FOLFOXIRI v neo-		HR = 0.68		HR = 0.70
		adjuvant FOLFIRI		p = 0.0006		p = 0.032
Gruenberger et	50	Neoadjuvant XE-		24.7**		38
al. (2008)		LOX or FOLFOX4		p = 0.002		p = 0.017
Small et al.	54	Neoadjuvant		12.5[a]	39%**	20[a]
(2009)		FOLFOX or FOLFIRI		p = ?	p = ?	p = ?

* DFS mentioned as event-free survival, ** DFS mentioned as recurrence-free survival,
*** DFS mentioned as progression-free survival, [a]Mean time instead of median,
[b]Results based on Nordlinger clinical record form

Neoadjuvant chemotherapy plus resection versus adjuvant chemotherapy plus resection
Two studies compared neoadjuvant to adjuvant chemotherapy in combination with resection of CRLM.[21;24] Lubezky et al. found an event-free survival benefit in the first year for the neoadjuvant group comparing to the adjuvant group.[24] This benefit disappeared in the second and third years. Reddy et al. found no significant difference in DFS between the neoadjuvant and adjuvant group.[21]

Comparing different types of neoadjuvant chemotherapy plus resection
One study compared different types of neoadjuvant chemotherapy.[26] Falcone et al. showed in a randomized trial that neoadjuvant FOLFOXIRI improves DFS compared to neoadjuvant FOLFIRI.

Neoadjuvant chemotherapy plus resection observation
The last two studies examined the effect of neoadjuvant chemotherapy following CRLM resection on the risk of recurrence and outcome.[25;27] Gruenberger et al. prospectively investigated the effect of neoadjuvant XELOX or FOLFOX.[27] They found a median recurrence free survival of 24.7 months in responding patients. Small et al. presented data from patients who had received FOLFOX or FOLFORI before CRLM resection.[29] The mean recurrence free survival was 12.5 months.

Overall survival
The median OS for all studies ranged from 20 to 65 months. The lowest 3-year OS was 38%, the highest 70%. Two studies measured the 5- year OS, which was 45% and 53%.[21;23]

Neoadjuvant chemotherapy plus resection versus resection only
Two of three studies in this group found no difference in OS between neoadjuvant chemotherapy and resection only.[21;23] Pinto et al. found an improved 3-year and 5-year OS in the neoadjuvant setting.[22]

Neoadjuvant chemotherapy plus resection versus adjuvant chemotherapy plus resection
In the study of Lubezky et al., no significant difference was found in OS between the adjuvant and neoadjuvant chemotherapy group.[24] In contrast, Reddy et al. showed a significantly higher OS for the adjuvant chemotherapy group.[21]

Comparing different types of neoadjuvant chemotherapy plus resection
Falcone et al. compared two types of neoadjuvant chemotherapy.[26] This study determined that neoadjuvant FOLFOXIRI significantly improves OS compared to FOLFIRI.

Neoadjuvant chemotherapy plus resection observation
The last two studies in the group (both observational studies) found a median OS of 38 months, and a mean OS of 20 months.[25; 27]

Adjuvant chemotherapy
After complete CRLM resection, it is possible to give patients adjuvant chemotherapy. Adjuvant chemotherapy could kill the occult tumour cells throughout the body and therefore reduce the risk of recurrence and prolong the DFS.[28;29] An improved DFS could predict a prolonged OS. We investigated whether adjuvant chemotherapy has a positive effect on DFS and OS.

Study results
We found seven studies that looked at DFS and OS after resection of CRLM with adjuvant chemotherapy. The main results of the studies are presented in Table 3.
All studies used FOLFOX and/or FOLFIRI as main adjuvant therapy. One study also included a few patients who recieved modern targeted therapies such as bevacizumab or cetuximab besides FOLFOX/FOLFIRI.[21] Another study also gave cetuximab and bevacizumab, but the main therapy remained FOLFOX/FOLFIRI.[30] A minority of patients recieved cetuximab or bevacizumab.
All seven studies looked retrospectively at the effect of adjuvant chemotherapy on DFS and OS.[21;24;30-34] One of these studies was a randomized controlled trial.[31]
The seven studies together had total a patient population of n=625. To compare the studies we looked at 3-year DFS and 3-year OS; all the studies included this measurement in their results.

Disease-free survival
The lowest 3-year DFS was 23% and the highest 3-year DFS of 50.8%. The median DFS varied from 14.1 months to 34.3 months.

Adjuvant chemotherapy plus resection versus resection only
Reddy et al. looked at patients with synchronous resectable liver metastases.[21] In patients with synchronous disease there are metastases at the time of diagnosis. This study shows that post-hepatectomy chemotherapy is associated with prolonged DFS: a 3-year DFS of 37% and a median DFS of 24 months. This study looked at neoadjuvant, perioperative and adjuvant chemotherapy compared to surgical treatment alone. They found a survival benefit of 39 months versus 99 months (95% CI: 65-144 months) for the adjuvant therapy group, although this was not significant.

Adjuvant chemotherapy plus resection versus neoadjuvant chemotherapy plus resection
In the study of Lubezky et al, 19 patients received adjuvant chemotherapy and 21 patients received neoadjuvant chemotherapy.[24] Most of the patients received FOLFIRI as adjuvant treatment (83%). The median DFS was 14.4 months and the 3-year DFS was 49%. Adjuvant chemotherapy showed similar results as neoadjuvant chemotherapy after three years.

Comparing different types of adjuvant chemotherapy plus resection
Chan et al. investigated the difference between two periods with different kinds of chemotherapy (with or without oxaliplatin and irinotecan).[30] They found a 3-year and 5-year DFS of 28.2% and 26.2%, respectively. They concluded that current chemotherapy gives better results on DFS compared to the older types of chemotherapy.

Ychou et al. compared adjuvant 5FU/FA with FOLFIRI.[33] No significant difference was found between these two treatments, but FOLFIRI seemed to have a better effect on DFS. The 3-year DFS was 41.6 %.

Two studies compared FOLFOX and FOLFIRI to older adjuvant chemotherapy in patients with CRLM. Liu et al. showed better results on DFS with the new FOLFOX/FOLFIRI regimens.[32] The 3-year DFS was 50.8% with a median DFS of 34.3 months. The other study [33] showed a better effect on DFS when FOLFOX or FOLFIRI was given, with a significantly benefit of FOLFOX over FOLFIRI.
Kim et al. looked at an oxaliplatin based regimen (FOLFOX) in the adjuvant setting.[34] They found improved DFS for surgery plus adjuvant chemotherapy compared to surgery alone. It has a 3-year DFS of 45.7% and a median DFS of 32.8 months.

Overall survival
The 3-year OS found in the studies ranged from 47.9% to 85.7%. The median OS ranged from 36 months to 76 months. Most of the studies showed an advantage of adjuvant chemotherapy compared to surgery alone. Several studies looked at the effect of modern chemotherapy compared to older chemotherapy.[30-33] Two of these studies showed no significant difference between the modern and older chemotherapy regimens.[31;33] Most of the studies showed a positive effect on OS compared to former studies.[21;30-32;35]

Perioperative chemotherapy
Perioperative chemotherapy is given before and after CRLM resection, whereby it is possible to combine benefits of both regimens. The preoperative setting of this type of chemotherapy could downsize the CRLM, which may facilitate resection of these CRLM. Moreover, it is possible to assess the response to the chemotherapy, which determines whether the same chemotherapy should also be given after CRLM resection.[36] It has also been shown that the pathologic response to preoperative chemotherapy can predict the outcome after CRLM resection.[37] The postoperative chemotherapy setting could be effective in eradicating dormant cancer cells in the remnant liver. Perioperative chemotherapy could improve progression free survival.[38]

Study results
We found two studies which investigated the effect of modern perioperative chemotherapy in combination with CRLM resection.[21;38] Both studies measured the 3-year DFS and 5-year OS. Nordlinger et al. used a prospective design, while Reddy et al. assessed retrospectively the influence of perioperative chemotherapy in combination with CRLM resection. The total number of patients in the included studies is n=776. The major results of these two studies are reported in table 4.

Disease-free survival
The median DFS after perioperative chemotherapy in combination with CRLM resection was 18.7 months in the study of Nordlinger et al.[38] This study randomized patients to either six cycles of FOLFOX before surgery and six cycles after surgery or to surgery alone. The intention-to-treat analyses did not show a significant difference in 3-year DFS between the two treatment arms. However, there was a significant improvement of 3-year DFS in eligible patients (42.4% versus 33.2%).
The study of Reddy et al. found a 3-year DFS of 30% in perioperative patients, which is not significantly different from patients who received no chemotherapy.[21]

Table 3 - Results of individual studies on the effect of adjuvant chemotherapy

Study	Number of patients	Treatment arms	3-year disease free survival	Median disease free survival	3- year overall survival (percentage))	Median overall survival (months)
Reddy et al.	412	Neoadjuvant v	27% v 37% v 30%		67% v 77% v 80%	53 (95% CI 46–60) v
(2009)		adjuvant v peri-	v 35%**		v 47%	76 (95% CI 47–104) v
		operative v CRLM	p = 0.85		5-year OS:	67 (95% CI 54–80) v 36
		resection only			45% v 60% v 55%	(95% CI 24–48)
					v 35%	
					p < 0.01	
Chan et al.	279	Irinotecan or	28.2% v 19.1%*	16.5 v 10.0*	56.5% v 25.2%	44.6 v 19
(2011)		oxaliplatin based	5-year DFS:		5-year OS:	
		chemotherapy v	26.2% v 15.6%*		45.2% v 18.9%	
		5-FU/leucoverin	p = 0.013		p < 0.0001	
Lubezky et al.	56	Neoadjuvant v	50% v 49%	14.8 v 14.4[a]	70% v 84%	
(2009)		adjuvant FOLFOX	p = ?	p = ?	p = ?	
		or FOLFIRI				
Ychou et al.	306	Adjuvant FOLFIRI v	41.6% v 40%	24.7 (95% CI 14.6-	72.7 (95% CI	
(2009)		adjuvant 5FU/FA	p = 0.44	30.4) v	63.2–78.5) v	
				21.6 (95% CI	71.6 (95% CI	
				18.7-38.9)	64.0–79.7)	
Liu et al.	52	Adjuvant FOLFOX	50.8% v 21.1%	34.3 v 14.2	85.7% v 51.8%	>57.7 v 49
(2010)		or FOLFIRI v	p = 0.022	p < 0.05	5-year OS: 54.0%	p < 0.05
		adjuvant 5-FU/			v 34.6%	
		leucoverin			p = 0.028	
Kim et al.	60	Adjuvant FOLFOX	45.7%**	32.8**	68.8%	62.8
(2011)			5-year DFS:	(95% CI 5.8–59.6)	5-year OS: 55.5%	(95% CI 44.1-81.3)
Kim et al.	156	Adjuvant FOLFOX v	39.2%**	23.4 v 14.1 v 16.3	p < 0.05	51.2 v 47.9 v 60
(2009)		adjuvant FOLFIRI v	p < 0.05	p = 0.088		p = 0.219
		adjuvant 5-FU	37.5% v 23% v		78% v 72% v 65%	
			22%			

*DFS mentioned as recurrence-free survival, ** DFS mentioned as relapse-free survival,
[a]Mean time instead of median

Overall survival

The median OS of the two studies varied from 61 to 67 months. The lowest 5-year OS was 51.2%, the highest 55%.
Nordlinger et al. found no significant difference in 5-year OS between the perioperative chemotherapy arm and to the arm which received surgery alone.[38]
Reddy et al. showed that perioperative chemotherapy was associated with improved five year OS with regard to neoadjuvant or no chemotherapy.[21] There was no significant difference in 5-year OS between perioperative and adjuvant chemotherapy.

Discussion

Our results provide evidence that adjuvant chemotherapy in combination with CRLM resection improves 3-year DFS and OS compared to CRLM resection only. We found no conclusive evidence that neoadjuvant or perioperative chemotherapy in combination with resection of CRLM improves DFS or OS.

Results of CRLM resection only in former studies show spread values of DFS and OS.[35;38-40] A meta-analysis of two studies which compared adjuvant 5-FU plus leucoverin after surgical resection of CRLM to resection only, found a median DFS after surgery of 18.8 months.[35] The 2 and 5-year DFS were respectively 40.2% and 27.7%. The median OS for the CRLM resection only arm was 47.3 months. The 3-year and 5-year OS were respectively 71.0% and 39.6%.

Because of the divergent DFS and OS results after CRLM resection only, it is difficult to compare our results to these study results. We found that five out of seven studies which used adjuvant chemotherapy reported higher 3-year OS compared to

Table 4 - Results of individual studies on the effect of perioperative chemotherapy

Study	Number of patients	Treatment arms	3-year disease free survival (percentage)	Median disease free survival (months)	5-year overall survival (percentage))	Median overall survival (months)
Nordlinger et al.	364	FOLFOX v CRLM	35.4% v 28.1*	18.7 v 11.7*	51.2% v 47.8%	61 v 54
(2008)		resection only	HR = 0.79		HR = 0.88	
			p = 0.058		p = 0.34	
Reddy et al.	412	Neoadjuvant v	27% v 37% v 30%		45% v 60% v 55%	53 (95% CI 46–60) v
(2009)		adjuvant v peri-	v 35%**		v 35%	76 (95% CI 47–104) v
		operative v CRLM			p < 0.01	67 (95% CI 54–80) v
		resection only	p = 0.85			36 (95% CI 24–48)

* DFS mentioned as progression-free survival, **DFS mentioned as recurrence-free survival

CRLM resection only, which had a 3-year OS of 71.0% based on the literature.[7;9;10;21;24] The 3-year OS of neoadjuvant chemotherapy was lower compared to the 3-year OS of CRLM resection only.

Based on our results, it appears that adjuvant chemotherapy in combination with CRLM resection improves 3-year DFS and OS compared to CRLM resection only. This positive trend was not seen with neoadjuvant or perioperative chemotherapy.

Most of the retrospective studies collected data from every patient, regardless of the duration and dose of chemotherapy which was given. The patients were treated according to the best judgement of their doctors, without using a standard protocol like those in randomized controlled trials. Therefore some kind of bias is likely due to possible personalized treatment of patients.

Due to the lack of randomized controlled trials in our review, we cannot conclude that adjuvant chemotherapy in combination with CRLM actually improves 3-year DFS and OS compared to CRLM resection only.

The only randomized controlled trial included in our study, compared FOLFOX in the perioperative setting with CRLM resection only and found no significant difference between these two treatment regimens. Despite the evidence from other types of studies, based on this RCT we cannot conclude that modern chemotherapy in the neoadjuvant, adjuvant or perioperative setting improves 3-year DFS or OS compared to CRLM resection only.

A possible reason for this finding is that our included studies did not separate their patients according to prognostic groups to determine if there is a survival benefit with modern chemotherapy. The prognosis of patients with CRLM is determined by several risk factors. Fong et al. defined the main risk factors predicting worse outcome and prognosis as: a positive plane of intersection, extrahepatic disease, lymphnode positive primary, DFS of less than 12 months and number of metastases in liver (>1).[41] Based on these five criteria for determining the clinical risk score (CRS), patients can be stratified into different prognostic groups. Our included studies did not separate their patients based on their Fong CRS. Therefore we could not calculate the Fong CRS in our results, while patients with different scores could have different DFS and OS. For this reason it is advisable to separate the patients based on these prognostic factors. It is possible that some groups will benefit from (neo)adjuvant or perioperative chemotherapy, which could result in an improved 3-year DFS and OS compared to surgery of CRLM alone. We recommend using these Fong prognostic factors for further investigations.

A very recent study only published as an abstract investigated the possible survival benefit of neoadjuvant chemotherapy in resectable CRLM stratified by the Fong CRS. They showed an improved median OS in the high Fong CRS group (3-5 risk factors) with neoadjuvant chemotherapy, compared to CRLM resection only. This improvement was not found in patients with a low clinical risk score (1-2 risk factors). The results of this approach seems promising, so further investigation is needed to determine whether neoadjuvant chemotherapy in combination with CRLM resection should be the standard treatment for patients with a high Fong CRS. Other limitations of the studies in our review are the following:

Most included studies were small, and the long term outcomes, such as 5-year and even 10-year OS rates, were not always reported. This could have been caused by a lack of power.

In the neoadjuvant setting it was not always clear if the study only included patients with resectable CRLM or also patients with initially unresectable CRLM. This could have influenced our interpretation of the results.

The included studies had differing patient inclusion criteria. Some studies only included the patients with the best predicted prognosis.

The timing of chemotherapy differed between studies and patients. Some of our studies found out that you must give it early so you can kill the micro metastases in an early state.[21;31]

The toxicity rates and the quality of life were not properly investigated in the treated patients. Chemotherapy has a major toxic effect on the human body and therefore causes many of adverse effects, which can influence the quality of life. Before a standard treatment can be established the survival benefit must be weighed against the adverse effects of chemotherapy in terms of quality of life. Some of our included studies looked at the side effects and found mild toxicities with modern chemotherapy, suggesting that quality of life will not decrease dramatically.[31;34]

Neoadjuvant chemotherapy can induce hepatic damage and thereby affect the postoperative outcome. One side effect on the normal liver parenchyma of mainly neoadjuvant therapy is chemotherapy-associated steatohepatitis, but CRLM resection can be safely performed in patients with more than 30% steathosis.[42-43] This effect is mostly due to irinotecan. Oxaliplatin can cause vascular changes like hepatic sinusoidal obstruction. This will increase morbidity after surgery and therefore predicts a worse outcome.[44] Prolonged neoadjuvant chemotherapy increases the risk of postoperative hepatotoxicity.[45]

Another disadvantage of neoadjuvant chemotherapy are the invisible residual CRLM. One study showed that in 80% of the CRLM which disappeared on the CT-scan images after neoadjuvant chemotherapy, the cancer was still present.[46] It is therefore difficult for the surgeon to identify the size of the CRLM. Finally, with neoadjuvant chemotherapy the postsurgical liver remnant must larger: least 30% instead of 20% without neoadjuvant chemotherapy.[47;48]

Limitations

We were unable to calculate the mean DFS and OS of all studies due to the lack of provided information by those studies. Furthermore, studies that also reported DFS as an endpoint were rare. Adjuvant chemotherapy did show a significant improvement of 3-year DFS and OS in most studies comparing to older regimens and surgery alone, but these were retrospective studies. This might have caused a certain selection bias in our interpretation of the results.

Conclusion

Overall, we can conclude that modern chemotherapy like FOLFOX or FOLFIRI should not be added to surgery of CRLM as a standard treatment with curative intent for advanced colorectal disease with liver metastases. Neoadjuvant chemotherapy did not show a significant difference in 3-year DFS and OS compared to CRLM resection only. Moreover neoadjuvant chemotherapy often causes hepatotoxicity.

Furthermore, we cannot conclude that neoadjuvant, adjuvant or perioperative chemotherapy deserve a place in the treatment of CRLM. Future studies could divide patients into different prognostic groups to determine if the group with the worst prognosis might have a significant survival benefit from chemotherapy.

More research is needed to determine which type of chemotherapy should be used and in which setting. Especially in the neoadjuvant setting, toxicity needs to be compared between FOLFOXIRI and FOLFIRI therapy.

Finally, the studies in our review that compared the neoadjuvant setting to adjuvant, showed a significant benefit for the adjuvant setting. Therefore, the effect of adjuvant chemotherapy should be examined in patients with high Fong CRS. It is also possible to investigate the effect of perioperative chemotherapy in patients with high Fong CRS, but the added value of perioperative chemotherapy is unknown. Although perioperative chemotherapy was significantly better than CRLM resection in only one retrospective study, there was no difference with adjuvant chemotherapy. This suggests that the benefit of perioperative chemotherapy is mostly due to the adjuvant chemotherapy setting. Moreover, perioperative chemotherapy results in higher costs and morbidity than adjuvant chemotherapy.

References

(1) *Ferlay J, e. Estimates of the cancer incidence and mortality in Europe in 2006. Ann Oncol 2007; 18: 581-92.*
(2) *http://cijfersoverkanker.nl/Dataset_1/img50643f8aaecc7*
(3) *Bray F, et al. Global cancer transitions according to the Human Development Index (2008-2030): a population-based study. Lancet Oncol 2012; 13: 790-801.*
(4) *Leonard GD, Brenner B, Kemeny NE. Neoadjuvant chemotherapy before liver resection for patients with unresectable liver metastases from colorectal carcinoma. J Clin Oncol 2005; 23: 2038-48.*
(5) *Shimada H, et al. Treatment for colorectal liver metastases: a review. Langenbecks Arch Surg 2009; 397: 973-83.*
(6) *Manfredi S, et al. Epidemiology and management of liver metastases from colorectal cancer. Ann Surg 2006; 244: 254-9.*
(7) *http://oncoline.nl/index.php?pagina=/richtlijn/item/pagina.php&id=24786&richtlijn_id=48*
(8) *Pawlik TM, Schulick RD, Choti MA. Expanding criteria for resectability of colorectal liver metastases. Oncologist 2008; 13: 51-64.*
(9) *Jatzko GR, et al. Hepatic resection for metastases from colorectal carcinoma--a survival analysis. Eur J Cancer 1995; 31A: 41-6.*
(10) *Yang S, et al. Utility of preoperative imaging in evaluating colorectal liver metastases declines over time. HPB (Oxford) 2010; 12: 605-9.*
(11) *Scheele J, Altendorf-Hofmann A. Resection of colorectal liver metastases. Langenbecks Arch Surg 1999; 384: 313-27.*
(12) *de Jong MC, et al. Rates and patterns of recurrence following curative intent surgery for colorectal liver metastasis: an international multi-institutional analysis of 1669 patients. Ann Surg 2009; 250: 440-8.*
ecurrent colorectal metastases: 143 cases. Association Francaise de Chirurgie. J Clin Oncol 1994; 12: 1491-6.
(14) *Van der Pool AE, et al. Local treatment for recurrent colorectal hepatic metastases after partial hepatectomy. J Gastrointest Surg 2009; 13: 890-5.*
(15) *http://www.oncoline.nl/colorectale-levermetastasen*
(16) *Chiappa A, et al. The management of colorectal liver metastases: Expanding the role of hepatic resection in the age of multimodal therapy. Crit Rev Oncol Hematol 2009; 72: 65-75.*
(17) *Bismuth H, et al. Resection of nonresectable liver metastases from colorectal cancer after neoadjuvant chemotherapy. Ann Surg 1996; 224: 509-20.*
(18) *Ismaili N. Treatment of colorectal liver metastases. World J Surg Oncol 2011; 9: 154.*
(19) *Folprecht G, et al. Neoadjuvant treatment of unresectable colorectal liver metastases: correlation between tumour response and resection rates. Ann Oncol 2005; 16: 1311-9.*
(20) *Ayez N, et al. Outcome of microscopic incomplete resection (R1) of colorectal liver metastases in era of neoadjuvant chemotherapy. Ann Surg Oncol 2012; 19: 1618-27.*
(21) *Reddy SK, et al. Timing of multimodality therapy for resectable synchronous colorectal liver metastases: a retrospective multi-institutional analysis. Ann Surg Oncol 2009; 16: 1809-19.*
(22) *Pinto MH, et al. Peri-operative chemotherapy for resectable colorectal liver metastasis: does timing of systemic therapy matter? J Surg Oncol 2012; 105: 511-9.*
(23) *Ayez N, et al. Is the clinical risk score for patients with colorectal liver metastases still useable in the era of effective neoadjuvant chemotherapy? Ann Surg Oncol 2011; 18: 2757-63.*
(24) *Lubezky N, et al. Is there a survival benefit to neoadjuvant versus adjuvant chemotherapy, combined with surgery for resectable colorectal liver metastases? World J Surg 2009; 33: 1028-34.*
(25) *Small RM, et al. Response to chemotherapy predicts survival following resection of hepatic colo-rectal metastases in patients treated with neoadjuvant therapy. J Surg Oncol 2009; 99: 93-8.*
(26) *Falcone A, et al. Phase III trial of infusional fluorouracil, leucovorin, oxaliplatin, and irinotecan (FOLFOXIRI) compared with infusional fluorouracil, leucovorin, and irinotecan (FOLFIRI) as first-line treatment for metastatic colorectal cancer: the Gruppo Oncologico Nord Ovest. J Clin Oncol 2007; 25: 1670-6.*
(27) *Gruenberger B, et al. Importance of response to neoadjuvant chemotherapy in potentially curable colorectal cancer liver metastases. BMC Cancer 2008; 8: 120.*
(28) *Chun YS, et al. Management of chemotherapy-associated hepatotoxicity in colorectal liver metastases. Lancet Oncol 2009; 10: 278-86.*
(29) *Wang CC, Li J. An update on chemotherapy of colorectal liver metastases. World J Gastroenterol 2012; 18: 25-33.*
(30) *Chan KM, et al. Outcomes of resection for colorectal cancer hepatic metastases stratified by evolving eras of treatment. World J Surg Oncol 2011; 9: 174.*
(31) *Ychou M, et al. A randomized phase III study comparing adjuvant 5-fluorouracil/folinic acid with FOLFIRI in patients following complete resection of liver metastases from colorectal cancer. Ann Oncol 2009; 20: 1964-70.*
(32) *Liu JH, et al. Adjuvant oxaliplatin- or irinotecan-containing chemotherapy improves overall survival following resection of metachronous colorectal liver metastases. Int J Colorectal Dis 2010; 25: 1243-9.*
(33) *Kim SY, et al. Resected colorectal liver metastases: does the survival differ according to postoperative chemotherapy regimen? J Surg Oncol 2009; 100: 713-8.*
(34) *Kim HR, et al. Efficacy of oxaliplatin-based chemotherapy in curatively resected colorectal cancer with liver metastasis. Oncology 2011; 81: 175-83.*
(35) *Mitry E, et al. Adjuvant chemotherapy after potentially curative resection of metastases from colorectal cancer: a pooled analysis of two randomized trials. J Clin Oncol 2008; 26: 4906-11.*
(36) *Brouquet A, Nordlinger B. Neoadjuvant and adjuvant therapy in relation to surgery for colorectal liver metastases. Scand J Gastroenterol 2012; 47: 286-95.*
(37) *Blazer DG, et al. Pathologic response to preoperative chemotherapy: a new outcome end point after resection of hepatic colorectal metastases. J Clin Oncol 2008; 26: 5344-51.*
(38) *Nordlinger B, et al. Perioperative chemotherapy with FOLFOX4 and surgery versus surgery alone for resectable liver metastases from colorectal cancer (EORTC Intergroup trial 40983): a randomised controlled trial. Lancet 2008; 371: 1007-16.*
(39) *Lupinacci R, Penna C, Nordlinger B. Hepatectomy for resectable colorectal cancer metastases--indicators of prognosis, definition of resectability, techniques and outcomes. Surg Oncol Clin N Am 2007; 16: 493-viii.*
(40) *Simmonds PC, et al. Surgical resection of hepatic metastases from colorectal cancer: a systematic review of published studies. Br J Cancer 2006; 94: 982-99.*
(41) *Fong Y, et al. Clinical score for predicting recurrence after hepatic resection for metastatic colorectal cancer: analysis of 1001 consecutive cases. Ann Surg 1999; 230: 309-18.*
(42) *Ayez N, et al. The role of neoadjuvant chemotherapy in patients with resectable colorectal liver metastases.*
(43) *Abdalla EK, Vauthey JN. Chemotherapy prior to hepatic resection for colorectal liver metastases: helpful until harmful? Dig Surg 2008; 25: 421-9.*
(44) *Benoist S, Nordlinger B. Neoadjuvant treatment before resection of liver metastases. Eur J Surg Oncol 2007; 33: S35-S41.*
(45) *Kishi Y, et al. Extended preoperative chemotherapy does not improve pathologic response and increases postoperative liver insufficiency after hepatic resection for colorectal liver metastases. Ann Surg Oncol 2010; 17: 2870-6.*
(46) *Benoist S, et al. Complete response of colorectal liver metastases after chemotherapy: does it mean cure? J Clin Oncol 2006; 24: 3939-45.*
(47) *Andreou A, et al. Recent advances in the curative treatment of colorectal liver metastases. Gastrointest Cancer Res 2011; 4: S2-S8.*
(48) *Azoulay D, et al. Resection of nonresectable liver metastases from colorectal cancer after percutaneous portal vein embolization. Ann Surg 2000; 231: 480-6.*

Opinion

Update on the potential of biochemical markers to predict malignant transformation in pheochromocytomas and paragangliomas

Tim Korevaar[a] and Ashley Grossman[b]
[a] Medical student, Erasmus MC University Medical Center Rotterdam Department of Endocrinology, the Netherlands
[b] Supervisor, Professor of Endocrinology, Oxford Centre for Diabetes, Endocrinology and Metabolism, University of Oxford, United Kingdom
Correspondence: Tim Korevaar, e-mail: 303864tk@student.eur.nl

Abstract

Pheochromocytomas and paragangliomas (PPGLs) are rare tumors that arise from the adrenal glands or sympathetic neuronal tissue, and are often catecholamine-secreting. Malignant transformation of these tumors occurs in 2% to 26% of patients, and may cause significantly lower overall survival rates. In patients with PPGLs it is impossible to identify malignancy without the presence of metastatic disease, which can occur as long as 20 years after initial surgery. Early identification of malignant disease would enable a more aggressive treatment approach, which could result in better disease outcome. Currently, no marker can objectively determine malignant potential. Tumor size and extra-adrenal location are considered the most reliable predictors. Biochemical markers, such as high dopamine/methoxytyramine, or a high norepinephrine and epinephrine to total catecholamine ratio, have often been correlated to malignancy. To update the current state of progress in assessing the malignant potential in patients with PPGLs, we reviewed publications on biochemical malignancy-predictors and summarized how various biochemical markers are linked to malignancy. We also investigated the confounding role of tumor characteristics and hereditary syndromes.

Introduction

Pheochromocytomas and paragangliomas (PPGLs) are rare tumors that arise from chromaffin cells of the adrenal medulla (pheochromocytoma) or from sympathetic neuronal tissue in extra-adrenal sites (paraganglioma). As their origin suggests, they usually secrete catecholamines (epinephrine or norepinephrine). However, dopamine-secreting and non-secreting PPGLs have also been described. Paragangliomas of the head and neck usually arise from parasympathetic tissue but often do not secret catecholamines, and are therefore usually assessed separately.

As defined by the WHO, PPGLs can be considered malignant only if there is frank loco-regional invasion or metastases at non-chromaffin sites distant from the primary neoplasm.[1] The prevalence of malignancy in PPGLs has been reported to range between 2% and 26%, and can occur as long as 20 years after primary surgery.[2,3] The most common metastatic sites for chromaffin-cell tumors are local lymph nodes, bone, liver and lung. The 5-year overall survival of patients without metastases has recently been reported as 89.3%,[4] whereas patients with metastatic disease exhibit 5-year overall survival rates ranging from 40% to 72%.[5,6] Moreover, in patients initially diagnosed with benign disease, the main decrease in survival is due to the occurrence of metastatic disease.[4] It is therefore of paramount importance to diagnose the malignant potential of these tumors before the appearance of metastases.

The ability to predict malignancy in these tumors could enable more aggressive treatment. Either surgical treatment (for example, using a transabdominal approach with or without locoregional lymphnode resection), and/or medical treatment (such as chemotherapy with cyclophosphamide, vincristine and dacarbazine or the inclusion in clinical trials with for example sunitinib or pazopanib) may increase survival rates.

At present, the best predictive markers for malignancy are tumor size, extra-adrenal location and genetic succinate dehydrogenase (SDH)-B mutations. For size, many data have shown a positive correlation with malignancy. A cut-off of 5cm diameter is often suggested, and patients with tumors larger than 5cm were found twice as likely to eventually present malignant disease.[7] In contrast, an overall survival analysis also using 5cm as a cut-off value showed a significantly higher, but clinically irrelevant chance of death.[8] Moreover, a 5-cm diameter cut-off resulted in a false-negative diagnosis in some 10–25% of patients; indeed, malignancy in PPGLS has been reported in tumors as small as 2.8 cm.[9] Regarding extra-adrenal location, out of 104 extra-adrenal tumors, those located either infra-diaphragmatic, paraaortic, or in the mediastinum showed metastases in approximately 65–70% of patients, compared to 44, 25, and 36% at bladder, adrenal or other sites, respectively.[8] Of 9 metastatic tumors treated with succinate dehydrogenase-B, 7 mutations were primarily located at extra-adrenal sites, which is in agreement with recent data on tumor location and poor survival in patients with SDH-B germline mutations.[10-12]

We described and analyzed biochemical factors which have been investigated to predict malignancy in PPGLs. To determine the most reliable serum markers that can be used to make an earlier prediction of malignant behavior, we explored the relationship between tumor characteristics and biochemistry in attempt to update and enhance previous findings.[13]

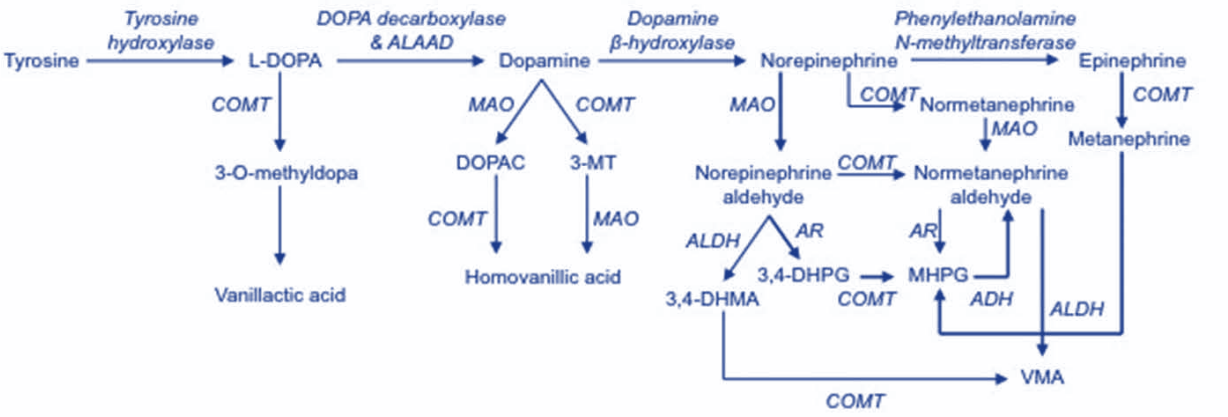


Figure 1 - Catecholamine synthesis and metabolism pathways. Abbreviations: 3-methoxytyramine (3-MT), 3,4-Dihydroxymandelic acid (3,4-DHMA), 3,4-Dihydroxyphenylglycol (3,4-DHPG), Alcohol dehydrogenase (ADH), Aldehyde dehydrogenase (ALDH), Aldehyde reductase (AR), Catechol-O-methyltransferase (COMT), 3,4-Dihydroxyphenylacetic acid (DOPAC), 3-Methoxy-4-hydroxyphenylglycol (MHPG), Monoamine oxidase (MAO) and Vanillylmandelic acid (VMA).

Biochemical markers

Increasing knowledge of the pathways of catecholamine synthesis and metabolism has enabled diagnosis of PPGLs with greater certainty. Recent studies have suggested that blood and urine metabolites might be used as indicators of the risk of malignancy, specifically in patients with genetically related syndromes such as multiple endocrine neoplasia type 2 (MEN 2), von Hippel-Lindau (VHL) syndrome, neurofibromatosis type 1 (NF1) and SDH-B and SDH-D mutations.[5] The pathways of catecholamine synthesis, starting from tyrosine and ending with epinephrine, and the various metabolic products, are shown in Figure 1.

Over the years numerous substances have been used to diagnose PPGLs, including dopamine (DA), 3-methoxytyramine (3-MT), norepinephrine (NE) and epinephrine (EPI), and their metabolites normetanephrine (NM) and metanephrine (MN), measured in both blood and urine. More recently, other markers such as aromatic L-amino acid decarboxylase (ALAAD), DA β-hydroxylase (DBH), as well as the classic vanillylmandelic acid (VMA), have been investigated as possible indicators of malignancy risk.

Chromogranin A (ChrA) is a soluble protein that is co-stored and co-secreted with catecholamines from vesicles in the adrenal medulla and sympathetic nerve endings during exocytosis. Elevated serum levels of ChrA suggest the presence of a chromaffin-secreting tumor or a neuroendocrine tumor in general. The expected positive correlation between ChrA and the size of PPGLS has been shown multiple times.[14-17] Although some reports have indicated a notable correlation of elevated ChrA with malignancy,[18,19] in their study of 39 tumors, van der Harst et al. did not find an association between ChrA and malignancy: 80% (8/10) of malignant tumors (median ChrA 151 U/l; range 17.9-864) but also 69% (20/29) of the benign tumors (median ChrA 146 U/l; range 13-2500) presented with elevated ChrA levels.[20] Furthermore, they showed positive correlations between ChrA levels and tumor volume (r=0.34;P=0.04) and especially tumor weight (r=0.67;P<0.01).

The lyase enzyme ALAAD is involved in the early stages of the catecholamine producing pathway. Because of its association with recurrence in neuroblastomas, van der Harst et al. collected data regarding ALAAD levels in PPGLs. They found elevations in 6/11 malignant tumors compared to 3/35 benign tumors, with median levels of 61 pg/ml (range 15.6-330.8) and 30 pg/ml (range 12-236), respectively.[20]

The metabolite VMA is produced in the later stages of catecholamine metabolism, and some studies have reported a significant difference in VMA excretion between benign and malignant tumors.[7,21,22] However, VMA is elevated in both benign and malignant tumors, but not in epinephrine-secreting PPGLs.[23] Also, it was found to correlate with both NE (r=0.64; P<0.01) and DA (r=0.67;P<0.01), but not with EPI.[20] Thus, a correlation between malignancy and VMA seems more likely to reflect the amount of NE or DA.

A significant difference in the elevation of urinary and/or plasma NE between benign and malignant PPGLs has been reported in a number of studies.[7,20,24,25] As shown in Figure 1, phenylethanolamine N-methyltransferase (PNMT) converts norepinephrine to epinephrine. However, PNMT is induced by glucocorticoids which, in response to stress, are transported from the adrenal cortex to the adrenal medulla. As both extra-adrenal and metastatic PPGLs lack the availability of locally-produced corticosteroids due to their anatomical location, higher NE levels might be expected due to the relative paucity of PNMT. Therefore, the predictive properties for malignant tumor transformation of NE are difficult to interpret. Furthermore, not all reports have shown higher NE levels in malignant PPGLs.[22,26]

As for EPI, reports which show significant differences in malignant PPGLs are sporadic [7], as most reports do not find significant differences.[20,22,24-27] On the other hand, EPI-positive malignant tumors have been associated with significantly better overall survival.[24]

While there is uncertainty as to whether EPI or NE can predict malignancy, two reports have claimed that the ratio of EPI/EPI+NE can be useful, and to date this claim has not been refuted.[20,25] It is noteworthy that these studies found size (defined as diameter measured post-operatively) to be predictive of malignancy (mean size ±SD; malignant 90.9mm ±31.1 (n=11) vs. benign 55.7mm ±28.6 (n=118);P<0.01) but also exhibited a better positive correlation with NE than with EPI (NE r=0.42;P<0.01 vs. EPI r=0.22;P=0.05).[20,25] This suggests that the ratio of EPI/EPI+NE may predict diameter/size rather than the actual malignancy per se.[17] Furthermore, the quantity and production of EPI and NE are PNMT-dependent. As mentioned above, PNMT itself and also the amount/presence of corticosteroids may particularly affect this ratio. In addition, the levels of corticosteroids, the availability of PNMT, the variable tumor production, and secretion of catecholamines will determine the ratio of MN to NM. It is, however, worth

mentioning that post-operative DA levels >100 pg/ml, NE levels >1900 pg/ml, ALAAD levels >56 mU/l and a ratio of EPI/total catecholamines <11% are all associated with significantly decreased metastasis-free survival.[20]

Tumors that predominantly or exclusively produce DA are rare. Even so, strong evidence suggests that high levels of DA are predictive of malignancy as this may represent '"premature"' catecholamine secretion due to profound tumor dedifferentiation.[20,22,28-30] Nonetheless, a recent report by Zelinka et al.[24] combined DA-secreting tumors with non-secretory tumors but found no difference between benign (21.3%) and malignant (29.3%) tumors. A reason for this may be the low number of patients with SDH-B mutations in their patient population since such tumors are associated with both a higher percentage of malignancy and DA-secreting/silent tumors.[14,31-34] The group of malignant DA-secreting/silent tumors did, however, show a significantly decreased survival in this study, with a 5-year survival of around 50%.[24]

Methyoxytyramine, a metabolite of dopamine, has recently been proposed as a more specific marker for dopamine production and a novel marker for malignant PPGLs.[35] In 365 patients, Eisenhofer et al. showed that methoxytyramine was the most accurate marker for metastatic disease out of the 18 catecholamine-related substances tested. When serum methoxytyramine levels were >3 nmol/l, the likelihood of metastases in adrenal tumors increased from 10% to 33% and in non-adrenal tumors from 36% to 79%. Metastatic tumors also exhibited 4.7-fold higher plasma methoxytyramine levels compared to non-metastatic tumors. Nonetheless, the sensitivity and specificity for the optimal cut-off value (2.0 nmol/l) were suboptimal at 57% and 85%, respectively. Also, higher levels of methoxytyramine were found in patients with SDH-B mutations and/or extra-adrenal disease. Interestingly, tumor size was a risk factor for metastatic disease independent of extra-adrenal disease or SDH-B mutations.

Table 1 - Percentage correctly classified

Test or test combination	MEN 2 and NF1 vs. VHL, SDHB and SDHD	VHL vs. SDHB and SDHD
Plasma O-methylated metabolites		
NMNa	47%	60%
MN	97%	50%
MTY	53%	78%
MN and NMN	99%	59%
MN and MTY	99%	79%
NMN and MTY	54%	78%
NMN and MN and MTY	100%	78%
Plasma catecholamines		
NE	55%	59%
EPI	81%	46%
DA	47%	61%
EPI and NE	84%	60%
EPI and DA	82%	59%
NE and DA	53%	65%
NE and EPI and DA	85%	70%
Urine metanephrins		
NMN	50%	66%
MN	98%	57%
NMN and MN	98%	64%
Urine catecholamines		
NE	57%	60%
EPI	92%	62%
DA	61%	59%
EPI and NE	94%	61%
EPI and DA	93%	62%
NE and DA	67%	62%
NE and EPI and DA	94%	69%

Discriminant analysis for classification of patients according to neurochemical profile. Abbreviations: MEN 2, multiple endocrine neoplasia type 2; NF1, neurofibromatosis type 1; VHL, von Hippel-Lindau disease; SDHB/SDHD, succinate dehydrogenase B/D.
a NMN, normetanephrine; MN, metanephrine; MTY, Methoxytyramine; NE, norepinephrine; EPI, epinephrine; DA, dopamine. Adapted with permission from Eisenhofer et al.[33]

Recent findings have shown different levels of biochemical entities in the various hereditary mutations which are known to cause PPGLs. Patients with MEN 2 and NF1 mutations have been found to have an "adrenergic" phenotype, whereas VHL, SDH-B and SDH-D rarely show increased EPI/MN. If they do, the levels are proportionally smaller than the elevations in NE.[33,36,37] Furthermore, mutations of SDHB in particular have been found to exhibit a "noradrenergic" and "dopaminergic" phenotype.[32,33] Eisenhofer et al.[33] reported on 173 patients with proven familial disease and found that the combined measurement of NE, EPI and DA has a high predictive value to distinguish those with MEN 2/NF1 mutations from those with VHL and SDH-B/SDH-D. Using such plasma and urine measurements, the percentages correctly classified were 64%-100%, as shown in Table 1.

Nevertheless, it remains unclear whether these biochemical findings only reflect the genetic effect on (anatomical) tumor characteristics or whether they truly predict tumor malignancy. Future studies should apply statistical methods to answer this question. Catecholamine/size ratios and multivariate logistic regression models appear to be very useful for distinguishing between size, weight and biochemical characteristics.[7,22] Furthermore, most published data do not separate syndromic PPGLS from sporadic PPGLS. Recently, other mutations such as SDH complex assembly factor 2 (SDHAF2), flavoprotein SDH complex subunit A (fp) (SDH-A), transmembrane protein 127 (TMEM127) and MYC-associated factor X (MAX) have been associated with PPGLs. It would appear that biochemical parameters as such are unable to predict malignancy in PPGLs, although they may be useful indicators.

Conclusions

Considerable data suggest various markers for malignancy in PPGLs. Currently, the most reliable predictors are still the size and location of the tumor. We believe that most of the proposed biochemical examination factors, such as high NE levels, high ChrA levels and a low EPI/EPI+NE ratio, simply reflect these tumor characteristics. The increased production of DA or methoxytyramine is uncommon. Nonetheless, it can be associated with malignant potential as it seems to predict a worse degree of tumor dedifferentiation or larger size compared to NE-producing tumors, which is reflected by their lower survival rates.

We conclude that no biochemical marker is currently able to predict metastatic transformation of PPGLs with sufficient accuracy. Therefore, we recommend that the clinical assessment of malignant potential, in any case of confirmed PPGLs, should consist of a carefully weighted combination of tumor characteristics, biochemistry, nuclear and conventional imaging properties and histological markers. We suggest that the principal baseline characteristic should still be size; other markers may be added to either raise or lower the suspicion of malignant disease accordingly. Specific biochemical situations such as a high NE, ChrA or DA/methoxytyramine, even without a large primary tumor and no radiographic distant disease, should determine further investigation.

References

1. Thompson LD Y.W., Kawashima A, Komminoth P, Tischler AS . 2004 Malignant adrenal phaeochromocytoma. In: DeLellis RA, Lloyd RV, Heitz PU, Eng C (eds) WHO classification of tumours: pathology and genetics-tumours of endocrine organs IARC, Lyon, pp 147-150
2. Eisenhofer G., Bornstein S.R., Brouwers F.M., et al. 2004 Malignant pheochromocytoma: current status and initiatives for future progress. Endocr Relat Cancer 11:423-436
3. Lehnert H., Mundschenk J., Hahn K. 2004 Malignant pheochromocytoma. Front Horm Res 31:155-162
4. Timmers H.J., Brouwers F.M., Hermus A.R., et al. 2008 Metastases but not cardiovascular mortality reduces life expectancy following surgical resection of apparently benign pheochromocytoma. Endocr Relat Cancer 15:1127-1133
5. Chrisoulidou A., Kaltsas G., Ilias I., et al. 2007 The diagnosis and management of malignant phaeochromocytoma and paraganglioma. Endocr Relat Cancer 14:569-585
6. Nomura K., Kimura H., Shimizu S., et al. 2009 Survival of patients with metastatic malignant pheochromocytoma and efficacy of combined cyclophosphamide, vincristine, and dacarbazine chemotherapy. J Clin Endocrinol Metab 94:2850-2856
7. Feng F., Zhu Y., Wang X., et al. 2011 Predictive factors for malignant pheochromocytoma: analysis of 136 patients. J Urol 185:1583-1590
8. Jimenez C., Ayala-Ramirez M., Feng L., et al. 2011 Clinical Risk Factors for Malignancy and Overall Survival in Patients with Pheochromocytomas and Sympathetic Paragangliomas: Primary Tumor Size and Primary Tumor Location as Prognostic Indicators. Journal of Clinical Endocrinology & Metabolism 96:717-725
9. Strong V.E., Kennedy T., Al-Ahmadie H., et al. 2008 Prognostic indicators of malignancy in adrenal pheochromocytomas: clinical, histopathologic, and cell cycle/apoptosis gene expression analysis. Surgery 143:759-768
10. Blank A., Schmitt A.M., Korpershoek E., et al. 2010 SDHB loss predicts malignancy in pheochromocytomas/sympathethic paragangliomas, but not through hypoxia signalling. Endocr Relat Cancer 17:919-928
11. Maier-Woelfle M., Brandle M., Komminoth P., et al. 2004 A novel succinate dehydrogenase subunit B gene mutation, H132P, causes familial malignant sympathetic extraadrenal paragangliomas. J Clin Endocrinol Metab 89:362-367
12. Neumann H.P., Bausch B., McWhinney S.R., et al. 2002 Germ-line mutations in nonsyndromic pheochromocytoma. N Engl J Med 346:1459-1466
13. Korevaar T.I., Grossman A.B. 2011 Pheochromocytomas and paragangliomas: assessment of malignant potential. Endocrine 40:354-365
14. Amar L., Baudin E., Burnichon N., et al. 2007 Succinate dehydrogenase B gene mutations predict survival in patients with malignant pheochromocytomas or paragangliomas. J Clin Endocrinol Metab 92:3822-3828
15. Hsiao R.J., Neumann H.P., Parmer R.J., et al. 1990 Chromogranin A in familial pheochromocytoma: diagnostic screening value, prediction of tumor mass, and post-resection kinetics indicating two-compartment distribution. Am J Med 88:607-613
16. Hsiao R.J., Parmer R.J., Takiyyuddin M.A., et al. 1991 Chromogranin A storage and secretion: sensitivity and specificity for the diagnosis of pheochromocytoma. Medicine (Baltimore) 70:33-45
17. Unger N., Hinrichs J., Deutschbein T., et al. 2012 Plasma and Urinary Metanephrines Determined by an Enzyme Immunoassay, but not Serum Chromogranin A for the Diagnosis of Pheochromocytoma in Patients with Adrenal Mass. Experimental and clinical endocrinology & diabetes : official journal, German Society of Endocrinology [and] German Diabetes Association 120:494-500
18. Rao F., Keiser H.R., O'Connor D.T. 2000 Malignant pheochromocytoma. Chromaffin granule transmitters and response to treatment. Hypertension 36:1045-1052
19. Szalat A., Fraenkel M., Doviner V., et al. 2011 Malignant pheochromocytoma: predictive factors of malignancy and clinical course in 16 patients at a single tertiary medical center. Endocrine 39:160-166
20. van der Harst E., de Herder W.W., de Krijger R.R., et al. 2002 The value of plasma markers for the clinical behaviour of phaeochromocytomas. European Journal of Endocrinology 147:85-94
21. Park J., Song C., Park M., et al. Predictive characteristics of malignant pheochromocytoma. Korean J Urol 52:241-246
22. John H., Ziegler W.H., Hauri D., et al. 1999 Pheochromocytomas: can malignant potential be predicted? Urology 53:679-683
23. Gerlo E.A., Sevens C. 1994 Urinary and plasma catecholamines and urinary catecholamine metabolites in pheochromocytoma: diagnostic value in 19 cases. Clin Chem 40:250-256
24. Zelinka T., Musil Z., Duskova J., et al. 2011 Metastatic pheochromocytoma: Does the size and age matter? Eur J Clin Invest
25. Plouin P.F., Chatellier G., Fofol I., et al. 1997 Tumor recurrence and hypertension persistence after successful pheochromocytoma operation. Hypertension 29:1133-1139
26. Mornex R., Badet C., Peyrin L. 1992 Malignant pheochromocytoma: a series of 14 cases observed between 1966 and 1990. J Endocrinol Invest 15:643-649
27. Axelrod J. 1966 Methylation reactions in the formation and metabolism of catecholamines and other biogenic amines. Pharmacol Rev 18:95-113
28. Proye C., Fossati P., Fontaine P., et al. 1986 Dopamine-secreting pheochromocytoma: an unrecognized entity? Classification of pheochromocytomas according to their type of secretion. Surgery 100:1154-1162
29. Schlumberger M., Gicquel C., Lumbroso J., et al. 1992 Malignant pheochromocytoma: clinical, biological, histologic and therapeutic data in a series of 20 patients with distant metastases. J Endocrinol Invest 15:631-642
30. Tippett P.A., McEwan A.J., Ackery D.M. 1986 A re-evaluation of dopamine excretion in phaeochromocytoma. Clin Endocrinol (Oxf) 25:401-410
31. Chen H., Sippel R.S., O'Dorisio M.S., et al. 2010 The North American Neuroendocrine Tumor Society consensus guideline for the diagnosis and management of neuroendocrine tumors: pheochromocytoma, paraganglioma, and medullary thyroid cancer. Pancreas 39:775-783
32. Timmers H.J., Kozupa A., Eisenhofer G., et al. 2007 Clinical presentations, biochemical phenotypes, and genotype-phenotype correlations in patients with succinate dehydrogenase subunit B-associated pheochromocytomas and paragangliomas. J Clin Endocrinol Metab 92:779-786
33. Eisenhofer G., Lenders J.W., Timmers H., et al. 2011 Measurements of plasma methoxytyramine, normetanephrine, and metanephrine as discriminators of different hereditary forms of pheochromocytoma. Clin Chem 57:411-420
34. Eisenhofer G., Timmers H.J., Lenders J.W., et al. 2011 Age at diagnosis of pheochromocytoma differs according to catecholamine phenotype and tumor location. J Clin Endocrinol Metab 96:375-384
35. Eisenhofer G., Lenders J.W., Siegert G., et al. 2012 Plasma methoxytyramine: a novel biomarker of metastatic pheochromocytoma and paraganglioma in relation to established risk factors of tumour size, location and SDHB mutation status. Eur J Cancer 48:1739-1749
36. Eisenhofer G., Lenders J.W., Linehan W.M., et al. 1999 Plasma normetanephrine and metanephrine for detecting pheochromocytoma in von Hippel-Lindau disease and multiple endocrine neoplasia type 2. N Engl J Med 340:1872-1879
37. Eisenhofer G., Walther M.M., Huynh T.T., et al. 2001 Pheochromocytomas in von Hippel-Lindau syndrome and multiple endocrine neoplasia type 2 display distinct biochemical and clinical phenotypes. J Clin Endocrinol Metab 86:1999-2008

Sorry, but you will have to lose weight before receiving your knee replacement

Kasper C. Roth[a], Gert J.H.J.M. Bessems[b]
[a] Medical student, Erasmus MC University Medical Center Rotterdam, the Netherlands
[b] MD, Orthopaedic Surgeon, Department of Orthopaedics, Erasmus MC University Medical Center Rotterdam, the Netherlands
Correspondence: Kasper Roth, e-mail: 328159kr@student.eur.nl

ABSTRACT

Background: **The upcoming 'obesity epidemic' is likely to result in a greater need for total knee replacements, as obesity is the most important risk factor for osteoarthritis of the knee. Often orthopaedic surgeons are reluctant to perform surgery on highly obese patients. The aim of this paper is to discuss the scientific and ethical grounds of the unwillingness of orthopaedic surgeons to operate on morbidly obese patients.**

Methods**: Literature on the scientific and ethical merits to demanding a morbidly obese patient to lose weight prior to knee replacement surgery was reviewed and an expert was interviewed.**

Results**: Morbidly obese patients can benefit from TKR in terms of decrease in pain and increase in ROM and walking distance. However, these benefits are of lesser magnitude than in the non-obese patients. Morbidly obese patients are at greater risk of developing deep infection, deep venous thrombosis, complications in wound healing and premature failure of their prosthesis. Obese patients do no tend to lose weight after receiving a TKR. Bariatric surgery prior to TKR can lead to a better outcome.**

Conclusion**: It is important to educate patients on the importance of losing weight prior to surgery. One also has to consider if the principle of doing good outweighs the principle of inflicting no harm on the patient.**

Introduction

Say, you are an orthopaedic surgeon and a forty-five year old woman presents herself at your out-patient clinic with severe disabling pain of her knee. During physical examination her weight is assessed at 150 kilograms and her body mass index (BMI) is 48. This categorizes her as a morbidly obese (MO) patient. Radiological assessment reveals narrowing of the joint, osteophytes, subchondral cysts and subchondral sclerosis. These radiological findings confirm your suspicion of extensive degenerative joint disease and osteoarthritis is diagnosed. Your patient could benefit from a total knee replacement, but will you be able to perform such an operation despite her morbid obesity?

Obesity is the most important risk factor for osteoarthritis of the knee. A morbidly obese patient has as much as a 32 times greater likelihood of requiring a total knee replacement (TKR) than a patient who is of normal weight.[1] This increase is caused directly by the increased mechanical load on the joint or indirectly by comorbidity or decreased mobility of the overweight patient.

The upcoming 'obesity epidemic' is expected to result in an increased incidence of osteoarthritis of the knee and therefore a greater need for TKR.[2]
Additionally, morbidly obese patients often present themselves with severe osteoarthritis years sooner than non-obese patients. Many surgeons will choose not to offer these morbidly obese patients treatment owing to an increased risk of complications as well as difficulty in performing surgery. Can this denial of treatment be justified?

At this moment, obese patients in need of joint replacement living in Suffolk, United Kingdom are out of luck. Patients with a BMI above 30 are denied this treatment.[3] Is this ethically acceptable?

The pain caused by osteoarthritis causes a decline in mobility and physical activity, which makes it hard for an obese patient to lose weight. On the other hand, surgeons demand that these patients lose weight before placing a TKR. This leads to a dilemma. Obese patients are not able to lose weight, because they are limited in their mobility due to the pain and need surgery. Obese patients are not allowed surgery, which will allow them pain-free mobility, because they have to lose weight first. This raises the following question: "Is it acceptable for doctors to demand obese patients to lose weight before granting them a total knee replacement?"

Scientific, medical view on the matter

To be able to answer this question it is important to differentiate between obesity and morbid obesity. The body mass index (BMI) is defined as the individual's weight (in kilograms) divided by the square of his or her height (in meters) and is used to assess how much someone's weight departs from what is normal or desirable for a person of their height. The BMI of a healthy individual ranges from 18,5 to 25. Individuals are considered obese when their BMI exceeds 30. When an individual's BMI rises above the 40, they are considered as morbidly obese. As mentioned before, some surgeons will choose not to operate on morbidly obese patients. Therefore, it is important to discuss whether we can draw a line and, if so, where to draw the line. Medically speaking, above what BMI should we decline patients a TKR?

In 2010 Samson et. al published a review, in which they studied the functional outcome of TKRs in morbidly obese patients. They showed that morbidly obese patients definitely can benefit from a TKR in terms of an increase of range of motion, a decrease in pain and an increase in walking distance. However, these benefits are of a lesser magnitude than in non-obese patients.[2]

Multiple studies have been published on the results of TKRs in obese patients compared to non-obese patients, showing significantly lower clinical scores in the group of obese patients. Patients without obesity had a greater range of motion and less pain. Surprisingly, no significant difference in satisfaction was found between the two groups.[5-7]

One of the great concerns orthopaedic surgeons have about operating obese patients, is the great increase in complications such as deep infections, but also the more superficial wound infections, deep venous thrombosis and poor wound healing. The odds for a deep prosthetic infection were 9 times greater in patients with morbid obesity.[8] The morbidly obese also had a significant higher incidence of wound infections, attributed to poor oxygenation of adipose tissue, increased wound tension and underlying endocrine disorders.[2]

In patients with (moderate) obesity no significant difference was found in the number of complications after receiving a TKR.[7,9]

Another possible concern of the orthopaedic surgeons is that the increased body-weight causes an increased load on a TKR and surrounding bone, which leads to premature loosening and increased failure rates of the prosthesis. Amin et al.[10] and Berend et al.[11] showed that in morbidly obese patients loosening of the prosthesis and higher rates of failure of the prosthesis were seen more often than in non-obese patients. The five-year survivorship of the prosthesis was 74.2% in the MO group compared to 100% in the non-obese group.[10] In nearly a third of the morbidly obese patients radiolucent lines were demonstrated, suggestive for possible failure.[11] Yet, current literature is discordant on the matter: other studies did not find a significant difference in survival of the prosthesis between morbidly obese and non-obese patients.[12-14]

The increase in loosening and premature failures is only seen in the morbidly obese patients and not in the moderate obese group.[15] A hypothesis that could explain this phenomenon is that lower activity levels seen in obese patients compensate for the higher loads on the tibial component, but once a patient becomes morbidly obese this lower level of activity cannot compensate for the increasing load across the implant-bone interface.

Also, a relationship exists between obesity and the age knee replacement is undertaken. The mean age of the patients with morbid obesity was thirteen years younger than that of non-obese patients.[16] Thus a MO-patient requires a TKR at a younger age and his or her prosthesis lasts shorter, causing a greater likelihood that he or she will require a revision or even multiple revisions.

An interesting finding is that, in general, obese patients don't lose weight after joint replacement. These patients often blame their inability to lose weight on their pain secondary to osteoarthritis restricting their activity levels and argue that they need an operation to eliminate the pain to be able to increase their activity levels and lose weight. It appears that lower preoperative activity levels are not the cause for the inability to lose weight or that the gain in mobility achieved by joint replacement does not result in weight loss.[17] This invalidates the argument that patients are unable to lose weight because of their movement impairment due to their osteoarthritis.

If patients with severe osteoarthritis lose weight, this will result in a reduction of pain, so weight loss can even be seen as a conservative treatment for osteoarthritis, which may even delay the necessity of a TKR for obese patients.[18] Poolman et al. suggest that these highly obese patients awaiting surgery for a knee replacement should first be submitted to a weight-loss programma prior to their TKA.[19]

Morbidly obese individuals with severe osteoarthritis, who are considered unsuitable to undergo orthopaedic surgery because of excess weight, can qualify for bariatric surgery prior to joint replacement as a solution for their obesity. Parvici et al. performed bariatric surgery in twenty morbidly obese patients prior to total knee replacement to enable weight loss before the orthopaedic surgery. The average time from bariatric surgery to arthroplasty was 23 months, during which patients underwent a successful reduction of a mean BMI of 49 to a mean BMI of 29. Total joint arthroplasty after bariatric surgery had an excellent outcome in these patients with an acceptable complication rate. All but one were satisfied with the result at follow-up.[20] However, bariatric surgery has a 22-35% risk of late complications and failure and a mortality rate of 0.28%.[21]

Summarized, morbidly obese patients have functional benefits from a total knee replacement. The functional outcome after TKR increases as the BMI decreases. Morbidly obese patients have a concerning, elevated risk of complications after TKR. Also, surgery becomes more difficult and has a longer duration in MO patients. Obese patients require TKR at a younger age: this applies even more to the morbidly obese. The survival of the prosthesis is significantly shorter for MO patients, but this is not significant in obese patients. Surprisingly, obese patients do not lose weight after TKR. Bariatric surgery could aid the morbidly obese in losing weight prior to undergoing joint arthroplasty.

Ethical aspects

54% of all doctors, who participated in a survey in England, agree to the fact that patients, who don't want to lose weight before surgery, don't deserve a TKR.[22] Can we justify this ethically?

A patient presents itself at your out-patient clinic with severe disabling pain of the knee. This patient can be treated by undergoing a TKR. As a doctor, you have the task to stimulate the well-being of your patient. Therefore, it would seem likely to just give every patient with degenerative joint disease a TKR. This, however, is a very simple and incorrect view on how to treat obese patients requiring arthroplasty. The reason why not all patients, for instance morbidly obese patients, receive a TKR, is because they have a high risk of not being better off after the operation. For instance, when a patient develops a deep periprosthetic infection after TKR he will have to stay in the hospital for a long period of time, will need antibiotics through infusion with the additional adverse effects and, if unlucky, will have to undergo another surgery to remove the prosthesis. In the worst case scenario, this could lead to needing a knee arthrodesis or even an amputation.

One has to consider if the principle of beneficence, doing good, outweighs the principle of nonmaleficence, stating "primum non nocere" which means "first, do not harm". There are many ways to interpret nonmaleficence. It could imply that if one cannot do good without also causing harm, then one should not act at all. This, however, makes action almost impossible, in a world where even the best actions may have some harmful results.

In such conflicting situations the "principle of double effect" could offer a more reasonable way of dealing with the situation. This states that an action, that is good in itself can have two effects: an intended good effect or an unintended, yet foreseen evil effect. Therefore, it is very important to weigh out the foreseen beneficial effects to the possible, foreseen harmful effects.[23]

The goal of the treating orthopaedic surgeon should be to maintain a patient's mobility, improve or maintain quality of life and provide pain relief while minimizing risk and complications of treatment.[4]

Furthermore, a patient needs to take responsibility for the choices they make. By choosing for an unhealthy lifestyle, consuming junk food frequently and rarely exercising, your tendency to become overweight is greater than that of a healthy individual. The greater the load a joint has to carry, the sooner it wears out and the sooner a total knee replacement is required. Does this mean an individual can be declined surgery, because he or she has contributed to the origin of the problem? Can you demand a patient to lose weight prior to surgery? When a doctor starts making such demands and declines a patient surgery, an individual's rights

to make his/her own choices are endangered. A doctor must not make judgements about an individual's lifestyle and should have a neutral attitude towards their patient. An individual should be free from coercion in deciding how to live their life.

In addition, a doctor must treat patients equally. Every patient has the right to receive proper medical treatment. If some patients don't succeed in losing weight prior to surgery, you can't leave them out in the cold.[24]

Lastly, an important issue is also the orthopaedic surgeon's attitude. To many orthopaedic surgeons it is very important to score as low as possible complication rates. It is possible that, unknowingly, orthopaedic surgeons keep this thought in mind when they are considering treatment options for morbidly obese patients and they could be influenced by this. In this aspect, the interests of the patient should always be put above those of the treating surgeon.

Expert opinion

To get a good view on the opinion of orthopaedic surgeons on the treatment of morbidly obese patients, I interviewed Prof. Dr. J.A.N. Verhaar, orthopaedic surgeon in Erasmus MC, specialised in total knee replacements. A short summary of his view on the treatment of osteoarthritis in the morbidly obese is as follows.

"The effect of obesity goes deeper than just the mechanical load on the joint. Inflammatory processes are amplified or maintained by adipocytokines secreted by adipose tissue. Obesity often causes osteoarthritis of the knee at relatively young age and these patients present themselves at your out-patient clinic demanding a total knee replacement as a solution for their problem."

Prof. Dr. Verhaar's patients often blame their obesity on the fact that they are limited in their mobility, but research and his experience have shown that patients do not tend to increase their daily activity and do not lose weight after undergoing TKR.[17] "Placing a TKR is not without risk. On the one hand, post-operatively risks exist that are caused by malplacement of the prosthesis, due to the difficulty in performing surgery in obese patients. On the other hand, morbidly obese patients have a higher risk of infection, due to poor healing of soft tissue and diabetes, which obese patients often have. Operation of MO patients also has a longer duration and post-operative mobilisation is laborious."

In his experience, orthopaedic surgeons actually decline patients surgery or demand them to lose weight first. These patients often present themselves at his out-patient clinic for a second opinion. When asked on his opinion on bariatric surgery, he answered that bariatric surgery can definitely be an important part of treatment, but only in a multidisciplinary context in which psychological treatment is also applied. "Obesity is a disorder which will only pass after specific treatment by a whole team of specialists. On top of that, if the obesity remains, the other non-orthopaedic health risks stay as well."

At this stage, Prof. Dr. Verhaar does not experience competition between orthopaedic surgeons or between different hospitals to score the lowest complication rates, but he believes we should definitely keep this issue in mind and it might become relevant someday. Health insurance companies already take reoperation percentages and infection rates into account, which are both elevated in high-risk patients, such as the morbidly obese. Currently, some patients don't fit in 'fast-track programs' and are declined treatment in private health institutions, because of insufficient funds they raise.

Prof. Dr. Verhaar concludes: "for multiple reasons a TKR is not the appropriate solution for a morbidly obese patient and will not lead to weight reduction. It is important to address the root cause and not fight the symptoms. Morbid obesity should be seen as a life-event that requires a multidisciplinary approach and cannot be resolved by just an orthopaedic surgeon."

First author's opinion

Based on the literature, I believe in the importance of losing weight prior to undergoing a total knee replacement. This applies even more to the morbidly obese patient. Ethically however, I do not believe it is justifiable to demand these patients to lose weight first. The willingness and motivation of the patient to lose weight must not come from the doctor. I expect patients to have a lot more difficulty in losing weight, if it is the doctor's wish and not their own.

From the surgeon's point of view, requiring a patient to lose weight first is an easy plausible excuse for the orthopaedic surgeon not to do the operation. In my opinion, an orthopaedic surgeon has certain duties before operating on a patient with obesity. Firstly, it is important to educate the patient about the effects of increased BMI on the outcome after a TKR and the higher complication rate obesity brings along. Next, all the possible means that could aid the patient in losing weight prior to surgery should be discussed. If the attempts to lose weight are unsuccessful, the surgeon needs to make a professional judgement about the executability of the operation. I do not believe it is possible to draw a line declining all patients above a certain BMI surgery. It is not black and white. There are many other factors that should also be taken into account. The possibility of bariatric surgery should be considered and shared with the patient.

Conclusion

A morbidly obese patient should not undergo a TKR before having been motivated to lose weight or having been informed of the higher complication rate and lower functional outcome among morbidly obese patients. The physician should not only motivate but also assist the patient in losing weight by offering (surgical) weight loss therapy. However, the patient has the freedom of choice whether to comply with the physician's recommendations. In conclusion, I believe an orthopaedic surgeon should not, as the title says, demand a patient to lose weight prior to surgery, but he/she should strongly advise and help motivate patients to lose weight before deciding to operate.

Literature

1. *Bourne R., Mukhi S., Zhu N., et al. Role of obesity on the risk for total hip or knee arthroplasty. Clin Orthop Relat Res. 2007; 465: 185-188.*
2. *Samson A.J., Mercer G.E., Campbell D.G. Total knee replacement in the morbidly obese: a literature review. ANZ J Surg. 2010; 80: 595-599.*
3. *The weighty matter of care for all. The Lancet. 2006; 367: 876.*
4. *Hamlin B.R. Treatment of knee arthrosis in the morbidly obese patient. Orthop Clin North Am. 2011; 42: 107-113, vii.*
5. *Jackson M.P., Sexton S.A., Walter W.L., et al. The impact of obesity on the mid-term outcome of cementless total knee replacement. J Bone Joint Surg Br. 2009; 91: 1044-1048.*
6. *Bordini B., Stea S., Cremonini S., et al. Relationship between obesity and early failure of total knee prostheses. BMC Musculoskelet Disord. 2009; 10: 29.*
7. *Yeung E., Jackson M., Sexton S., et al. The effect of obesity on the outcome of hip and knee arthroplasty. Int Orthop. 2011; 35: 929-934.*
8. *Dowsey M.M., Choong P.F. Obese diabetic patients are at substantial risk for deep infection after primary TKA. Clin Orthop Relat Res. 2009; 467: 1577-1581.*
9. *Hamoui N., Kantor S., Vince K., et al. Long-term outcome of total knee replacement: does obesity matter? Obes Surg. 2006; 16: 35-38.*
10. *Amin A.K., Clayton R.A., Patton J.T., et al. Total knee replacement in morbidly obese patients. Results of a prospective, matched study. J Bone Joint Surg Br. 2006; 88: 1321-1326.*
11. *Berend K.R., Lombardi A.V., Jr., Mallory T.H., et al. Early failure of minimally invasive unicompartmental knee arthroplasty is associated with obesity. Clin Orthop Relat Res. 2005; 440: 60-66.*
12. *Dewan A., Bertolusso R., Karastinos A., et al. Implant durability and knee function after total knee arthroplasty in the morbidly obese patient. J Arthroplasty. 2009; 24: 89-94, 94 e81-83.*
13. *Spicer D.D., Pomeroy D.L., Badenhausen W.E., et al. Body mass index as a predictor of outcome in total knee replacement. Int Orthop. 2001; 25: 246-249.*

14. Krushell R.J., Fingeroth R.J. Primary Total Knee Arthroplasty in Morbidly Obese Patients: a 5- to 14-year follow-up study. J Arthroplasty. 2007; 22: 77-80.
15. Namba R.S., Paxton L., Fithian D.C., et al. Obesity and perioperative morbidity in total hip and total knee arthroplasty patients. J Arthroplasty. 2005; 20: 46-50.
16. Changulani M., Kalairajah Y., Peel T., et al. The relationship between obesity and the age at which hip and knee replacement is undertaken. J Bone Joint Surg Br. 2008; 90: 360-363.
17. Donovan J., Dingwall I., McChesney S. Weight change 1 year following total knee or hip arthroplasty. ANZ J Surg. 2006; 76: 222-225.
18. Bliddal H., Leeds A.R., Stigsgaard L., et al. Weight loss as treatment for knee osteoarthritis symptoms in obese patients: 1-year results from a randomised controlled trial. Ann Rheum Dis. 2011; 70: 1798-1803.
19. Poolman R.W., Wagensveld B.A., Artrose van de knie: eerst afvallen? Ned Tijdschr Geneeskd. 2013;157: A6043
20. Parvizi J., Trousdale R.T., Sarr M.G. Total joint arthroplasty in patients surgically treated for morbid obesity. J Arthroplasty. 2000; 15: 1003-1008.
21. Kulkarni A., Jameson S.S., James P., et al. Does bariatric surgery prior to lower limb joint replacement reduce complications? Surgeon. 2011; 9: 18-21.
22. Campbell D. Doctors back denial of treatment for smokers and obese. The Observer. 2012 29 April. Available from: http://www.guardian.co.uk/society/2012/apr/28/doctors-treatment-denial-smokers-obese.
23. Ascension Health. Key Ethical Principles. [cited 2013 Jan, 18]; Available from: http://www.ascensionhealth.org/index.php?option=com_content&view=article&id=47&Itemid=171.
24. De Kweste, praktijkboek ethiek voor de gezondheidszorg. In: Ethiek rond obesitas: Inez de beaufort, Medard Hilforst, Sofie Vandamme, Suzanne van de Vathorst; 2008;

Placebos in clinical practice: management of medically unexplained symptoms

Nico Jansen [a,b]
[a] *Medical Student, Erasmus MC University Medical Center Rotterdam, the Netherlands*
[b] *Research Master Student, Department of Neuroscience, Erasmus MC University Medical Center Rotterdam, the Netherlands*
Correspondence: Nico Jansen, e-mail: n.jansen@erasmusmc.nl

Placebos in clinical practice

The double-blind, randomized, placebo-controlled trial is the gold standard in clinical research to investigate the efficacy of a certain intervention.[1] A placebo is regarded as an inert substance or intervention, it is therefore by definition unable to elicit an effect.[2] Placebos show effects in a wide variety of clinical trials, which include analgesic effects, but also in disorders ranging from depression, Parkinson's disease and anxiety disorder to irritable bowel syndrome (IBS).[2-4] This should impel us to reconsider the effects of a placebo. One should focus on how it affects the patient, thereby leaving the "traditional" view of the effects attributable to the inert substance or intervention.

Obviously, the use of placebos should never prevent or delay the use of a preferred initial treatment regimen. However, as there is often no proven primary care treatment for patients with medically unexplained symptoms this is a patient population likely to benefit from placebo treatment.[5] This has also been demonstrated recently in trials including patients with IBS, which will be discussed later.[3,4] Medically unexplained symptoms can be defined as "physical symptoms that prompt the sufferer to seek healthcare but remain unexplained after an appropriate medical examination".[6] This group includes IBS, unexplained back pain, chronic fatigue syndrome, fibromyalgia, interstitial cystitis and a variety of other conditions. They form a heterogeneous group, but report a lower quality of life, general and mental health than other chronically ill patients. When compared to the general population, this group consumes more health care resources leading to higher economic and health care costs.[7]

In this essay, I postulate the implementation of placebos in clinical practice with the focus on medically unexplained symptoms. However, as placebos can elicit the (equal) beneficial effects of an intervention without the side effects under some circumstances, their implementation might be considered broader.[8]

Placebos: how to study their effects?

Most randomized controlled trials (RCTs) involving placebo treatment lack an additional group controlling for the placebo treatment, which limits the possibility to attribute a response to the placebo effect.[1,8] This response might for instance reflect the natural course of the disease and fluctuations in (subjective) symptoms.[2] Hence, to unravel its clinical importance it is important to include no-treatment groups, as clinical research now lags behind research into the mechanisms of placebos.[2] However, although a meta-analysis that included no-treatment groups found placebo effects to be small, trials studying placebo mechanisms found placebo effects to be much larger.[1,9-12] An example of these contrasting findings is provided by a meta-analysis by Vase et al., who found significantly different decreases in pain scores when clinical studies using placebo controls were compared to studies of placebo analgesic mechanisms.[9] This inconsistency might be due to a more representative clinical context in trials investigating placebo mechanisms, which supports the use of these trials and representative clinical trials to fully appreciate the effects of placebos.[2]

Neurobiological and psychological basis

The effects of placebos can roughly be regarded in a neurobiological and psychological approach. I will concisely discus both.

An important contribution to the neurobiological support of the effects of the placebo comes from research on its analgesic capacities. Two trials found antagonism of placebo analgesia by the opioid antagonist naloxone.[13,14] This finding proclaims the involvement of endogenous opioids in placebo analgesia. This was later strengthened by trials demonstrating the reversal of respiratory insufficiency, bradycardia and β-adrenergic activity, all conditioned effects of placebo treatment, by administration of naloxone.[15,16] These finding are in line with the evidence of involvement of the "opioidergic descending pain controlling system" in placebo analgesia, originating in the frontal lobe and modulating nociceptive input in the dorsal horn of the spinal cord.[17]

Reward expectations are also likely to play an important role in the placebo effect.[8] The functioning of the reward system, in which endogenous opioids and the neurotransmitter dopamine are thought to play a crucial role (mainly through their activity in the nucleus accumbens), might also partially explain why not all individuals respond to placebos.[8]

As earlier mentioned, placebos show effects in a variety of clinical disorders, which can be a consequence of the physiologic effects on the endocrine, cardiovascular, respiratory and immune system.[2] In the scope of this essay, these effects will not be discussed further.

Two important psychological factors of placebo treatment in clinical practice are conditioned associations of the patient and expectations, both of the patient and the physician.[2,18,19] The expectations of the patient are the result of both conditioned associations, e.g. intervention leads to curation, as well as expectations of the physician.[19] The introduction of the intervention and the first consultation with the doctor are thought to be critical to allow conditioned associations and expectancies of the patient, thereby leading to (in)significant placebo effects. This is subsequently essential for the patients' expectations during follow-up.[2]

Many factors are known to modulate the expectations of the patient, leading to profound placebo effects. Among these are attributing positive effects to the treatment by the physician, the type of treatment (physical intervention is more effective than prescription of a substance), physician-patient relationship, the extent of the medical ritual, but (famously) also the color, size and quantity of the placebo pill.[2,18-20]

Placebos in IBS

Medically unexplained symptoms (MUS) are prevalent and form a major health burden. The patient-physician relationship seems especially important in these disorders and may even maintain and potentiate the symptoms.[21] To illustrate the impact of the placebo treatment in MUS, I will discuss a RCT studying placebo responses in IBS.[3]

In this trial, patients where randomly assigned to one of three groups: one waiting list group, one acupuncture group and one acupuncture group with an "intensive" physician-patient interaction. After three and six weeks, the intervention groups showed significant improvement in adequate relief when compared to the waiting list group. However, the "intensive interaction" had an even more profound effect on adequate relief, which made the authors conclude that factors contributing to the placebo effect can be combined to achieve a graded response and that the physician-patient interaction is the most robust component.[3]

True informed consent: achievable?

A placebo does not contain biological activity, its effects are attributable to non-specific effects of the treatment and the physician-patient interaction. It has been argued that the term "placebo effect" is therefore often misinterpreted, as it is not a result of an inert intervention, but of a specific treatment.[18] When we consider this definition, it is reasonable to conclude that the "placebo effect" is always present in clinical practice to a large extent, as the non-specific effects of treatment and physician-patient interaction can naturally not been excluded. Therefore, it has been argued to replace the term "placebo effect" by the term "positive care effect".[18] As mentioned earlier, these effects can be very potent, which has ethically far-reaching implications as it raises the question to what extent informed consent is already achievable in the clinical context.

However, transparency about all beneficial aspects of an intervention is not advisable, as physicians should not be imposed to start their consultation by proclaiming the effects of their "friendly bedside manner". The inevitable "use" of the placebo effect in clinical practice I will therefore name the "passive use" of placebos.

The effects of placebos in clinical practice are obviously impossible to ignore and it is tempting to "actively" use them. Yet, misguidance of the patient is of course the way in which placebos "operate" and their active use can consequently affect the patients' trust and may disrupt the relationship with the physician.

Inconsistency in current thinking

The effects of placebos are evident and recent studies have revealed that their effects are more profound than previously thought. Physicians should be informed about these effects and ethical implications for clinical practice should be considered carefully.

As you may have noticed, there is an inconsistency in current thinking about the placebo effect.

On one hand, physicians make "passive" use of it in clinical practice, which undermines informed consent in a broader sense.[2] The contemporary view on this matter is not self-evident, as patients might demand for improvements in physician-patient interaction when they knew about the power of this interaction. One could even argue that the patients' trust might be severely damaged when they notice the importance of the communication skills of the physician in their treatment.[19]

On the other hand, we constrain the "active" use of placebos, as it disturbs informed consent. Strictly speaking, the active use of the placebo effect is already abundant in clinical practice, as communication skills are "tools" of the physician to achieve beneficial effects in patients.[19] Moreover, these communication skills are actually the most robust components in the placebo effect, as demonstrated earlier.[3]

Placebos in future clinical practice

I propose that, as the ethical differences between "active" and "passive" use of the placebo effect in clinical practice are not clear-cut, introduction of placebos in clinical practice should be discussed. Evidence-based implementation of placebos should be possible in the future, as the use of communication skills by physicians is already possible. Patients with MUS form an important target group, given the lack of effective interventions in this group.

Unmistakably, the best way to implement placebos should be studied and extensively (ethically) discussed. Open-label use of placebos is already accepted by the FDA and seems to be effective in subgroups of IBS.[4] Publicity campaigns have been proposed as a possibility to introduce the use of "misguiding" placebos in clinical practice to the public at large.[19]

Making use of the placebo effect, or positive care effect, in clinical practice is at odds with informed consent. However, when restricted to carefully chosen subgroups, this effect might be very desirable. A little more restriction of informed consent, when compared to the profound placebo effect resulting from physician-patient interaction in clinical practice nowadays, may be allowed to achieve this effect.

References

1. *Price D.D., Finniss D.G., Benedetti F. A comprehensive review of the placebo effect: recent advances and current thought. Annual review of psychology. 2008; 59: 565-590.*
2. *Finniss D.G., Kaptchuk T.J., Miller F., et al. Biological, clinical, and ethical advances of placebo effects. Lancet. 2010; 375: 686-695.*
3. *Kaptchuk T.J., Kelley J.M., Conboy L.A., et al. Components of placebo effect: randomised controlled trial in patients with irritable bowel syndrome. BMJ. 2008; 336: 999-1003.*
4. *Kaptchuk T.J., Friedlander E., Kelley J.M., et al. Placebos without deception: a randomized controlled trial in irritable bowel syndrome. PloS one. 2010; 5: e15591.*
5. *Smith R.C., Lyles J.S., Gardiner J.C., et al. Primary care clinicians treat patients with medically unexplained symptoms: a randomized controlled trial. Journal of general internal medicine. 2006; 21: 671-677.*
6. *Richardson R.D., Engel C.C., Jr. Evaluation and management of medically unexplained physical symptoms. The neurologist. 2004; 10: 18-30.*
7. *Edwards T.M., Stern A., Clarke D.D., et al. The treatment of patients with medically unexplained symptoms in primary care: a review of the literature. Mental health in family medicine. 2010; 7: 209-221.*
8. *Enck P., Benedetti F., Schedlowski M. New insights into the placebo and nocebo responses. Neuron. 2008; 59: 195-206.*
9. *Vase L., Riley J.L., 3rd, Price D.D. A comparison of placebo effects in clinical analgesic trials versus studies of placebo analgesia. Pain. 2002; 99: 443-452.*
10. *Hrobjartsson A., Gotzsche P.C. Unsubstantiated claims of large effects of placebo on pain: serious errors in meta-analysis of placebo analgesia mechanism studies. Journal of clinical epidemiology. 2006; 59: 336-338; discussion 339-341.*
11. *Hrobjartsson A., Gotzsche P.C. Is the placebo powerless? An analysis of clinical trials comparing placebo with no treatment. The New England journal of medicine. 2001; 344: 1594-1602.*
12. *Pollo A., Amanzio M., Arslanian A., et al. Response expectancies in placebo analgesia and their clinical relevance. Pain. 2001; 93: 77-84.*
13. *Levine J.D., Gordon N.C., Fields H.L. The mechanism of placebo analgesia. Lancet. 1978; 2: 654-657.*
14. *Benedetti F. The opposite effects of the opiate antagonist naloxone and the cholecystokinin antagonist proglumide on placebo analgesia. Pain. 1996; 64: 535-543.*
15. *Benedetti F., Amanzio M., Baldi S., et al. Inducing placebo respiratory depressant responses in humans via opioid receptors. The European journal of neuroscience. 1999; 11: 625-631.*
16. *Pollo A., Vighetti S., Rainero I., et al. Placebo analgesia and the heart. Pain. 2003; 102: 125-133.*
17. *Eippert F., Finsterbusch J., Bingel U., et al. Direct evidence for spinal cord involvement in placebo analgesia. Science. 2009; 326: 404.*
18. *Weeks R.E., Newman E. Behavioral factors in the placebo response. Neurological sciences : official journal of the Italian Neurological Society and of the Italian Society of Clinical Neurophysiology. 2011; 32 Suppl 1: S3-8.*
19. *Blease C. The principle of parity: the 'placebo effect' and physician communication. Journal of medical ethics. 2012; 38: 199-203.*
20. *Rief W., Bingel U., Schedlowski M., et al. Mechanisms involved in placebo and nocebo responses and implications for drug trials. Clinical pharmacology and therapeutics. 2011; 90: 722-726.*
21. *Page L.A., Wessely S. Medically unexplained symptoms: exacerbating factors in the doctor-patient encounter. Journal of the Royal Society of Medicine. 2003; 96: 223-227.*

Instructions for EJM authors

General

The instructions that follow have several purposes. First, we want to make life easy for you, the authors, and for the editors and peer reviewers, the layout (prepress) people, and the journal readers.

The Guidelines for the Storyline will help you to organize your article in a logical, credible and readable way. This will help you - it tells you what goes where—and, thus, save you time. It will help the editors and peer reviewers—they will easily see the credibility and relevance of your work— and, thus, save them from writing rejection letters. And, it will help readers to quickly and easily read and understand your work and see its value.

The section entitled Formatting Instructions will help you as well; the basic idea is to keep the formatting as simple as possible, so you can focus on content and not get involved with layout. The language editor and the prepress people will also be able to more efficiently do their jobs. Please follow these instructions.
Please be aware that we will have to return papers that do not conform to these instructions to the authors.

What you can enter

Research articles - Research articles describe one study or analysis, usually from a fourth-year elective research project or one of the masters programs. Number of words: max. 3500 + 4 figures or tables.
Extended abstracts - Extended abstracts consist of a condensed presentation of complete final or temporary results of a study. Number of words: 350 words + 1 figure or table.
Reviews - Second year students can submit their review written in the second year elective course. Number of words: 1500 + 3 figures or tables.
Opinion papers - These are papers that reflect the opinion of the author on a scientific topic. The author should be clear where evidence ends and personal opinion starts. A paper typically has a length of about 1000 words.
Comments - In this section editors, or faculty staff, as well students are invited to write a short critical comment on a paper, putting it into perspective for a broader medical public readership. Number of words: 350.
Letters to the editor - The editorial board encourages students to write a letter to the editor to comment on published papers, or on the journal in general. These will be published on the website of the journal. Letters should not exceed 200 words and may be abbreviated by the editor.
Research projects - Here researchers or teachers describe ongoing research projects at the Erasmus MC University Medical Centre for which they want to invite students to participate. Number of words: 350.

The review process

Papers may be submitted to the editorial office. Please indicate which author will act as corresponding author. We expect this author to maintain contact with the other authors and to speak and decide on their behalf.

Each paper will be assigned to a team consisting of a managing editor and an associate editor. Each submitted paper will be checked for compliance with the author instructions. If this is not the case, the paper may be returned to the author.

When the paper is taken into review, it will be sent out to two external reviewers, a student and a staff member of Erasmus MC. Based upon these reviewers comments, their recommendations and the opinion of the editorial team, a decision will be made: reject, major revision, minor revision, accept with or without minor changes.

The paper will then be returned to the corresponding author, along with the recommendation. We try to return papers within 3 weeks after submission.
When a paper is rejected, it cannot be resubmitted, but we encourage resubmissions when we recommend major or minor changes to a paper. Resubmitted paper will be reviewed again by the same reviewers and editorial team.

Before a paper can be accepted for publication, we will need a statement that the staff member that supervised your work agrees with the submission of your paper. Moreover, we need a signed Copyright Transfer Agreement (CTA) and a signed Conflict of Interest statement. When your research project involves patients or volunteers, we need a statement in the paper that the research protocol has been reviewed by a Medical Ethics Committee. Failure to provide this information at an early stage of the submission may impair the review process.

When a paper is accepted for publication, it will often be forwarded to our language editing and restructuring editors. They will each in turn give recommendations and ask the author adapt the paper accordingly. When this phase is completed, the paper will be forwarded to the publisher. Page proofs will be sent to the author for a final check.

Formatting instructions

Entry format - Papers should be submitted by email, to ejm@erasmusmc.nl. Word 2003 files are preferred for the initial submission. The file should include all figures and tables.
Title page - The title page should clearly identify the authors, the institute where the research project was carried out, as well as the staff member who supervised the project. The corresponding author name (first name and family name), email address, student id, should be clearly indicated. In case of multiple authors, state functions and departments only in superscript in alphabetical order. Example:
First name A.G. Family name[a] and First name W.F. Family name[a]
Supervisor: First name R. Lastname[b]
[a] Medical students, Erasmus MC University Medical Center Rotterdam, the Netherlands
[b] Dept. of Internal Medicine, Erasmus MC University Medical Center Rotterdam, the Netherlands
Correspondence: First name A.G. Family name, email: Firstname-Familyname@me.com.

Structure - Please use the following sections in all papers (except in comments and opinion papers): Abstract, Introduction, Methods, Results, Discussion, References, Tables, Figures.
References - Number references in order of appearance. References should have the following format:
Rothwell, P. M. Medical and surgical management of symptomatic carotid stenosis. Int J Stroke. 2006; 1: 140-149. (I.e. year;vol:ppp-ppp) In case of more than 3 authors, name the first 3 and insert "et al.". Limit the number of references to 30. References should appear in the text as follows: "… treatment is of proven benefit.[1]"
Tables and figures - Tables and illustrations (both numbered in Arabic numerals) should be prepared on separate pages. Number tables and figures separately and consecutively. Tables require a heading and figures a legend, also prepared on a separate page and should be formatted with a text editor (example). Figures should be submitted electronically. B/w half-tone and color illustrations must

have a final resolution of 300 dpi after scaling, line drawings one of 800-1,200 dpi (jpg and tiff is an acceptable format). Please note that all color-figures will be converted to gray tones. Please adapt graphs to suit this format, i.e. make use of dotted and dashed lines and hatched bars instead of colored items.. The final submission should contain figures as JPG or TIFF files.

Page layout
- Standard margins
- no headers or footers
- no columns
- left align (ragged right)
- font: 12pt Arial
- single line spacing
- main headings 14 pt bold; subheading 12 point italic
- indent every paragraph, except after headings, tables, bulleted lists or figures

Other formatting
- number all tables and figures sequentially
- place tables and figures at the end of article; insert captions at correct locations in body text
- no text boxes
- no footnotes or end notes
- do not submit figures with text as drawing objects (they cannot be edited)
- limit the use of italics and do not use italics for simple emphasis; do not italicize quotations; quotation marks are sufficient
- do not use italics for commonly understood Latin expressions such as “in vitro”
- use italics for other foreign words, such as expressions in Dutch
- no “sub-paragraphs”
- no hyphenation (afbreking)

Language
US English spelling and punctuation

Advice to the reviewers of EJM

for the convenience of our future contributors and our readers, we publish here the Advice we give to our reviewers.

In the process of reviewing a paper, please refer to the following points:

- Your first step should be to evaluate your relationship with the authors. To ensure the credibility of the process, reviewers should not have a conflict of interest with the authors. If this is a case, the paper should be appointed to other reviewers. Please keep us informed whether conflict of interest is an issue for you as an appointed reviewer.
- Is this work relevant and interesting for EJM?
- Are the objectives appropriate and clearly stated?
- Are the data valid?
- Are the conclusions valid and properly supported?
- Is the already existing work described adequately?
- Paper structure/organization; is this logical?
- Does abstract clearly convey meaning of the paper?
- Is the paper well written and can be easily understood? (Please keep in mind that students don't have the experience to reed throughout the paper very quickly and to understand everything in a research paper at the first glance)
- Are all sections really needed, or could they be shortened?
- Is the science reliable? Please, be aware of ethical issues such as plagiarism!

Comments should be detailed and specific. Mentoring the authors includes helping authors improve their paper under review even if these papers will/could not be accepted for publication in our journal. By careful reviewing, you will help improving the quality of papers published elsewhere too. Avoid vague complaints and provide appropriate citations if authors are unaware of the relevant work.

Please consider a manuscript received for reviewing as a confidential document and do not discuss the content of this paper with others. To maintain the validity of this process, you should never contact the authors about the paper under review.

The review process serves two important goals: providing guidance to the authors to improve the quality of their paper, and providing the editor or editorial board with valuable recommendations regarding the acceptance or rejection of the peer-reviewed papers (along the whole spectrum of major revision- minor revision-rejection). So it is important that you give comments to the authors, and to the editor in separate sections. Please use the provided form, because this makes life easier for you, the editor and the authors.

EJM is committed to rapid editorial decisions and publication. We request that reviewers return their comments within the time indicated at invitation. If any unanticipated difficulties arise that may prevent you from submitting the review on time, contact us by sending an email to the editor at erasmusjournalofmedicine@gmail.com, or the editorial office at ejm@erasmusmc.nl. You are welcome to contact us if you have any questions.

For more information about guidelines for the review process, please visit our website: www.erasmusmc.nl/ejm. We also recommend you to view the presentations of the EJM workshop on our website. Here you can find instructions about how to scan through a paper and grab its essence, and how to structure your comments to the authors and to the editor.

October 2013, Editorial board of Erasmus Journal of Medicine.

Erasmus Journal of Medicine