Quality of care in rheumatology

Translating evidence into practice

Nienke Lesuis

Quality of care in rheumatology

Translating evidence into practice

Nienke Lesuis

Colofon
ISBN: 978-90-9029850-4
Lay-out and printing: Nicole Nijhuis (Gildeprint, Enschede)
Cover illustration: Aikaterini Chatzidionysiou
Cover design: Ruben Janssen (Burorub grafisch ontwerp & illustratie, Nijmegen)

Publication of this thesis was financially supported by the Sint Maartenskliniek and the Radboud University Nijmegen, the Netherlands.

Quality of care in rheumatology

Translating evidence into practice

Proefschrift

ter verkrijging van de graad van doctor
aan de Radboud Universiteit Nijmegen
op gezag van de rector magnificus prof. dr. J.H.J.M. van Krieken,
volgens het besluit van het college van decanen in het openbaar te verdedigen op
maandag 3 oktober 2016, om 16.30 uur precies.

door

Nienke Lesuis
geboren op 21 september 1986
te Oosterhout (Noord-Brabant)

Promotoren:	Prof. dr. M.E.J.L. Hulscher
	Prof. dr. R.F. van Vollenhoven (Karolinska Instituut, Zweden)
Copromotor:	Dr. A.A. den Broeder (Sint Maartenskliniek)
Manuscriptcommissie:	Prof. dr. J.W.A. Smit
	Prof. dr. T.P.M. Vliet-Vlieland (Leids Universitair Medisch Centrum)
	Prof. dr. P. Verschueren (Universitair Ziekenhuis Leuven, België)

CONTENTS

GENERAL INTRODUCTION

PROLOGUE

This thesis deals with quality of care in rheumatology, focussing on the potential gap between evidence and daily practice of two important aspects of daily care in rheumatology: the use of laboratory tests in the diagnostic process of rheumatic diseases and adherence of rheumatologists to optimal care recommendations in the treatment of rheumatoid arthritis (RA). The behaviour of rheumatologists will be the central focus of both themes and this introduction will deal with the general background of this thesis whereas a specific introduction to both themes will be provided before the start of each theme.

WHAT IS OPTIMAL QUALITY OF CARE?

Quality of care in itself is a rather abstract term, but more practical descriptions do exist. One of the most used descriptions, developed around 1980 by Donabedian, distinguishes structures, processes and outcomes of care[1]. The structure of care describes aspects of the setting in which care is delivered, such as the number of rheumatologists or the presence of a treatment protocol. Next, the process of care describes the actions of the health care professionals, for example, whether the protocol is followed. Finally, the outcome reflects the effect of the given care in terms of mortality, morbidity and health status. It is believed that more desirable outcomes are obtained if the structure of care provides the opportunity to deliver the most optimal care processes.

Around 1990, the Institute of Medicine (IOM) defined quality of care as 'the degree to which health services for individuals and populations increase the likelihood of desired health outcomes and are consistent with current professional knowledge'[2]. This definition incorporates two different aspects of quality of care: are current standards of care adhered to by health professionals and does this improve patients outcomes. The last aspect is also the domain of clinical trials, for example testing the effect of new drugs on clinically important outcomes. This means that standards of care continually change as trials provide new insights on the best care that should be provided to patients. The next step, professionals adhering to standards of care, will be the focus of this thesis.

In addition to the definition of quality of care, the IOM also formulated six criteria that pertain to quality of care. Care should be i) safe; ii) effective; iii) patient-centred; iv) timely; v) efficient; and vi) equitable[2]. In recent years 'transparency' and 'verifiability' are often used as additional criteria of quality of care and in its most recent quality of care statement the Dutch Rheumatology associations has incorporated all seven criteria[3].

IS QUALITY OF DAILY CARE CURRENTLY OPTIMAL?

The literature on the two themes of this thesis suggests that rheumatologists' behaviour regarding test ordering and guideline adherence is not yet optimal[4;5]. With regard to diagnostic tests, clinicians in general often order too many tests (test overuse)[4], whereas recommendations for all kinds of diseases are often not applied (non-adherence to guidelines)[6]. Although not extensively studied within rheumatology, existing studies on this topic suggest the same pattern[5;7-9]. This implies that also in rheumatology a gap exists between evidence and daily practice, and therefore the first question this thesis aims to answer is 'do rheumatologists provide evidence-based care in daily practice?'.

HOW CAN WE IMPROVE QUALITY OF DAILY CARE?

If a gap between evidence and practice exists, the next step would be to bridge this gap by finding effective interventions to close this gap. Before commenting on different effective interventions, it is first important to realize what factors influence test ordering behaviour and guideline adherence of physicians as this information is needed to develop effective intervention strategies[10;11]. This subject has been extensively studied outside rheumatology, leading to different checklists, frameworks and taxonomies. With regard to guideline adherence, these different studies have been summarized in a systematic review by Flottorp *et al*, providing a comprehensive checklist of 54 determinants in seven domains (guideline factors; health professional factors; patient factors; professional interactions; incentives and recourses; capacity for organizational change; social, political and legal factors)[12]. For test ordering behaviour a similar review exists, classifying determinants into five categories (diagnostic, therapeutic and prognostic, patient-related, doctor-related, and policy and organization-related factors)[13].

These kind of checklists can be used when planning an intervention in order to choose the one that best fits with the specific local situation[11]. Many different types of interventions exist and a useful overview is provided by the Effective Practice and Organization of Care (EPOC) group. This Cochrane review group is specialized in undertaking reviews on all types of interventions that aim to improve health professional practice. According to their taxonomy 19 different types of interventions targeted at healthcare workers can be recognized[14]. Of those interventions, education, audit and feedback, and reminders are much used interventions to improve care. According to the different EPOC reviews on those three types of interventions, they result in a small to moderate improvement of the desired behaviour[15-18]. Interventions specifically tailored to the local situation have also been subject of a EPOC review concluding that they are more effective than non-tailored interventions[11].

All in all, test ordering behaviour and guideline adherence of physicians can be influenced by a large array of factors and many different types of interventions exist to improve quality of care. Unfortunately, these types of studies are scarce within rheumatology despite evidence that quality of care in rheumatology is not always optimal. Therefore, the second and third question of these thesis are 'what factors influence whether evidence-based care is provided?' and 'how can the provision of evidence-based care be improved?' respectively.

AIM AND OUTLINE OF THIS THESIS

In summary, this thesis aims to describe current quality of care with regard to laboratory test use and RA guideline adherence of rheumatologists, also aiming to explore underlying determinants and assess the effectiveness of different interventions to further improve quality of rheumatologic care.

These aims are incorporated in the two themes of this thesis: theme 1 comprises the **chapters 1** to **6** and will describe overuse of different laboratory tests commonly used by rheumatologists and possible interventions to counter this overuse. The **chapters 7** to **10** form the second theme and they describe current guideline adherence in RA and potential methods to enhance uptake of guidelines in daily practice. Both themes start with a separate introduction of the topics (**chapter 1** and **chapter 7**), also including a more specific outline of the content of these chapters.

REFERENCES

(1) Donabedian A. The quality of care. How can it be assessed? *JAMA* 1988; 260(12):1743-1748.

(2) Committee of quality of health care in America. Crossing the quality chasm: a new health system for the 21st century. National Academy press 2001. Available at http://www.nap.edu/read/10027/chapter/1#xix

(3) Den Broeder AA. Memo strategie kwaliteitsbeleid Nederlandse Vereniging van Reumatologie 2014.

(4) Zhi M, Ding EL, Theisen-Toupal J, Whelan J, Arnaout R. The landscape of inappropriate laboratory testing: a 15-year meta-analysis. *PLoS One* 2013; 8(11):e78962.

(5) MacLean CH, Louie R, Leake B, McCaffrey DF, Paulus HE, Brook RH et al. Quality of care for patients with rheumatoid arthritis. *JAMA* 2000; 284(8):984-992.

(6) McGlynn EA, Asch SM, Adams J, Keesey J, Hicks J, DeCristofaro A et al. The quality of health care delivered to adults in the United States. *N Engl J Med* 2003; 348(26):2635-2645.

(7) Kiely P, Williams R, Walsh D, Young A. Contemporary patterns of care and disease activity outcome in early rheumatoid arthritis: the ERAN cohort. *Rheumatology (Oxford)* 2009; 48(1):57-60.

(8) Harrold LR, Harrington JT, Curtis JR, Furst DE, Bentley MJ, Shan Y et al. Prescribing practices in a US cohort of rheumatoid arthritis patients before and after publication of the American College of Rheumatology treatment recommendations. *Arthritis Rheum* 2012; 64(3):630-638.

(9) Benhamou M, Rincheval N, Roy C, Foltz V, Rozenberg S, Sibilia J et al. The gap between practice and guidelines in the choice of first-line disease modifying antirheumatic drug in early rheumatoid arthritis: results from the ESPOIR cohort. *J Rheumatol* 2009; 36(5):934-942.

(10) Morgan DJ, Brownlee S, Leppin AL, Kressin N, Dhruva SS, Levin L et al. Setting a research agenda for medical overuse. *BMJ* 2015; 351:h4534.

(11) Baker R, Camosso-Stefinovic J, Gillies C, Shaw EJ, Cheater F, Flottorp S et al. Tailored interventions to address determinants of practice. *Cochrane Database Syst Rev* 2015; 4:CD005470.

(12) Flottorp SA, Oxman AD, Krause J, Musila NR, Wensing M, Godycki-Cwirko M et al. A checklist for identifying determinants of practice: a systematic review and synthesis of frameworks and taxonomies of factors that prevent or enable improvements in healthcare professional practice. *Implement Sci* 2013; 8:35.

(13) Whiting P, Toerien M, de S, I, Sterne JA, Dieppe P, Egger M et al. A review identifies and classifies reasons for ordering diagnostic tests. *J Clin Epidemiol* 2007; 60(10):981-989.

(14) Effective Practice and Organisation of Care (EPOC). EPOC Taxonomy 2015. Oslo: Norwegian Knowledge Centre for the Health Services. Available at: http://epoc.cochrane.org/epoc-taxonomy

(15) Ivers N, Jamtvedt G, Flottorp S, Young JM, Odgaard-Jensen J, French SD et al. Audit and feedback: effects on professional practice and healthcare outcomes. *Cochrane Database Syst Rev* 2012; 6:CD000259.

(16) Shojania KG, Jennings A, Mayhew A, Ramsay CR, Eccles MP, Grimshaw J. The effects of on-screen, point of care computer reminders on processes and outcomes of care. *Cochrane Database Syst Rev* 2009;(3):CD001096.

(17) Forsetlund L, Bjorndal A, Rashidian A, Jamtvedt G, O'Brien MA, Wolf F et al. Continuing education meetings and workshops: effects on professional practice and health care outcomes. *Cochrane Database Syst Rev* 2009;(2):CD003030.

(18) O'Brien MA, Rogers S, Jamtvedt G, Oxman AD, Odgaard-Jensen J, Kristoffersen DT et al. Educational outreach visits: effects on professional practice and health care outcomes. *Cochrane Database Syst Rev* 2007;(4):CD000409.

"The lab test results are back — now it’s time to roll the dice."

Theme 1

Use of diagnostic laboratory tests in rheumatic diseases

"The lab test results are back — now it's time to roll the dice."

1

INTRODUCTION TO THEME 1

Use of diagnostic laboratory tests in rheumatic diseases

WHAT IS OPTIMAL DIAGNOSTIC TEST USE?

The use of laboratory tests during the diagnostic process of patients with rheumatic complaints is common practice among rheumatologists and physicians in general. Often the tests used are perceived as 'perfect' and its results seen as 'the truth'. Unfortunately, this is often not the case and the diagnostic value of a laboratory test depends on many factors. To begin with, the test should be of good technical quality and be able to discriminate between healthy and sick individuals, or between clinically important outcomes. This is often expressed as the sensitivity and specificity of a test. The sensitivity of a test is the percentage of individuals with a certain disease, having a positive test result; whereas the specificity is the percentage of individuals without the disease, having a negative test result. A perfect test would have a sensitivity and specificity of 100%, but no such tests exist. Thus false-negative (negative test when individual has a disease) and false-positive (positive test in healthy individual) will inevitably occur.

Besides no test having a 100% sensitivity and specificity, there is another problem with these test characteristics: they do no tell what a physician needs to know. Sensitivity and specificity assume that it is already known if an individual has a disease or not. But for a physician seeing a patient, the question is the other way around: does the positive or negative test in my patient means that he has or has not the disease? Exactly this question can be answered by calculating the predictive value of a test. In contrast to the sensitivity and specificity, the positive and negative predictive value (PPV and NPV respectively) take into account the pre-test probability of the disease in question. The pre-test probability on disease is an estimation of the chance for an individual person to have the disease, based on information gathered before performing the test. For healthy individuals the pre-test probability is similar to the prevalence of the disease in the general population, whereas disease-specific symptoms can increase this chance.

The effect of pre-test probability on the interpretation of lab results is illustrated in table 1. In this table two scenarios are described, using anti-CCP testing (sensitivity and specificity of 90%) for rheumatoid arthritis (RA) as an example. As can be seen in table 1, the PPV differs between the scenarios (90% in scenario 1 and only 32% in scenario 2). As the sensitivity and specificity in both scenarios are similar, the differences are caused by the differences in pre-test probability on RA between the scenarios (50% and 5% in scenario 1 and 2 respectively). This principle is known as Bayes Theorem and provides a mathematical framework for analyzing how diagnostic probability is influenced by pretest probability, the test characteristics (sensitivity/specificity), and the outcomes of the test. Bayes theorem has important clinical consequences. In an optimal situation (scenario 2), a physician can trust the laboratory result to give a fairly definitive answer on the question 'has my patient RA?'.

Table 1: Influence of pre-test probabilities of disease on the diagnostic value of a test

Scenario 1 (pre-test probability on rheumatoid arthritis = 50%)			
	RA present	**RA absent**	
Test positive	450	50	500
Test negative	50	450	500
	500	500	1000
Positive predictive value	450 / 500 = 90%		
Negative predictive value	450 / 500 = 90%		
Scenario 2 (pre-test probability on rheumatoid arthritis = 5%)			
	RA present	**RA absent**	
Test positive	45	95	140
Test negative	5	855	860
	50	950	1000
Positive predictive value	45 / 140 = 32%		
Negative predictive value	855 / 860 = 99%		

However, in the more realistic situation (scenario 1) the physician can still not answer this question as the chance on RA in his patient is only 32% if the test result is positive.

In addition to the positive and negative predictive value, a good laboratory test should also have additional value. This means that the chance on disease after the test (PPV) has to be substantially higher or lower than the chance on disease before the test (pre-test probability). This is the case in both scenarios of the RA example, although the absolute PPV of 32% in scenario 2 still limits the actual diagnostic value. The criterion of additional value also implies that diagnostic testing in scenarios with a very high or very low pre-test probability on disease is not useful as the pre- and post-test probabilities will be similar. In general it is recognized that diagnostic tests have the most optimal PPV and NPV if the pre-test probability on disease lies between 30 and 60%[1].

Finally, a good laboratory tests also needs to have consequences for treatment or prognosis and needs to be cost-effective. A test which is followed by the same decision regardless of the test result is essentially a wasted test. Often this has to do with a too high or low pre-test probability on disease, leading to no or not enough additional value of the test to have consequences for treatment or prognosis.

All in all it can be concluded that perfect diagnostic laboratory tests do not exist and that physicians need to take into account the uncertainty associated with the use of these tests. This is a huge challenge for many physicians that will be described next.

IS THE USE OF DIAGNOSTIC LABORATORY TESTS IN DAILY PRACTICE CURRENTLY OPTIMAL?

That correct use of diagnostic laboratory tests constitutes a challenge for many physicians is reflected in studies on inappropriate use of laboratory tests. In the most recent meta-analysis on this topic it was found that around 20% of the laboratory tests can be classified as overutilization[2]. Keeping the previous example on RA in mind, overuse will mimic the second scenario (low pre-test probability) leading to a higher rate of false-positives. Furthermore, inappropriate testing leads to higher costs and a higher patient burden due to uncertainty and additional testing, making this a real and significant problem in medicine.

Assuming that rheumatologists are not very different from other medical specialists with regard to the use of diagnostic laboratory tests, overutilization of laboratory testing will probably also present in rheumatology care. Although no review on the use of the full range of laboratory tests in rheumatology is available, some studies on specific laboratory use in rheumatology are present. Those mainly focus on anti-CCP testing or immunologic laboratory tests such as Antinuclear Antibodies (ANA)[3-5]. Especially the latter test has gained extra attention within rheumatology in recent years, confirming that ANA testing is commonly overused by rheumatologists[7-11]. The Choosing Wisely campaign, launched in 2012 by the ABIM foundation (Advancing Medical Professionalism to Improve Health Care; United States of America), seems to have had a large role in this[6]. This campaign aims to 'advance a national dialogue on avoiding wasteful or unnecessary medical tests, treatments and procedures' by publishing top five lists of 'things that physicians and patients should question' (www.choosingwisely.org). As a result, such a list was published in 2013 by the American College of Rheumatology (ACR), including ANA testing as 'a thing to question'[6]. However, rheumatologists use far more laboratory tests than ANA alone and for those tests much less is known about their use in daily practice, although it is likely that overuse is also present for those tests seeing the evidence from other specialties.

WHAT FACTORS INFLUENCE DIAGNOSTIC LABORATORY TEST ORDERING IN DAILY PRACTICE?

As ordering diagnostic test is such an integral part of healthcare, many studies have tried to identify determinants that influence the use of laboratory tests. A review from 2007 groups these determinants into diagnostic, therapeutic and prognostic, patient, physician and policy-related factors[12]. A second review focussing on physician-related determinants only, divided these determinants into modifiable and non-modifiable factors. In this last review practice location, practice setting, age, gender and specialisation of the physician were identified as non-modifiable determinants. Determinants such as physician experience and

knowledge were found to be modifiable determinants[13]. Despite these reviews, the exact relation between determinants and laboratory testing often remains unclear as different studies have different results. For example, a higher physician age can both lead to fewer or more laboratory tests ordered[12]. Unfortunately, determinant studies on laboratory test overuse in rheumatology are virtually non-existent, so insight is needed in the relation between potential determinants and laboratory testing by rheumatologists.

HOW CAN DIAGNOSTIC LABORATORY TEST USE IN DAILY PRACTICE BE IMPROVED?

Many different studies have tried to decrease laboratory test overuse by performing different types of interventions. The most recent systematic review on this topic classifies these interventions into educational interventions, audit and feedback, system-based interventions such as Clinical Decision Support Systems (CDSS), and incentive or penalty interventions[14]. Almost 30% of the interventions were multifaceted with educational, audit and feedback, and a system-based intervention being the most frequently performed combination of interventions. This review observed a large variation in the effectiveness of interventions, with effects on laboratory test use ranging between 99.7% reduction to a 27.7% increase in test volume. When looking at the separate intervention categories, educational interventions had the highest median relative reduction in test volume (34.5%, interquartile range (IQR) 16.5% to 49.0%). Audit and feedback or system-based interventions had similar median relative reduction rates (22%, IQR 8.6% to 34.6% and 22.2%, IQR 3.6% to 68.3% respectively), while incentive or penalty interventions only gave a relative reduction of 5.8%. Finally, it seemed that multifaceted interventions were more effective than single-component interventions with relative reductions of 32.7% and 21.4% respectively. Based on these results this review concludes that all type of interventions can reduce laboratory test ordering, but the effect range is large and much heterogeneity between studies is present[14]. Not many intervention studies to decrease laboratory test use within rheumatology exist. A few intervention studies on ANA overuse have been performed, all introducing some kind of clinical guideline or algorithm to prevent second line testing if the ANA result is negative. These studies reported positive effects – i.e. a decrease in the number of second line tests - whereas specificity of the tests increased[15-17].

OUTLINE OF THEME 1: USE OF DIAGNOSTIC LABORATORY TESTS IN RHEUMATIC DISEASES

As described in this chapter, studies on diagnostic laboratory test overuse amongst rheumatologists are still lacking although initiatives to counter overuse, such as the Choosing Wisely campaign, are employed. In order to provide more insight into this topic, the first part of this thesis will describe laboratory test overuse, explore determinants of this overuse and assess the effectiveness of different interventions to counter overuse.

Chapter 2 and **3** both focus on Antinuclear Antibody (ANA) testing as this is a commonly used test by rheumatologists despite the limited place it should have within rheumatology. In **Chapter 2** we describe the results of a simple intervention (education and feedback) on the ANA use by rheumatologists in three different hospitals in the Netherlands. **Chapter 3** is a continuation of the previous chapter, now exploring various determinants of ANA use by rheumatologist. This is done using both a quantitative (questionnaire study on the influence of personality, thinking styles, cognitive bias and numeracy) and a qualitative approach (focus group meeting with rheumatologists).
In **chapter 4** a non-peer reviewed article is presented, describing the Dutch version of the Choosing Wisely campaign for rheumatology. As part of the Choosing Wisely campaign the American College of Rheumatology published a list of 'five things that physicians and patients should question' and this example was followed in 2014 by the Dutch Society of Rheumatology. The final Dutch top 5 list and its development are described in this chapter.
Chapter 5 provides a closer look on two other commonly used tests in rheumatology: creatine kinase (CK) and thyroid stimulating hormone (TSH). Both tests are often recommended in the diagnostic workup of fibromyalgia (FMS), but hardly any underlying evidence for these recommendations exists. Therefore we assessed the prevalence of abnormal CK and TSH, and the prevalence of related diagnoses in a cohort of patients with suspected FMS.
The final chapter, **chapter 6,** describes the effects of an intervention aimed at reducing use of diagnostic laboratory tests. Similar to CK and TSH testing, the evidence for widespread use of complement, cryoglobuline, gammaglobuline and M protein testing in rheumatology is scarce. To discourage the use of these tests, an automatic reminder was incorporated in the electronic health record, reminding the rheumatologist of the limited evidence for the use of these tests if they ordered one. Using routine laboratory tests such as C-reactive protein as a control group, the effect of this reminder is assessed.

REFERENCES

(1) Landewé RBM. Aanvullend onderzoek bij diagnostiek. In: Bijlsma JWJ, van Laar JM, editors. Leerboek reumatologie en klinische immunologie. Houten: Bohn Stafleu van Loghum; 2013. 45-56.

(2) Zhi M, Ding EL, Theisen-Toupal J, Whelan J, Arnaout R. The landscape of inappropriate laboratory testing: a 15-year meta-analysis. *PLoS One* 2013; 8(11):e78962.

(3) Suarez-Almazor ME, Gonzalez-Lopez L, Gamez-Nava JI, Belseck E, Kendall CJ, Davis P. Utilization and predictive value of laboratory tests in patients referred to rheumatologists by primary care physicians. *J Rheumatol* 1998; 25(10):1980-1985.

(4) Abeles AM, Abeles M. The clinical utility of a positive antinuclear antibody test result. *Am J Med* 2013; 126(4):342-348.

(5) Bhagat M, Sehra ST, Shahane A, Kwan M. Utility of immunologic testing in suspected rheumatologic disease. *Curr Allergy Asthma Rep* 2014; 14(1):405.

(6) Yazdany J, Schmajuk G, Robbins M, Daikh D, Beall A, Yelin E et al. Choosing wisely: the American College of Rheumatology's Top 5 list of things physicians and patients should question. *Arthritis Care Res (Hoboken)* 2013; 65(3):329-339.

(7) Mohammadi S, Shaik I, Chevli P, Gonzalez-Ibarra F, Sohini S, Acharya S et al. Improper use of Antinuclear Antibody (ANA) test can result in misdiagnosis, increased patient anxiety, and wasted health care resources [abstract]. *Arthritis Rheumatol* 2014; 66[S10]:S1-S1402.

(8) Bulbin D, Meadows A, Kelsey S, Harrison H, Denio A. Choosing subserologies wisely: an opportunity for rheumatologic healthcare resource savings [abstract]. *Arthritis Rheumatol* 2014; 66[S10]:S1-S1402.

(9) Bulbin D, Meadows A, Denio A, Kirchner H, Kelsey S, Harrison H. Do rheumatologists (and other specialists) practice what we preach? A study of serology ordering patterns with attention to subserologies when the Antinuclear Antibody by enzyme linked immunosorbent assay is negative; and the clinical significance of these positive subserology results [abstract]. *Arthritis Rheum* 2013; 65[S10]:S1-S1331.

(10) Sheth T, Alcid D. Choosing not so wisely: the tale of Antinuclear Antibody testing [abstract]. *Arthritis Rheumatol* 2014; 66[S10]:S1-S1402.

(11) Davis L, Goldstein B, Tran V, Keniston A, Yazdany J, Hirsh J et al. Applying Choosing Wisely: identifying inefficiences of Antinuclear Antibody subserology testing in a safety net health system [abstract]. *Arthritis Rheum* 2013; 65[S10]:S1-S1331.

(12) Whiting P, Toerien M, de S, I, Sterne JA, Dieppe P, Egger M et al. A review identifies and classifies reasons for ordering diagnostic tests. *J Clin Epidemiol* 2007; 60(10):981-989.

(13) Sood R, Sood A, Ghosh AK. Non-evidence-based variables affecting physicians' test-ordering tendencies: a systematic review. *Neth J Med* 2007; 65(5):167-177.

(14) Kobewka DM, Ronksley PE, McKay JA, Forster AJ, van WC. Influence of educational, audit and feedback, system based, and incentive and penalty interventions to reduce laboratory test utilization: a systematic review. *Clin Chem Lab Med* 2015; 53(2):157-183.

(15) Tampoia M, Fontana A, Di SF, Maggiolini P, Pansini N. Application of a diagnostic algorithm in autoantibody testing: assessment of clinical effectiveness and economic efficiency. *Clin Chim Acta* 2003; 333(2):181-183.

(16) Tampoia M, Brescia V, Fontana A, Zucano A, Morrone LF, Pansini N. Application of a combined protocol for rational request and utilization of antibody assays improves clinical diagnostic efficacy in autoimmune rheumatic disease. *Arch Pathol Lab Med* 2007; 131(1):112-116.

(17) Bonaguri C, Melegari A, Ballabio A, Parmeggiani M, Russo A, Battistelli L et al. Italian multicentre study for application of a diagnostic algorithm in autoantibody testing for autoimmune rheumatic disease: conclusive results. *Autoimmun Rev* 2011; 11(1):1-5.

"The lab test results are back — now it's time to roll the dice."

2

Choosing Wisely in daily practice: an intervention study on Antinuclear Antibody testing by rheumatologists

Nienke Lesuis
Marlies EJL Hulscher
Ester Piek
Hatice Demirel
Nicole van der Laan-Baalbergen
Inger Meek
Ronald F van Vollenhoven
Alfons A den Broeder

Arthritis Care Res (Hoboken) 2016;68:562-9. Doi 10.1002/acr.22725

ABSTRACT

Objective

To assess the effect of a simple intervention on Anti-Nuclear Antibody (ANA) test overuse by rheumatologists.

Methods

This was an explorative, pragmatic before and after controlled implementation study among rheumatologists working at three rheumatology departments of secondary and tertiary care centers in the Netherlands. The intervention was given in all study centers separately and combined education with feedback.

Six outcome measures describe the intervention effects: the ANA/new patient ratio (APR), difference with the target APR, percentage of positive ANA tests, percentage of repeated ANA testing, percentage of ANA associated diseases and APR variation between rheumatologists.

Outcomes were compared between the pre- and post-intervention period (both 12 months) using (multilevel) logistic regression or F-testing. Results are reported together for center 1 and 2, and separately for center 3 because ANA tests could not be linked to an individual rheumatologist in center 3.

Results

The APR decreased from 0.37 to 0.11 after the intervention in center 1 and 2 (odds ratio (OR) 0.19, 95%-confidence interval (95%-CI) 0.17 to 0.22, p-value <0.001) and from 0.45 to 0.30 in center 3 (OR 0.53, 95% CI 0.45 to 0.62, p <0.001). The percentage of repeated ANA requests in all centers and the APR variation center 1 and 2 decreased significantly. Only in center 3 the percentage of ANA associated diseases increased significantly.

Conclusion

A simple intervention resulted in a relevant and significant decrease in the numbers of ANA tests requested by rheumatologists, together with an improvement on three other outcome measures.

Trial registration number (ClinicalTrials.gov): NCT02409251

SIGNIFICANCE AND INNOVATIONS

- Overuse of laboratory testing is a widespread problem in medicine with an estimated 21% of laboratory tests being requested inappropriately
- Overuse of tests leads to an increase in health care costs and increased numbers of false positive results, causing unnecessary anxiety in both patients and doctors.
- The American College of Rheumatology recently named Anti-Nuclear Antibody (ANA) testing as one of the 'top five things to avoid' in the Choosing Wisely campaign
- A relatively simple intervention, teaching rheumatologists' how to correctly use ANA testing, can lead to a sizable reduction in the number of ANA requests

INTRODUCTION

The number of laboratory tests performed has steadily risen over the past years and it is estimated that 21% of laboratory tests are requested inappropriately.(1) This is a real and significant problem in medicine, as inappropriate testing leads to a higher patient burden due to uncertainty and additional testing, higher false-positive rates and higher costs.(1) In recent years this problem has gained more attention and the internationally expanding Choosing Wisely campaign is a good example of the effort taken to decrease overuse.(2) As part of the Choosing Wisely campaign the American College of Rheumatology (ACR) published a list of 'five things that physicians and patients should question'. The first item on this list is: 'do not test Anti-Nuclear Antibody (ANA) sub-serologies without a positive ANA and clinical suspicion of immune-mediated disease'.(3)

ANA testing is often used as a screening test for various rheumatic diseases. However, a large review on this subject concluded that the false-positive rate associated with ANA testing severely limits its usefulness as a screening test. ANA testing can be useful in case of a reasonable clinical suspicion on Systemic Lupus Erythematosus (SLE), Systemic Sclerosis (SSc), polymyositis (PM), dermatomyositis (DM), Mixed Connective Tissue Disorder (MCTD) and Sjögren Syndrome (SS).(4,5) Despite this advice, several publications that have described ANA and/or ANA sub serology use in clinical practice conclude that these tests are frequently overused.(6-11)

Combining current literature and the Choosing Wisely advice we argue that in order to avoid 'broad testing of auto antibodies' (3) one should first decrease ANA testing itself. For this reason we developed an intervention to improve ANA test requests, thus decreasing ANA overuse. This paper describes the effects of this intervention on ANA requests done by rheumatologists.

MATERIAL AND METHODS

Study design and participants

An explorative, pragmatic, before- and after-controlled implementation study on ANA overuse was performed among rheumatologists working at the outpatient clinic departments of different hospitals in the Netherlands. The rheumatologists of seven rheumatology departments of secondary and tertiary care centers were approached and asked to participate in this study. If a center decided to participate, all rheumatologists working at this center at study start were eligible for participation. Rheumatologists in training (trainees) were not eligible for participation as they often work only for a short period of time in the same hospital, leading to incomplete data. Furthermore, we could never be sure if the decision to order an ANA test was their own decision as trainees work under the supervision of a rheumatologist. Rheumatologists not giving their consent or not working the full pre- and post-intervention period as a rheumatologist at a participating hospital were excluded. Consent from all rheumatologists was sought at the moment of introduction of the study during regular staff meetings.

Although the intervention was aimed at rheumatologists, patient data were needed to assess the intervention effects. Therefore, all patients with an ANA test requested by a participating rheumatologist during the study period were included. Patients with an ANA test requested during a clinical admission were excluded.

The pre- and post-intervention period both lasted 12 months, with the time needed to prepare the intervention included in the pre-intervention period (see also 'statistical analysis and reporting of results').

When reporting this study, the Standards for Quality Improvement Reporting Excellence (SQUIRE) guidelines were followed.(12)

Intervention

The intervention was an improved version of an intervention used before in one of the study centers to optimize requesting of Magnetic Resonance Imaging scans of the sacroiliac joints. (13)

The intervention in this study consisted of a one-hour group session, combining an educational meeting with feedback (14), and was given separately in all participating centers. Six months after the intervention, a one-hour booster session was held. During both sessions, feedback on ANA testing was provided, followed by background information on ANA testing, a short recommendation on when to request an ANA test – or not - in daily practice and a target ANA/new patient ratio (APR) to reach after the intervention. The target APR was calculated by doubling the number of patients with an ANA-associated diagnosis in the pre-intervention period and dividing this by the number of new patients seen in the pre-intervention period.

This calculation follows Bayes theorem which assumes that a test performs optimally when the pre-test probability on disease is approximately 50% (see also 'outcome measures' for an explanation of Bayes theorem).
Both the intervention and the booster session were planned during regular meeting hours, in order to make it easy for rheumatologists to attend. Attendance was monitored during the sessions, and rheumatologists not present received an email with alternative times on which the session was held again. Although rheumatologists in training could not participate in this study, they were allowed to attend the intervention and booster session as part of their rheumatology training. All sessions were given by a resident of rheumatology (NL), with additional support from an experienced rheumatologist and epidemiologist (AdB). A complete description of the intervention and the PowerPoint slides used during the intervention can be found in appendices 1 and 2 respectively.

Outcome measures

The diagnostic value of a test in clinical practice depends primarily on the difference between pre- and post-test probabilities, the latter being expressed as the positive or negative predictive value of a test (PPV and NPV respectively). In contrast to the sensitivity and specificity of a test, the PPV and NPV are highly influenced by the pre-test probability of the disease. This application of 'Bayes theorem' results in a low diagnostic value of a test in case of a low pre-test probability of the disease of interest, even if the sensitivity and specificity are high (see appendix 3 for more a more elaborate explanation including example calculations).(15)
When applying Bayes theorem to clinical practice, overuse of a test is essentially characterized by suboptimal patient selection by the requesting physician. This results in a too low pre-test probability and consequently a low diagnostic value of the test. To estimate the pre-test probability for all individual patients intensive chart review would be needed. As this was deemed to labor intensive, we defined the following six, readily available, outcome measures that are all closely related to pre-test probability on disease: the ANA/new patient ratio (APR), the difference with the target APR, percentage of positive ANAs, percentage of repeated ANA testing within one year, percentage of ANA associated diseases and APR variation between rheumatologists (table 1). All these outcomes were measured on clinic level, but for the latter outcome an extra step was needed. In order to assess variation between rheumatologists, we first calculated the APR for all single rheumatologists in one center which was followed by calculating the mean APR out of these individual APRs. The standard deviation around this mean APR on clinic level was then used as the outcome measure to assess variation between rheumatologists. As a result the mean APR will differ slightly from the APR described as the first outcome measure.

***Table 1**: outcome measures*

Outcome measure	Calculation	Expected situation before the intervention (overuse)	Expected change after the intervention (less overuse)
APR (ratio)	Number of ANA tests requested divided by the number of new patients seen at the rheumatology outpatient clinic	High in comparison with the population at risk	Decrease (only ANA testing in the population at risk due to improved patient selection)
Difference with the target APR* (%)	Absolute difference between the actual and target APR divided by the target APR	High in comparison with the population at risk (i.e. an actual APR well above the target APR)	Decrease (only ANA testing in the population at risk due to improved patient selection)
Positive ANA tests (%)‡	Number of ANA tests with a positive result divided by the total number of ANA tests	Few positive tests, the percentage of positive ANA tests will approach the percentage of positive ANA tests in the healthy population	Increase (only ANA testing in the population at risk due to improved patient selection)
Repeated ANA testing within one year (%)Ξ	Number of patients with more than one ANA test done divided by the total number of patients with an ANA test done	High in comparison to the population at risk. Repeated ANA testing without a change in clinical picture is not recommended as the pre-test probability will be very high or low (4;6)	Decrease (repeated ANA testing only in those with a change in clinical picture, i.e. improved patient selection)
ANA associated diseases (%)	Number of patients with a diagnosis of SLE, SSc, PM/DM, MCTD or SS divided by the total number of patients	Few patients with an associated diagnosis, the percentage of associated diagnosis will approach the incidence in the normal population	Increase (only ANA testing in the population at risk due to improved patient selection)
APR variation between rheumatologists	Standard deviation around the mean APR	Variation between rheumatologists that cannot be explained by case mix differences	Decrease (more similar behavior between rheumatologists)

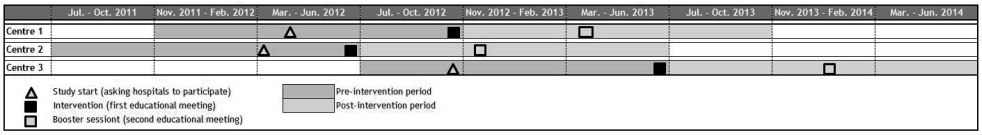
APR: ANA/patient ratio, ANA: Anti-Nuclear Antibody, SLE: systemic lupus erythematosus, SSc: systemic sclerosis, PM/DM: polymyositis/dermatomyositis, MCTD: mixed connective tissue disease, SS: Sjögren's syndrome. *Due to the differences in the number of patients with an ANA-associated diagnosis between the centers, the target differed between the centers and therefore this outcome will be reported for all centers separately. ‡Tests results reported as 'weakly positive' are regarded as negative. ΞAs we had no access to laboratory databases outside the participating centers, only repeated testing within one center could be assessed.

Study time frame

The study started in February 2012 with asking seven different hospitals to participate. In the hospitals that decided to participate, the study started between March and October 2012. The intervention in those hospitals took place consecutively between June 2012 and

June 2013, the booster session taking place six months thereafter from November 2012 to January 2014 (figure 1).

Figure 1: *study time frame per participating center*



Data sources

In order to assess the outcome measures, data were obtained from two different sources. First, laboratory databases from the participating hospitals were used to retrieve ANA data. These included the number of ANA tests performed, test results, test dates, gender and age of the patient in whom the ANA was tested, and the name of the rheumatologist who requested the ANA. However, this last variable proved not to be reliable in all participating centers due to organizational issues. Therefore, outcome measures requiring these data (APR variation) were omitted for those centers.

Secondly, information on patient diagnosis and number of new patients seen was obtained from administrative hospital databases. After combining both datasets locally, patient data were analyzed anonymously.

All patient data were retrieved from the relevant sources on three time points. Firstly, data on the twelve month period before study start was obtained from a participating center directly after the rheumatologists decided to participate (study start). This data was used to prepare the feedback provided to the rheumatologists during the intervention. Secondly, two months before the booster session data was collected to prepare the feedback on first three months of the post-intervention period. Thirdly, the final data on the full study period (24 months; pre- and post-intervention) to assess and analyze the outcome measures was obtained twelve months after the intervention (end of study). This means that both pre- and post-intervention data were obtained retrospectively.

Characteristics of the rheumatologist study population, including demographic (age, gender) and practice (work experience, patient contact, PhD) data were collected at study start.

Ethical approval

This study was presented to the local medical ethical board (Commissie Mensgebonden Onderzoek [CMO]), but according to Dutch Act on Medical Research Involving Human Subjects, this study did not need ethical approval (CMO reference number 2015-1653). All participating hospitals approved the study, and all participating rheumatologist gave consent. In addition, it was made clear to all participating rheumatologists that they could stop with this study at any time without providing a reason.

All patient data (consisting of only age, sex, ANA testing result, and administrative diagnosis) were retrieved within the hospital by matching two datasets locally, after which the data was anonymized. As this data cannot be traced back to an individual patient (no name, initials or other identifying information), no written informed consent was needed from the patients according to Dutch Data Protection Act. The study was registered at ClinicalTrials.gov (NCT02409251).

Statistical analysis & reporting of results

All analyses were done using STATA version 13. Depending on the type of variable descriptive statistics are presented as percentages with the accompanying absolute numbers or as means with standard deviations. All outcome measures, except for the difference with the target APR, were compared between the pre- and post-intervention period. The exception for the target APR was made because the difference between the actual and target APR was expressed as a percentage of the target APR (table 1). This made testing for statistical significance between the pre- and post-intervention difference difficult. Therefore we chose to deviate from the study protocol and only describe the results of this outcome measure without further statistical testing.

Because the intervention took place on rheumatologist level and some outcomes were measured at patient level, four out of six outcomes (APR, % positive ANA tests, % repeated ANA requests and % ANA associated diseases) had to be analyzed using mixed model multilevel logistic regression. However, in one center the ANA tests could not reliably be attributed to individual rheumatologists. This made it impossible to perform multilevel logistic regression analyses including the data from this center. Therefore it was decided to analyze (using logistic regression) and report those four outcome measures separately for this center. All results from the (multilevel) regression analyses are reported as odds ratios (OR) with the corresponding 95% confidence interval (95% CI) and p-value.

For the remaining outcome measure (practice variation) another analysis technique was chosen. Practice variation was defined as the standard deviation around the mean APR (based on the APR of individual rheumatologists). Testing of its statistical significance was done using the F-test (two-sided, $\alpha = 0.05$). All analyses were done on an intention-to-treat basis.

As mentioned before, the time needed to prepare the intervention was included in the pre-intervention period. During this preparation period the participating rheumatologists knew about the study, but the intervention was yet to come (see figure 1). As we did not know if this knowledge could already influence requesting behavior of the rheumatologists, a post-hoc sensitivity analysis was done with and without the preparation period included in the pre-intervention period.

A second post-hoc sensitivity analysis was done in order to assess whether the effects in the post-intervention period were different before and after the booster session.

RESULTS

Setting & participants

Of the seven hospitals approached, three decided to participate (43%). These included one general (center 1), one specialized (center 2) and one academic center (center 3). A total of 30 rheumatologists worked at these centers and 29 of them could be included (see figure 2). With 316 rheumatologists registered in the Netherlands, almost 10% of the total population of rheumatologists was included in this study.

One of rheumatologists in center 3 was lost to follow-up after the intervention due to a change of jobs. Attendance at the intervention and booster session was high, with only two rheumatologists from center 3 not attending the sessions for unknown reasons despite repeated invitations for an alternative time. In table 2 the baseline characteristics of the three participating centers and their rheumatologists are given.

***Figure 2**: study flow chart*

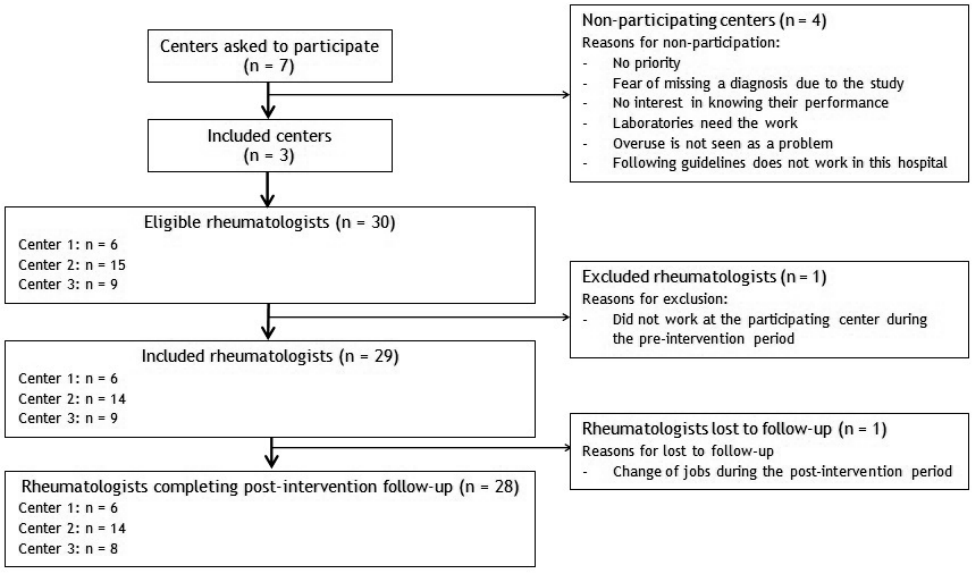


***Table 2**: characteristics of the study population*

	Centre 1	Centre 2	Centre 3
Type of center (hospital)	General hospital	Specialized hospital	University hospital
Included rheumatologists, no	6	14	9
Demographic and practice characteristics of included rheumatologists'			
Age, mean ±SD years	40.2 ± 4.3	48.8 ± 9.8	46.8 ± 8.5
Females, % (no./total no.)	83.3 (5/6)	43.0 (6/14)	66.7 (6/9)
Work experience as a rheumatologist, mean ±SD years	4.7 ± 2.7	12.2 ± 10.3	12.3 ± 7.3
Patient contact per week, mean ±SD no	60.8 ± 9.3	58.9 ± 17.0	49.7 ± 16.2
Completed or ongoing PhD, % (no./total no.)	33.3 (2/6)	71.4 (10/14)	77.9 (7/9)

Intervention effect on the ANA tests requested

The intervention resulted in a significant decrease of the APR in all centers, with the APR decreasing from 0.37 to 0.11 (OR 0.19 95% CI 0.17 to 0.22) in center 1 and 2; and from 0.45 to 0.30 (OR 0.53, 95% CI 0.45 to 0.62) in center 3. Also significant improvement was seen in the percentage of repeated ANA requests in all centers, the percentage of ANA associated diagnosis in center 3 and APR variation between rheumatologists in center 1 and 2. The percentage of positive ANA tests did not increase in any of the centers. Before the intervention all centers had an APR well above the target APR (absolute difference between +200% and +850% of the target APR). After the intervention the difference had decreased (+83% to +175%), but none of the centers reached the target APR. All results are described in table 3 and the APR variation between rheumatologists, including the target APR, is further illustrated in figure 3.

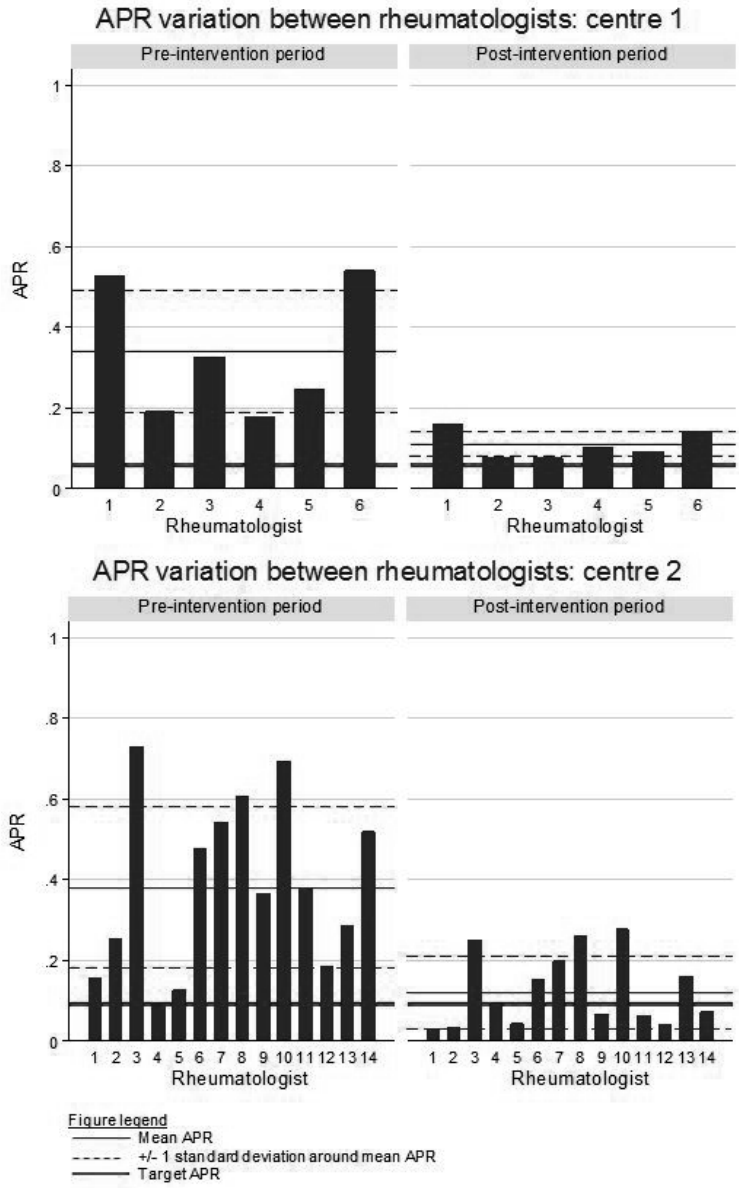
As mentioned in the methods section, two post-hoc sensitivity analyses were performed. No differences were observed compared to the primary analyses, except for the percentage of repeated ANA requests being lower in the post-intervention period after the booster session compared to the period before this session (1.7% versus 9.8%; OR 0.15, 95% CI 0.04 to 0.68).

***Table 3**: pre- and post-intervention results on the six outcome measures*

	Pre-intervention	Post-intervention	OR (95% CI)	P-value
APR (no./total no.)				
Center 1 and 2	**0.37 (1899/5150)**	**0.11 (475/4204)**	**0.19 (0.17 to 0.22)**	**<0.01**[‡]
Center 3	**0.45 (619/1364)**	**0.30 (341/1121)**	**0.53 (0.45 to 0.62)**	**<0.01**[†]
Difference with target APR, % (95% CI)				
Center 1 (target APR: 0.04)	+850% (775 to 900%)	+175% (150 to 250%)	n/a	n/a
Center 2 (target APR: 0.06)	+517% (483 to 550%)	+83% (67 to 117%)	n/a	n/a
Center 3 (target APR: 0.15)	+200% (180 to 220%)	+100% (80 to 120%)	n/a	n/a
Positive ANA tests, % (no./total no.)				
Center 1 and 2	24.6 (466/1896)	24.4 (115/471)	1.05 (0.82 to 1.36)	0.69[‡]
Center 3	37.0 (228/616)	37.7 (127/337)	1.03 (0.78 to 1.35)	0.84[†]
Repeated ANA requests, % (no./total no.)				
Center 1 and 2	**6.9 (122/1767)**	**1.3 (6/469)**	**0.15 (0.07 to 0.35)**	**<0.01**[‡]
Center 3	**9.0 (50/554)**	**3.7 (12/327)**	**0.38 (0.20 to 0.73)**	**<0.01**[†]
ANA associated diagnosis, % (no./total no.)				
Center 1 and 2	6.8 (127/1870)	8.2 (38/462)	1.38 (0.93 to 2.03)	0.11[‡]
Center 3	**17.1 (105/614)**	**25.2 (78/309)**	**1.63 (1.17 to 2.28)**	**<0.01**[†]
APR variation between rheumatologists* (mean APR)				
Center 1 and 2	**0.20 (0.37)**	**0.08 (0.12)**	**n/a**	**<0.01**[Δ]
Center 3	n/a	n/a	n/a	n/a

OR: odds ratio, 95% CI: 95% confidence interval. *Due to the quality of the data this could only be assessed for center 1 and 2. [‡]Mixed model multilevel logistic regression. [†]Logistic regression. [Δ]F-test, two-sided. Significant values are in bold.

***Figure 3**: APR of individual rheumatologists in center 1 and 2*



DISCUSSION

To our knowledge this is the first study attempting implementation of the ANA advice from the Choosing Wisely campaign by trying to improve rheumatologists' requesting behavior with a relative simple intervention. Our study shows a significant and relevant decrease in the numbers of ANA tests requested by rheumatologists in the post-intervention period. This effect was accompanied by improvement on three other outcomes on appropriate ANA testing (less repeated ANA requests, more ANA associated diagnosis and less APR variation between rheumatologists). As not all outcome measured changed, it seems that ANA requesting behavior has improved, but is not yet optimal.

Strengths & limitations

Besides being one of the first studies on this topic, the main strengths of this study are the inclusion of three different types of hospitals, the relatively large number of included rheumatologists, the use of a simple intervention, the adequate follow-up duration, and the inclusion of outcome measures based on the theoretical framework of Bayes theorem. These measures have the additional advantage of being easy to retrieve from hospital databases, thereby aiding future replication of our results and further implementation in clinical practice.

However, this study has some limitations. Firstly, in center 3, ANA tests could not be reliable attributed to a single rheumatologist within this center. As a result, no individualized feedback could be provided during the intervention and non-included health care providers (such as residents) could not be excluded from the dataset. This probably resulted in the lower intervention effects seen in center 3. Nevertheless, also in center 3 we found a significant and relevant intervention effect. Secondly, we only studied patients where an ANA test had been ordered, so we are unable to comment on the decision that was made in many other patients not to request an ANA test. The literature suggests there is considerable overuse of ANA testing rather than underuse, and therefore we chose to focus on the patients where such overuse could be present (i.e. where ANA testing had been done). Furthermore, assessing whether not requesting an ANA test was correct would call for very labor intensive chart review making this less feasible.

Thirdly, participation of hospitals was on a voluntary basis, making our conclusion about the intervention effects only generalizable to centers willing to participate in this kind of implementation projects. Also, within the included centers participation was voluntary and some selection bias could have occurred with the drop out of one rheumatologist in center 3. Again this bias would be small and lead to more conservative estimates.

Lastly, due to our study design we are not able to infer a definite causal relation between our intervention and the results afterwards because other events in the same time period

might have attributed to the observed results. However, we are not aware of any events during the study that could have influenced rheumatologists' ANA requesting behavior with this magnitude. The only exception could be the blinding of the rheumatologists, as they only did know about the study just before the intervention and not during the full study period. However, sensitivity analysis yielded no difference between including or excluding this period in the pre-intervention period.

Relation to other studies

As mentioned in the introduction, several groups have studied ANA overuse; however, we are only aware of two other intervention studies. In these studies, a diagnostic algorithm was implemented; according to the algorithm ANA sub serology was not done by the laboratory if the ANA was negative, even if both tests were requested simultaneously by the physician.(16;17) This led to a decrease in ANA (sub serology) testing in both studies. Notably, this diagnostic algorithm was already in use in all study centers for several years; despite this algorithm overuse was still present.

With regard to the effects observed in this study, we were surprised by the extend of APR decrease, especially since intervention effects in other studies on overuse of diagnostic tests are not often as large as ours.(18-20) A reason for the observed effect, especially in center 1 and 2, could be the presence of the second author (AdB) during the interventions. He is probably seen as the informal leader of the rheumatology department of both center 1 and 2 and he has played a crucial role in the close collaboration that exists between these two centers, which is known to aid implementation.(21)

Unfortunately, we can only compare two outcome measures, other than the APR, with other studies. These two studies have yielded similar results to our pre-intervention results with regard to repeated ANA testing and ANA-associated diagnoses. As this were not intervention studies, further comparison is not possible. Furthermore, it is noteworthy that not all measures in this study changed after the intervention. The percentage of positive ANA tests did not change in any of the centers, no center reached the target APR although they came closer, and only in center 3 the percentage of ANA associated diagnosis increased. These results may seem counterintuitive, especially the lack of increase in the percentage of ANA positive tests after the intervention. However, we think that this is (still) caused by suboptimal patient selection of the rheumatologists in combination with a relatively low specificity of the ANA test itself. Or in other words, although the extent of ANA overuse has decreased after the intervention, rheumatologists still request too many ANA tests and did not come close enough to the target APR to see an increase in the percentage of ANA positivity.

Practical implications & future research

Despite its limitations, we assume that our study accurately represents the beneficial effects of a simple intervention on ANA overuse in centers willing to change their behavior. Possible gains are less false-positive test results, less patient burden and less costs. Unfortunately, we were not able to perform a formal cost-analysis. However, if we assume that without the intervention the APR would have remained similar, around 1200 extra ANA tests would have been requested in the post-intervention period. As not only the costs from the ANA test itself (approximately 10 euro per test) are saved, but also costs from potential subsequent testing and treatment, it seems that savings gained with this simple and cheap intervention are substantial.

Despite our positive results and the possible gains, replication of our results would be warranted before this intervention is widely implemented. Given the fact that overuse is still present in the participating centers, improvements of our intervention could also be explored.

Although both the Choosing Wisely advice and this study were aimed at rheumatologists, decreasing ANA overuse can also be relevant for other specialties such as internal medicine, neurology and primary care physicians. The latter could be especially important as the setting in which ANA tests are requested most (primary care versus secondary or tertiary care), might differ between countries. Furthermore, overuse of laboratory tests is not limited to ANA tests, making it worthwhile to apply the same type of intervention also to other tests. This might help many more physicians to avoid doing unnecessary tests and to choose even more wisely.

ACKNOWLEGDMENTS

We would like to thank all participating rheumatologists at the different study centers for their willingness to attend the intervention and boosters sessions, and receive feedback on their own behavior. Furthermore, we would like to thank Arjen de Boer, Siep Windhorst, Patrick van Baaren, Gideon Lansbergen, Sietske Wijnberg and Gertrude van de Wiel for their effort to provide all necessary data. Reinier Akkermans we want to thank for his valuable help during the statistical analyses.

CONTRIBUTORS

All authors contributed to the study design. NL and AdB designed and executed the intervention in all participating hospitals. NL and EP participated in data collection. NL, AdB and MH participated in the data analysis. NL, AdB, MH and RvV assisted in the interpretation of the data. The manuscript was written by NL under the supervision of AdB, MH and RvV. All authors have revised the draft version of the manuscript, and read and approved the final version of the manuscript. NL is guarantor.

COMPETING INTEREST DECLARATION AND FUNDING

All authors have completed the Unified Competing Interest form at www.icmje.org/coi_disclosure.pdf (available on request from the corresponding author) and declare that RvV has received grants and personal fees from AbbVie, BMS, GSK, Pfizer, Roche, UCB, Biotest, Crescendo, Janssen, Lilly, Merck and Vertex outside the submitted work; all other authors reported no competing interests. This study received no external funding.

DATA SHARING STATEMENT

Patient level data, full dataset, technical appendix and statistical code are available at a reasonable request from the corresponding author. Consent from the patients was not obtained but the presented data are anonymized and risk of identification is low.

REFERENCES

(1) Zhi M, Ding EL, Theisen-Toupal J, Whelan J, Arnaout R. The landscape of inappropriate laboratory testing: a 15-year meta-analysis. PLoS One 2013; 8:e78962.

(2) Levinson W, Kallewaard M, Bhatia RS, Wolfson D, Shortt S, Kerr EA. 'Choosing Wisely': a growing international campaign. BMJ Qual Saf 2015; 24:167-74.

(3) Yazdany J, Schmajuk G, Robbins M, Daikh D, Beall A, Yelin E et al. Choosing wisely: the American College of Rheumatology's Top 5 list of things physicians and patients should question. Arthritis Care Res (Hoboken) 2013; 65:329-39.

(4) Solomon DH, Kavanaugh AJ, Schur PH. Evidence-based guidelines for the use of immunologic tests: antinuclear antibody testing. Arthritis Rheum 2002; 47:434-44.

(5) Bhagat M, Sehra ST, Shahane A, Kwan M. Utility of immunologic testing in suspected rheumatologic disease. Curr Allergy Asthma Rep 2014; 14:405.

(6) Abeles AM, Abeles M. The clinical utility of a positive antinuclear antibody test result. Am J Med 2013; 126:342-8.

(7) Mohammadi S, Shaik I, Chevli P, Gonzalez-Ibarra F, Sohini S, Acharya S et al. Improper use of Antinuclear Antibody (ANA) test can result in misdiagnosis, increased patient anxiety, and wasted health care resources [Abstract]. Arthritis Rheumatol 2014; 66:S1-S1402.

(8) Bulbin D, Meadows A, Kelsey S, Harrison H, Denio A. Choosing subserologies wisely: an opportunity for rheumatologic healthcare resource savings [Abstract]. Arthritis Rheumatol 2014; 66:S1-S1402.

(9) Sheth T, Alcid D. Choosing not so wisely: the tale of Antinuclear Antibody testing [Abstract]. Arthritis Rheumatol 2014; 66:S1-S1402.

(10) Davis L, Goldstein B, Tran V, Keniston A, Yazdany J, Hirsh J et al. Applying Choosing Wisely: identifying inefficiences of Antinuclear Antibody subserology testing in a safety net health system [Abstract]. Arthritis Rheum 2013; 65:S1-S1331.

(11) Bulbin D, Meadows A, Denio A, Kirchner H, Kelsey S, Harrison H. Do rheumatologists (and other specialists) practice what we preach? A study of serology ordering patterns with attention to subserologies when the Antinuclear Antibody by enzyme linked immunosorbent assay is negative; and the clinical significance of these positive subserology results [Abstract]. Arthritis Rheum 2013; 65:S1-S1331.

(12) Ogrinc G, Mooney SE, Estrada C, Foster T, Goldmann D, Hall LW et al. The SQUIRE (Standards for QUality Improvement Reporting Excellence) guidelines for quality improvement reporting: explanation and elaboration. Qual Saf Health Care 2008; 17:i13-i32.

(13) Vossen MH, den Broeder AA, Hendriks-Roelofs F, van der Heijde DM, Reijnierse M. Improvement in deployment of MRI of the sacroiliac joints in patients suspected for spondyloarthritis using a targeted intervention: a case study. Rheumatology (Oxford) 2013; 52:933-8.

(14) Effective Practice and Organisation of Care (EPOC). EPOC Taxonomy 2002. Oslo: Norwegian Knowledge Centre for the Health Services. Available at: http://epoc.cochrane.org/epoc-taxonomy

(15) Baird G. The laboratory test utilization management toolbox. Biochem Med (Zagreb) 2014; 24:223-34.

(16) Bonaguri C, Melegari A, Ballabio A, Parmeggiani M, Russo A, Battistelli L et al. Italian multicentre study for application of a diagnostic algorithm in autoantibody testing for autoimmune rheumatic disease: conclusive results. Autoimmun Rev 2011; 11:1-5.

(17) Tampoia M, Fontana A, Di SF, Maggiolini P, Pansini N. Application of a diagnostic algorithm in autoantibody testing: assessment of clinical effectiveness and economic efficiency. Clin Chim Acta 2003; 333:181-3.

(18) Solomon DH, Hashimoto H, Daltroy L, Liang MH. Techniques to improve physicians' use of diagnostic tests: a new conceptual framework. JAMA 1998; 280:2020-7.

(19) Ivers N, Jamtvedt G, Flottorp S, Young JM, Odgaard-Jensen J, French SD et al. Audit and feedback: effects on professional practice and healthcare outcomes. Cochrane Database Syst Rev 2012; 6:CD000259.

(20) Kobewka DM, Ronksley PE, McKay JA, Forster AJ, van WC. Influence of educational, audit and feedback, system based, and incentive and penalty interventions to reduce laboratory test utilization: a systematic review. Clin Chem Lab Med 2015; 53:157-83.

(21) Marshall M, Pagel C, French C, Utley M, Allwood D, Fulop N et al. Moving improvement research closer to practice: the Researcher-in-Residence model. BMJ Qual Saf 2014; 23:801-5.

APPENDIX 1: COMPLETE DESCRIPTION OF THE INTERVENTION AND BOOSTER SESSION

Intervention

The intervention (an one hour, educational session) consisted of three parts, but started with a short introduction about the study and it was made clear to the rheumatologists that all results presented were strictly confidential and that in no way the information could be used outside the study (for example by the management of the centers). This was done in order to create a non-judgmental environment.

In the first part of the intervention feedback on ANA requesting behavior over the last year was presented to the rheumatologists. The feedback started with information on the number of ANA tests, the results of the ANA tests requested, the number of repeated ANA tests within one year and the number of ANA associated diagnosis. All results were presented as percentages, except for the number of ANA tests requested. This was given as an absolute number, accompanied by the ANA/new patient ratio (number of ANA tests divided by the number of new patients; APR). In this way, the number of ANAs was put in perspective with the number of new patients. In addition, the variation in the APR between the individual rheumatologists and/or centers was discussed.

The feedback was provided in three different ways. First, all outcomes from all participating centers were presented. This made it possible for the rheumatologists to compare their group performance with group level performance of their peers in the other participating hospitals. Secondly, a non-anonymous bar graph picturing the individual APR of all rheumatologists in one center was presented. Also, an anonymous bar graph picturing the same information of the rheumatologists in the other centers was given. Finally, all rheumatologists received an individual feedback report after the intervention. This report contained all results on individual level together with (non-)anonymous peer information as described above.

Despite our efforts to provide a similar intervention to all participating hospitals, the feedback component differed slightly between the hospitals on two points due to organizational issues. When the intervention was given for the first time, data from the other participating hospitals was not yet available. Therefore, the rheumatologists in the first participating hospital (center 2) did not receive information about their peers in other hospitals. Furthermore, it was not possible to obtain data on the level of the individual rheumatologists in all participating centers. As a consequence, the rheumatologists in this hospital (center 3) received their feedback only on hospital level and information on variation between rheumatologists could not be provided.

The second part of the intervention consisted of general background information on different topics, all deemed relevant for correct use of ANA tests. We started with providing generic background on the relation between pre-test probability and diagnostic test results (also called Bayes theorem). This was done because literature suggests that physicians fail to apply knowledge about pre-test probability of a disease and test characteristics when deciding to request a particular test.[22] Bayes theorem was explained and further illustrated by calculating the positive and negative predictive values in a situation with overuse (pre-intervention data) and without overuse (hypothetical numbers reflecting an 'ideal' situation).[15] An example of this calculation can be found in appendix 3. After this general background the focus was shifted back to the main topic and more information on ANA testing was given, including test characteristics and current literature on correct ANA use. The second part of the intervention ended with an introduction about the nature and effects of cognitive bias. This topic was included because other studies suggest that medical decisions, including requesting laboratory tests, can be influenced by cognitive bias.[22] Cognitive bias can be explained as a type of error in our thinking that occurs when we are processing and interpreting information around us. These errors are often a result of our attempt to simplify information processing, but this simplification can lead to inaccurate decisions. Many different types of cognitive biases are known, but during the intervention only four of them were described (confirmation bias, availability bias, primacy error and insensivity to sample size). These four were judged to be most important when deciding on requesting a laboratory test such as an ANA test.

The third and final part of the intervention consisted of two simple recommendations on how to apply the first two parts in daily practice. Firstly, a short guideline on when – not - to request an ANA test was given (see table 1). Secondly, in every center the rheumatologists were given a target APR to reach after the intervention. This target was calculated for all centers separately, using the same method and assuming that a test performs optimally when the pre-test probability on disease is around 50%. In order to create an indication of the number of patients needed to create a pre-test probability of 50%, the number of patients with an ANA-associated disease in the pre-intervention period was doubled. This number was then divided by the number of new patients seen in the pre-intervention period, as to calculate the APR. For example, if in the pre-intervention the numbers of newly seen patients, ANA tests and ANA associated diagnoses were 5000, 1000 and 50 respectively, the pre-test probability on disease would be 0.05% (50/1000) with an APR of 0.2 (1000/5000). The target after the intervention would be 100 ANA tests (50*2), giving an APR of 0.02 (100/5000) and a pre-test probability of 50% (50/100). Again, to account for differences in number of patients, the target was given as the APR instead of the absolute number of ANA tests.

Table 1: recommendations and targets provided during the intervention

	Guideline/target description
Short guideline on ANA use in daily practice	Only request an ANA test in case of a reasonable clinical suspicion on an ANA-associated disease (SLE, SSc, PM/DM, SS or MCTD). Do not request an ANA test in the following situations: - Patients with rheumatoid arthritis, fibromyalgia, Raynaud phenomenon without other complaints, undifferentiated arthralgia - Patients recently tested for ANA positivity (<1 year) - As an 'yearly follow-up' - At the start of TNF-blockers - Patients already known to you with changing complaints but no reasonable suspicion on an ANA-associated disease
Target APR to reach after the intervention	Centre 1: 0.04 Centre 2: 0.06 Centre 3: 0.15

SLE: systemic lupus erythematosus, SSc: systemic sclerosis, PM/DM: polymyositis/dermatomyositis, SS: Sjögren's syndrome, MCTD: mixed connective tissue disease.

Booster session

The booster session included the same components as the intervention, now using post-intervention data from the first three months as feedback.

APPENDIX 2: POWERPOINT SLIDES USED DURING THE INTERVENTION AND BOOSTER SESSION

INTERVENTION

Rheumatologists & Antinuclear Antibody testing: how to make them choose wisely?

An example of the presentation used during the intervention in three centres in the Netherlands

Authors: Nienke Lesuis (resident of rheumatology; MD) and Alfons A den Broeder (rheumatologist-epidemiologist; MD, PhD)
Sint Maartenskliniek, Nijmegen, the Netherlands

Content

- Introduction
- How are we doing now?
  - Methods
  - Results
- How can we choose more wisely?
  - Background information on laboratory testing in general
  - Background information on ANA testing
  - Background information on cognitive bias
- In daily clinical practice…
- Take home message

Content

- **Introduction**
- How are we doing now?
  - Methods
  - Results
- How can we choose more wisely?
  - Background information on laboratory testing in general
  - Background information on ANA testing
  - Background information on cognitive bias
- In daily clinical practice…
- Take home message

Introduction

Why Antinuclear Antibodies (ANA)?

- Laboratory tests are an important part of daily clinical practice and suboptimal use is common
- Suboptimal use mainly leads to overuse of laboratory tests
- Overuse can have important negative consequences
  - Increased number of false-positives, leading to:
    - More overuse of other (laboratory) tests / unnecessary treatment
    - Decreased quality of life (patients)
    - Increased health care costs

Zhi et al. PLOS one november 2013; van Walraven et al. JAMA August 2012.

Content

- Introduction
- **How are we doing now?**
  - **Methods**
  - **Results**
- How can we choose more wisely?
  - Background information on laboratory testing in general
  - Background information on ANA testing
  - Background information on cognitive bias
- In daily clinical practice…
- Take home message

How are we doing now? - *methods*

General

Main goal: improving the quality of our health care

Study goal: to assess the effect of a simple training session and booster session on the ANA tests requested by rheumatologists.

- Number of ANA requests & their results
- Patient characteristics & diagnosis
- Practice variation between rheumatologists

"The lab test results are back—now it's time to roll the dice."

Content

- Introduction
- **How are we doing now?**
  - Methods
  - **Results**
- How can we choose more wisely?
  - Background information on laboratory testing in general
  - Background information on ANA testing
  - Background information on cognitive bias
- In daily clinical practice…
- Take home message

How are we doing now? - *results*

Summary

- Too many ANA tests requested
  - Numbers, results & diagnoses
- Too much practice variation
  - Large differences between rheumatologists

How are we doing now? - *results*

Summary

- **Too many ANA tests requested**
  - **Numbers, results & diagnoses**
- Too much practice variation
  - Large differences between rheumatologists

How are we doing now? - *results*

Too many ANA tests requested: numbers & results

Numbers
- 1000 ANAs in twelve months (1-1-2010 to 1-1-2011)

Results
- 86% negative
- 14% positive

How are we doing now? - *results*

Too many ANA tests requested: diagnoses

Top 3 diagnoses in the total population of patients in whom an ANA was requested

1. Undifferentiated arthralgia/myalgia 30%
2. Fibromyalgia 11%
3. Rheumatoid arthritis 10%

ANA associated diseases: 5%

How are we doing now? - *results*

	Disease +	**Disease -**	
ANA +	45	95	140
ANA -	5	855	860
	50	950	1000

How are we doing now? - *results*

Pre-test probability of the disease: 50 / 1000 = 5.0%

	Disease +	**Disease -**	
ANA +	45	95	140
ANA -	5	855	860
	50	950	**1000**

How are we doing now? - *results*

Pre-test probability of the disease: 50 / 1000 = 5.0%
Sensitivity: 45 / 50 = 90% Specificity: 855 / 950 = 90%

	Disease +	**Disease -**	
ANA +	**45**	95	140
ANA -	5	855	860
	50	950	1000

How are we doing now? - *results*

Pre-test probability of the disease: 50 / 1000 = 5.0%
Sensitivity: 45 / 50 = 90% **Specificity: 855 / 950 = 90%**

	Disease +	**Disease -**	
ANA +	45	95	140
ANA -	5	**855**	860
	50	**950**	1000

How are we doing now? - *results*

Pre-test probability of the disease: 50 / 1000 = 5.0%
Sensitivity: 45 / 50 = 90% Specificity: 855 / 950 = 90%
PPV: 45 / 140 = 32% NPV: 855 / 860 = 99%

	Disease +	**Disease -**	
ANA +	**45**	95	**140**
ANA -	5	855	860
	50	950	1000

How are we doing now? - *results*

Pre-test probability of the disease: 50 / 1000 = 5.0%
Sensitivity: 45 / 50 = 90% Specificity: 855 / 950 = 90%
PPV: 45 / 140 = 32% **NPV: 855 / 860 = 99%**

	Disease +	**Disease -**	
ANA +	45	95	140
ANA -	5	**855**	**860**
	50	950	1000

How are we doing now? - *results*

Pre-test probability of the disease: 50 / 1000 = 5.0%
Sensitivity: 45 / 50 = 90% Specificity: 855 / 950 = 90%
PPV: 45 / 140 = 32% NPV: 855 / 860 = 99%

	Disease +	**Disease -**	
ANA +	45	95	140
ANA -	5	855	860
	50	950	1000

PPV & NPV depend on the population assessed (pre-test probability), sensitivity & specificity do not …

How are we doing now? – *results (ideal situation)*

Pre-test probability of the disease: 500 / 1000 = 50.0%

	Disease +	**Disease -**	
ANA +			**500**
ANA -			**500**
			1000

How are we doing now? – *results (ideal situation)*

Pre-test probability of the disease: 500 / 1000 = 50.0%
Sensitivity: 90% Specificity: 90%

	Disease +	**Disease -**	
ANA +	**450**	**50**	500
ANA -			500
			1000

How are we doing now? – *results (ideal situation)*

Pre-test probability of the disease: 500 / 1000 = 50.0%
Sensitivity: 90% **Specificity: 90%**

	Disease +	**Disease -**	
ANA +	450	50	500
ANA -	**50**	**450**	500
	500	500	1000

How are we doing now? – *results (ideal situation)*

Pre-test probability of the disease: 500 / 1000 = 50.0%
Sensitivity: 90% Specificity: 90%
PPV: 450 / 500 = 90% NPV: 855 / 860 = 99%

	Disease +	**Disease -**	
ANA +	**450**	50	**500**
ANA -	50	450	500
	500	500	1000

How are we doing now? – *results (ideal situation)*

Pre-test probability of the disease: 500 / 1000 = 50.0%
Sensitivity: 90% Specificity: 90%
PPV: 450 / 500 = 90% **NPV: 450 / 500 = 90%**

	Disease +	**Disease -**	
ANA +	450	50	500
ANA -	50	**450**	**500**
	500	500	1000

How are we doing now? - *results*

Summary

- Too many ANA tests requested
  - Numbers, results & diagnoses
- **Too much practice variation**
  - **Large differences between rheumatologists**

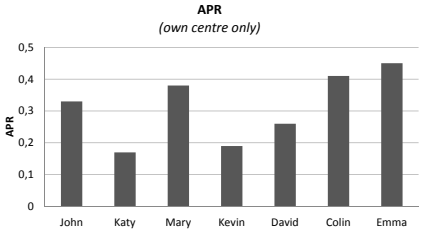
How are we doing now? - *results*

Practice variation

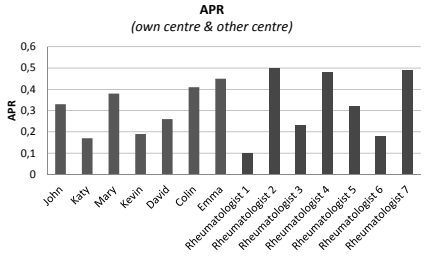
- 1000 ANA requests by 7 rheumatologists in 12 months
  - 900 unique patients (so, 100 repeated requests)
- 3200 new outpatient clinic patients seen in those 12 months
  - ANA/new patient ratio (APR): 1000 / 3200 = 0.31
- Variation between rheumatologists:
  - APR between 0.17 and 0.45 (standard deviation = 0.11)

→ Patient: chance of getting an ANA test seems to be heavily dependent of your rheumatologist…

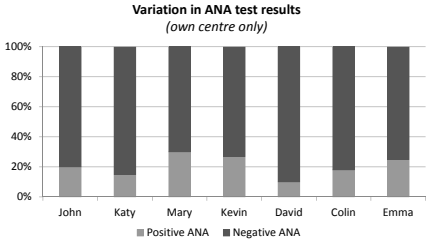
How are we doing now? - *results*



How are we doing now? - *results*



How are we doing now? - *results*



Content

- Introduction
- How are we doing now?
  - Methods
  - Results
- **How can we choose more wisely?**
  - **Background information on laboratory testing in general**
  - **Background information on ANA testing**
  - **Background information on cognitive bias**
- In daily clinical practice…
- Take home message

How can we choose more wisely? – *laboratory testing*

Use of laboratory tests in general

- Positive likelyhood ratio (pLR) of 2-5: test is useful (↑ post-test probability)
- Negative likelyhood ratio (nLR) of 0.2-0.5: test is useful (↓post-test probability)

- Pre-test probability: combination of the clinical picture & prevalence of the disease

- Low pre-test probability & a good pLR → post-test probability is still low. Don't use the test.
- Moderate pre-test probability (± 50%) & a good pLR/nLR → post-test probability will clearly in- or decrease. Use the test
- High pre-test probability & a good pLR → test result will not make a difference. Don't use the test.

Solomon et al. Arthritis & Rheumatism August 2002.

How can we choose more wisely? – *ANA testing*

ANA – test techniques & results

- Different techniques give different results
- Test characteristics depend on the disease of interest and control group used
  - SLE: sens 93% spec 57% pLR 2,2 nLR 0,1
  - Sjögren: sens 48% spec 52% pLR 0,99 nLR 1,01
- In general: sensitivity >> specificity
  - Many false-positives

→ ANA testing is only useful to exclude certain (rheumatic) diseases

Solomon et al. Arthritis & Rheumatism August 2002.

How can we choose more wisely? – *ANA testing*

Test characteristics – positive results

- Healthy population
  - **1:40** **25-30%** *(Own centre: 28%)*
  - 1:80 10-15%
  - >1:160 5%
- Relatives
  - >1:40 25-30%

→ ANA testing is not useful as a screening test for rheumatic diseases

Solomon et al. Arthritis & Rheumatism August 2002.

How can we choose more wisely? – *Cognitive bias*

Cognitive bias

- Errors in our thinking that occur when we are processing and interpreting information around us, leading to inaccurate decisions and wrong interpretations → irrationality
  - Heuristics (shortcuts in decision making)
  - Limited information processing capacity
  - Emotions
  - …
- Unconscious & hard to recognize in your own thinking
- Influencing diagnostic & therapeutic decisions

Bornstein et al. Journal of Evaluation in Clinical Practice 2000.

How can we choose more wisely? – *Cognitive bias*

Confirmation bias

- Selective collection and interpretation of evidence in order to confirm a hypotheses ('cherry picking')
  - Being right feels better than being wrong…
  - Severe sicca complaints, arthritis and a diagnosis of Sjögren's syndrome (diagnosed by a colleague somewhere else) → ANA to 'confirm'

How can we choose more wisely? – *Cognitive bias*

Availability bias

- The tendency to overestimate the likelihood of events with greater "availability" in memory (influenced by how recent the memories are or how unusual or emotionally charged they may be)
  - Your last patient was a young women with severe, newly diagnosed SLE → ANA in the next young women presenting with arthralgia

How can we choose more wisely? – *Cognitive bias*

Primacy error/anchoring

- The tendency to rely too heavily, or "anchor," on one trait or piece of information when making decisions. Usually the first piece of information that we acquire ('you never have a second chance for a first impression)
  - Two groups of doctors were asked to estimate the probability on a pulmonary embolism, both groups were given a irrelevant percentage as 'anchor':
    - Case description & low anchor (1%) → estimated probability of 23% on pulmonary embolism
    - Case description & high anchor (90%) → estimated probability of 53% on pulmonary embolism

Brewer et al. Medical Decision Making March-April 2007.

How can we choose more wisely? – *Cognitive bias*

Insensitivity to sample size

- The tendency to not take into account sample size and/or not take into account underlying prevalence of a disease
  - Bilateral wrist pain in an older patient with hypothyroidism: carpal tunnel syndrome or osteoarthritis?
  - Side effects seen in two patients generalising to all patients?

How can we choose more wisely? – *Cognitive bias*

What should we do with cognitive bias?

- Know that these biases exist
  "from unskilled & unaware to unskilled & aware"
- Protect yourself:
  - Recognition of cognitive bias in your own thinking (difficult!)
  - Adhere to guidelines (diagnostic, treatment)

For the interested reader: 'Irrationality' by David Sutherland and 'Thinking, Fast and Slow' by Daniel Kahneman

Content

- Introduction
- How are we doing now?
  - Methods
  - Results
- How can we choose more wisely?
  - Background information on laboratory testing in general
  - Background information on ANA testing
  - Background information on cognitive bias
- **In daily clinical practice…**
- Take home message

In daily clinical practice

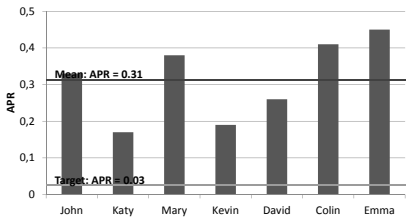
When should I use ANA testing?

- ANA in case of:
  - A reasonable suspicion on an ANA-associated disease (systemic lupus erythematosus, systemic sclerosis, polymyositis/dermatomyositis, Sjögren's syndrome, mixed connective tissue disease)
- No ANA in case of:
  - RA, fibromyalgia, Raynaud without other complaints, arthralgia
  - Recently (< 1 year) tested (in another centre)
  - Yearly 'follow-up'
  - Start of TNF blockers
  - Patient already known by you without a change in clinical picture

Solomon et al. Arthritis & Rheumatism August 2002.

In daily clinical practice

How many ANA tests should we do?



Content

- Introduction
- How are we doing now?
  - Methods
  - Results
- How can we choose more wisely?
  - Background information on laboratory testing in general
  - Background information on ANA testing
  - Background information on cognitive bias
- In daily clinical practice…
- **Take home message**

Take home message

- ANA testing is only useful in case of a reasonable suspicion on an ANA-associated disease
- Always use clinical context when interpreting ANA test results
- Protect yourself against cognitive bias, use a guideline

BOOSTER SESSION

Rheumatologists & Antinuclear Antibody testing: how to make them choose wisely?

An example of the presentation used during the booster session in three centres in the Netherlands

Authors: Nienke Lesuis (resident of rheumatology; MD) and Alfons A den Broeder (rheumatologist-epidemiologist; MD, PhD)
Sint Maartenskliniek, Nijmegen, the Netherlands

Content

- What was it all about?
- How are we doing now?
  - Results
- How can we choose even more wisely?
  - Background information on laboratory testing in general
  - Background information on ANA testing
  - Background information on cognitive bias
- In daily clinical practice…
- To be continued…

Content

- **What was it all about?**
- How are we doing now?
  - Results
- How can we choose even more wisely?
  - Background information on laboratory testing in general
  - Background information on ANA testing
  - Background information on cognitive bias
- In daily clinical practice…
- To be continued…

What was it all about?

General

Main goal: improving the quality of our health care

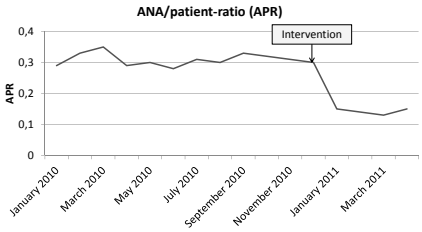
Study goal: to assess the effect of a simple training session and booster session on the ANA tests requested by rheumatologists

- Number of ANA requests & their results
- Patient characteristics & diagnosis
- Practice variation between rheumatologists

Content

- What was it all about?
- **How are we doing now?**
  - **Results**
- How can we choose even more wisely?
  - Background information on laboratory testing in general
  - Background information on ANA testing
  - Background information on cognitive bias
- In daily clinical practice…
- To be continued…

How are we doing now? - *results*



How are we doing now? - *results*

How are we doing now? - *results*

Summary

- Decrease in the number of ANA tests requested
  - Numbers, results & diagnoses
- Decrease in practice variation
  - Smaller differences between rheumatologists

How are we doing now? - *results*

Summary

- **Decrease in the number of ANA tests requested**
  - **Numbers, results & diagnoses**
- Decrease in practice variation
  - Smaller differences between rheumatologists

How are we doing now? - *results*

Decrease in ANA tests requested: numbers & results

Numbers

- 150 ANAs in four months (compared to 330 in the same period last year)

Results

- 85% negative
- 15% positive

How are we doing now? - *results*

Decrease in ANA tests requested: diagnoses

Top 3 diagnoses in the total population of patients in whom an ANA was requested [percentage in the pre-intervention period]

1. Undifferentiated arthralgia/myalgia 28% [30%]
2. Fibromyalgia 12% [11%]
3. Rheumatoid arthritis 10% [10%]

ANA associated diseases: 7% [5%]

How are we doing now? - *results*

Pre-test probability of the disease: 11 / 150 = 7.0%
Sensitivity: 10 / 11 = 90% Specificity: 125 / 139 = 90%
PPV: 10 / 24 = 42% **NPV: 125 / 126 = 99%**

	Disease +	**Disease -**	
ANA +	**10**	14	**24**
ANA -	1	**125**	**126**
	11	139	150

How are we doing now? - *results*

	Disease +	**Disease -**	
ANA +	10	14	24
ANA -	1	125	126
	11	139	150

How are we doing now? - *results*

Pre-test probability of the disease: 11 / 150 = 7.0%

	Disease +	**Disease -**	
ANA +	10	14	24
ANA -	1	125	126
	11	139	**150**

How are we doing now? - *results*

Pre-test probability of the disease: 11 / 150 = 7.0%
Sensitivity: 10 / 11 = 90% Specificity: 125 / 139 = 90%

	Disease +	**Disease -**	
ANA +	**10**	14	24
ANA -	1	125	126
	11	139	150

How are we doing now? - *results*

Pre-test probability of the disease: 11 / 150 = 7.0%
Sensitivity: 10 / 11 = 90% **Specificity: 125 / 139 = 90%**

	Disease +	**Disease -**	
ANA +	10	14	24
ANA -	1	**125**	126
	11	**139**	150

How are we doing now? - *results*

Pre-test probability of the disease: 11 / 150 = 7.0%
Sensitivity: 10 / 11 = 90% Specificity: 125 / 139 = 90%
PPV: 10 / 24 = 42% NPV: 855 / 860 = 99%

	Disease +	**Disease -**	
ANA +	**10**	14	**24**
ANA -	1	125	126
	11	139	150

How are we doing now? – *results*
(pre-intervention)

Pre-test probability of the disease: 50 / 1000 = 5.0%
Sensitivity: 45 / 50 = 90% Specificity: 855 / 950 = 90%
PPV: 45 / 140 = 32% NPV: 855 / 860 = 99%

	Disease +	**Disease -**	
ANA +	45	95	140
ANA -	5	855	860
	50	950	1000

How are we doing now? - *results*

Pre-test probability of the disease: 11 / 150 = 7.0%
Sensitivity: 10 / 11 = 90% Specificity: 125 / 139 = 90%
PPV: 10 / 24 = 42% **NPV: 125 / 126 = 99%**

	Disease +	Disease -	
ANA +	**10**	14	**24**
ANA -	1	**125**	**126**
	11	139	150

PPV & NPV depend on the population assessed (pre-test probability), sensitivity & specificity do not …

How are we doing now? - *results*

Summary

- Decrease in the number of ANA tests requested
  - Numbers, results & diagnoses
- **Decrease in practice variation**
  - **Smaller differences between rheumatologists**

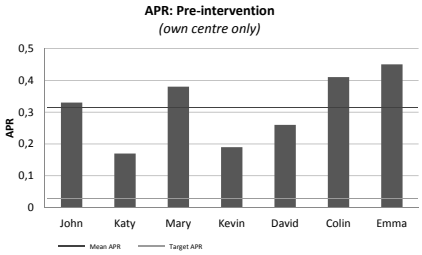
How are we doing now? - *results*

Practice variation

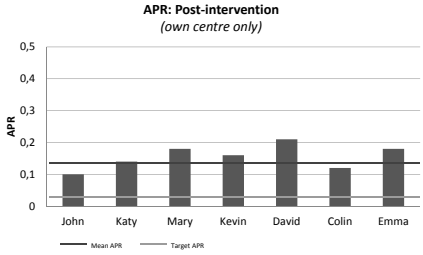
- 150 ANA requests by 7 rheumatologists in 4 months
  - 145 unique patients (so, 5 repeated requests)
- 1100 new outpatient clinic patients seen in those 4 months
  - ANA/new patient ratio (APR): 150 / 1100 = 0.14
- Variation between rheumatologists:
  - APR between 0.10 and 0.21 (standard deviation = 0.04)

→ Patient: chance of getting an ANA test seems to be dependent of your rheumatologist… (but less than before the intervention)

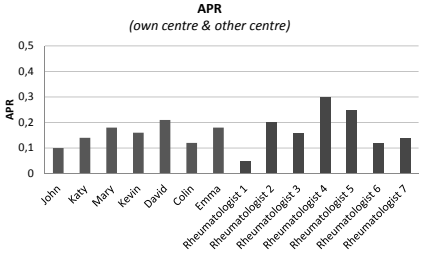
How are we doing now? - *results*



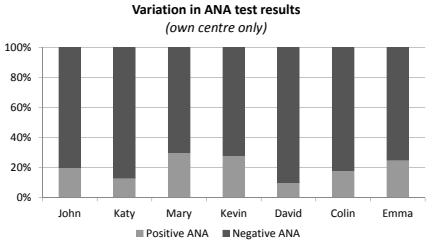
How are we doing now? - *results*



How are we doing now? - *results*



How are we doing now? - *results*



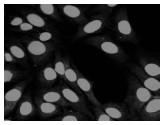
How are we doing now? - *results*

Summary

- Decrease in the number of ANA tests requested
- % positive ANA tests and diagnosis almost unchanged
- Practice variation decreased
- Target APR: approximately reached by some rheumatologists

Content

- Introduction
- How are we doing now?
  - Methods
  - Results
- **How can we choose more wisely?**
  - **Background information on ANA testing**
  - Background information on cognitive bias
- In daily clinical practice…
- Take home message



How can we choose more wisely? – *ANA testing*

ANA – test techniques & results

- Different techniques give different results
- Test characteristics depend on the disease of interest and control group used
  - SLE: sens 93% spec 57% pLR 2,2 nLR 0,1
  - Sjögren: sens 48% spec 52% pLR 0,99 nLR 1,01
- In general: sensitivity >> specificity
  - Many false-positives

→ **ANA testing is only useful to exclude certain (rheumatic) diseases**

Solomon et al. Arthritis & Rheumatism August 2002.

How can we choose more wisely? – *ANA testing*

Test characteristics – positive results

- Healthy population
  - **1:40** **25-30%** *(Own centre: 28%)*
  - 1:80 10-15%
  - >1:160 5%
- Relatives
  - >1:40 25-30%

→ **ANA testing is not useful as a screening test for rheumatic diseases**

Solomon et al. Arthritis & Rheumatism August 2002.

Content

- Introduction
- How are we doing now?
  - Methods
  - Results
- **How can we choose more wisely?**
  - Background information on ANA testing
  - **Background information on cognitive bias**
- In daily clinical practice…
- Take home message

Principles of Magic #2
People want to believe. The first impulse is to believe. Doubting is secondary.

How can we choose more wisely? – *Cognitive bias*

Casus babies

A certain city is served by two hospitals. In the larger hospital **45 babies** are born each day, and in the smaller hospital about **15 babies** are born each day. For a period of one year, the two hospitals recorded the **days on which more the 60% of the babies born were boys**. At the end of the year, **which hospital recorded more such days**?

Larger hospital
Small hospital
About the same (i.e., within 5% of each other)

Inventory of Cognitive Bias in Medicine (Hershberger et al. Acad Med 1994)

How can we choose more wisely? – *Cognitive bias*

Casus babies

A certain city is served by two hospitals. In the larger hospital **45 babies** are born each day, and in the smaller hospital about **15 babies** are born each day. For a period of one year, the two hospitals recorded the **days on which more the 60% of the babies born were boys**. At the end of the year, **which hospital recorded more such days**?

Larger hospital
Small hospital
About the same (i.e., within 5% of each other)

Insensitivity to sample size: statistic variation increases as sample size decreases

Inventory of Cognitive Bias in Medicine (Hershberger et al. Acad Med 1994)

How can we choose more wisely? – *Cognitive bias*

Casus phenylketonuria

Testing for phenylketonuria is very effective in identifying and preventing serious morbidity and mortality in newborns. The **sensitivity is 99.9% and specificity is 99.9%.** About **one out of every 14.000** live births is afflicted. If an infant tests positive, the likelihood of disease is:

High
Moderate
Low

Inventory of Cognitive Bias in Medicine (Hershberger et al. Acad Med 1994)

How can we choose more wisely? – *Cognitive bias*

Casus phenylketonuria

Testing for phenylketonuria is very effective in identifying and preventing serious morbidity and mortality in newborns. The **sensitivity is 99.9% and specificity is 99.9%.** About **one out of every 14.000** live births is afflicted. If an infant tests positive, the likelihood of disease is:

High
Moderate
Low

The positive predictive value depends on your test population…
Calculate!

Inventory of Cognitive Bias in Medicine (Hershberger et al. Acad Med 1994)

How can we choose more wisely? – *Cognitive bias*

Casus phenylketonuria

Pre-test probability of the disease = incidence = 1 : 14000
Sensitivity = 99.9% Specificity = 99.9%

	Disease +	Disease -	
Test +			
Test -			

How can we choose more wisely? – *Cognitive bias*

Casus phenylketonuria

Pre-test probability of the disease = incidence = 1 : 14000
Sensitivity = 99.9% Specificity = 99.9%

	Disease +	Disease -	
Test +			
Test -			
	13	**179987**	180000*

*Number of babies born in the Netherlands every year

How can we choose more wisely? – *Cognitive bias*

Casus phenylketonuria

Pre-test probability of the disease = incidence = 1 : 14000
Sensitivity = 99.9% Specificity = 99.9%

	Disease +	Disease -	
Test +	**12.99**		
Test -	**0.01**		
	13	179987	180000*

*Number of babies born in the Netherlands every year

How can we choose more wisely? – *Cognitive bias*

Casus phenylketonuria

Pre-test probability of the disease = incidence = 1 : 14000
Sensitivity = 99.9% **Specificity = 99.9%**

	Disease +	Disease -	
Test +	12.99	**180**	182.99
Test -	0.01	**179807**	179807.01
	13	179987	180000*

*Number of babies born in the Netherlands every year

How can we choose more wisely? – *Cognitive bias*

Casus phenylketonuria

Pre-test probability of the disease = incidence = 1 : 14000
Sensitivity = 99.9% Specificity = 99.9%
NPV = 179807 / 179807.01 = 100%

	Disease +	Disease -	
Test +	12.99	180	182.99
Test -	0.01	**179807**	**179807.01**
	13	179987	180000*

*Number of babies born in the Netherlands every year

How can we choose more wisely? – *Cognitive bias*

Casus phenylketonuria

Pre-test probability of the disease = incidence = 1 : 14000
Sensitivity = 99.9% Specificity = 99.9%
PPV = 12.99 / 182.99 = 7% NPV = 179807 / 179807.01 = 100%

	Disease +	Disease -	
Test +	**12.99**	180	**182.99**
Test -	0,01	179807	179807.01
	13	179987	180000*

*Number of babies born in the Netherlands every year

How can we choose more wisely? – *Cognitive bias*

Casus atrial fibrillation

Coumadin or aspirin is indicated for embolus prophylaxis in patients with chronic atrial fibrillation (although coumadin is believed to be better). The last three patients you have treated with coumadin have developed complications secondary to the drug. You have never seen any complication related to aspirin. You are now seeing in referral a patient with atrial fibrillation who was sent to you for your recommendation regarding embolic prophylaxis. Which drug do you recommend?

Coumadin
Aspirin

Inventory of Cognitive Bias in Medicine (Hershberger et al. Acad Med 1994)

How can we choose more wisely? – *Cognitive bias*

Casus atrial fibrillation

Coumadin or aspirin is indicated for embolus prophylaxis in patients with chronic atrial fibrillation (although coumadin is believed to be better). The last three patients you have treated with coumadin have developed complications secondary to the drug. You have never seen any complication related to aspirin. You are now seeing in referral a patient with atrial fibrillation who was sent to you for your recommendation regarding embolic prophylaxis. Which drug do you recommend?

Coumadin
Aspirin

Availability bias: the tendency to overestimate the likelihood of events with greater "availability" in memory

Inventory of Cognitive Bias in Medicine (Hershberger et al. Acad Med 1994)

How can we choose more wisely? – *Cognitive bias*

Casus boy or girl

You are a paediatrician. The last four patients you have seen in your office today have been girls. The next patient you are to see is more likely to be:

Girl
Equal change of being a boy or girl
Boy

Inventory of Cognitive Bias in Medicine (Hershberger et al. Acad Med 1994)

How can we choose more wisely? – *Cognitive bias*

Casus boy or girl

You are a paediatrician. The last four patients you have seen in your office today have been girls. The next patient you are to see is more likely to be:

Girl
Equal change of being a boy or girl
Boy

Gamblers fallacy: misconceptions of chance, chance is commonly but erroneously viewed as a self-correcting process.

Inventory of Cognitive Bias in Medicine (Hershberger et al. Acad Med 1994)

How can we choose more wisely? – *Cognitive bias*

What should we do with cognitive bias?

- Know that these biases exist
  "from unskilled & unaware to unskilled & aware"
- Protect yourself:
  - Recognition of cognitive bias in your own thinking (difficult!)
  - Adhere to guidelines (diagnostic, treatment etc)

For fun: http://www.youtube.com/watch?v=3RsbmjNLQkc&feature=player_detailpage

For the interested reader: 'Irrationality' by David Sutherland and 'Thinking, Fast and Slow' by Daniel Kahneman

Content

- What was it all about?
- How are we doing now?
  - Results
- How can we choose even more wisely?
  - Background information on laboratory testing in general
  - Background information on ANA testing
  - Background information on cognitive bias
- **In daily clinical practice…**
- **To be continued…**

In daily clinical practice

When should I use ANA testing?

- ANA in case of:
  - A reasonable suspicion on an ANA-associated disease (systemic lupus erythematosus, systemic sclerosis, polymyositis/dermatomyositis, Sjögren's syndrome, mixed connective tissue disease)
- No ANA in case of:
  - RA, fibromyalgia, Raynaud without other complaints, arthralgia
  - Recently (< 1 year) tested (in another centre)
  - Yearly 'follow-up'
  - Start of TNF blockers
  - Patient already known by you without a change in clinical picture

Solomon et al. Arthritis & Rheumatism August 2002.

To be continued…

- Final post-intervention measurement (12 months pre-intervention)
- Keep on doing less ANA tests!

APPENDIX 3: BAYES THEOREM

Often the value of diagnostic tests is displayed as the sensitivity and specificity of the test, but the real value in clinical practice is highly influenced by the pre-test probability of the disease for which the test is requested. In contrast to the sensitivity and specificity, the positive and negative predictive value take into account the pre-test probability of the disease and are therefore more insightful when judging the clinical value of a test. This can be best illustrated with an example (see also table 1). A new test has been developed to diagnose Sjögren's syndrome (SS) and has a sensitivity and specificity of both 90%. Due to this good test characteristics, this test is used to screen 1000 healthy people in a general hospital (pre-test probability of 5% on SS). Given these characteristics, 140 people will have a positive test, but only 45 of them have SS. This leads to a PPV of only 32% (45/140). In the same way the NPV can be calculated, which is 99% (855/860). In this situation, a rheumatologist can rule out SS in case of a negative test, but if a person tests positive, there is still uncertainty about the diagnosis.

Because of these results, the hospital decides to be more selective when using this test, so now only people with complaints compatible with SS are tested (pre-test probability of 50%). With the same number of people and the same test characteristics, a PPV of 90% and a NPV of 99% are found, giving much more certainty to the rheumatologist whether or not to diagnose SS. This discrepancy in the PPV between the two scenarios is caused by the much larger percentage of false-positive results compared to the true positives in the first setting (pre-test probability 5%), compared to the second setting (pre-test probability 50%). In general it is recognized that diagnostic tests have the most optimal PPV and NPV if the pre-test odds on disease lies between the 30 and 60%.

***Table 1**: 2 x 2 table for scenarios with different pre-test probabilities*

Scenario 1 (pre-test probability on SS = 50%)			
	Sjögren's syndrome present	**Sjögren's syndrome absent**	
Test positive	450	50	500
Test negative	50	450	500
	500	500	1000
Scenario 2 (pre-test probability on SS = 5%)			
	Sjögren's syndrome present	**Sjögren's syndrome absent**	
Test positive	45	95	140
Test negative	5	855	860
	50	950	1000

Re: Lesuis et al. Choosing Wisely in daily practice: an intervention study on Antinuclear Antibody testing by rheumatologists.

Robert Ferrari

Dear Editor,

I commend Lesuis et al.[1] for taking a further step in the effort to effect change in how physician's choose tests, in this case examining the often overused anti-nuclear antibody (ANA) test. Lesuis et al. showed that an educational intervention could reduce the number of ANA tests ordered by rheumatologists. The usefulness of this approach in other scenarios may be an issue, however. In Canada, for example, the Canadian Rheumatology Association (CRA) has developed a list of 5 tests, procedures, or therapies that have evidence indicating they may be not adding value and, in some instances, may be harmful.[2] Among the list of five items developed by the CRA for Choosing Wisely Canada is the following: "Don't order anti-nuclear antibodies (ANA) as a screening test in patients without specific signs or symptoms of systemic lupus erythematosus (SLE) or another connective tissue disease (CTD)."[2] In a subsequent study of our membership (data unpublished), we found that more than 80% of rheumatologists surveyed were already following this approach. The problem in Canada appears not to be the rheumatologists. In Alberta, Canada, for example approximately 60,000 ANA tests are done each year. For a population of 4 million in Alberta, this means, for example, that 1 in 65 Albertans have their ANA tested every year at a cost of 3 million dollars US annually, not including costs for additional physician visits, referrals, and investigations associated with a positive result. Local experience indicates that most of these tests are not ordered by specialists.[3] Rheumatologists in Canada routinely report receiving referrals based solely on a positive ANA.

Although education could be an approach used with all physicians, disseminating that education widely may be a challenge. Another approach may instead be modifying the serology ordering form to encourage evidence-based ANA testing. Various classification criteria have been developed to increase the sensitivity and specificity of a test in terms of an SLE diagnosis. Although clinical diagnoses do not always meet classification criteria, the chances of a patient having a diagnosis of SLE meeting formally declared classification criteria are high (i.e., the criteria were developed by looking at patients with the diagnosis and those without). The Systemic Lupus Erythematosus Collaborating Clinics (SLICC) classification criteria consider not merely ANA positivity, but also serositis, oral ulcers, arthritis, photosensitivity, cytopenia, renal involvement, neurological disorders, and specific rashes.[4] Having at least four positives from these categories signifies a high likelihood of correctly classifying a patient as having SLE. Using this knowledge, one could develop a serology ordering form that asks the ordering physician to consider the following:

ANA (Anti-Nuclear Antibody) test. At least two of the following criteria must be met for this test to be completed.

- Lupus rash
- Oral ulcers
- Physician-observed swelling of two or more joints OR tender joints with morning stiffness
- Serositis, pleurits, or pericarditis
- Evidence of renal disease
- Evidence of neurologic disease
- Cytopenia
- Anti-phospholipid antibody positive
- Non-scarring alopecia
- Low complement levels

That is, of a patient does not have at least 2 of these criteria present, positive ANA testing will certainly not be helpful. It has recently been shown in an Alberta practice setting that using these *a priori* minimum criteria before ordering ANA tests can greatly reduce the number of tests ordered without missing important diagnoses.[3]

1. Lesuis N, Hulscher ME, Piek E, Demirel H, van der Laan-Baalbergen N, Meek I, et al. Choosing Wisely in daily practice: an intervention study on Antinuclear Antibody testing by rheumatologists. Arthritis Care Res 2015 Sep 28. doi: 10.1002/acr.22725.
2. Chow SL, Thorne C, Bell MJ, Ferrari R, Bagheri Z, Boyd T, et al. Choosing Wisely Canada. Canadian Rheumatology Association: five things physicians and patients should question. J Rheumatol 2015;42:682-9.
3. Ferrari R. Evaluation of the Canadian Rheumatology Association Choosing Wisely recommendation concerning anti-nuclear antibody (ANA) testing. Clin Rheumatol 2015;34:1551-6.
4. Petri M, Obai AM, Alarcón GS, Gordon C, Merril JT, Fortin PR, et al. Derivation and validation of the Systemic Lupus International Collaborating Clinics classification criteria for systemic lupus erythematosus. Arthritis Rheum 2012;64:2677-86.

Comment on "Choosing Wisely in daily practice: an intervention study on Antinuclear Antibody testing by rheumatologists": reply

Nienke Lesuis; Marlies EJL Hulscher; Ester Piek; Hatice Demirel; Nicole van der Laan-Baalbergen; Inger Meek; Ronald F van Vollenhoven; Alfons A den Broeder

Dear editor,

We thank dr. Ferrari for his interest in and valuable comments on our recently published study on Antinuclear Antibody (ANA) overuse by rheumatologists.

Firstly, dr. Ferrari notes that in Canada ANA overuse among rheumatologists seems not to be widespread as 80% of surveyed rheumatologists reports to follow the Canadian Choosing Wisely advice to only order an ANA test in case of specific signs and symptoms of systemic lupus erythematosus (SLE) or another connective tissue disease (CTD).[1] This seems to concern self-reported adherence to the Choosing Wisely advice, and we are not certain that the same results would be found if ANA test use in those rheumatologists would actually be measured. This because self-reported results have been proven before to overrate actual results [2]. Therefore, we think that ANA overuse might still be a problem among Canadian rheumatologists.

We agree with dr. Ferrari that ANA overuse might also be a major problem amongst general practitioners (GP), stressing the need for an intervention that is easy to implement in a primary care setting. Dr. Ferrari suggested to modify serology ordering forms by making it obligatory for ordering physicians to complete a list of clinical criteria and only if two or more of those criteria are present, the ANA test will be completed.[1] Although we support the general idea, this approach has some drawbacks. Firstly, the checklist focuses on SLE only, while ANA testing can also be useful in other CTDs such as systemic sclerosis or in other patient populations (for example autoimmune hepatitis).[3] For this reason we chose to include a more general comment in our educational meeting ('do not use ANA testing without a reasonable suspicion on an ANA-associated disease'). Of note, we have tested the effect of computerized reminders in our ordering system, repeating the intervention advice every time a rheumatologist wanted to order an ANA test or another rarely indicated test in rheumatology such as complement levels. This resulted in a large decrease in the number of orders for those tests (manuscript in preparation). So, we agree with dr. Ferrari that the modification of ordering forms can be a very effective way of promoting evidence-based test ordering but the exact content of such modifications can be debated. In addition, to our opinion such a modification should always be combined with education as to explain the rationale of the planned modification to its users.

Finally, we want to comment on the implicit suggestion of dr. Ferrari that the proposed modification of order forms could also be used to reduce ANA overuse by GPs. Although this might be possible, we think that it could be argued that ANA testing has no place at all in

primary care. This because of the very low incidence of ANA associated diseases and the at best modest test characteristics of ANA testing, resulting in the need to refer a patient with relevant CTD complaints irrespective of the outcome of an ANA test.

1. Ferrari R. Re: Lesuis et al. Choosing Wisely in daily practice: an intervention study on Antinuclear Antibody testing by rheumatologists. Arthritis Care Res (Hoboken) 2015.
2. Davis DA, Mazmanian PE, Fordis M *et al*. Accuracy of physician self-assessment compared with observed measures of competence. A systematic review. JAMA 2006;296(9):1094-1102.
3. Solomon DH, Kavanaugh AJ, Schur PH *et al*. Evidence-based guidelines for the use of immunologic tests: Antinuclear Antibody testing. Arhritis Rheum 2002;47(4):434-44.

"The lab test results are back — now it's time to roll the dice."

3

Choosing Wisely in daily practice: a mixed methods study on determinants of Antinuclear Antibody testing by rheumatologists

Nienke Lesuis
Alfons A den Broeder
Ronald F van Vollenhoven
Johanna E Vriezekolk
Marlies EJL Hulscher

Accepted for publication in the Scandinavian Journal of Rheumatology

ABSTRACT

Objectives

To explore the relation between Antinuclear Antibody (ANA) overuse and rheumatologist-related factors before and after an intervention aimed at reducing ANA overuse.

Methods

In this mixed methods study we performed surveys among rheumatologists (n=20) before and after the ANA intervention (education and feedback). We identified clinician-related determinants of ANA overuse (demographic characteristics, cognitive bias, numeracy, personality, thinking styles and knowledge) by multivariate analysis. Two focus group meetings with rheumatologists were held six months after the intervention to explore self-reported determinants.

Results

Questionnaires were completed by all rheumatologists and eight participated in the focus groups. Rheumatologist with more work experience and a less extravert personality ordered more ANA tests before the intervention (ß 0.01, 95% confidence interval (95%-CI) 0.003 to 0.02, p= 0.01; ß -0.11, -0.21 to -0.01, p= 0.04 respectively; R^2 47%). After the intervention, female rheumatologists changed less than their male colleagues with regard to the number of ordered ANA tests (ß 0.15, 95%-CI 0.03 to 0.26, p= 0.02; R^2 25%). During the focus groups seven themes were identified that influenced improvement in ANA overuse: determinants related to the intervention and study, individual health professionals, patients, professional interactions, incentives and resources, capacity for organizational change and, social, political and legal factors.

Conclusions

We identified several determinants that together explained a sizable part of the variance observed in the ANA outcomes at baseline and in the change in ANA outcomes afterwards. Furthermore, the focus groups yielded additional factors suggesting a complex interplay of determinants influencing rheumatologists' ANA ordering behavior.

Trial registration

ClinicalTrials.gov, https://clinicaltrials.gov/ct2/home, NCT02409251

INTRODUCTION

Approximately 20% of laboratory tests ordered by physicians are ordered inappropriately, leading to higher costs, higher false-positive rates and a higher patient burden.[1] In the past few years the internationally expanding Choosing Wisely campaign has resulted in the American and Canadian rheumatology associations having included a statement on Antinuclear Antibody (ANA) use in their Choosing Wisely lists.[2, 3] Furthermore, a research agenda on medical overuse was published and recommended to identify determinants of overuse.[4] Although ANA overuse has been described in several publications, its determinants are not yet known.[5-8]

We previously conducted a pragmatic implementation study aiming to decrease ANA overuse.[9] Using the data of this study, we now explore the relationship between ANA overuse and potential physician-related determinants of this overuse.

MATERIALS AND METHODS

Study design and participants

This was a mixed methods study, combing quantitative and qualitative methods. The study was embedded in a multicenter (n= 3), pragmatic, before-and after controlled implementation study assessing the effect of education and feedback on ANA overuse. The implementation study took place in the Netherlands and lasted 24 months (12 month pre- and post-intervention period).[9]

Rheumatologists working in one of the three participating hospitals (general, specialized and academic hospital) from the implementation study were eligible for participation in this study. Those working the full study period at a study hospital of whom individual data on ANA overuse was available could be included.

Outcome measures & data collection

ANA overuse

To cover different aspects of ANA overuse the following four outcome measures were used: the ANA/patient ratio (APR; number of ANA tests ordered divided by the number of new patients seen at the rheumatology outpatient clinic), percentage of positive ANA tests, percentage of repeated ANA testing within one year and the percentage of ANA associated diseases (systemic lupus erythematosus, systemic sclerosis, polymyositis/dermatomyositis, mixed connective tissue disease or Sjögren's syndrome). All outcome measures were calculated at rheumatologist level.

The ANA outcomes calculated over the pre-intervention period were used to explore baseline associations with the determinants, whereas the difference between the pre- and post-intervention period outcomes (ANA outcome change scores) were used to analyze improvement in ANA use.

Determinants of ANA overuse

The following potential determinants for ANA overuse were assessed: age, gender, PhD degree, years of work experience, cognitive bias, personality traits, thinking styles, numeracy and ANA knowledge (last five questionnaires further explained in supplement 1). These determinants were assessed at baseline using web-based questionnaires (invitation send between March and June 2012; reminder two weeks after the first invitation).

In addition, questionnaires assessing modifiable determinants (cognitive bias, numeracy and ANA knowledge) were administered a second time, three months after the intervention. The difference between the baseline score and post-intervention score (determinant change score) was then used in the analyses with the ANA outcome change scores.

Self-reported determinants of improvement in ANA overuse

To explore self-reported determinants of improvement in ANA overuse we organized two 1.5 hour focus group sessions in the largest study center (n= 14). Both meetings were planned after regular working hours, took place at the study hospital and were led by a female psychologist-researcher (JV, PhD). A minimum of 4 and maximum of 8 participants could participate in one session. The main question during the focus group was "why did you change your behavior after the intervention". Both sessions were audiotaped and transcribed verbatim (see also supplement 2 for a more detailed methods description).

Data analysis

All analyses were performed using STATA 13.1. Descriptive statistics are presented as percentages with the accompanying absolute numbers, or as means with standard deviations. Associations between ANA baseline or change outcomes and the determinants were analyzed using multiple linear regression. Results from these analyses are reported as regression coefficients (ß) with the corresponding 95% confidence interval (95%-CI), p-value and explained variance (R^2). Before these analyses, variance, floor/ceiling effects and co linearity were assessed, and all determinants were tested separately with linear regression (univariate analyses).

The transcripts of the focus groups were analyzed by NL and JV using thematic content-analysis. The results of the qualitative analysis were discussed with the other researchers (AdB, MH and RvV) and compared to existing frameworks on this subject (supplement 2) [10-12].

Ethical approval

This study was presented to the local medical ethical board (CMO region Arnhem-Nijmegen) and this study was exempt from ethical approval (CMO number 2015-1653). All participating hospitals approved the study, and all participating rheumatologist gave consent.

All patient data, needed to calculate ANA outcomes, were retrieved within the study hospitals by matching two datasets locally, after which data was anonymized and provided to the research group. The study was registered at ClinicalTrials.gov (NCT02409251).

RESULTS

Setting and participants

All rheumatologists (n=29) working at the three study hospitals were eligible for participation. However, the nine rheumatologists from the university hospital had to be excluded because no data on individual ANA overuse was available. The remaining 20 rheumatologists completed all questionnaires on both time points (mean age 46.2 ± 9.3 years, 55% female, 60% PhD degree or pursuing a PhD, mean working experience 9.9 ± 9.3 years). Eight of the rheumatologists from the specialized hospital participated in the focus groups.

***Table 1**: Outcomes on ANA overuse and questionnaire scores*

	Pre-intervention (baseline; n = 20)	Post-intervention (n = 20)	Change score (n = 20)
ANA overuse outcomes			
ANA/new patient ratio	0.37 ± 0.19	0.12 ± 0.08	-0.25 ± 0.14
% positive ANA tests	25.7 ± 14.6	24.9 ± 16.9	-0.8 ± 16.2
% repeated ANA testing within 1 year	10.9 ± 10.4	1.3 ± 3.7	-9.5 ± 9.4
% ANA associated diseases*	7.1 ± 4.5	9.6 ± 10.2	2.5 ± 7.4
Questionnaire scores			
Cognitive bias	13.0 ± 4.3	13.4 ± 3.9	0.4 ± 2.4
Numeracy**	6.6 ± 1.0	n/a	n/a
ANA knowledge	3.3 ± 2.3	5.2 ± 1.5	1.9 ± 1.8
Personality			
Extraversion	3.33 ± 0.72	n/a	n/a
Neuroticism	2.80 ± 0.53	n/a	n/a
Openness to experience	3.63 ± 0.50	n/a	n/a
Consciousness	3.67 ± 0.52	n/a	n/a
Agreeableness	3.80 ± 0.31	n/a	n/a
Thinking styles			
Rational	77. ± 8	n/a	n/a
Experiential	64.6 ± 8.1	n/a	n/a

* Systemic lupus erythematosus, systemic sclerosis, polymyositis/dermatomyositis, mixed connective tissue disease or Sjögren's syndrome. **This determinant was excluded from further analysis due to marked ceiling effects

Results on ANA outcomes and determinants

In table 1 descriptive results on ANA outcomes and determinants are shown.

In the multivariate regression analyses, three out of four ANA baseline outcome measures (APR, repeated ANA and ANA associated diagnosis) and all ANA change outcomes were associated with one or more baseline determinants (table 2). No associations between ANA change outcomes and determinant change scores (ANA knowledge and cognitive bias) were found.

***Table 2**: Multivariate associations between ANA outcomes and determinants*

ANA outcome	ß(95% CI)	P-value
Associations between ANA baseline outcomes and baseline determinants		
ANA/new patient ratio		
Work experience, in years	0.01 (0.003 to 0.02)	0.01
Extraversion	-0.11 (-0.21 to -0.01)	0.04
Explained variance (%)	*47*	
% of ANA associated diagnoses		
Work experience, in years	0.22 (0.03 to 0.42)	0.03
Extraversion	2.87 (0.38 to 5.35)	0.03
Explained variance (%)	*39*	
% of repeated ANA tests		
Female gender	-10.2 (-18.92 to -1.53)	0.02
Explained variance (%)	*25*	
Associations between ANA change outcomes and baseline determinants		
ANA/patient ratio		
Female gender	0.15 (0.03 to 0.26)	0.02
Explained variance (%)	*25*	
% of positive ANA tests Female gender *Explained variance (%)*	 -16.53 (-29.93 to -3.12) *23*	 0.02
% of ANA associated diagnoses		
Extraversion	6.61 (2.72 to 10.49)	<0.01
Agreeableness	-15.05 (-23.99 to -6.10)	<0.01
Explained variance (%)	*47*	
% of repeated ANA tests		
Female gender	10.33 (2.78 to 17.88)	0.01
Explained variance (%)	*28*	

During the focus group meetings seven major themes were identified as being of importance to ordering less ANA tests after the intervention. These themes related to 1) the intervention and study, 2) individual health professionals,3) patients, 4) professional interactions, 5) incentives and resources, 6) capacity for organizational change and 7) social, political and legal factors (table 3). A full explanation of all themes including quotes is provided in supplement 2.

Table 3 Themes, subthemes and groups of self-reported determinants of improvement in ANA overuse as identified during the focus group meetings

Theme	Subtheme	Group	Quote
1. *Intervention & study factors*	Study characteristics		*"What helped me enormously was actually the first talk, in which a few things that of course I already knew, became really clear. But that had simply drifted away."*
	Study participation		
	Quality of the intervention		
	Content of the intervention		
	Recommended behavior during the intervention	Feasibility	*"An enthusiastic researche, that has a mutually encouraging effect. And then you're more likely to adjust your behavior than when some idiot says: shouldn't you be ordering fewer tests?"*
	Research team	Familiarity with the research team	
		Trust in the research team	
		Enthusiasm of the research team	
2. *Individual health professional factors*	(Previous) clinical experience		*"The fact that it's in actually a useless screening instrument was an extra eye-opener."*
	Knowledge & awareness	Awareness regarding the importance of clinical judgment and valid reasons to order a test	
		Awareness regarding own practice and its consequences	*"The fact that someone says to you, 'you don't need to order those things': what a relief…"*
		Awareness regarding false assumptions on test properties and previous behavior	
	Cognitions (including attitudes)	Agreement with the proposed problem and solution	
		Expected outcome	
		Professional pride & curiosity	
		Fear of uncertainty	
		Responsibility	
		Changing attitudes on ANA testing	
	Professional behavior	Nature of the new behavior	
		Confidence in new behavior	

3.	*Patient factors*	Patient numbers		-
4.	*Professional interactions*	Team process	Experience with peer to peer coaching	*"I actually found it quite a challenge and so it's become a bit of a game to order as few ANA tests as possible (…). After all, you want to be top of the class, don't you."*
			Atmosphere	
			Willingness to change	
			Peer to peer contact	
5.	*Incentives and resources*	Financial incentives and disincentives	Healthcare costs	-
		Nonfinancial incentives and disincentives	Time constraints	
		Assistance for clinicians	Electronic Health Record as barrier	
			Electronic Health Record as aid	
6.	*Capacity for organizational change*	Regulations, rules, policies	Department's personnel policy	
			Department's care policy	*"It's actually the continuation of a trend that has been going on for years. That we're simply taking decisions with respect to the diagnosis and treatment, based on evidence."*
			Organizations' research policy	
7.	*Social, political and legal factors*	Economic constraints on the health care budget	Healthcare costs	*"With all the commotion surrounding the cost of healthcare, I think we need to look critically at our ordering behavior, that's part and parcel of it too."*

DISCUSSION

This study shows that baseline rheumatologist characteristics such as gender, work experience and personality were associated with the ANA outcomes before and a change in outcomes after the intervention. Furthermore, the focus groups yielded many additional potential explanations for the observed improvement in ANA overuse after the intervention.

The main strength of this study is the use of a mixed methods design, making it possible to capture more determinants of ANA overuse than with quantitative analysis alone. Conversely, as the study sample was relatively small and the focus groups took place in only one center, not all determinants might be captured and our results may not be generalizable to other

centers. However, many of the themes identified during the focus groups are also found in previous studies.[10-12]

Although studies on determinants of ANA overuse are scarce, two reviews summarize the reasons for diagnostic test ordering in general. The relation between female gender, work experience and test ordering has been observed before, but so far no consistent pattern has emerged. For example, more work experience was associated with a decrease, increase or no change in the number of tests ordered.[11, 12]
Also in our study determinants seemed not always to be consistently associated with ANA outcome measures. For example, at baseline more work experience was both associated with more ANA tests and more ANA associated diagnoses. However, if principles on optimal test use are applied one would expect more tests to lead to less related diagnoses (Bayes theorem). That our results are not following this principle, might be caused by two rheumatologists in our sample, who have many years of work experience and also preferably see patients with systemic auto-immune diseases. Finally, three out of the seven relations found, included personality: less extravert rheumatologists ordered more ANA tests; more extravert rheumatologists had more ANA associated diagnoses; and more extravert and less agreeable rheumatologists changed more after the intervention regarding ANA associated diagnoses. The first two associations are in accordance with each other and follow Bayes theorem. Still, the reasons behind these observations are hard to give, although the original descriptions of both extraversion (enthusiastic, assertive, confident) and agreeableness (altruistic, cooperative) seem to fit with the direction of the associations.[13] As we are aware of only one other study on this topic, observing the same association [14], this would be an interesting direction for further research.

In addition to the questionnaires, the focus groups yielded many possible determinants which could be grouped into seven themes, matching an existing framework on this subject. [10] Interestingly, some of the themes yielded from the focus groups seemed contradictory to the determinants assessed with questionnaires. This inconsistency was most profound in the knowledge domain. Focus group participants stressed the importance of the knowledge that was refreshed by the intervention and the new awareness on different aspects of ANA testing. Although this seems to follow the increased scores on the ANA knowledge questionnaire after the intervention it did not translate into any association with ANA outcomes. This observation is also known from studies on guideline adherence.[11, 15]

In summary, the baseline determinants work experience, personality and gender are associated with ANA overuse at baseline and improvement in ANA overuse after the intervention. The focus groups yielded many more potential determinants suggesting a

complex interplay of factors affecting ANA ordering behavior. Future research could focus on quantifying the relationship between the factors mentioned in the focus group and actual behavior. Furthermore, we hope that our results might help many more physicians to avoid doing unnecessary tests and to choose even more wisely.

ACKNOWLEGDEMENTS

We would like to thank all rheumatologists who participated in this study for their willingness to complete the questionnaires and special thanks for those who attended the focus group meetings. Reinier Akkermans we want to thank for his valuable help during the statistical analyses.

FUNDING

This study received no external funding.

CONTRIBUTORS

All authors contributed to the study design. NL and AdB designed and executed the intervention in all participating hospitals. NL and JV participated in data collection. NL and JV participated in the data analysis. NL,JV, AdB, MH and RvV assisted in the interpretation of the data. The manuscript was written by NL under the supervision of JV, AdB, MH and RV. All authors have revised the draft version of the manuscript, and read and approved the final version of the manuscript. NL is guarantor.

COMPETING INTEREST DECLARATION

All authors have completed the Unified Competing Interest form at www.icmje.org/coi_disclosure.pdf (available on request from the corresponding author) and declare that RvV has received grants and personal fees from AbbVie, BMS, GSK, Pfizer, Roche, UCB, Biotest, Crescendo, Janssen, Lilly, Merck and Vertex outside the submitted work; all other authors reported no competing interests.

DATA SHARING STATEMENT

Patient level data, full dataset, technical appendix and statistical code are available at a reasonable request from the corresponding author. Consent from the patients was not obtained but the presented data are anonymized and risk of identification is low.

REFERENCES

(1) Zhi M, Ding EL, Theisen-Toupal J, Whelan J, Arnaout R. The landscape of inappropriate laboratory testing: a 15-year meta-analysis. PLoS One 2013;8:e78962.

(2) Chow SL, Carter TJ, Bell MJ, Ferrari R, Bagheri Z, Boyd T, et al. Choosing wisely: the Canadian Rheumatology Association's list of 5 items physicians and patients should question. J Rheumatol 2015;42:682-9.

(3) Yazdany J, Schmajuk G, Robbins M, Daikh D, Beall A, Yelin E, et al. Choosing wisely: the American College of Rheumatology's Top 5 list of things physicians and patients should question. Arthritis Care Res (Hoboken) 2013;65:329-39.

(4) Morgan DJ, Brownlee S, Leppin AL, Kressin N, Dhruva SS, Levin L, et al. Setting a research agenda for medical overuse. BMJ 2015;351:h4534.

(5) Abeles AM, Abeles M. The clinical utility of a positive antinuclear antibody test result. Am J Med 2013;126:342-8.

(6) Sheth T, Alcid D. Choosing not so wisely: the tale of Antinuclear Antibody testing [abstract]. Arthritis Rheumatol 2014;66:S1-S1402.

(7) Davis L, Goldstein B, Tran V, Keniston A, Yazdany J, Hirsh J, et al. Applying Choosing Wisely: identifying inefficiences of Antinuclear Antibody subserology testing in a safety net health system [abstract]. Arthritis Rheum 2013;65:S1-S1331.

(8) Bulbin D, Meadows A, Kelsey S, Harrison H, Denio A. Choosing subserologies wisely: an opportunity for rheumatologic healthcare resource savings [abstract]. Arthritis Rheumatol 2014;66:S1-S1402.

(9) Lesuis N, Hulscher M, Piek E, Demirel H, van der Laan-Baalbergen N, Meek I, et al. Choosing Wisely in daily practice: an intervention study on Antinuclear Antibody testing by rheumatologists. Arthritis Care Res (Hoboken) 2015 published on 28th September 2015. doi:10.1002/acr.22725.

(10) Flottorp SA, Oxman AD, Krause J, Musila NR, Wensing M, Godycki-Cwirko M, et al. A checklist for identifying determinants of practice: a systematic review and synthesis of frameworks and taxonomies of factors that prevent or enable improvements in healthcare professional practice. Implement Sci 2013;8:35.

(11) Sood R, Sood A, Ghosh AK. Non-evidence-based variables affecting physicians' test-ordering tendencies: a systematic review. Neth J Med 2007;65:167-77.

(12) Whiting P, Toerien M, de S, I, Sterne JA, Dieppe P, Egger M, et al. A review identifies and classifies reasons for ordering diagnostic tests. J Clin Epidemiol 2007;60:981-9.

(13) John O, Naumann L, Soto C. Paradigm shift to the integrative big-five trait taxonomy: history, measurement and conceptual issues. In: John O, Robins R, Pervin L, editors. Handbook of personality: theory and research. New York: Guilford Press; 2008:114-58.

(14) Ornstein S, Johnson A, Markert G, Afrin L. Association between family medicine residents' personality and laboratory test-ordering for hypertensive patients. J Med Educ 1987;62:603-5.

(15) Harrold LR, Harrington JT, Curtis JR, Furst DE, Bentley MJ, Shan Y, et al. Prescribing practices in a US cohort of rheumatoid arthritis patients before and after publication of the American College of Rheumatology treatment recommendations. Arthritis Rheum 2012;64:630-8.

SUPPLEMENT 1: TABLE ON QUESTIONNAIRES USED TO MEASURE ANA DETERMINANTS

Table 1 Questionnaires used to measure cognitive bias, personality traits, thinking styles, numeracy and ANA knowledge

Determinant	Rationale	Questionnaire	Number of subscales	Score range
Cognitive bias	Cognitive bias is an error in our thinking that occurs when people try to simplify information. Many different types of cognitive biases are known and they can influence medical decision making.[1, 2]	Inventory for Cognitive Bias in Medicine (ICBM). [3]	1	0 to 22 (higher score indicates less cognitive bias)
Personality	Personality traits play an important part in our behavior and can also influence medical decision making.[4]	Big Five Inventory (BFI; Dutch version). [5]	5 (extraversion, neuroticism, openness, conscientiousness and agreeableness)	1 to 5 on every subscale (higher scores indicating a stronger personality trait on the specific subscale)
Thinking styles	Two different thinking styles are distinguished: rational ('deliberate') and experiential ('automatic') thinking style. These styles can influence medical decision making.[6, 7]	Rational Experiential Inventory (REI; Dutch version). [8]	2 (rationality and experientality)	20 to 100 on every subscale (a higher score indicating a more rational/ experiential thinking style)
Numeracy	Numeracy and risk literacy are important when interpreting and acting on risk information. These skills are important in many health decisions but physicians seem to struggle with numeracy.[9]	Berlin Numeracy Test (BNT; Dutch version)	1	0 to 7 (a higher score indicating a higher level of numeracy)
Knowledge	Knowledge is often mentioned as an important determinant of behavior change.[10]	A self-developed questionnaire	1	-5.4 to 10 (negative scores due to correction for guessing). A higher score indicates more knowledge.

REFERENCES

(1) Brewer NT, Chapman GB, Schwartz JA, Bergus GR The influence of irrelevant anchors on the judgments and choices of doctors and patients. Med Decis Making 27; 203-11.

(2) Whiting P, Toerien M, de S, I *et al.* A review identifies and classifies reasons for ordering diagnostic tests. J Clin Epidemiol 60; 981-9.

(3) Hershberger PJ, Part HM, Markert RJ, Cohen SM, Finger WW Development of a test of cognitive bias in medical decision making. Acad Med 69; 839-42.

(4) Ornstein S, Johnson A, Markert G, Afrin L Association between family medicine residents' personality and laboratory test-ordering for hypertensive patients. J Med Educ 62; 603-5.

(5) Denissen JJ, Geenen R, van Aken MA, Gosling SD, Potter J Development and validation of a Dutch translation of the Big Five Inventory (BFI). J Pers Assess 90; 152-7.

(6) Sladek RM, Bond MJ, Phillips PA Why don't doctors wash their hands? A correlational study of thinking styles and hand hygiene. Am J Infect Control 36; 399-406.

(7) Sladek RM, Bond MJ, Huynh LT, Chew DP, Phillips PA Thinking styles and doctors' knowledge and behaviours relating to acute coronary syndromes guidelines. Implement Sci 3; 23.

(8) Witteman C, Bercken van den J, Claes L, Godoy A Assessing rational and intuitive thinking styles. Eur J Psychol Assess 25; 39-47.

(9) Wegwarth O, Schwartz LM, Woloshin S, Gaissmaier W, Gigerenzer G Do physicians understand cancer screening statistics? A national survey of primary care physicians in the United States. Ann Intern Med 156; 340-9.

(10) Flottorp SA, Oxman AD, Krause J *et al.* A checklist for identifying determinants of practice: a systematic review and synthesis of frameworks and taxonomies of factors that prevent or enable improvements in healthcare professional practice. Implement Sci 8; 35.

SUPPLEMENT 2: METHODS AND RESULTS ON THE SELF-REPORTED DETERMINANTS OF IMPROVEMENT IN ANA OVERUSE

METHODS

To assess self-reported determinants of improvement in ANA overuse we organized focus group sessions. For pragmatic reasons, only in the largest center two 1.5 hour focus group meetings were held. Here, all 14 eligible rheumatologists received an invitation (email) explaining the purpose of the meeting. Both meetings were planned after regular working hours, took place at the study hospital and were led by a female psychologist-researcher (JV, PhD). A minimum of 4 and maximum of 8 participants could participate in one session. Another female researcher (NL, MD) was present during the sessions to take field notes, but had no active role in the discussion. Both researchers worked at the study center and were known by the participating rheumatologists. No other people were present during the meetings except for the participants and the researchers.
During the focus group JV used an interview guide, which was prepared by NL, JV and AdB. Due to time constraints, pilot testing was not possible. The focus group started with a short introduction about its content, after which introductory questions were asked. This was followed by the main question "why did you change your behavior after the intervention, i.e. why are you now ordering less ANA tests compared to the period before the intervention". Both sessions were audiotaped and transcribed verbatim.

The transcripts of the focus groups were analyzed by NL and JV using thematic content-analysis. Firstly, the transcripts were read by NL to identify text parts relevant to the identification of (sub) themes (open coding). Secondly, all these parts were separately double coded by two NL and JV (axial coding). Next, they compared codes and discrepancies between them were resolved by discussion. Thirdly, all codes were grouped into (sub) themes by NL and JV (selective coding). Finally, the results of the coding process were discussed with the other researchers (AdB, MH and RvV) and compared to existing frameworks on this subject [1-3], if necessary adaptations were made.

RESULTS

During the focus group meetings seven major themes were identified as being of importance to ordering less ANA tests after the intervention. These themes related to 1) the intervention and study, 2) individual health professionals,3) patients, 4) professional interactions, 5)

incentives and resources, 6) capacity for organizational change and 7) social, political and legal factors (table 1). All these themes will be discussed below.

Theme 1: intervention and study factors. This theme comprised the subthemes study characteristics, study participation, quality of the intervention, content of the intervention, recommended behavior during the intervention and the research team. The majority of the codes were related to the content or quality of the intervention and the research team. For example, it was stressed by the participants that the content of the intervention was correct and refreshed existing knowledge, aiding their behavior change.

"And of course you agree with the content, otherwise you wouldn't do it…"

"What helped me enormously was actually the first talk, in which a few things that of course I already knew, became really clear. But that had simply drifted away."

Participants mentioned that the enthusiasm of and trust in the research team, and the fact that the research team was known to the rheumatologists, were important factors for changing their ordering behavior.

"An enthusiastic researcher, that has a mutually encouraging effect. And then you're more likely to adjust your behavior than when some idiot says: shouldn't you be ordering fewer tests?"

Theme 2: individual health professional factors. This theme comprised the subthemes clinical or previous clinical experience, knowledge & awareness, cognitions and professional behavior. Awareness was frequently mentioned as an important prerequisite for behavior change. Knowing or refreshment of existing knowledge on an ANA not being a perfect test was important for behavioral change, similar to being self-aware about own ordering behavior.

"The fact that it's in actually a useless screening instrument was an extra eye-opener."

Furthermore, participants indicated that their thoughts on ANA testing changed due to the intervention. Ordering less ANA tests was now seen as delivering high quality healthcare and not ordering an ANA test was being regarded as 'cool'. Moreover, the rheumatologists felt relieved by the fact that it can also be okay not to order an ANA test.

"So in terms of content, you have a feeling that you're now producing better work."

"I've been really good. In the past 3 months, I've only ordered one ANA…"

"The fact that someone says to you, 'you don't need to order those things': what a relief…"

Finally, it was mentioned that participants broke their former routines and they felt confident with the new behavior.

Theme 3: patients. It was mentioned by one rheumatologist that after the intervention fewer patients with a possible ANA associated disease were seen at the study clinic, leading to less ANA orders. The other participants did not directly recognize this determinant and no other patient factors were mentioned.

Theme 4: professional interactions. All determinants in this theme related to team processes. For example, participants felt safe within the team, they were willing to change their behavior and they were triggered by the non-anonymous feedback given during the intervention leading to competition between participants.

"I actually found it quite a challenge and so it's become a bit of a game to order as few ANA tests as possible (…). After all, you want to be top of the class, don't you."

Theme 5: incentives and resources. Financial (dis)incentives, nonfinancial (dis)incentives and assistance for clinicians were identified as subthemes. For example, potential cost savings by ordering less ANA tests, time constraints and the electronic health record system were referred to as influencing factors to change ordering behavior.

Theme 6: capacity for organizational change. Participants felt that the department had created a safe learning environment. In addition, the ANA study was regarded by the participants as fitting with existing policies on both hospital- and department level.

"It's actually the continuation of a trend that has been going on for years. That we're simply taking decisions with respect to the diagnosis and treatment, based on evidence."

Theme 7: social, political and legal issues. Within this theme participants only mentioned that societal concerns on (rising) healthcare costs influenced their behavior.

"With all the commotion surrounding the cost of healthcare, I think we need to look critically at our ordering behavior, that's part and parcel of it too."

Table 1 *Full coding of the focus groups*

Subtheme	Group	Code
Theme 1: Intervention & study factors		
Study characteristics		Nice and unique research
Study participation		Participating in the ANA study
		Participation has no negative consequences
		Completing questionnaires
Quality of the intervention		The content of the intervention was correct
Content of the intervention		Insight gained into own decision making process regarding ANA testing
		Intervention refreshed existing knowledge
		The relation between pre-test probability on disease, test characteristics and ANA use was explained
Recommended behavior during the intervention	Feasibility	During the intervention practical tools were given on how to decrease ANA use
Research team	Familiarity with the research team	The local opinion leader was part of the research team
		The researcher was known by the participants
	Trust in the research team	The local opinion leader was trusted
		The researchers was trusted by the participants
	Enthusiasm of the research team	The researcher was enthusiastic
Theme 2: Individual health professional factors		
(Previous) clinical experience		Previous experience with 'wrong' ANA ordering behavior by others
		Work experience
Knowledge & awareness	Awareness regarding the importance of clinical judgement and valid reasons to order a test	Awareness about the importance of history taking and physical examination
		Awareness about situations in which an ANA test has added value
		Awareness that you need to justify the ANA test to yourself
		Awareness that you need to have different policies for different ANA results in place beforehand
		Awareness about the value of thinking longer before ordering an ANA test
	Awareness regarding own practice and its consequences	Awareness about the importance of the relation between one's own ANA ordering behavior and the ANA test results
		Awareness about one's own ANA ordering behavior
		Awareness about one's own reasons to order an ANA test
	Awareness regarding false assumptions on test properties and previous behavior	Awareness about the ANA not being a perfect test
		Awareness about ANA testing giving false security
		Awareness about ANA test results being confusing
		Awareness about former routines being wrong

Cognitions (including attitudes)	Agreement with the proposed problem and solution	Agreement with the proposed problem of ANA tests being overused by the participants
		Agreement with the fact that ANA testing is not useful to do in all patients with rheumatic complaints
		Agreement with the content of the intervention
	Expected outcome	After the intervention the quality of care will improve
	Professional pride & curiosity	Proudness about one's own achievement
		Willingness to perform well
		Willingness to deliver care of good quality
		Seeing critical thinking as a challenge
		Curious about the results of the ANA study
	Fear of uncertainty	A need for control
	Responsibility	Feeling personally addressed by the feedback
	Changing attitudes on ANA testing	No ANA is cool
		Legitimately not doing something
Professional behavior	Nature of the new behavior	Breaking with former routines
		Behaviour equivalent to new way of thinking
		Reversing former routines
	Confidence in new behavior	Confirmation that you can legitimately not do something
		Notice that recording a patient's history and physical examination are enough to rule out ANA-related disease
Theme 3: Patient factors		
Patient numbers		Less patients seen at the clinic with a possible ANA associated disease
Theme 4: Professional interactions		
Team process	Experience with peer to peer coaching	Used to give and receive feedback to/from other rheumatologists within their team
	Atmosphere	Feeling safe within the team of rheumatologists
	Willingness to change	As a team being willing to change
		Broad support for the ANA study within the team of rheumatologists
	Peer to peer contact	Being able to discuss difficult patient cases with colleagues
		Competition among rheumatologists due to the ANA study

Theme 5: Incentives and resources		
Financial incentives and disincentives	Healthcare costs	One's own concern about healthcare costs
		Potential cost savings that can be achieved by ordering less ANA tests
Nonfinancial incentives and disincentives	Time constraints	Saving time by not ordering an ANA test
Assistance for clinicians	Electronic Health Record as barrier	Since the introduction of the EHR ordering an ANA costs more time
	Electronic Health Record as aid	The checklist, included in the EHR, on symptoms associated with rheumatic diseases helps with deciding (not) to order an ANA test
Theme 6: Capacity for organizational change		
Regulations, rules, policies	Department's personnel policy	Feedback is also discussed during the annual appraisal
		Presence of a safe learning environment
	Department's care policy	The ANA study fits with the general policy within the study clinic to apply evidence-based-medicine in daily practice
	Organisation's research policy	The ANA study matches with other studies performed at the study clinic
		The ANA study matches the study clinic policy
		Research being carried out on this topic
Theme 7: Social, political and legal factors		
Economic constraints on the health care budget	Healthcare costs	Importance of societal concerns on healthcare costs

REFERENCES

(1) Flottorp SA, Oxman AD, Krause J, Musila NR, Wensing M, Godycki-Cwirko M, et al. A checklist for identifying determinants of practice: a systematic review and synthesis of frameworks and taxonomies of factors that prevent or enable improvements in healthcare professional practice. Implement Sci 2013;8:35.

(2) Sood R, Sood A, Ghosh AK. Non-evidence-based variables affecting physicians' test-ordering tendencies: a systematic review. Neth J Med 2007;65:167-77.

(3) Whiting P, Toerien M, de S, I, Sterne JA, Dieppe P, Egger M, et al. A review identifies and classifies reasons for ordering diagnostic tests. J Clin Epidemiol 2007;60:981-9.

"The lab test results are back — now it's time to roll the dice."

4

Dutch Society for Rheumatology - Choosing Wisely

Nienke Lesuis
Alfons A den Broeder

Nederlands Tijdschrift voor Reumatologie 2015;2:56-7

The Dutch Society for Rheumatology (NVR) has issued a top 5 list of wise choices as part of the 'Choosing Wisely' campaign in the Netherlands. This campaign has been set up by the Federation of Medical Specialists and the Netherlands Organization for Health Research and Development (ZonMw) and was inspired by the American 'Choosing Wisely' campaign. The aim of the campaign is to provide support for medical specialists and patients in taking decisions about appropriate care. One of the elements of the Choosing Wisely campaign is the formulation of 'Wise Choices' by the medical specialist organizations. This list of evidence-based recommendations is intended as a practical guide that will enable doctors and patients to decide together on the most appropriate care for the individual patient. Lists of wise choices have already been issued for various specialisms including internal medicine, neurosurgery, radiology, urology and orthopedics.

The NVR has now also prepared a list of five recommendations on patient-centered diagnostics and care. These recommendations were drafted by the NVR's Quality Committee and assessed by the Council. In selecting the topics covered in the list, it was decided to focus on procedures that either occur frequently (measuring both ESR and CRP during follow-up of rheumatoid arthritis) or have a significant impact (doses of biological therapies that exceed the registered dose). The NVR hopes that this list will encourage its members to think about patient-centered care in rheumatology and that these five recommendations will be followed in daily practice so that patient-centered care will actually start to be delivered.

More information on the Choosing Wisely Netherlands campaign (including the publication of Wise Choices by the medical specialist organizations) is available at http://www.kwaliteitskoepel.nl/verstandig-kiezen/. More information about the American Choosing Wisely campaign can be found at http://www.choosingwisely.org/ and http://www.choosingwisely.org/societies/american-college-of-rheumatology/.

RECOMMENDATION 1

During the follow-up for patients with rheumatoid arthritis the preferred option is to request either a CRP test alone or an ESR test alone.

There is a high correlation between CRP and ESR, which means that there is no value in carrying out both tests at the same time in patients with rheumatoid arthritis (RA). Using the test characteristics (CRP has higher specificity than ESR), costs (ESR is cheaper than CRP) and other factors (CRP is less subject to influence by the age and sex of the patient), one of the tests can be excluded so that only either ESR or CRP is tested during the follow-up of RA patients.

Source: Crowson et al. Which measure of inflammation to use? A comparison of erythrocyte sedimentation rate and C-reactive protein measurements from randomized clinical trials of golimumab in rheumatoid arthritis. J Rheumatol 2009. Wolfe F. Comparative usefulness of C-reactive protein and erythrocyte sedimentation rate in patients with rheumatoid arthritis. J Rheumatol 1997.

RECOMMENDATION 2

Prescribing doses of biological therapies that are higher than the registered dose is not appropriate.

Biological therapies, like other medications, are registered for a dose that has maximum effect at patient group level. Higher doses do little or nothing to increase effectiveness and are accompanied by an increase in side effects. In addition, such treatment is not cost-effective. If a patient does not respond within 4 to 6 months of starting treatment, switch to a different anti-rheumatic therapy: this has a greater likelihood of producing a response than increasing the biological dose.

Source: Bongartz et al. Anti-TNF antibody therapy in rheumatoid arthritis and the risk of serious infections and malignancies: systematic review and metaanalysis of rare harmful effects in randomized controlled trials. JAMA 2006. Pavelka et al. Increasing the infliximab dose in rheumatoid arthritis patients: a randomised, double blind study failed to confirm its efficacy. Ann Rheum Dis 2009.

RECOMMENDATION 3

Only request Lyme diagnostics if the patient has joint symptoms consistent with Lyme disease.

Joint symptoms with Lyme disease often present as asymmetrical monoarthritis or oligoarthritis, particularly of the knee. This means that Lyme diagnostics are only appropriate in rheumatology where patients have monoarthritis or oligoarthritis involving the knee and other causes have been excluded, or where patients have large joint monoarthritis or oligoarthritis and there are indications for Lyme borreliosis in the patient's history (for example, a tick bite). Non-specific joint symptoms or tiredness in combination with a tick bite (or suspected bite) do not constitute indications for Lyme diagnostics.

Source: Dutch College of General Practitioners guideline for arthritis 2009 https://www.nhg.org/standaarden/volledig/nhg-standaardartritis#note-5). Textbook on rheumatology and clinical immunology 2013.

RECOMMENDATION 4

Only request an ANA test if history-taking and physical examination suggest a reasonable chance of an ANA-related disease.

25-30% of the healthy population has a positive ANA (low titer) and this high percentage of false positives means that ANA is not suitable as a screening test for rheumatic conditions. ANA testing only has value as part of the diagnostic process where a reasonable suspicion (based on history, physical examination and some additional tests) of an ANA-related disease exists. These relatively rare diseases (within the field of rheumatology) are: systemic lupus erythematosus, systemic sclerosis, polymyositis/dermatomyositis, mixed connective tissue disease and Sjögren's syndrome.

Source: Solomon et al. Evidence-based guidelines for the use of immunologic tests: antinuclear antibody testing. Arthritis & Rheumatism 2002.

RECOMMENDATION 5

The preferred option is to prescribe a traditional NSAID such as ibuprofen, naproxen or diclofenac - if necessary in combination with PPI - rather than a selective NSAID (etoricoxib and celecoxib).

There is no difference in effectiveness at patient group level between traditional NSAIDs and selective NSAIDs. The risk of gastric complications is also equally high for users of traditional NSAIDs combined with a proton pump inhibitor as for users of a selective NSAID. The differences in other side effects are minimal. The cost of a coxib is significantly higher than for a traditional NSAID.

"The lab test results are back — now it's time to roll the dice."

5

The value of routine creatine kinase and thyroid stimulating hormone testing in patients with suspected fibromyalgia: a cross-sectional study

Nienke Lesuis
Judith van Vliet
Nadine Boers
Nathan den Broeder
Hans Cats
Marlies EJL Hulscher
Aad Verrips
Alfons A den Broeder

Rheumatology (Oxford) 2016;55:1273-6. Doi 10.1093/rheumatolog/kew046

ABSTRACT

Objectives

To examine the prevalence of abnormal creatine kinase (CK) and thyroid stimulating hormone (TSH) values and previously unknown myopathy or thyroid disease in patients with suspected fibromyalgia (FMS).

Methods

All adult patients with suspected FMS, referred to the study hospital between November 2011 and April 2014 could participate. Patients with a history of myopathy or a previous diagnosis of thyroid disorder were excluded. Outcome measures were the percentages of abnormal CK and TSH values and the final diagnosis in those patients.

Results

373 patients were included in this study (94% female, mean age 42 years). 7.5% (95%-CI 5.2% to 10.6%) of them had an abnormal CK according to the local reference values. Applying the European Federation of the Neurological Societies guideline this changed to 0.5% (95%-CI 0.2% to 1.9%). In none of these patients hyperCKemia related myopathy was diagnosed and the final diagnosis was FMS in 89% of the patients.

3.5% (95%-CI 2.1% to 5.9%) of the patients had an elevated TSH and 1.4% (95-CI 0.6% to 3.1%) a lowered TSH, with one patient having an somewhat lowered Free Thyroid Hormone level. The final diagnosis was FMS in all these patients.

Conclusions

Abnormal CK and TSH values are rare in patients with suspected FMS, and do not result in an alternative diagnosis. Therefore, it seems that routine testing of CK and TSH levels in patients with suspected FMS referred to secondary care does not contribute to the diagnostic process.

INTRODUCTION

Fibromyalgia syndrome (FMS) is a chronic pain syndrome mainly affecting women. The prevalence is around 2% and due to this high prevalence as well as its large impact on patients' quality of life, FMS is a major health issue [1, 2]. In addition, due to disputes about aetiology, pathogenesis and classification, FMS still is a somewhat controversial disease [3, 4].

FMS can be diagnosed using the preliminary 2010 American College of Rheumatology (ACR) diagnostic criteria. One of those criteria is the absence of another disease that causes the complaints [5]. Although not specifically included in these diagnostic criteria, numerous blood tests have been recommended as routine screening for patients with suspected FMS to exclude alternative diagnoses [6]. However, it is still unclear whether these tests contribute to the diagnostic process in patients with suspected FMS.

Both creatine kinase (CK) and thyroid stimulating hormone (TSH) are frequently mentioned tests in the routine diagnostic work-up of FMS due to the presumed similarities between FMS and myopathies or hypothyroidism [4, 6-10]. However, the routine use of CK and TSH testing is, to our knowledge, not adequately supported by data. For example, there is no clear data to suggest that there is indeed a higher pre-test chance of myopathy or thyroid disease in patients with suspected FMS compared to the general population. Also, data is absent on the presumed increased chance of abnormal CK and TSH values in patients with suspected FMS compared to healthy controls. Finally, the added value of CK and TSH testing has not been assessed.

Our aim was therefore to explore the diagnostic value of CK and TSH testing in patients with suspected FMS. More specifically, we aimed to determine the prevalence of abnormal CK and TSH values and the prevalence of previously unknown myopathy or thyroid disease in patients with suspected FMS.

PATIENTS AND METHODS

Study design

This cross-sectional study is embedded in a study examining the prevalence of myotonic dystrophy type 2 among patients with suspected FMS. Details on the methods of the original study are described in a separate manuscript (under review). The methods relevant to our study will be described below.

The local ethical committee approved this study (CMO Nijmegen Arnhem: 3655109111) and all patients provided informed consent.

Setting and participants

All consecutive patients with suspected FMS referred to the outpatient clinic of the rheumatology department at the Sint Maartenskliniek (specialized hospital in rheumatology, orthopaedics and rehabilitation; the Netherlands) between November 2011 and April 2014 were eligible for participation. Exclusion criteria were age <18 years, an established other diagnosis responsible for the pain and currently receiving Cognitive Behavioural Therapy (as participation in the original study might interfere with its goals). In addition, we applied an extra exclusion criterion in our current study: having an established diagnosis of myopathy or thyroid disorder.

Outcome measures and data collection

Outcome measures were the percentage of patients with abnormal CK and TSH values and the final diagnosis in those patients.

Baseline patient characteristics were retrieved from the patients' charts (age, gender, medical history, final diagnosis, and if the patient was referred to the study centre as a second opinion). The latter characteristic was defined as the patient being seen in the last year by another secondary care specialist for the same complaints as presented during the first visit at the study centre. The final diagnosis (ICD-9 code) was determined by the treating rheumatologist using a protocolized diagnostic approach (history taking, physical examination, a neuromuscular questionnaire and the 2010 ACR FMS diagnostic criteria).

Serum CK and TSH were determined in all patients at the clinical laboratory of the study centre. The reference values used in the study centre for a normal CK in men and women were <200 U/l and <170 U/l respectively. As there is debate on the reference values for CK, we also used the hyperCKemia values proposed by the European Federation of the Neurological Societies (EFNS).[11] According to this guideline hyperCKemia is present when CK ≥504 U/l for non-black men and ≥325 U/l for non-black women [11]. CK levels may vary within individuals (for example after physical exercise), therefore CK testing was repeated if the first test result was abnormal. A normal TSH was defined as a TSH between 0.4 and 4.0 mE/l. Free Thyroid Hormone (FT4) was assessed when TSH was abnormal (reference values for normal FT4: 8 to 22 pmol/l). In case of abnormal TSH or FT4 values, no repeated testing was performed as TSH and FT4 values show no relevant day to day variation.[12]

Data analysis

Statistical analysis was performed using STATA version 13.1. All outcome measures are given as percentages or means and include the 95%-confidence interval (95%-CI) or standard deviation (SD), as appropriate. A post-hoc sensitivity analysis was performed excluding patients seen as a second opinion.

RESULTS

Setting and participants

All 398 patients included in the original study were considered for participation in this study. We had to exclude 23 patients because of known previous myopathy or thyroid disorder (figure 1). Due to missing data on CK and TSH values in 2 patients, 373 patients were included in the final analysis (mean age of 42 years ±11 years, 94% female). Of the patients with a final ICD-9 diagnosis of fibromyalgia, 92% fulfilled the ACR 2010 criteria (table 1).

Figure 1 *Study flowchart*

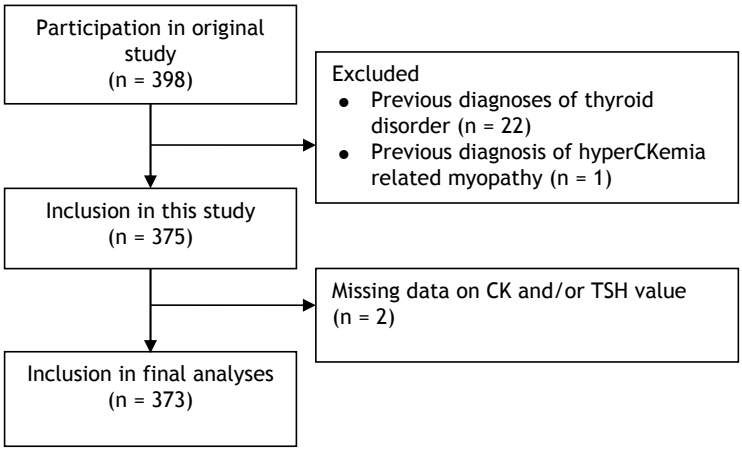


Table 1 *Characteristics of the patient population*

Characteristics	Study population (n = 373)
Female sex, *n(%)*	354 *(94%)*
Mean age, *years (range)*	42 ± 11 *(18 to 75)*
Second opinion, *n(%)*	107 *(29%)*
Final clinical diagnosis, *n(%)*	
Fibromyalgia	*358 (95%)*
Other*	*17 (5%)*

*Other diagnosis were mono-arthritis; polyarthropathy or polyarthritis unspecified; ankylosing spondylitis/Bechterev's syndrome; spondylosis; osteo-arthrosis; bursitis/enthesiopathy/synovitis; osteoporosis; arthropathy/arthralgia; hypermobility syndrome or Ehlers-Danlos syndrome and lumbago or neuralgia, neuritis or radiculitis, unspecified.

Results of CK and TSH testing

The mean CK in our study population was 96 ± 50 U/L (range: 23 to 470 U/L) and the mean TSH was 1.8 ± 1.4 mE/L (range: 0.1 to 17.6mE/L).

28 (7.5%; 95%-CI 5.2% to 10.6%) patients had an elevated CK according to the reference standard used at the study centre (range: 171 to 470 U/l). Using the EFNS reference standard, 2 (0.5%; 95%-CI 0.2% to 1.9%) patients had an abnormal CK (357 and 470 U/l respectively). In both of these patients a repeated CK test was normal. No diagnosis of hyperCKemia related myopathy was given in any of the 28 patients with an abnormal CK, and their final diagnoses were fibromyalgia (n=25), osteoporosis (n=1), arthropathy/arthralgia (n=1) and ankylosing spondylitis (n=1).

18 patients had an abnormal TSH value, with 13 (3.5%; 95%-CI 2.1% to 5.9%) patients having an elevated TSH and 5 (1.4%; 95%-CI 0.6% to 3.1%) a lowered TSH. One patient with an elevated TSH had a slightly reduced FT4 (7.7 pmol/l). This was interpreted by the treating rheumatologist as subclinical hypothyroidism unrelated to the FMS complaints, and no further action was taken. The final diagnosis in all patients with abnormal TSH values was fibromyalgia.

A sensitivity analysis excluding all second opinion patients yielded similar results to the original analyses.

DISCUSSION

This study suggests that relevant abnormal CK and TSH values and a final diagnosis of underlying thyroid disease or hyperCKemia related myopathy are rare in patients with suspected FMS. Therefore, it seems that routine testing of CK and TSH in secondary care patients with suspected FMS does not contribute positively to the diagnostic process.

To our knowledge this is the first study assessing the diagnostic value of two commonly used test in suspected FMS. Some strong points of our study are the prospective design, the well-defined patient population and the relatively large sample size. However, this study has some limitations. Firstly, we were not able to compare the CK and TSH results against a gold standard. For example, muscle biopsies could have been taken to serve as the golden standard for myopathies. However, this was not deemed feasible in the context of our study. Instead, we used a combination of history taking, physical examination, a neuromuscular checklist and the ACR 2010 FMS criteria as, to our view, create a reasonable surrogate golden standard.

Secondly, there was a relatively high proportion of second opinion patients in our study, probably higher than in other rheumatology departments. However, the sensitivity analyses

with exclusion of 2nd opinion patients showed similar results. Therefore this does not seem to hamper the validity of our findings.
Thirdly, patients may already have had their CK and TSH tested by the general practitioner. This could have caused a selection bias leading to underestimation of the prevalence of abnormal CK and TSH testing and associated diseases in our study. This because patients that were already diagnosed with thyroid disorder or hyperCKemia related myopathy would probably not have been referred to the rheumatologist for a FMS work-up. Although a valid concern, even if the majority of patients had received CK and TSH testing by the general practitioner this would not invalidate the generalisability of our results to other secondary care rheumatology departments. However, we would encourage the execution of a similar study in primary care.

In spite of widespread use of CK and TSH testing in suspected FMS and recommendations on this topic in some guidelines, our results do not contradict existing evidence. Although this may seem counterintuitive, this has to do with the earlier mentioned lack of data on CK and TSH testing in suspected FMS. Publications suggesting that routine CK or TSH testing is relevant in the diagnostic work-up of FMS [4, 6, 7] base this recommendation on the presumed similarity of symptoms between FMS and thyroid disease or myopathies. However, these recommendations are not based on prevalence data as provided in our study and also seem to ignore the very low prevalence of clinically relevant myopathies.
Furthermore, for both CK and TSH studies on normal values in healthy controls are available. Regarding CK, the median value in healthy controls was 84 U/l and 122 U/l for women and men respectively [13]. With regard to TSH, the prevalence of abnormal TSH ranged from 7.3% to 10.4% [14-16]. As our results come close to these results in the normal population, they support our conclusion of not using routine CK and TSH testing in suspected FMS patients. Based on these studies routine testing in suspected FMS patients would be just as irrational as routine testing in the whole general population.
Finally, with regard to CK testing there are some additional limitations. Several studies claim that CK in general is not a good test due to its low specificity. There is a wide variation in serum CK levels in the healthy population, dependent on physiological factors like sex, race and recent physical exercise [13, 17]. Therefore, the reference values for serum CK are subject to debate.[11] In our study the use of either strict or liberal reference values had a large impact on number of patients with abnormal values (28 versus 2 patients respectively), with the two highest CK values turning out to be false positives after repeated testing. Furthermore, elevated serum CK can reflect a muscular disorder but can also occur in other conditions such as hypothyroidism, drug use, alcoholism, muscle trauma, infections, and malignancies [18-20].

In summary, it seems that routine CK and TSH testing did not contribute to the diagnostic process in any of the studied patients. Therefore, we recommend against the routine use of CK and TSH testing in patients with suspected FMS seen at a secondary care centre. However, elective testing in patients with signs and or symptoms suggestive of muscular or thyroid disease should still be done and be followed by appropriate diagnostic or therapeutic steps.

FUNDING

This study received no external funding.

CONFLICT OF INTEREST STATEMENT

None of the authors have disclosures.

REFERENCES

(1) Wolfe F, Ross K, Anderson J, Russell IJ, Hebert L. The prevalence and characteristics of fibromyalgia in the general population. Arthritis Rheum 1995; 38:19-28.

(2) Hoffman DL, Dukes EM. The health status burden of people with fibromyalgia: a review of studies that assessed health status with the SF-36 or the SF-12. Int J Clin Pract 2008; 62:115-26.

(3) Wolfe F. Fibromyalgia wars. J Rheumatol 2009; 36:671-8.

(4) Wolfe F, Hauser W. Fibromyalgia diagnosis and diagnostic criteria. Ann Med 2011; 43:495-502.

(5) Wolfe F, Clauw DJ, Fitzcharles MA, Goldenberg DL, Katz RS, Mease P et al. The American College of Rheumatology preliminary diagnostic criteria for fibromyalgia and measurement of symptom severity. Arthritis Care Res (Hoboken) 2010; 62:600-10.

(6) Hauser W, Wolfe F. Diagnosis and diagnostic tests for fibromyalgia (syndrome). Reumatismo 2012; 64:194-205.

(7) Udd B, Meola G, Krahe R, Wansink DG, Bassez G, Kress W et al. Myotonic dystrophy type 2 (DM2) and related disorders report of the 180th ENMC workshop including guidelines on diagnostics and management 3-5 December 2010, Naarden, The Netherlands. Neuromuscul Disord 2011; 21:443-50.

(8) Neeck G, Riedel W. Thyroid function in patients with fibromyalgia syndrome. J Rheumatol 1992; 19:1120-2.

(9) Wilson J, Walton JN. Some muscular manifestations of hypothyroidism. J Neurol Neurosurg Psychiatry 1959; 22:320-4.

(10) Hochberg MC, Koppes GM, Edwards CQ, Barnes HV, Arnett FC, Jr. Hypothyroidism presenting as a polymyositis-like syndrome. Report of two cases. Arthritis Rheum 1976; 19:1363-6.

(11) Kyriakides T, Angelini C, Schaefer J, Sacconi S, Siciliano G, Vilchez JJ et al. EFNS guidelines on the diagnostic approach to pauci- or asymptomatic hyperCKemia. Eur J Neurol 2010; 17:767-73.

(12) Ehrenkranz J, Bach PR, Snow GL, Schneider A, Lee JL, Ilstrup S et al. Circadian and Circannual Rhythms in Thyroid Hormones: Determining the TSH and Free T4 Reference Intervals Based Upon Time of Day, Age, and Sex. Thyroid 2015; 25:954-61.

(13) Lilleng H, Abeler K, Johnsen SH, Stensland E, Loseth S, Jorde R et al. Variation of serum creatine kinase (CK) levels and prevalence of persistent hyperCKemia in a Norwegian normal population. The Tromso Study. Neuromuscul Disord 2011; 21:494-500.

(14) Ittermann T, Glaser S, Ewert R, Felix S, Volzke H, Dorr M. Serum thyroid-stimulating hormone levels are not associated with exercise capacity and lung function parameters in two population-based studies. BMC Pulm Med 2014; 14:145.

(15) Kvetny J, Ellervik C, Bech P. Is suppressed thyroid-stimulating hormone (TSH) associated with subclinical depression in the Danish General Suburban Population Study? Nord J Psychiatry 2015; 69:282-6.

(16) Hollowell JG, Staehling NW, Flanders WD, Hannon WH, Gunter EW, Spencer CA et al. Serum TSH, T(4), and thyroid antibodies in the United States population (1988 to 1994): National Health and Nutrition Examination Survey (NHANES III). J Clin Endocrinol Metab 2002; 87:489-99.

(17) Brewster LM, Mairuhu G, Sturk A, van Montfrans GA. Distribution of creatine kinase in the general population: implications for statin therapy. Am Heart J 2007; 154:655-61.

(18) Reijneveld JC, Notermans NC, Linssen WH, Wokke JH. Benign prognosis in idiopathic hyper-CK-emia. Muscle Nerve 2000; 23:575-9.

(19) Simmons Z, Peterlin BL, Boyer PJ, Towfighi J. Muscle biopsy in the evaluation of patients with modestly elevated creatine kinase levels. Muscle Nerve 2003; 27:242-4.

(20) Prelle A, Tancredi L, Sciacco M, Chiveri L, Comi GP, Battistel A et al. Retrospective study of a large population of patients with asymptomatic or minimally symptomatic raised serum creatine kinase levels. J Neurol 2002; 249:305-11.

"The lab test results are back — now it's time to roll the dice."

6

The effects of an educational meeting and subsequent computer reminders on the ordering of laboratory tests by rheumatologists: an interrupted time series analysis

Nienke Lesuis
Nathan den Broeder
Nadine Boers
Steven Teerenstra
Marlies EJL Hulscher
Ronald F van Vollenhoven
Alfons A den Broeder

Accepted for publication in Clinical and Experimental Rheumatology

ABSTRACT

Objectives

To examine the effects of an educational meeting and subsequent computer reminders on the number of ordered laboratory tests.

Methods

Using interrupted time series analysis we assessed whether trends in the number of laboratory tests ordered by rheumatologists between September 2012 and September 2015 at the Sint Maartenskliniek (the Netherlands) changed following an educational meeting (September 2013) and introduction of computer reminders into the Computerized Physician Order Entry System (July 2014). The analyses were done for the set of tests on which both interventions had focussed (intervention tests; complement, cryoglobulins, immunoglobins, myeloma protein) and a set of control tests unrelated to the interventions (alanine transferase, anti-cyclic citrullinated peptide, C-reactive protein, creatine, haemoglobin, leukocytes, mean corpuscular volume, rheumatoid factor and thrombocytes).

Results

At study start 101 intervention tests and 7660 control tests were ordered per month by the rheumatologists. After the educational meeting both the level and trend of ordered intervention and control tests did not change significantly. After implementation of the reminders the level of ordered intervention tests decreased with 85.0 tests (95%-CI -133.3 to -36.8, $p < 0.01$), the level of control tests did not change following the introduction of reminders.

Conclusions

In summary, an educational meeting alone was not effective in decreasing the number of ordered intervention tests, but the combination with computer reminders did result in a large decrease of those tests. Therefore, we recommend using computer reminders additionally to education if reduction of inappropriate test use is aimed for.

INTRODUCTION

Excessive use of diagnostic laboratory tests is a major problem in healthcare and has recently been recognized as an important field for research [1]. Not only does it constitute a waste of resources, it also results in an increased rate of false positives which may lead to further unnecessary testing, unnecessary treatment, and increased anxiety in both patients and physicians [1, 2]. Still, many clinicians order irrelevant laboratory tests despite the available information about unnecessary test utilization [2, 3].

A number of approaches have been used to reduce inappropriate testing using methods such as discouraging or not automatically fulfilling test orders, reducing availability of testing, giving feedback, raising awareness through education and the use of computer reminders [4]. Education, feedback and reminders are much used intervention strategies and although results differ between studies they have shown to be effective in different settings, including the reduction of unnecessary test orders [5-8].

A substantial proportion of diagnostic laboratory tests that were ordered at the rheumatology department of the study centre were tests that are not, or only very rarely, indicated for use in patients with a suspected rheumatic disease. This concerned the following tests: complement, cryoglobulins, immunoglobulins, myeloma protein (M protein) and Anti-Nuclear Antibodies (ANA). These tests were all ordered relatively frequently, are expensive, and overuse could easily lead to false positives and associated over-treatment. Therefore, the aim of this study was to examine the effect of two interventions on the number of ordered tests. Of note, ANA testing has been subject to a specific intervention which results are described elsewhere and was therefore not included in this study [9].

METHODS

Study design and setting

This is a controlled trial, using an interrupted time series design, on the effect of an educational meeting and subsequent introduction of computer reminders on the number of diagnostic tests ordered by 26 clinicians (15 rheumatologists, 7 residents, 4 physician assistants and nurse practitioners) from the rheumatology department at the Sint Maartenskliniek (specialized clinic for rheumatology, orthopaedics and rehabilitation medicine; the Netherlands).

As this was a quality improvement project performed by researchers working at the study centre, no formal ethical approval was needed. Furthermore, no informed consent of the patients was asked as no individual patient data were needed to assess the effectiveness of the interventions.

Interventions

An educational meeting regarding the intervention tests (complement, cryoglobulins, immunoglobulins and M protein; table 1) took place in September 2013. This meeting consisted of a 1-hour educational meeting, presented by two experienced medical immunologists working at the external laboratory where the intervention tests were performed. During the meeting background information on the intervention tests was provided and the rare indications in daily rheumatology practice for these tests were explained. Clinicians were advised not to order test tests unless specific signs or symptoms were present. At the time of the educational meeting 26 clinicians (15 rheumatologists, 4 physician assistants and 7 residents) were working at the rheumatology department, all were invited for the meeting.

Table 1 *Intervention and control tests used in this study*

Intervention tests	Pop-up text
Complement 3 (C3) Complement 4 (C4)	There is no indication within rheumatology for complement testing. One exception is C3 testing in the follow up of systemic lupus erythematosus to assess the risk of nephritis or neuro-psychiatric systemic lupus erythematosus although the evidence is limited. When clinical signs of complement deficiency disease are present, do not test for complement but refer to the internal medicine department.
Cryoglobulins	Cryoglobulins testing is only indicated in cutaneous vasculitis and/or mononeuritis, and when there are signs of hyperviscosity syndrome.
Immunoglobulin A (IgA) Immunoglobulin G (IgG) Immunoglobulin M (IgM)	There is no indication within rheumatology for measurement of immunoglobulins. There is a limited association with the presence of a rheumatic disease, and abnormal results have no clinical consequences. Only in the analysis of recurring infections and when monitoring gamma globulin therapy, testing is recommended.
M protein	There is no indication within rheumatology for assessment of monoclonal gammopathy. In case of a suspected haematologic malignancy referral to the haematology or internal medicine department is necessary, independent of a M protein result. Therefore, testing by rheumatologists is not needed
Control tests	
Alanine transaminase (ALT)	Leukocytes
Anti-cyclic citrullinated peptide (anti-CCP)	Mean corpuscular volume (MCV)
C-reactive protein (CRP)	Rheumatoid factor (RF)
Creatine	Thrombocytes
Haemoglobin	

In July 2014 computer reminders were incorporated into the Computerized Physician Order Entry system, which is linked to the Electronic Health Record (EHR) used at the study centre (EZIS 5.2, Chipsoft). The reminders functioned as follows: whenever a clinician tried to order one of the intervention tests a pop-up message appeared explaining in which specific rheumatology-related situation the test was indicated or not (table 1). If clinicians still wanted to order the test, a text field was shown asking the reason for ordering the test. The intervention tests could only be ordered if a reason was entered although there was no check whether the reason was valid. Both the educational meeting and the computer reminders were available for all clinicians ordering laboratory tests at the rheumatology department. The development of the reminders was a collaboration between the laboratory and rheumatology department.
The content of each reminder was created by 3 rheumatologists, including an expert on systemic diseases and the coordinator of the laboratory.

Outcome measures and data collection

The primary outcome of this study was the number of intervention and control tests ordered by clinicians working at the study clinic. Data on the outcomes are measured at equally spaced (monthly) intervals over the study period (September 2012 to September 2015). The control tests were a set of simple routine tests at which no intervention was targeted (table 1). These were included to control for time trends in clinicians' ordering behaviour that were unrelated to the interventions.
Secondary outcomes were the percentage of abnormal intervention test results and the percentage of valid reasons provided with the intervention test orders after implementation of the reminders, as judged by two experts.
Data on the number of intervention and control tests ordered by included clinicians were collected retrospectively using the local laboratory database. Data was collected for the twelve months before any intervention (pre-intervention period; September 2012 to September 2013), 10 months after the educational meeting but before implementation of the computer reminders (post-intervention period 1; September 2013 to July 2014), and the 14 months after computer reminders were implemented (post-intervention period 2; July 2014 to September 2015).

Data analysis

Statistical analyses were performed using STATA version 13.1. Depending on the type of variable, descriptive statistics are presented as percentages with the accompanying absolute numbers or as means.
To assess the impact of the two interventions on the numbers of tests ordered segmented regression analysis of interrupted time series analysis (ITSA) was used. Using ITSA we were

able to detect whether or not our interventions had a significantly greater effect than any underlying secular trend [10]. The segmented linear regression models included two change points (education and reminders) in order to estimate changes in the level and trend of the number of ordered tests after the change points. Separate regression models were run for the intervention and control test data. Sensitivity analyses assessing the impact of possible autocorrelation between measurements of consecutive months were performed. Of note, due to the order of implementation of the interventions, the effect of education is relative to the pre-invention period and the estimated effect of reminders is the additional effect of reminders on top of the educational intervention. Results are reported as regression coefficients (level and trend) with the corresponding 95% confidence interval (95% CI) and p-value.

A chi square test was used to compare the number of abnormal results over the three periods. Outcomes were defined as statistical significant if $p<0.05$.

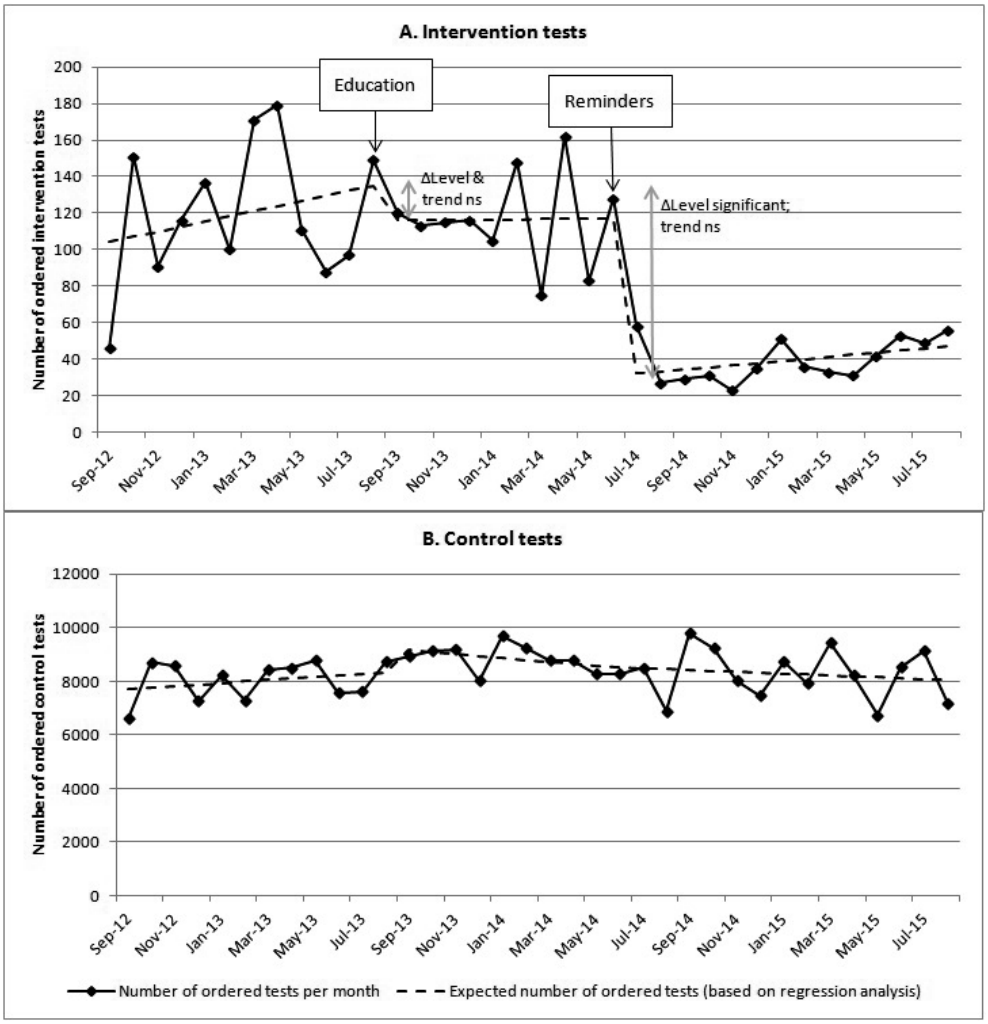
RESULTS

23 out of the 26 clinicians attended the educational meeting (15 out of 15 rheumatologists, all 4 of the physician assistants and 4 out of 7 residents), which was received positively and discussed critically.

At the start of the pre-intervention period 101 intervention tests were ordered per month (level) and during the pre-intervention period the number of ordered intervention tests increased with 2.8 test per month (non-significant; trend). After the educational meeting the number of intervention tests ordered remained stable, with no significant changes in level or trend. After implementation of the reminders, the level of ordered intervention tests decreased significantly with 85.0 tests (95%-CI -133.3 to -36.8, $p <0.01$) although the difference in trend was non-significant (1.06, 95%-CI -6.2 to 8.3, $p= 0.77$). With regard to the control tests, 7660 tests were ordered at the start of the pre-intervention period. None of the two interventions resulted in a significant change in either the level or the trend of ordered control tests. The results are also graphically depicted in figure 1.

In the analysis of intervention and of control test, the autocorrelation function and the Durbin-Watson tests (up to order 10) showed little indication of autocorrelation between consecutive months. Sensitivity analyses modelling autocorrelation (up to order 10) were performed. These did not substantially improve fit and led to the same conclusion for the interventions.

Figure 1 *Number of intervention and control test orders per month*



Similar to the numbers of intervention tests ordered, the percentage of abnormal test results did not change (pre-intervention period 17.7% and 17.4% after the educational meeting, p= 0.83). Comparable results were obtained for the period following the reminders with the percentage of abnormal intervention test results being 18.4% (p= 0.73).

Finally, all intervention tests ordered after the implementation of the reminders included a clearly worded reason. However, only 34% of those reasons were judged to be valid. In addition, this did not make any difference regarding the percentage of abnormal test results in this subgroup with18.8% being abnormal (p= 0.89).

DISCUSSION

In this study we observed that an educational meeting alone was not effective, however addition of targeted computer reminders did give a substantial decrease in the number of orders for intervention tests while the number of control test orders did not change.
The strengths of our study are the use of an interrupted time series design, the inclusion of control tests and the stepwise independent use of two different and relatively simple interventions. However, some limitations are also present. Firstly, the educational meeting and the implementation of computer reminders were available to all clinicians, meaning that after implementation we did not have a control group of intervention tests without being subject to the intervention. However, we did include a selection of routine laboratory tests as a control group to see whether a change in the number of intervention tests ordered could be caused by an overall change in test ordering behaviour. Secondly, when only looking at the intervention tests, we are not able to infer a definite causal relation between our intervention and the results afterwards because other events in the same time period might have attributed to the observed results. However, as no changes were observed in the number of ordered control tests after the implementation of the computer reminders, it is very unlikely that another event in the same time period caused the observed decrease in the number of intervention tests. Furthermore, the number of patients seen at the study clinic was relatively stable over the full study period (data not shown), excluding this as a reason for the observed results. Finally, a steady increase in the number of ordered intervention tests is visible near the end of the observation period. Although the difference in trend between the pre-intervention period and the period after intervention 2 was non-significant, the increase may continue past the end of this study. Therefore we plan to replicate our analysis in the future to extend our current follow-up period.

The lack of effect from the educational meeting alone in our study corresponds with previous research suggesting that education is a necessary but on its own insufficient intervention to reduce test ordering. Strengthening this conclusion is the fact that nearly all of the departments clinicians (23 out of 26) attended the educational meeting, meaning that the lack of effect from the educational meeting is not likely to be a result of a lack of attendance. A factor that could explain the lack of effect from the educational meeting is that it was provided by two immunologists from another hospital. This may have reduced the acceptance as one of our previous studies regarding a similar intervention showed that a familiar intervention team was an important factor for clinicians in changing their ordering behaviour (manuscript under preparation). Furthermore, our study confirms previous observations that education has a stronger effect when combined with other interventions [4, 11]. In addition, the effect of the computer reminders corresponds to other studies.

These studies also found the reminders to be effective, although this usually concerned smaller effects than observed in our study [5].
Other than the combination between education and computer reminders, three other factors are likely to have contributed to the effect of reminders observed in our study. Firstly, previous studies have shown that reminders are more successful if they interrupt in the practitioner's routine and are delivered at the time of decision making [12]. Both factors were incorporated with our reminder system, as reminders were shown directly after selecting one of the intervention tests in the CPOE system. Secondly, the intervention tests in this study were not ordered very frequently, even before any intervention. This means that the risk of pop-up fatigue, a weakness of computer reminders where overly frequent pop-ups get ignored after some time, was limited [11, 13]. However, pop-up fatigue may not have been completely avoided as the number of ordered intervention tests is slowly increasing again near the end of the study period. Thirdly, the pop-up text was created in collaboration by 3 of the department's rheumatologists and a laboratory coordinator, which may have increased the acceptance of the pop-ups compared to a situation where someone from outside the clinic was responsible for the pop-ups.

In addition to the number of tests ordered, we also assessed the percentage of abnormal test results. One would expect that a reduction in the number of unnecessary tests would also increase the proportion of correctly ordered tests and therefore an increase in the percentage of abnormal results. Contrary to this, the percentage of abnormal test results did not change, which may seem counterintuitive. However, this can be explained by the observation that only 34% of the reasons provided with the intervention test orders were valid, making it likely that patient selection by clinicians is still suboptimal. This is further supported by the fact that the percentage of abnormal test results did not differ between the intervention tests ordered with a valid reason and those without one, which means that a higher percentage of correctly ordered tests would not necessarily increase the percentage of abnormal results. This lack of difference in the percentage of abnormal results could indicate that even for patients with a valid reason for testing the test may not be useful in the majority of patients. In other words, overuse of intervention tests by clinicians is still present, although to a lesser extent than before the interventions.

In summary, the educational meeting alone was not effective in decreasing the number of ordered intervention tests, but the combination with computer reminders did result in a large decrease of those tests. Therefore, we recommend using computer reminders additionally to education if reduction of inappropriate test use is aimed for.

FUNDING

This study received no external funding.

CONFLICT OF INTEREST STATEMENT

Dr. van Vollenhoven reports grants from AbbVie, BMS, GSK, Pfizer, Roche, UCB, personal fees from AbbVie, Biotest, BMS, Crescendo, GSK, Janssen, Lilly, Merck, Pfizer, Roche, UCB, Vertex, outside the submitted work; the other authors have nothing to disclose.

REFERENCES

(1) Morgan DJ, Brownlee S, Leppin AL *et al*. Setting a research agenda for medical overuse. BMJ 2015;351:h4534.

(2) Zhi M, Ding EL, Theisen-Toupal J, Whelan J, Arnaout R The landscape of inappropriate laboratory testing: a 15-year meta-analysis. PLoS One 2013;8:e78962.

(3) Axt-Adam P, van der Wouden JC, van der Does E Influencing behavior of physicians ordering laboratory tests: a literature study. Med Care 1993;31:784-94.

(4) Janssens PM Managing the demand for laboratory testing: options and opportunities. Clin Chim Acta 2010;411:1596-602.

(5) Shojania KG, Jennings A, Mayhew A, Ramsay CR, Eccles MP, Grimshaw J The effects of on-screen, point of care computer reminders on processes and outcomes of care. Cochrane Database Syst Rev 2009:CD001096.

(6) Kobewka DM, Ronksley PE, McKay JA, Forster AJ, van WC Influence of educational, audit and feedback, system based, and incentive and penalty interventions to reduce laboratory test utilization: a systematic review. Clin Chem Lab Med 2015;53:157-83.

(7) Ivers N, Jamtvedt G, Flottorp S *et al*. Audit and feedback: effects on professional practice and healthcare outcomes. Cochrane Database Syst Rev 2012;6:CD000259.

(8) Forsetlund L, Bjorndal A, Rashidian A *et al*. Continuing education meetings and workshops: effects on professional practice and health care outcomes. Cochrane Database Syst Rev 2009: CD003030.

(9) Lesuis N, Hulscher ME, Piek E *et al*. Choosing Wisely in daily practice: An intervention study on Antinuclear Antibody testing by rheumatologists. Arthritis Care Res (Hoboken) 2015 published on 28 september 2015. doi:10.1002/acr.22725.

(10) Wagner AK, Soumerai SB, Zhang F, Ross-Degnan D Segmented regression analysis of interrupted time series studies in medication use research. J Clin Pharm Ther 2002;27:299-309.

(11) Baird G The laboratory test utilization management toolbox. Biochem Med (Zagreb) 2014;24: 223-34.

(12) Shojania KG, Jennings A, Mayhew A, Ramsay C, Eccles M, Grimshaw J Effect of point-of-care computer reminders on physician behaviour: a systematic review. CMAJ 2010;182:E216-E225.

(13) Baysari MT, Westbrook JI, Richardson K, Day RO Optimising computerised alerts within electronic medication management systems: A synthesis of four years of research. Stud Health Technol Inform 2014;204:1-6.


TREAT
TO
TARGET
Don't move... I'm sure I will
hit bull's eye this time!

Theme 2

Guideline adherence in the treatment of rheumatoid arthritis


TREAT
TO
TARGET
Don't move... I'm sure I will
hit bull's eye this time!

7

INTRODUCTION TO THEME 2

Quality in rheumatoid arthritis care

Sehrash Mahmood
Nienke Lesuis
Lilian HD van Tuyl
Piet van Riel
Robert Landewé

Best Practice & Research Clinical Rheumatology 2015;29: 664-679

ABSTRACT

While most rheumatology practices are characterized by strong commitment to quality of care and continuous improvement to limit disability and optimize quality of life for patients and their families, the actual step towards improvement is often difficult. This because there are still barriers to be addressed and facilitators to be captured before a satisfying and cost-effective practice management is installed. Therefore this review aims to assist practicing rheumatologists with quality improvement of their daily practice, focusing on care for rheumatoid arthritis (RA) patients.

First we define quality of care as 'the degree to which health services for individuals and populations increase the likelihood of desired health outcomes and are consistent with current professional knowledge'. Often quality is determined by the interplay between structure, processes and outcomes of care, which is also reflected in corresponding indicators to measure quality of care. Next, a brief overview is given of the current treatment strategies used in RA, focusing on the tight control strategy, since this strategy forms the basis of international treatment guidelines. Adherence to tight control strategies leads, also in daily practice, to better outcomes in patients with regard to disease control, functional status and work productivity. Despite evidence in favor of tight control strategies, adherence in daily practice is often challenging. Therefore, the next part of the review focuses on possible barriers and facilitators of adherence, and potential interventions to improve quality of care. Many different barriers and facilitators are known and targeting these can be effective in changing care, but these effects are rather small to moderate. With regard to RA, few studies have tried to improve care, such as a study aiming to increase the number of disease activity measures done by a combination of education and feedback. Two out of the three studies showed markedly positive effects of their interventions, suggesting that change is possible. Finally, a simple step-by-step plan is described, which could be used by rheumatologists in daily practice wanting to improve their RA patient care.

INTRODUCTION

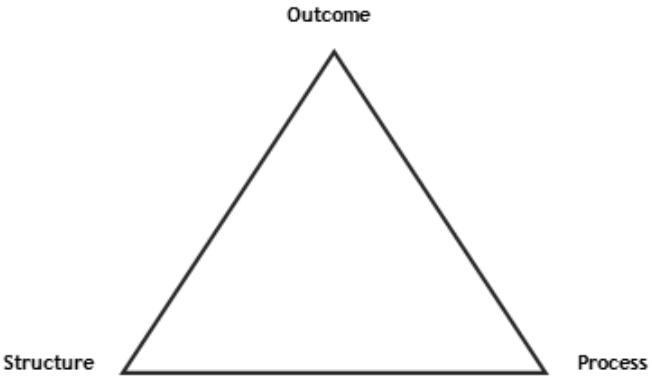
Musculoskeletal disorders, such as gout, osteoarthritis and rheumatoid arthritis (RA) are considered to be among the most burdensome medical conditions (1). This has led to the execution of many randomized controlled trials (RCTs) that have provided evidence for the best therapeutic interventions for these diseases. Despite this constant stream of evidence based recommendations, the translation into daily practice is often suboptimal (1).

While many practicing rheumatologists will agree that quality of care is an important aspect in rheumatology, the actual step to improve quality of care is often difficult, since rheumatologists do not know where and how to start, and there are no clear strategies available how to approach improvement of quality of care in their clinical practice. This review, with the goal of assisting practicing rheumatologists with their own quality improvement of care, aims to fill this gap. It starts with a brief general introduction on quality of care and how to measure this. Thereafter, the focus will shift to RA and we will discuss what optimal RA care is, how we can measure whether quality demands are met or not, and how this could be improved. In the latter part, two case descriptions of successful quality improvement projects in RA will be discussed. Finally, we will give practical recommendations to rheumatologists who want to further improve their own performance.

A. WHAT IS QUALITY OF CARE AND HOW CAN YOU MEASURE IT?

Quality of care in itself is a rather abstract term, but more practical descriptions do exist. One of the most used descriptions, developed around 1980 by Donabedian, distinguishes structures, processes and outcomes of care (2). The structure of care describes aspects of the setting in which care is delivered, such as the number of rheumatologists or the presence of a treatment protocol. Next, the process of care describes the actions of the health care professionals, for example, whether the protocol is indeed followed. Finally, the outcome reflects the effect of the given care in terms of mortality, morbidity and health status. It is believed that more desirable outcomes are obtained if the structure of care provides the opportunity to deliver the most optimal care processes (fig 1).

***Figure 1**: the Donabedian Triad. Donabedian hypothesized that all elements are linked to each other (3)*



Around 1990, the Institute of Medicine (IOM) defined quality of care as 'the degree to which health
services for individuals and populations increase the likelihood of desired health outcomes and are consistent with current professional knowledge.' Furthermore the IOM formulated the following six criteria that pertain to quality of care; Care should be i) safe; ii) effective; iii) patient-centered; iv) timely; v) efficient; and vi) equitable (4),. When using these criteria it is important to take into account the different perspectives of the stakeholders (patients or health insurers, for example) (5).

Knowing how to describe quality of care is a prerequisite for its measurement. Often quality indicators are used to assess quality of care. A quality indicator is 'a measurable element of practice performance for which there is evidence or consensus that it can be used to assess the quality, and hence change the quality of care provided'(6). Quality indicators are often grouped using the before mentioned quality definition by Donabedian, thus providing structure-, process- and outcome indicators. Outcome indicators reflect the result of the care that was provided by the healthcare provider, while process indicators reflect the actual care given to patients ('what is done'). Structure indicators, on the other hand, describe organisational aspects ('what is available') (7). How these indicators are used within rheumatology will be described later in this review; we will now first describe what optimal care in RA is.

B. WHAT IS OPTIMAL RHEUMATOID ARTHRITIS CARE?

The treatment of RA has substantially improved during the last two decades. Until 1990 the use of Disease Modifying Anti-Rheumatic Drugs (DMARDs) was limited, due to the belief that DMARDs were too toxic to use for a non-life threatening disease such as RA (8-10). Obviously, these assumptions changed and in the following decades the importance of DMARDs, both synthetic and biological, in the management of RA has become more obvious. This has resulted in treatment strategies such as a step-up approach and combination therapy, as well as treat-to-target strategies (8-10).

Improvements in care for RA patients have not only been the result of an increase in the number of effective therapeutic options, but also because of broader insight into the course of the disease and its prognosis. For example, it became clear that active RA is associated with a high risk of cardiovascular morbidity and mortality. Furthermore, it was found that RA treatment should start as early as possible (in the so called 'window of opportunity') in order to prevent the occurrence of irreversible joint damage or at least to halt progression of the disease (11, 12) Other terminology used in this context is 'hit hard, hit early' (intensive treatment early in the disease course) and 'tight control'. Although tight control is the mainstay of optimal clinical RA care, it is not the only part of good RA care. Shared care with specialized nurses or physician assistants, cardiovascular risk management and the management of comorbidities are some examples of other important aspects of RA care. As we cannot cover all these aspects and the tight control principle currently forms the basis of major treatment guidelines, we will now focus on this strategy (13, 14).

Tight control, also called treat to target, can be defined as 'frequent assessment of disease activity combined with an objective structured protocol to make treatment changes that maintain low disease activity or remission at an agreed target' (15). Recently, an international task force provided an update of the 2010 treat to target recommendations (14). These recommendations describe a generic principle or strategy, not necessarily advocating a particular type of intervention, that should be adhered to in order to reach disease remission or low disease activity in RA patients. The four overarching principles and ten recommendations focus around shared decision making, the importance of setting a treatment target, measuring disease activity, changing treatment until the desired goal is reached and maintaining the treatment goal thereafter (14).

Various studies have proven the effectiveness of the tight control regime, with the TICORA (Tight Control of Rheumatoid Arthritis) study being one of the first to show the beneficial effects of tight control. In the TICORA study, patients in the tight control group had a

significantly better disease outcome after 18 months as compared to the control group with regard to the EULAR good response criteria (82% vs 44%, p<0.0001) and the mean decrease in Disease Activity Score (DAS; -3.5 vs -1.5, p<0.0001) (11).

After the TICORA study, several studies have replicated these findings and in 2010 a meta-analysis on the effects of tight control was published. This meta-analysis concluded that patients treated according to tight control principles had significantly better DAS-28 responses as compared to patients treated with usual care (mean difference = 0.59; p < 0.001) (16). In addition, they also compared tight control with- and without protocolized treatment adjustments. These comparisons showed a beneficial effect of protocolized treatment adjustments, with the DAS-28 decreasing 0.66 points more (95% CI 0.72 to 1.11; P < 0.0001) if a specific treatment protocol was used. Also, an improvement in functionality and a decrease in joint damage was observed (16).

Although tight control studies so far have focussed on reaching remission or low disease activity, secondary analyses have shown that lower disease activity is also associated with improved work productivity, less comorbidity and lower cardiovascular risk (17). This may imply that applying tight control in daily practice benefits RA patients with regard to disease control and important other aspects of their lives such as work.

In summary, due to the complexity of RA and the increasing treatment arsenal, it can be difficult for rheumatologists to provide optimal RA care in all patients. However, it seems that using tight control based treatment strategies could assist rheumatologists in achieving low disease activity or remission in the majority of their patients, ensuring better clinical outcomes and promoting better work productivity, less comorbidity and lower cardiovascular risk (17).

C. HOW CAN WE MEASURE WHETHER OPTIMAL RHEUMATOID ARTHRITIS CARE IS PROVIDED?

As mentioned in the first section, quality of care can be assessed using predefined quality indicators for the structure, processes and outcomes of care (6,7). With regard to RA, a broadly accepted set of quality indicators is lacking. However, several groups around the world have made an attempt to develop sets of RA quality indicators. In the following paragraphs some of these indicator sets will be discussed.

Dutch researchers have described one of the first sets, designed to monitor RA disease course in the Dutch Rheumatoid Arthritis Monitoring (DREAM) cohort. This indicator set consists of 10 process-, 5 structure- and 3 outcome-indicators and is divided into different subcategories. These subcategories are: the measurement of disease activity, structural damage, functionality, follow-up frequency, intensification of pharmacological therapy, prerequisites for measuring disease activity and patients' disease activity (for example the percentage of RA patients in remission one year after diagnosis) (18).
Two other groups in Europe have also developed sets of quality indicators. Firstly, the National Health Service in England (NHS) has developed quality indicators for RA, along with indicators for other diseases, in order to standardize improvements in the delivery of primary care (19). Management of RA in primary care may include; checking for cardiovascular risk and blood pressure, checking the risk for osteoporosis and checking for signs of depression. During an annual meeting in primary care the effects of the disease upon a person's life can be assessed, for example by monitoring the side-effects of medication or assessing the psychological situation of the patient. The NHS indicator sets reflects this care and comprises one structural-, one outcome-, and two process indicators, subdivided in two domains. The first domain is 'records'(19). This domain documents whether the primary care physician establishes and maintains a register for patients of 16 years and older with RA. The second domain is 'ongoing management'. In this domain, for example the documentation of the percentage of patient who had an annual face to face meeting with general practitioner in the preceding 12 months is documented (19).

The second European indicator set is developed by the European Musculoskeletal Conditions Surveillance and Information Network (EUMUSC.NET) and contains 14 indicators (one outcome, two structural and 11 process indicators) (20). To our knowledge, the EUMUSC.NET has not divided these indicators in domains. Therefore, we decided to divide this extensive list of indicators into six domains, namely; organisation, screening, pharmacological treatment, non-pharmacological treatment, monitoring and outcome. These indicators are also stated in table 1, where they can be compared with the American College of Rheumatology (ACR) indicators.

Table 1: indicators from the ACR and EUMUSC.net subdivided in domains

Subject	Indicator*	Source*
Organisation	Rheumatology practices should provide information (written or website) on how a patient can contact the practice for urgent consultations (in case of flares/worsening of the disease, serious side effects).	EUMUSC.net
	If a patient is referred to a physician for a new diagnosis of rheumatoid arthritis, then the patient should be seen by the physician within 3 months.	EUMUSC.net
	If a patient presents with suspected rheumatoid arthritis then he/she should be referred to and seen by a specialist (preferably a rheumatologist) for confirmation of diagnosis within 6 weeks after the onset of symptoms.	EUMUSC.net
	If a patient is diagnosed with RA and there are joint damage/soft tissue problems that may be solved by surgery then the patient should be assessed by an orthopaedic surgeon within 3 months.	EUMUSC.net
Education	If a patient is newly diagnosed with RA, then, he or she should be given individually tailored education by relevant health professionals about the natural history, treatment, and self -management of the disease within 3 months.	EUMUSC.net
Screening	Percentage of patients 18 years and older with a diagnosis of rheumatoid arthritis who have documentation of a tuberculosis (TB) screening performed and results interpreted within 6 months prior to receiving a first course of therapy using a biologic disease-modifying anti-rheumatic drug (DMARD	EUMUSC.net; ACR
Non-pharmacological treatment	If a patient is newly diagnosed with RA then a referral to a relevant health professional for instruction on an individualized exercise program including advice for physical activity, range of motion, muscle strengthening- and aerobic exercises should be provided within 3 months.	EUMUSC.net
Pharmacological treatment	If a patient has rheumatoid arthritis and is being treated with a DMARD and there is evidence of increased disease activity or there is evidence of progression of RA bony damage over a 6-month period of time, then one of the following should be done: change DMARD dose or route of administration, change DMARD, add an additional DMARD, start or increase dose of glucocorticoids or provide local glucocorticoid injection(s), unless the patient refuses or all of the above are contraindicated	ACR
	A rheumatologist should intensify disease modifying medication when disease activity is moderate or high (according EULAR recommendations)	EUMUSC.net
	If a patient has RA, then he/she should have a treatment plan developed between him/her and his/her clinician/ health professionals at each visit.	EUMUSC.net

Monitoring	If a patient is diagnosed with RA, then a rheumatologist and/or relevant health professionals from the multidisciplinary team should assess and document the following variables: 1) a measure of disease activity such as composite scores like DAS 28 or any of its variants CDAI or S-DAI, 2) structural damage (using the best available method, e.g. x-ray, MRI, ultrasound), 3) functional status, (e.g.HAQ), and 4) labour force participation. The assessment and documentation should occur at baseline and thereafter at appropriate time intervals, at least annually for 1, 3 and 4 examination.	EUMUSC.net
	If a patient has a diagnosis of rheumatoid arthritis, then baseline radiographs of the hands or feet should be performed within 3 months of the initial diagnosis and every 3 years.	EUMUSC.net
	Percentage of patients 18 years and older with a diagnosis of rheumatoid arthritis who have an assessment and classification of disease activity at least once within 12 months.	ACR
	Percentage of patients 18 years and older with a diagnosis of rheumatoid arthritis who have an assessment and classification of disease prognosis at least once within 12 months.	ACR
	Percentage of patients 18 years and older with a diagnosis of rheumatoid arthritis who have been assessed for glucocorticoid use and, for those on prolonged doses of prednisone > 10 mg daily (or equivalent) with improvement or no change in disease activity, documentation of glucocorticoid management plan within 12 months.	ACR
	Percentage of patients 18 years and older with a diagnosis of rheumatoid arthritis for whom a functional status assessment was performed at least once within twelve months.	
Outcome	If a patient is diagnosed with active RA (i.e. DAS 28 over 3.2) then the disease activity should be low (i.e. DAS28 below 3.2) 6 months after treatment has started	EUMUSC.net
Follow up	If a patient is diagnosed with RA and the target (=remission or low disease activity) is not attained then follow up visit should be scheduled by a rheumatologist within 3 months and when the target is attained a rheumatologist or a specialized nurse in rheumatology should schedule follow up visits at least once a year.	EUMUSC.net

*Indicators are reported in exact same words as used by corresponding authors of each published set.

In the United States, the Arthritis Foundation and the ACR have also developed sets of indicators The extensive set from the Arthritis Foundation comprises 27 process indicators and they can be divided in 17 domains being: time to referral, history and examination, regular follow up, radiographs of hand and feet, radiographs of cervical spine, DMARDs, folic acid with methotrexate (MTX), osteoporosis prophylaxis, use of glucocorticoids, exercise, assistive devices, surgery, baseline and follow-up studies, methotrexate transminitis (increases in aminotransferases), informing patients about risks (such as risks regarding the use of non-steroidal anti-inflammatory drugs (NSAIDs), DMARDs, glucocorticoids and narcotics), reproductive issues and finally vaccines (21).The set developed by the ACR includes five process indicators and one outcome indicator. We have grouped these indicators in the following domains; screening, pharmacological treatment and monitoring, as to our knowledge no domains were proposed by the research team. The indicators from this set are also stated in table 1 for illustrative purposes (22).

An international task force developed a set of 10 quality indicators, using the Measurement of Efficacy of Treatment in the Era of Outcome in Rheumatology (METEOR) database. This set consists of seven process indicators and three outcome indicators; time to diagnosis, antibodies and radiographic assessment, frequency of visits, disease activity assessment, functional status assessment, remission of disease activity (clinical remission), low disease activity, level of functional limitation, time to first DMARD and type of first DMARD (23).

Finally, the Australian Rheumatoid Association has proposed a set of three process indicators. These indicators cover measurement of disease activity and co-morbidities (24).

Looking at these seven sets, obviously the majority of these indicators include process measures. Of the 82 indicators, only 9 were outcome indicators, (18-24). The majority report the ‘number of times’ a certain outcome is measured (process), rather than the actual outcomes themselves. When thinking back to the triad suggested by Donabedian, which links process, structures and outcome of care to each other, an imbalance between the different types of indicators in the current sets is quite apparent (2, 7). Furthermore, the availability of many sets from which to choose may further jeopardise the implementation in daily practice. In conclusion, a better-balanced (more outcome, less process) and more widely accepted indicator set would be instrumental in achieving uniform measurement of RA care. In the meantime, a rheumatologist willing to measure the quality of his own practice should choose one of the available indicator sets which best reflects what one wants to measure.

D. IS OPTIMAL CARE DELIVERED TO RHEUMATOID ARTHRITIS PATIENTS?

As described in this review, the use of tight control strategies is beneficial to RA patients and major treatment guidelines have embraced the tight control strategy (11, 13, 15). Unfortunately, the existence of these guidelines and the underlying evidence for their efficacy seems to be insufficient to ensure application of tight control in daily practice. This issue has been addressed by several studies, with rather underwhelming results. Here we will briefly summarize some of these studies.

Benhamou *et al*, assessed the potential gap between daily practice and recommendations on first DMARD prescription in RA in the French multicenter ESPOIR cohort (25). This cohort included early RA patients between 2002 and 2005, and during this period two guidelines on DMARD treatment in early RA were introduced: the national guideline by the French Society of Rheumatologists (STPR guideline) and the international set of management recommendations by the European League Against Rheumatism (EULAR guideline). Benhamou *et al* observed that first DMARD prescriptions in early RA were performed according to STPR recommendations in 58% of the patients, while 54% of the prescriptions adhered to the EULAR guidelines. As both guidelines were presented at international conferences at the end of the ESPOIR inclusion period, the authors concluded that the potential gap between evidence and practice was substantial (25).

Around the same time another European study assessed treatment patterns in early RA patients (ERAN cohort) (15). In this cohort 97% of the patients were prescribed a DMARD; however, median time between onset of the RA symptoms and DMARD prescription was 8 months. Most often the first DMARD was prescribed as monotherapy (91%) and the addition of a second DMARD later in the treatment course was observed in 48%. Despite the high percentage of DMARD users, only 33% of the patients were in DAS-28 remission after three years (15).

Also in the United States the prescribing practices of rheumatologists were assessed. In contrast to the ESPOIR and ERAN cohort, this study also included biologic DMARD (bDMARD) prescriptions and compared adherence before- and after the publication of the ACR treatment recommendations (26). In this study, 43% of the MTX monotherapy users with moderate disease activity and poor prognosis received care according to the ACR guideline; in MTX monotherapy users with high disease activity this was 51%. In patients using multiple conventional synthetic DMARDs (csDMARDs), 43% and 51% of those with moderate or high disease activity respectively, received care consistent with the ACR guideline. Interestingly, the publication of the ACR guideline did not result in improved guideline adherence in patients with active disease (26).

The results of these three studies may seem rather disappointing. However, a Dutch study on guideline adherence in the DREAM remission induction cohort yielded more positive results (27). In this early RA cohort, adherence to adequate monitoring of disease activity (DAS-28 assessed at least every 3 months) and a predefined treatment protocol was assessed. The researchers observed that adequate monitoring of disease activity took place in 88% of the visits and in 69% of the visits the rheumatologist adhered to the treatment recommendations. According to the authors these results point to the feasibility of using a tight control strategy in daily practice (27).

Finally, in another study using data from the ESPOIR cohort it was observed that adherence to tight control strategies in daily practice may have real benefits for patients (28). It was found that early RA patients who were not treated according to the 2007 EULAR recommendations on early RA were at an increased risk of radiographic progression at 1 year and functional impairment at 2 years (odds ratio (OR) 1.98;1.08 to 3.62) and OR 2.36;1.17 to 4.67 respectively) (28). In addition two studies presented at the 2014 ACR annual meeting concluded that RA disease control was better in patients in whom a tight control strategy was actively applied than in patients in whom this was not the case (29, 30).

Overseeing the literature described in this section, we can conclude that application of a tight control strategy in daily practice is feasible, but general adherence is not yet optimal (14, 24-29). In addition we saw that suboptimal adherence may have negative consequences for patients with regard to disease control, radiographic progression and functional status.

E. HOW CAN YOU IMPROVE RHEUMATOID ARTHRITIS CARE?

We have learned from various studies described in the previous section that optimal RA care is not always delivered to patients. The main question now is: how can we improve RA care? In order to answer this question we will look to the field of implementation science, an area specifically focusing on bridging the gap between evidence and clinical practice. First we will describe which factors can influence the successful uptake of evidence by physicians (also called 'barriers and facilitators' to implementation). Thereafter, potential interventions to change clinical practice are described.

Factors influencing adherence to evidence-based recommendations

Many studies have made an attempt to understand barriers and facilitators associated with implementing change and consequently different checklists, frameworks and taxonomies have been developed. In 2013 a systematic review by Flottorp *et* al was published, describing the development of a comprehensive checklist (TICD checklist), integrating previous

checklists, frameworks and taxonomies (31). In this review 'barriers and facilitators to implementation' are called 'determinants of practice', and the developed checklist consists of 54 of such determinants. The determinants are subdivided in seven domains, namely: guideline factors; health professional factors; patient factors; professional interactions; incentives and recourses; capacity for organizational change; social, political and legal factors (table 2). According to the authors this checklist can be used as a screening tool to identify determinants that need further investigation before a particular change can be implemented. The underlying idea is that this can facilitate the development and evaluation of interventions tailored to specific determinants (31). This hypothesis will be discussed further when potential interventions to change practice are described.

***Table 2**: Overview of the domains and determinants of the TICD checklist*

Subdomain	Determinant
Domain 1: Guideline factors	
Recommendation	Quality of evidence supporting the recommendation
	Strength of recommendation
	Clarity
	Cultural appropriateness
	Accessibility of the recommendation
	Source of the recommendation
	Consistency with other guidelines
Recommended clinical intervention	Feasibility
	Accessibility of the intervention
Recommended behavior	Compatibility
	Effort
	Trialability
	Observability
Domain 2: Individual health professional factors	
Knowledge and skills	Domain knowledge
	Awareness and familiarity with the recommendation
	Knowledge about own practice
	Skills needed to adhere
Cognitions (including attitudes)	Agreement with the recommendation
	Attitudes towards guidelines in general
	Expected outcome
	Intention and motivation
	Self-efficacy
	Learning style
	Emotions
Professional behavior	Nature of the behavior
	Capacity to plan change
	Self-monitoring or feedback

Domain 3: Patient factors	
n/a	Patient needs
	Patient beliefs and knowledge
	Patient preferences
	Patient motivation
	Patient behavior
Domain 4: Professional interactions	
n/a	Communication and influence
	Team processes
	Referral processes
Domain 5: Incentives and resources	
n/a	Availability of necessary resources
	Financial incentives and disincentives
	Nonfinancial incentives and disincentives
	Information system
	Quality assurance and patient safety systems
	Continuing education system
	Assistance for clinicians
Domain 6: Capacity for organizational change	
n/a	Mandate, authority, accountability
	Capable leadership
	Relative strength of supporters and opponents
	Regulations, rules, policies
	Priority of necessary change
	Monitoring and feedback
	Assistance for organizational changes
Domain 7: Social, political and legal factors	
n/a	Economic constraints on the health care budget
	Contracts
	Legislation
	Payer or funder policies
	Malpractice liability
	Influential people
	Corruption
	Political stability

As described in table 2, many factors can influence the translation of research into practice, with the diffusion and dissemination of innovations facilitated or hindered by these factors. The TICD provides a comprehensive overview of these factors, but for the practicing rheumatologist in search of feasible tools to improve his own practice, this checklist can be difficult to apply. In our opinion, and partly based on own experience, we would suggest

that successful implementation of change in daily practice relies mainly on three domains: 1) knowledge about the desired change, 2) motivation to realize the change and 3) being able to apply the new behavior.

Potential interventions to improve care

Recommendations or guidelines can assist rheumatologists in providing optimal clinical care to RA patients, but we have also seen that implementation of guidelines is often difficult (1, 14, 24, 25). In the previous section we have discussed studies that have identified many factors that may influence the successful implementation of change. Now the question remains what types of interventions could be applied to improve care and how effective they are.

Before answering this question we will first summarize the types of interventions that do exist. The Effective Practice and Organization of Care (EPOC) group is a Cochrane review group specialized in undertaking reviews on all types of interventions that aim to improve health professional practice (http://epoc.cochrane.org). To aid researchers and clinicians in classifying their interventions, the EPOC has developed a taxonomy which can be used as a framework for characterizing interventions (32). Interventions are first divided into three different categories reflecting the healthcare system: interventions targeted at i) healthcare organizations; ii) healthcare workers; or iii) specific types of practice. Each of these categories can be further divided into different subcategories, but since this review focuses specifically on rheumatologists we will only discuss the second category: interventions targeted at healthcare workers. Audit and feedback, educational meetings and reminders are a few examples of these interventions (31). A full overview of the EPOC taxonomy on interventions targeted at healthcare workers can be found in table 3.

As can be seen in table 3, many different types of interventions could be used when implementing, for example, a new RA treatment guideline. As a consequence, many studies have been done on the effectiveness of these interventions. An in depth discussion regarding all these studies is beyond the scope of this article. But in a systematic review from 2004 the available evidence was summarized (32). This review concluded that the overall quality of the studies was poor and that the intervention effects varied considerably both across and within interventions, making it impossible to give evidence-based recommendations on when to use which type of intervention (33). Since the publication of this review, many additional studies have been conducted and new systematic reviews were published as well. We will describe the results of the reviews on the three most performed interventions: educational meetings, audit and feedback, and reminders.

***Table 3**: EPOC taxonomy of interventions targeted at healthcare workers*

Intervention	Description*
Audit and feedback	A written, electronic or verbal summary given to healthcare providers, regarding their performance over a specified period of time
Clinical incident reporting	A system where critical incidents can be reported
Monitoring the performance of the delivery of healthcare	Monitoring of services provided by individuals/healthcare organizations
Communities of practice	A group with common interest who aim at to increase knowledge in this area by interacting on an ongoing basis
Continuous quality improvement	An ongoing process to review and improve care including: involvement of healthcare teams, analysis of a process or system, a structured process improvement method or problem solving approach, and use of data analysis to assess changes.
Educational games	In order to improve standards of care, games can be used as an educational strategy.
Educational materials	Distribution of educational materials to support clinical care.
Educational meetings	Educational meetings such as courses, conferences or workshops
Educational outreach visits, or academic detailing	Providing information by a trained person, during personal visits, with the aim of changing practice
Clinical practice guidelines	Systematically developed statements to assist healthcare providers and patients
Inter-professional education	Continuing education for health professionals, involving participants from different professions.
Local consensus processes	Formal or informal consensus processes, for example; agreeing to a clinical protocol or promoting the implementation of guidelines.
Local opinion leaders	Promoting good clinical care by identifying and using local opinion leaders
Managerial supervision	Routine supervision visits by health staff
Patient-mediated interventions	Engage input from patients, to change professional practice
Public release of performance data	Release of performance data, to inform the public about the practice of healthcare providers
Reminders	Manual of computerized reminders prompting healthcare providers to perform an action
Routine patient-reported outcome measures	Reporting and administration of patient reported outcomes measures
Tailored interventions	Development of interventions to change practice, based on assessment of barriers for a specific clinical setting

*The description is derived from the EPOC taxonomy of interventions targeted at healthcare workers.

Educational meetings

The EPOC systematic review on educational meetings was published in 2006 and included 81 trials with over 11,000 included health professionals (34). When the researchers compared 'no intervention' with 'any intervention including educational meetings as a component', the risk difference (RD) for compliance with the desired practice was 6% (interquartile range 1.8 to 15.9). For interventions with educational meetings alone, similar results were obtained. In addition, analyses were done on determinants of success. It was observed that higher attendance and a mix between interactive and didactic educational meetings were associated with a higher RD. In contrast, educational meetings targeted at complex behavior or at perceived less serious outcomes seemed to be less effective (34).

Audit and feedback

Another EPOC review focused on audit and feedback and this review included 140 different studies (35). Comparing 'interventions with audit and feedback' as an essential component to 'no intervention' resulted in an increase of 4.3% of the desired behavior (interquartile range 0.5% to 16%). Determinants of success were; low baseline performance; the audit and feedback being provided by a colleague, being provided more than once ("booster" session, or regular feedback), or being delivered in verbal and written format; and finally when it includes specific targets or action plans (35).

Reminders

On-screen computer reminders are often used as an intervention and their effectiveness has also been evaluated by EPOC (36). Reminders resulted in a median improvement of 4.2% (interquartile range 0.8% to 18.8%) when all process outcomes were taken together. When looking at the separate outcomes, the median improvement in adherence was 3.3% for medication ordering, 3.8% for vaccinations and 3.8% for test ordering. A few studies also reported clinical outcomes (for example blood pressure) and the median absolute improvement for these outcomes was 2.5%. In this review none of the specific reminders or contextual factors were significantly associated with the magnitude of the observed effects (36).

Tailored interventions

When looking at the results from the above mentioned reviews and keeping the previous section in mind (determinants of successful implementation), one might wonder if interventions specifically tailored to facilitators or barriers of implementation are more effective than non-tailored interventions. The EPOC tried to answer this question as well, but studies directly comparing 'tailored interventions' to 'non-tailored interventions' or 'no intervention at all' were scarce, and a definitive conclusion was not possible. In spite of

this scarcity the conclusion was that tailored interventions can be effective, but that their effect is variable and tends to be small to moderate at best (37).

In summary, this section described different interventions that could be used when trying to improve quality in clinical practice. The effects of any single one of these different interventions are often small to modest at best, but some evidence exist that a combination of different interventions is more effective (33). Despite the lack of evidence on which specific intervention should be used in which situation, the reviews may provide some guidance as to how to choose an intervention for a desired change in daily practice. Based on the information in this section we would advise to practicing rheumatologists willing to change their practice that, after measuring a chosen indicator set, a simple barrier analysis is done using the before mentioned themes (knowledge, motivation and being able to change) followed by a intervention tailored to these barriers. If possible, a combination of different intervention types is preferred above a single component intervention.

F. WHAT INITIATIVES EXIST TO SPECIFICALLY IMPROVE RHEUMATOID ARTHRITIS CARE?

Several studies have tried to specifically improve RA care by introducing different types of interventions. To our knowledge there is no complete overview of these studies available, so we will discuss three different studies that describe some kind of intervention designed to improve care.

The first example is a pilot study aiming at improvement of disease activity and medication prescription in RA patients by implementing nurse-led DAS-28 measurements (38). As stated in this study, the assessment of a combined disease activity index like the DAS-28 is often perceived as too complex to regularly collect and calculate in a busy day to day clinical practice. Therefore, the researchers hypothesized that delegating DAS-28 measurement to a well-trained specialized nurse could be of value, since this would save time for the rheumatologist and enhances adherence (38). In this study three rheumatologists were randomized to the control group and four to the intervention group. In both groups the DAS-28 was performed and calculated by a specialized nurse, but only in the intervention group the DAS-28 score was provided to the rheumatologist (before the patient visited the rheumatologist). This was combined with a general advice to the rheumatologists to adjust DMARD therapy in case of active disease (DAS-28 $\geq$3.2). In order to investigate whether this intervention was effective, the change in DAS-28 score and the number of medication changes were used as primary outcome measures. After 18 months, a decrease

of 0.66 and 0.69 DAS28-points was found in the usual care group and in the intervention group respectively (p=0.7). In addition, no significant differences in number of medication changes were observed between the two groups (33% in the usual care group vs 35% in the intervention group; p=0.99) (38). According to the authors, one of the reasons for the failure of this intervention was the absence of a strict treatment protocol. Therefore they concluded that nurse-led care may be useful in making DAS28 assessments more feasible for use in daily practice, but that rheumatologists should be encouraged to change medication when necessary, using the individual DAS-28 values provided to them (38).

Another example of an attempt to change clinical practice is the Metrix study (39). This study assessed the effect of an educational intervention on rheumatologists' practice behavior in RA patients. In this study 20 rheumatologists participated, and they all had to perform a prospective chart audit of 50 consecutive RA patients at study start. During this chart audit they had to collect information on patient demographics, current and previous DMARD use and measurements done during the visit such as tender and swollen joint counts. After the chart audit the rheumatologists were randomized in an intervention group and a control group (10 in both groups) (39). Rheumatologists in the control group received no further interventions and also did not receive their results from the chart audit. In contrast, the rheumatologists in the intervention group received the chart audit results as feedback, had to attend monthly web-based conferences and a journal club. During the web-based conferences and journal club sessions, topics such as the value of systematic assessment in RA and the value of using tight control strategies were discussed. After 6 months the intervention group collected more global assessments (a 13% increase post-intervention for patient global and an increase of 9% in the physician global; $p < 0.05$) and Health Assessment Questionnaires (increase of 5%, $p < 0.05$), whereas the control group did not show any changes. Furthermore, a significant increase of 32% was seen in presence of calculable composite scores (any version of CDAI, SDAI or DAS), again no change in the control group was observed (39). In patients with active disease (either defined by SDAI or by DAS) therapy was changed more often in the intervention group than in the control group (66% vs 36% for patients with SDAI between 3.3 and 11; 57% vs 38% for patients with a DAS between 2.4 and 3.6 respectively) (39). The results of this study show that the combination of feedback and educational meetings can improve daily clinical practice. Interestingly, the rheumatologists in the control group who also reviewed 50 of their own charts did not change at all. This implies that in order to change practice, only reviewing your own work, without receiving its results and additional education, is not enough (39).

In the third and final study Ledwich and colleagues determined the effects of an Electronic Health Record (EHR) best practice alert (BPA) on vaccination rates in patients with

a rheumatic disease using an immunosuppressive drug (40). In this study the BPA was a clinical reminder on influenza and pneumococcal vaccinations. When an adult patient using immunosuppressive drugs visited the clinic during the influenza season, the BPA appeared prominently on screen to remind the rheumatologist about an influenza vaccination. For pneumococcal vaccinations the BPA only appeared in patients using an immunosuppressive drug if the patient did not received a vaccination before. With both BPAs rheumatologists were also able to document why a vaccination was not given. After implementation of the BPA vaccination rates increased significantly with influenza vaccination rates increasing from 47% to 65%, and pneumococcal vaccinations increasing from 19% to 41%. Based on these results the authors stated that the BPA is an effective tool for improving quality of care for patients receiving immunosuppressive drugs (40).

Together these three studies show that initiatives are started within rheumatology evaluating strategies to improve quality of care for patients with RA or rheumatic diseases in general. Of these three studies, the latter two have demonstrated beneficial effects with regard to improving care, whereas the first study suggests that monitoring alone without a strict treatment protocol is not effective enough to truly change practice behavior. Overseeing these studies we can say that small steps are taken to improve the quality of RA care, but we have not reached our goal yet.

Our conclusions from the previous paragraph are based on results of published studies. However, it is likely that many initiatives made to improve quality of care, never appear in international peer reviewed journals (publication bias). We have information on two of such quality improvement initiatives, only previously described in the Dutch literature.

Both of the described initiatives come from the Sint Maartenskliniek (SMK) in the Netherlands. In this specialized clinic for rheumatology, orthopedic surgery and rehabilitation medicine, the rheumatology department has implemented nurse-led DAS-28 assessments for all RA patients visiting the outpatient clinic, starting in 2010. Since then, RA patients arrive at the clinic one hour before their visit with the rheumatologist. Upon arrival, blood is drawn for routine laboratory testing and the patient is seen by a specialized nurse. During this visit the DAS-28 and HAQ are performed and the current medication of the patient is discussed in order to identify any side effects or changes that have occurred since the last visit. All the information gathered during the visit with the nurse is provided to the rheumatologist prior to his or her consultation with the patient, increasing the efficiency (41). At the same time a locally developed RA treatment protocol also became available to the rheumatologists. This protocol included a strict, tight control-based, set of treatment recommendations, explicitly stating which DMARD should be given in what order and when. For example, according to

this protocol, patients with newly diagnosed RA treatment should start with methotrexate and hydroxychloroquine combination therapy. When this first combination fails, etanercept and methotrexate should be given. To date, 82% of the RA patients combine their visit with the rheumatologist with a nurse-led DAS-28 assessment. This has resulted in DAS-28 measurements being available in 85% of the visits. Most importantly, 72% of all RA patients (irrespective of disease duration) have a DAS-28-CRP below 2.9, and 60% below 2.4..

After the implementation of nurse-led DAS-28 assessments in the department, it became clear that to obtain reliable DAS-28 scores across the different health care professionals (nurses, physician assistants, residents and rheumatologists) who performed the DAS-28, an acceptable level of agreement in the DAS-28 scores is mandatory. Furthermore, since rheumatologists did not perform the DAS-28 themselves anymore, they had to learn to rely on the nurses. Therefore, an interactive and competitive DAS-28 training was designed to increase inter-observer agreement and to improve mutual confidence between rheumatologists and nurses. In this so called 'DAS-28 battle' – using elements of serious gaming - rheumatologists and nurses were first trained by an experienced rheumatologist to perform the joint counts needed to calculate the DAS-28. Next, the participants were divided into small groups and were asked to perform a tender and swollen joint count in 4 different patients. In every group of patients, one 'fake patient' was present (usually a partner of a real RA patient) and these persons served as 'healthy controls'. This extra twist was added to assess the number of false positive joints (joint scored as swollen in a healthy control). Blood tests were available, so DAS-28 scores could be calculated and compared immediately. Measurement error and number of false positive joints were calculated per team, and the team with the best score on both items was awarded the 'Golden Hand' (42). Since 2009 two DAS-28 battles have been organized, and during the first battle the measurement error ranged from 0.16 to 0.44 (mean of 0.31) and the number of false positive joints varied between 0 and 4 per team. Three years later the battle was repeated, giving a similar mean measurement error but the variation between the teams had decreased (measurement error between 0.29 and 0.36). Due to changes in the health care professionals working at the SMK, the positive effects of the battle on both inter-observer agreement and trust, and the enthusiasm of the participants, the DAS-28 battle will be organized again in the near future.

G. PRACTICAL IMPLICATIONS

In this review we have provided an overview of the current status of quality of care in RA. In the first section we saw that defining and measuring quality of care can be challenging and that different types of quality indicators exist. Next, we described current treatment strategies used in RA, which are based on tight control. Different sets of indicators to measure RA care were discussed thereafter, which was followed by describing different studies assessing adherence to tight control recommendations. Unfortunately, adherence turned out to be suboptimal in most cases. In the subsequent section we have introduced the field of implementation science that has addressed this issue before, and examples of effective interventions were given. Finally, some of these interventions that are already applied in RA and have led to improvements in care provided to patients were described.
In our opinion, these data show that evidence for the most effective RA treatment is available and that rheumatologists are willing to use this evidence in order to treat their patients to the best they can, but they need to be assisted in doing so. So, how could rheumatologists be assisted in improving their own RAcare? In this final section we will address this theme from two different points of view: the researcher's view and the practicing rheumatologist's view.

The researcher's view

We have seen that different groups around the world have developed RA quality indicators. However, none of these indicators sets are universally accepted, making it difficult to use them in research. While the indicator set formed by the international taskforce of METEOR attempts a more international approach, we are still not there (23). Therefore, we need a set of clear and internationally accepted indicators to gain more insight in the processes and outcomes of RA care. In addition, this indicator set should also incorporate outcome indicators next to process indicators. For the development of a universally applicable indicator set, based on international consensus, a joint, international taskforce under the auspices of international societies (such as ACR and EULAR) is needed. This indicator set may not only help in measuring care, but also in improving care and in the evaluation of improvement interventions.
In addition, far more attention should be given to the translation of evidence into practice. Here we refer to Buchbinder *et al*, who have recently stated that 'investment in discovery research is essentially wasted if implementation research is ignored'(1). Therefore we would make an urgent call to all stakeholders involved in rheumatology research to invest in studies trying to find effective interventions that improve the quality of care for RA patients.

Finally, policy makers and developers of guidelines or practice recommendations should be more aware of the fact that only disseminating guidelines does not suffice to ensure uptake of recommendations in clinical practice. Therefore, any new or updated version of a guideline should be accompanied by an implementation plan or at least some recommendations on how to implement the guideline in daily clinical practice. The AGREE (Appraisal of Guidelines Research & Evaluation) tool is a helpful aid when developing a guideline as the AGREE gives recommendations to develop a high quality guideline. Besides recommendations on topics such as clarity and presentation of the guideline, AGREE also stresses the need for an implementation plan as a supplement with a guideline. In addition, additional materials could be useful, such as a summary document, educational tools or computer support. These additional materials should be provided with the guidelines in order to enhance their use (43). Furthermore, it might be needed that financial incentives are available to facilitate implementation of change.

The rheumatologist's view

In this review we have tried to answer the question 'how can a rheumatologist improve his or her own practice'. Unfortunately the literature is inconclusive and research in rheumatology is scarce. This problem has been addressed in the previous paragraphs and we will now propose some simple steps that rheumatology practices could use to improve their quality of care.

When changing current practice, a first step would be to define a manageable goal. For example: 'treat your RA patients in such a way that you achieve low disease activity in 60% of all your RA patients after one year'. Of note, the goal of 60% in this example is arbitrary and not based on evidence. Unfortunately, we do not know what such a percentage should be, but based on clinical trials this could be a feasible goal to start implementation with.

Next, it is necessary to check if have enough resources are available to reach the pre-set goal. For example, see if an up to date local RA treatment guideline is available and if not, try to see that such a guideline will become available. In our experience, guidelines are easier to use in daily practice if they include brief and specific descriptions of what to do in specific situations rather than elaborating on the underlying evidence. For example, providing a step-by-step description on what to do in a patient with active disease. When such a new or updated guideline is finished, all relevant stakeholders (nurses, residents, rheumatologists, pharmacists, etcetera) should be informed about this. If necessary, additional actions such as an educational meeting may be needed to improve implementation of the guideline. Apart from a clear treatment guideline, additional resources might be needed. As we saw in the TICD checklist by Flottorp, many potential barriers to adherence exist (31). However, for daily practice it is not feasible to do a full barrier analysis. Therefore we recommend to only focus on issues that have a potentially large impact on the results and that can

relatively easy be implemented. For example, if a specialized nurse or physician assistant is available, 'shared care' between them and the rheumatologists might be an option.

After all necessary actions are implemented the next step would be to check at specified intervals if the pre-set goal has been met. Often, appropriate information for this check is not readily available from existing systems such as the EHR. One solution could be to conduct a chart review after, for example, six months (medical audit) and collecting data on disease activity and on what has been done in response. Such a chart review can be very labor intensive if done in many patients, but for feedback purposes a sample from the total patient population is often enough. Using the local treatment protocol, a few aspects of RA care could be checked for ('Is disease activity measured during every routine visit?'; 'Is DMARD medication changed in response to active disease?' and 'Is low disease activity present?'). Individual data from the patient's charts can then be aggregated in order to see if the pre-set goal has been reached. This chart review would probably most useful if all rheumatologists in one practice are involved and individual results are compared. Of note, a safe learning environment is critical when comparing non-anonymized performances among rheumatologists.

Nearly always such a chart review will reveal that not all the care is in accordance with the guidelines or the pre-set goal. Additional measures may be needed to further improve the quality of care and reach the pre-set goal. When finally those additional measures have resulted in meeting your goal, a new cycle starts and continuous evaluation will be necessary to maintain quality improvement.

The above mentioned steps are also known as the Plan, Do, Check, Act (PDCA)-cycle or the Define, Measure, Analyse, Improve, Control (DMAIC) cycle (44) and closely resemble the tight control strategy used in RA treatment itself, as this cycle also involves goal setting, measuring and acting if the goals has not yet been reached. The application of such strategies in both patients and rheumatologists may result in a better translation of evidence into practice, and consequently guarantee the best possible care we could provide to our patients.

REFERENCES

1. Buchbinder R, Maher C, Harris IA. Setting the research agenda for improving health care in musculoskeletal disorders. Nature reviews Rheumatology. 2015.
2. Donabedian A. Quality of care. How can it be assessed? JAMA. 1988;260(12).
3. Nash D CJ, Skoufalos A, Horowitz M. Conceptualizations and Definitions of Quality. Health Care Quality: The Clinician's Primer: American College of Physician Executives; 2012. p. 412.
4. Medicine Io. Crossing the Quality Chasm: A New Health System for the 21st Century. Washington (DC)2001.
5. Markhorst J, Martirosyan L, Calsbeek H, Braspenning J. Stakeholders' perspectives on quality indicators for diabetes care: a qualitative study. Qual Prim Care. 2012;20(4):253-61.
6. Lawrence MO, F. Indicators of quality in health care European Journal of General Practice. 1997;3:103-8.
7. Donabedian A. Explorations in quality assessment and monitoring. Ann Arbor, Michigan: Health Administration Press; 1980.
8. Lee SJ, Kavanaugh A. Pharmacological treatment of established rheumatoid arthritis. Best practice & research Clinical rheumatology. 2003;17(5):811-29.
9. Fries JF. Current treatment paradigms in rheumatoid arthritis. Rheumatology. 2000;39 Suppl 1:30-5.
10. Abasolo Alcazar L. Triple therapy in rheumatoid arthritis. Reumatologia clinica. 2014;10(5):275-7.
11. Bakker MF, Jacobs JW, Verstappen SM, Bijlsma JW. Tight control in the treatment of rheumatoid arthritis: efficacy and feasibility. Annals of the rheumatic diseases. 2007;66 Suppl 3:iii56-60.
12. van Nies JA, Krabben A, Schoones JW, Huizinga TW, Kloppenburg M, van der Helm-van Mil AH. What is the evidence for the presence of a therapeutic window of opportunity in rheumatoid arthritis? A systematic literature review. Ann Rheum Dis. 2014;73(5):861-70.
13. Smolen JS, Aletaha D, Bijlsma JW, Breedveld FC, Boumpas D, Burmester G, et al. Treating rheumatoid arthritis to target: recommendations of an international task force. Annals of the rheumatic diseases. 2010;69(4):631-7.
14. Smolen JS, Breedveld FC, Burmester GR, Bykerk V, Dougados M, Emery P, et al. Treating rheumatoid arthritis to target: 2014 update of the recommendations of an international task force. Annals of the rheumatic diseases. 2015.
15. Kiely P, Williams R, Walsh D, Young A, Early Rheumatoid Arthritis N. Contemporary patterns of care and disease activity outcome in early rheumatoid arthritis: the ERAN cohort. Rheumatology. 2009;48(1):57-60.
16. Schipper LG, van Hulst LT, Grol R, van Riel PL, Hulscher ME, Fransen J. Meta-analysis of tight control strategies in rheumatoid arthritis: protocolized treatment has additional value with respect to the clinical outcome. Rheumatology. 2010;49(11):2154-64.
17. Stoffer MA, Schoels MM, Smolen JS, Aletaha D, Breedveld FC, Burmester G, et al. Evidence for treating rheumatoid arthritis to target: results of a systematic literature search update. Annals of the rheumatic diseases. 2015.
18. van Hulst LT, Fransen J, den Broeder AA, Grol R, van Riel PL, Hulscher ME. Development of quality indicators for monitoring of the disease course in rheumatoid arthritis. Ann Rheum Dis. 2009;68(12):1805-10.
19. Employers N. 2014/15 GENERAL MEDICAL SERVICES (GMS) CONTRACT QUALITY AND OUTCOMES FRAMEWORK (QOF). NHS England Gateway reference: 01264. 2014:120-2.
20. Petersson IF, Strombeck B, Andersen L, Cimmino M, Greiff R, Loza E, et al. Development of healthcare quality indicators for rheumatoid arthritis in Europe: the eumusc.net project. Ann Rheum Dis. 2014;73(5):906-8.

21. MacLean CH, Saag KG, Solomon DH, Morton SC, Sampsel S, Klippel JH. Measuring quality in arthritis care: methods for developing the Arthritis Foundation's quality indicator set. Arthritis Rheum. 2004;51(2):193-202.

22. National Committee for Quality Assurance PCfPI, American College of Rheumatology. Rheumatoid Arthritis Physician Performance Measurement Set American Medical Association and National Committee for Quality Assurance. 2008.

23. Navarro-Compan V, Smolen JS, Huizinga TW, Landewe R, Ferraccioli G, da Silva JA, et al. Quality indicators in rheumatoid arthritis: results from the METEOR database. Rheumatology. 2015;54(9):1630-9.

24. Conditions NCfMAaM. National indicators for monitoring osteoarthritis, rheumatoid arthritis and osteoporosis. Canberra: Australian Institute of Health and Welfare 2006.

25. Benhamou M, Rincheval N, Roy C, Foltz V, Rozenberg S, Sibilia J, et al. The gap between practice and guidelines in the choice of first-line disease modifying antirheumatic drug in early rheumatoid arthritis: results from the ESPOIR cohort. The Journal of rheumatology. 2009;36(5):934-42.

26. Harrold LR, Harrington JT, Curtis JR, Furst DE, Bentley MJ, Shan Y, et al. Prescribing practices in a US cohort of rheumatoid arthritis patients before and after publication of the American College of Rheumatology treatment recommendations. Arthritis and rheumatism. 2012;64(3):630-8.

27. Vermeer M, Kuper HH, Bernelot Moens HJ, Hoekstra M, Posthumus MD, van Riel PL, et al. Adherence to a treat-to-target strategy in early rheumatoid arthritis: results of the DREAM remission induction cohort. Arthritis research & therapy. 2012;14(6):R254.

28. Escalas C, Dalichampt M, Combe B, Fautrel B, Guillemin F, Durieux P, et al. Effect of adherence to European treatment recommendations on early arthritis outcome: data from the ESPOIR cohort. Annals of the rheumatic diseases. 2012;71(11):1803-8.

29. Maksymowych W ØM, Elkayam O, Landewé R, Homik J et al. Impact of Failure to Adhere to Treat-to-Target of Rheumatoid Arthritis in Real World Practice: Data from the International Rheumatoid Arthritis Biomarker Program. Abstract Number: 2912. 2014.

30. Kuusalo L, Puolakka K, Kautiainen H, Blafield H, Eklund KK, Ilva K, et al. Impact of physicians' adherence to treat-to-target strategy on outcomes in early rheumatoid arthritis in the NEO-RACo trial. Scandinavian journal of rheumatology. 2015:1-7.

31. Flottorp SA, Oxman AD, Krause J, Musila NR, Wensing M, Godycki-Cwirko M, et al. A checklist for identifying determinants of practice: a systematic review and synthesis of frameworks and taxonomies of factors that prevent or enable improvements in healthcare professional practice. Implementation science: IS. 2013;8:35.

32. Collaboration C. Effective Practice and Organisation of Care (EPOC). EPOC Taxonomy; 2015. http://epoc.cochrane.org/epoc-taxonomy: Cochrane Collaboration, 2015.

33. Grimshaw JM, Thomas RE, MacLennan G, Fraser C, Ramsay CR, Vale L, et al. Effectiveness and efficiency of guideline dissemination and implementation strategies. Health technology assessment. 2004;8(6):iii-iv, 1-72.

34. Forsetlund L, Bjorndal A, Rashidian A, Jamtvedt G, O'Brien MA, Wolf F, et al. Continuing education meetings and workshops: effects on professional practice and health care outcomes. The Cochrane database of systematic reviews. 2009(2):CD003030.

35. Ivers N, Jamtvedt G, Flottorp S, Young JM, Odgaard-Jensen J, French SD, et al. Audit and feedback: effects on professional practice and healthcare outcomes. The Cochrane database of systematic reviews. 2012;6:CD000259.

36. Shojania KG, Jennings A, Mayhew A, Ramsay CR, Eccles MP, Grimshaw J. The effects of on-screen, point of care computer reminders on processes and outcomes of care. The Cochrane database of systematic reviews. 2009(3):CD001096.

37. Baker R, Camosso-Stefinovic J, Gillies C, Shaw EJ, Cheater F, Flottorp S, et al. Tailored interventions to address determinants of practice. The Cochrane database of systematic reviews. 2015;4:CD005470.

38. van Hulst LT, Creemers MC, Fransen J, Li LC, Grol R, Hulscher ME, et al. How to improve DAS28 use in daily clinical practice?--a pilot study of a nurse-led intervention. Rheumatology. 2010;49(4):741-8.

39. Pope J, Thorne C, Cividino A, Lucas K. Effect of rheumatologist education on systematic measurements and treatment decisions in rheumatoid arthritis: the metrix study. The Journal of rheumatology. 2012;39(12):2247-52.

40. Ledwich LJ, Harrington TM, Ayoub WT, Sartorius JA, Newman ED. Improved influenza and pneumococcal vaccination in rheumatology patients taking immunosuppressants using an electronic health record best practice alert. Arthritis and rheumatism. 2009;61(11):1505-10.

41. den Broeder A. A vdEE, Schasfoort E, Vogels I, van Eerd J. Implementatie van klinimetrieondersteuning in de reumatologische praktijk. Nederlands Tijdschrift voor Reumatologie. 2011:26-31.

42. Straten V LN, den Broeder A. Naar een eenduidige uitvoering van de DAS28: de 'DAS28-battle'. Nederlands Tijdschrift voor Reumatologie. 2012:10-3.

43. Collaboration TA. Appraisal of guidelines for research & evaluation. London: St George's Hospital Medical School, 2001.

44. Nicolay CR, Purkayastha S, Greenhalgh A, Benn J, Chaturvedi S, Phillips N, et al. Systematic review of the application of quality improvement methodologies from the manufacturing industry to surgical healthcare. The British journal of surgery. 2012;99(3):324-35.


TREAT
TO
TARGET
Don't move... I'm sure I will
hit bull's eye this time!

8

Practice what you preach? An exploratory multilevel study on rheumatoid arthritis guideline adherence by rheumatologists

Nienke Lesuis
Alfons A den Broeder
Marlies EJL Hulscher
Ronald F van Vollenhoven

RMD Open 2016;2:e000195. Doi 10.1136/rmdopen-2015-000195

ABSTRACT

Objectives

To assess variation in, and determinants of rheumatologist guideline adherence in patients with rheumatoid arthritis (RA) in daily practice.

Methods

In this retrospective observational study, guideline adherence in the first year of treatment was assessed for seven pre-defined parameters on diagnostics, treatment and follow-up in all adult RA patients with a first outpatient clinic visit at the study centre between September 2009 and March 2011. Variation in guideline adherence was assessed on parameter and rheumatologist level. Determinants for guideline adherence were assessed in patients (demographic characteristics, rheumatoid factor (RF) and/or anti-cyclic citrullinated (aCCP) peptide antibody positivity, ESR, erosive disease, comorbidity and the number of available DMARD treatment options) and rheumatologists (demographic and practice characteristics, guideline knowledge and agreement, outcome expectancy, cognitive bias, thinking style, numeracy and personality).

Results

A total of 994 visits in 137 RA patients were reviewed. Variation in guideline adherence between parameters was present (adherence between 21% and 72%), with referral to the physician assistant as lowest scoring parameter and referral to a specialized nurse as highest scoring one. Variation in guideline adherence between rheumatologists was also present (adherence between 22% and 100%). Patient sex, the number of DMARD options, presence of erosions, comorbidity, RF/aCCP positivity, type of patient and the rheumatologists' scientific education status were associated with adherence to one or more guideline parameters.

Conclusions

Guideline adherence varied considerably between the guideline parameters and rheumatologists, showing that there is room for improvement. Guideline adherence in our sample was related to several patient- and rheumatologist determinants.

INTRODUCTION

Many guidelines and recommendations on optimal care for patients with rheumatoid arthritis (RA) have been developed to help clinicians choosing the best diagnostic and therapeutic strategies for their RA patients. All major RA treatment guidelines are now based on tight control principles, where monitoring of disease activity and changing treatment if a pre-set target is not reached are essential.[1-3] Adherence to these tight control principles, preferably combined with the use of a specific treatment guideline, results in lower disease activity and less functional damage compared to usual care.[4-6]

In view of the evidence supporting the benefit of adhering to protocolized tight control strategies, it is disappointing that current guideline adherence is still suboptimal as observed in multiple studies on this topic.[7-13] Unfortunately, these studies focus on DMARD-related treatment recommendations only, disregarding the fact that other aspects of RA care are also important. Furthermore, these studies are not performed in daily practice, but in pre-defined cohorts using subsets of RA patients. Therefore, the first aim of this study is to gain more insight in guideline adherence of rheumatologists in daily practice, using a broader set of guideline adherence parameters than before.

Our second study aim is to gain insight into determinants of guideline adherence. In order to improve guideline adherence, it is first necessary to understand the determinants that influence adherence. Knowledge on these determinants could then be used to develop targeted interventions as evidence suggests that this leads to better intervention effects. [14] Although knowledge on determinants of guideline adherence is not yet available from studies within rheumatology, studies outside rheumatology suggest the importance of various determinants, explaining the observed variation in guideline adherence between both hospitals and physicians.[15] Examples of such determinants are knowledge and cognitions of individual health care professionals and patient factors.[15]

All in all, data on RA guideline adherence in daily routine and its association with potential determinants is still lacking. Therefore, we aimed to 1) assess RA guideline adherence in daily clinical practice, 2) assess variation in guideline adherence on parameter and rheumatologist level, and 3) explore the impact of rheumatologist- and patient-related determinants on guideline adherence.

MATERIAL AND METHODS

Study design

An explorative, retrospective observational multi-level cohort study was performed. Guideline adherence is behaviour executed by a rheumatologist, but it is measured in patients

who visit the hospital. Hence this study has three different levels: outpatient clinic visits (level 1) are nested within patients (level 2), who are in turn nested within rheumatologists (level 3). This is also reflected in the data collection and measurement: guideline adherence is measured on patient- or visit level (data collection on visit level), whereas the possible determinants of guideline adherence were measured either on rheumatologist- or patient-level.

Setting

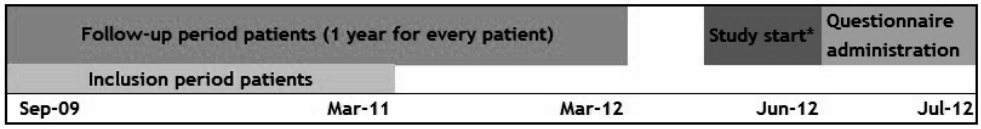
This study was conducted at the rheumatology department of the Sint Maartenskliniek, a large clinic specialized in rheumatology, rehabilitation medicine and orthopaedics in the Netherlands. In this centre a local, tight control based, RA treatment guideline was put into use in 2007. At the same time supportive actions were undertaken to aid rheumatologists in following the new guideline. Firstly, specialized nurses were available to provide patient education, discuss disease coping and to assess disease activity before the visit with the rheumatologist (nurse led assessment of the Disease Activity Score in 28 joints (DAS28)). Secondly, after a referral to the physician assistant (PA) by the rheumatologist, patients were seen in an alternating fashion by the PA and rheumatologist in order to share care between them. The PAs can independently make treatment decisions, but they work under the supervision of a rheumatologist and, at the time of this study, were not allowed to prescribe medication.

Participants

All 14 rheumatologists working at the study centre between September 2009 and July 2012 were eligible for participation. Rheumatologists who did not work the full period were excluded; no other exclusion criteria were set. Consent from all participants was sought by explaining the study during a regular staff meeting.

We included all patients of 18 years and older diagnosed with RA (ICD-9 code 714.0), treated by one of the included rheumatologists and having had a first outpatient clinic visit at the study centre between September 2009 and March 2011. Patients with both new and established RA could be included, as long as their first visit to the study centre took place during the given time period. If patients were seen as second opinions, they were only included if treatment was fully taken over by the study centre. After inclusion, all visits in the first year of treatment at the study clinic were used to assess guideline adherence (figure 1). This means that the follow-up period lasted until March 2012.

Figure 1 *Study time frame*



Guideline adherence measures & data collection

As guideline adherence is multidimensional and cannot be expressed by a single outcome measure, we defined a set of seven different parameters to measure various quality aspects of RA care. These parameters are based on the quality indicators stated in the Dutch national RA treatment guideline.[16] As the local RA guideline used in the study centre is an adapted version of the Dutch national guideline, the selected parameters were adapted accordingly. This resulted in a set of seven guideline adherence parameters concerning three main themes (diagnostics, treatment and follow-up & shared care). All parameters are reported as dichotomous outcomes ('yes' or 'no'), but depending on the type of parameter this is done at either visit- or patient level. All guideline adherence parameters are described in table 1. Online supplement 1 provides a more extensive version of this table, including corresponding treatment recommendations.

As mentioned before, all parameters were measured during the first year of treatment at the study centre. So, after a patient was included (between September 2009 and March 2011), all visits in the next year were used to measure guideline adherence.

Table 1 *Guideline adherence parameters*

Guideline adherence parameter	Level of measurement
Radiographs of hands, feet and thorax ordered within the first three visits, in patients with a disease duration ≤1 year	Patients
Prescription of conventional and biological DMARDs in agreement with the local preferential sequence	Patients
Referral to a specialized nurse within the first three visits	Patients
Referral to a PA or NP within the first year of treatment	Patients
Therapy change[ǂ] in case of moderate to high disease activity*	Visits
Regular outpatient clinic visits combined with a nurse led DAS28 assessment	Visits
Correct intervals between regular outpatient clinic visits	Visits

DMARD: Disease Modifying Anti-Rheumatic Drug; DAS28: Disease Activity Score in 28 joints; PA: physician assistant; NP: nurse practitioner. [ǂ]Therapy change included the intensification of DMARD therapy (dosage increase, shortening of the interval, adding a new DMARD and/or biological, switching to another DMARD and/or biological), starting or increasing corticosteroids (dose), local corticosteroid injections. *DAS28 >3.2 or corresponding judgement from the rheumatologist if a DAS28 was not available.

To calculate the different parameters, the following data from every visit in the first year of treatment was collected: date and type of the visit, name of treating rheumatologist, presence of a nurse led DAS28 assessment, DAS28 score (using erythrocyte sedimentation rate (ESR)), functional status by Health Assessment Questionnaire (HAQ), clinical judgement of disease activity, radiographs that were ordered, current medication use (conventional and biologic Disease Modifying Anti-rheumatic Drugs (DMARD), glucocorticoids and/or Non-Steroidal Anti Inflammatory Drugs (NSAIDs)), referral to a specialized nurse and referral to a PA. Using pre-defined algorithms, the seven guideline parameters were calculated using the above mentioned data.

Determinants of guideline adherence & data collection

Determinants of guideline adherence were assessed on two different levels: patient and rheumatologist level. On patient level, eight determinants were collected at baseline: age, gender, type of patient (new or second opinion), rheumatoid factor (RF) and/or anti-cyclic citrullinated peptide antibody (aCCP) positivity, ESR, presence of erosive disease, relevant comorbidity and the number of available DMARD treatment options. The last determinant provides the number of conventional and biological DMARDs that the patients has not yet used, but could be prescribed in the future in case of treatment failure on the current DMARD.

On rheumatologist level, information on five demographic and practice determinants was collected (age, gender, PhD, years of work experience as a rheumatologist and percentage of direct patient contact per week at the outpatient clinic). Furthermore, all participating rheumatologists were asked to complete self-developed questionnaires on guideline knowledge and agreement, and outcome expectancy. In addition, existing and validated questionnaires on cognitive bias, thinking styles, numeracy and personality traits were administered.[17-20] Some of the included questionnaires expressed their score on ≥1 subscale, resulting in 14 determinants being calculated from seven questionnaires (table 2 and online supplement 2). All questionnaires were web-based, of which the invitation was send to the rheumatologists in July 2012. After two weeks reminders were sent to all rheumatologist who had not yet completed the questionnaires.

Data sources

All data needed to calculate the guideline adherence parameters were retrospectively retrieved from paper hospital charts, using a patient list generated from the administrative hospital database. During this chart review, the patient-related determinants were also collected. All data collected during chart review was written on paper case report forms, after which the data was entered in an electronic database and anonymized. All rheumatologist-related determinants were collected using the questionnaires mentioned

in the previous section and scores were also entered in an electronic database. For the purpose of the study, anonymizing the rheumatologist data was not possible.

***Table 2** Questionnaires used to measure rheumatologist-level determinants*

Determinant	Questionnaire	Number of scales	Score range
Cognitive bias	Inventory for Cognitive Bias in Medicine (ICBM)[17]	1	0 to 22 (higher scores indicating less cognitive bias)
Personality	Big Five Inventory (BFI; Dutch version) [19]	5 (extraversion, neuroticism, openness, conscientiousness and agreeableness)	1 to 5 on every subscale (higher scores indicating a stronger personality trait on the specific subscale)
Thinking styles	Rational Experiential Inventory (REI; Dutch version)[18]	2 (rationality and experientality)	20 to 100 on every subscale (higher scores indicating a more rational/ experiential thinking style)
Numeracy	Berlin Numeracy Test (BNT; Dutch version) [20]	1	0 to 7 (a higher score indicating a higher level of numeracy)
Knowledge	Self-developed questionnaire	2 (general and specific knowledge)	General knowledge: 0 to 10. Specific knowledge: -5.2 to 10* (higher scores indicating more guideline knowledge)
Guideline agreement	Self-developed statements	2 (general and specific agreement with the guideline)	1 to 5 on every subscale (a higher score indicating a higher level of agreement)
Outcome expectancy	Self-developed statement	1	1 to 5 (a higher score indicating a higher level of outcome expectancy)

*Negative scores possible due to correction for guessing

Statistical analysis

Results on primary outcome measures (guideline adherence parameters) are reported as percentages with the accompanying absolute numbers. For the questionnaire scores and remaining variables means and SD or medians and interquartile ranges (IQR) are provided, as appropriate.

Because of the hierarchical structure of our study (repeated measures on the same patient and patient nested within rheumatologists) we performed linear or logistic multilevel regression analysis when analysing the relation between the guideline adherence parameters and determinants. Depending on the type of parameter (outcomes on patient or visit level), two or three levels were included in the analyses. For the parameters radiographs ordered, preferential DMARD order, referral to a specialized nurse and referral to a PA (patient level), the parameter had the same score for every visit within one patient. For example, patients should be referred to a PA within the first year of treatment, meaning that this parameter is

scored only once per patient taking into account all visits during the study period. Multilevel analysis for these parameters only accounted for clustering within rheumatologists (two level model). For the other parameters (therapy change, nurse led DAS28 assessments and correct visit intervals; visit level) multilevel analysis also accounted for clustering within patients. This extra level was added because the parameter score per visit could differ within patients. For example, nurse led DAS28 assessments were either done or not done during the various visits.

Multilevel analysis started with adding all patient determinants to the model. Then, one by one the least significant determinant was deleted from the model until all remaining determinants were significant ($p < 0.05$). Next, the rheumatologist determinant with the highest correlation was added to the multilevel regression model, and if significant, included in the final regression model. This process was repeated with the rheumatologist determinant with the second highest correlation. Depending of the p-value of this determinant in the model, the analysis stopped (final model) or another determinant was added. This method was chosen because the number of rheumatologists was relatively small compared to the numbers of rheumatologist determinants.

Only parameters and determinants with enough variation between rheumatologists were analysed for associations between them. In case of floor or ceiling effects a determinant was omitted from further analysis. Results are presented as odds ratios (OR) with the corresponding 95% confidence interval (95%-CI), p-value and explained variation (%). Explained variation was calculated using the method described by Snijders and Bosker.[21] Statistical analyses were performed using STATA version 13.0, except the multilevel analysis, this was done using SAS version 9.2.

Ethical approval

This study was approved by the local research committee at the study centre (RR-105-PP). Although no written informed consent was obtained from the rheumatologists, they were informed beforehand about this study and asked if they would participate. It was made clear to them that they could stop with this study at any time, without providing a reason. As this was a quality assessment performed in the hospital where the first two authors of this study worked, no written informed consent was needed from the patients. In addition, data collection was done by the first author and directly after chart review all patient data was anonymized.

RESULTS

Rheumatologist and patient characteristics

All 14 eligible rheumatologists (46.2% female; mean age 47.6 ± 10.0 years) participated in this study. All questionnaires were returned by all rheumatologists, except for the questionnaires on guideline knowledge & agreement and outcome expectancy, which were not completed by one rheumatologist.

According to the hospital database 241 patients with an ICD-code of RA were seen for the first time at the study clinic between September 2009 and March 2011. 61 patients were excluded because they turned out to be second opinion patients of which treatment was not taken over by the study centre. An additional 43 patients were excluded because charts were missing (n= 9), patients were not seen by an included rheumatologist (n= 11), chart review revealed another diagnosis than RA (n= 16) or they had deceased (n= 7). The remaining 137 RA patients (67.2% female; mean age 58.9 ± 14.1 year),with a total of 994 visits, were included in this study. Roughly half of the patients had not been seen by a rheumatologist before (46%), the remaining patients had been treated before by a rheumatologist outside the study centre, and visited the study centre for a second opinion. This led to a combination of new and established RA, as reflected in the median disease duration (0; IQR 0 to 7 years). In table 3 the baseline characteristics of both rheumatologists and patients are stated.

Guideline adherence parameters

Adherence to the different guideline adherence parameters varied between 21% and 72% (table 4). The best scoring indicator was 'referral to a specialized nurse', with 72% of the patients being referred to such a nurse. Ordering of radiographs and changing therapy in case of active disease was done in approximately two thirds of the patients or visits respectively. The remaining parameters had adherence percentages between 20% and 40% (PA referral, DMARD prescription, intervals between visits and nurse led DAS28 assessment). As described in table 4, not all guideline parameters apply to all patients or visits. This applies for example to the parameter 'therapy change in case of active RA'. A DAS28 and/or clinical judgment was available in 622 visits (63%) and in 285 of these visits active disease (DAS28 >3.2 or a corresponding judgment from the treating rheumatologist) was present (46%). In 191 of those visits (67%) the rheumatologists decided to change medication (parameter therapy change in case of active disease).

In addition to the aforementioned variation between guideline adherence parameters, variation was also observed between rheumatologists. The largest difference between the rheumatologists was seen in the parameter concerning radiograph ordering, with adherence percentage of individual rheumatologists between 22% and 100%. The least variation was

seen in the parameter on correct intervals between visits, with adherence percentages varying between 11% and 43%.

Table 3 *Baseline characteristics of included rheumatologists and patients*

Characteristic	Results
Rheumatologists (n = 14)	
Age, in years†	45.2 (39.5 to 56.7)
Female gender ***(%)***	46.2
PhD degree or pursuing a PhD ***(%)***	69.2
Experience as rheumatologist, in years†	6.9 (3.6 to 19.9)
Patient contact per week† ***(%)***	60.0 (45.0 to 70.0)
Guideline knowledge* ***[0-10]; [-5.2 to 10]***	
General	8.1 (1.0)
Specific	6.2 (1.8)
Guideline agreement* ***[0-5]***	
General	4.8 (0.5)
Specific	4.5 (0.5)
Outcome expectancy* ***[0-5]***	3.9 (0.8)
Cognitive bias* ***[0-22]***	12.5 (4.2)
Thinking styles* ***[0-100]***	
Rational	79.5 (9.2)
Experiential	63.7 (7.5)
Numeracy* ***[0-7]***	6.6 (1.1)
Personality* ***[0-5]***	
Extraversion	*3.4 (0.7)*
Neuroticism	2.8 (0.4)
Openness to experience	3.7 (0.6)
Consciousness	3.7 (0.4)
Agreeableness	3.8 (0.3)
Patients (n = 137)	
Age, in years*	58.9 (14.1)
Female gender ***(%)***	67.2
Disease duration, in years†	0 (0 to 7)
RF and/or aCCP positive ***(%)***	85.4
Erosions (%)	38.3
ESR† ***(mm/h)***	25 (12 to 36)
Comorbidity ***(%)***	66.4
Number of available DMARD treatment options†‡	15 (14 to 15)

*Mean (standard deviation). †Median (interquartile range). RF: rheumatoid factor; aCCP: anti-cyclic citrullinated peptide antibody; ESR: erythrocyte sedimentation rate; DMARD: Disease Modifying Anti-Rheumatic Drug. ‡Includes both conventional and biological DMARD treatment options.

Table 4 *Guideline adherence percentages*

Guideline adherence parameter	Adherence percentage
Patient level (n= 137)	
Radiographs of hands, feet and thorax ordered within the first three visits, in patients with a disease duration ≤1 year	66 (53/80)
Prescription of DMARDs* in agreement with the local preferential sequence	23 (29/126)
Referral to a specialized nurse within the first three visits	72 (98/137)
Referral to a PA or NP within the first year of treatment	21 (29/137)
Visit level (n= 994)	
Therapy change‡ in case of moderate to high disease activity	67 (191/285)
Regular outpatient clinic visits combined a nurse led DAS28 assessment	37 (253/690)
Correct intervals between regular outpatient clinic visits	32 (160/502)

DMARD: Disease Modifying Anti-Rheumatic Drug; DAS28: Disease Activity Score in 28 joints; PA: physician assistant; NP: nurse practitioner. *conventional and biological DMARDs. ‡Therapy change include starting or increasing dosage of a conventional DMARD or oral corticosteroids, starting a biological DMARD and intramuscular or -articular injections with corticosteroids.

Determinants of guideline adherence

All guideline adherence parameters showed enough variation and no floor/celling effects were present, therefore all parameters were included in the multilevel analyses. However, five determinants (general guideline knowledge, general guideline agreement, specific guideline agreement, outcome expectancy and numeracy) were not included in the analyses due to lack of variation in the scores and/or ceiling effects (table 3).

For the remaining determinants eight associations with five different parameters were found (table 5). The preferential order of DMARD prescriptions was adhered less to in case of more available treatment options. Furthermore, referral to the specialized nurse was less likely if patients had erosive disease and comorbidity at baseline. Females, aCCP and/or RF positive patients and second opinions had less visits combined with a nurse led DAS28 assessment. Correct intervals between visits were also less likely if a patient was seen as a second opinion.

Only one parameter was associated with a rheumatologist-related determinant: rheumatologists with a PhD degree or pursuing a PhD were more likely to refer their patients to a PA. Personality, thinking styles and cognitive bias did not impact rheumatologists adherence to any of the guideline adherence parameters.

The explained variance of the models was low to moderate. The lowest explained variance (2.5%) was seen in the model on correct intervals between visits, and the highest (12.0%) in the model on PA referral.

Table 5 *Multivariate associations between guideline adherence parameters and patient- (P) and rheumatologist (R) related determinants*

	Guideline adherence parameter	Odds ratio (95% CI)	P-value
Prescription of DMARDs in agreement with the local preferential sequence			
P	Number of treatment options	0.78 (0.63 to 0.97)	0.03
	Explained variance (%)	5.2	
Referral to a specialized nurse within the first three visits			
P	Presence of erosive disease	0.68 (0.16 to 0.93)	0.03
P	Comorbidity	0.68 (0.13 to 1.00)	0.05
	Explained variance (%)	9.8	
Referral to a PA or NP within the first year of treatment			
R	PhD degree or pursuing a PhD	4.14 (1.33 to 12.86)	0.01
	Explained variance (%)	12.0	
Regular outpatient clinic visits combined a nurse led DAS28 assessment			
P	Female gender	0.63 (0.41 to 0.97)	0.04
P	RF and/or aCCP positivity	0.43 (0.28 to 0.66)	<0.01
P	Seen by a rheumatologist before (second opinion)	0.41 (0.22 to 0.77)	0.01
	Explained variance (%)	7.9	
Correct intervals between regular outpatient clinic visits			
P	Seen by a rheumatologist before (second opinion)	0.56 (0.37 to 0.85)	0.01
	Explained variance (%)	2.5	

RF: rheumatoid factor; aCCP: anti-cyclic citrullinated peptide antibody; ESR: erythrocyte sedimentation rate; DMARD: Disease Modifying Anti-Rheumatic Drug.

DISCUSSION

Our results show that guideline adherence percentages varied considerably between parameters, suggesting suboptimal guideline adherence on at least some guideline recommendations. Furthermore, adherence also varied between rheumatologists and a part of this variation could be explained by several rheumatologist- and patient-related determinants.

Besides being one of the first studies in rheumatology assessing guideline adherence in daily practice, other strengths of this study are the inclusion of a wide range of guideline adherence parameters and the multi-level association analyses between these parameters and determinants. However, our study has some limitations. First, being a retrospective study with chart review as the main data source, it is possible that information has been missed due to the fact that not everything was well documented in the charts. However, the advantage of our retrospective design is that guideline adherence could not have been influenced by the study itself. Secondly, the sample size was, with only 14 participating

rheumatologists, relatively small. Nonetheless, if we have missed associations due to a lack of power, these associations are probably not very strong. Thirdly, this study was conducted in only one centre in the Netherlands, probably hampering generalizability. Nevertheless, our observation that guideline adherence is suboptimal is most likely to be generalizable as other groups before concluded the same. Only our estimates on the degree of guideline adherence might be less generalizable. Furthermore, due to the single centre design we were not able to assess the influence of organizational factors on guideline adherence. As the study centre already implemented some supportive actions to increase adherence, results in a centre without these actions might be different. Lastly, the single-centre design and the homogeneous population within this centre might have attributed to the fact that we had to exclude some of our determinants due to ceiling effects or lack of variation.

In our study, guideline adherence varied between 21% and 72% and as no absolute norms on optimal guideline adherence exist, we can only use relative norms to judge if guideline adherence in this study was optimal. Firstly, the adherence percentages of our best scoring parameters can be used as a relative norm. So, the observed level of adherence to the three highest scoring indicators (radiograph ordering, specialized nurse referral, therapy change; adherence 66% to 72%) was probably optimal. Furthermore, aiming for 100% adherence is not feasible due to for example patient comorbidity or medication side effects.
Secondly, we can compare our results with other studies. However, since previous studies have primarily focused on therapy recommendations (DMARD prescription and therapy change in case of active disease), this makes comparison with existing data impossible for all our parameters. With regard to DMARD prescriptions, the 23% guideline adherence we found seems to be on the lower end of the spectrum. Another study on this subject observed adherence percentages to the American College of Rheumatology (ACR) DMARD treatment guidelines of 24% to 90%, depending on the type of DMARD used, disease activity and prognosis.[11] However, the lower adherence percentages in our study could probably be explained by the more strict definition we used. For example, the ACR guideline names methotrexate the first choice DMARD, with combination therapy depending on disease activity, prognosis and disease duration. In contrast, according to our local guideline all new patients should be started on methotrexate and hydroxycloroquine combination therapy.
With regard to therapy change in case of active disease we can compare our results with two previous studies. One study by Fransen *et al* in patients with established RA used the same DAS28 threshold (3.2) at which therapy should be changed as our study, observing an adherence percentage of 20%.[9] Although Fransen *et al* only looked at DMARD therapy change whereas we included also corticosteroid use, the adherence of 67% we found is substantially higher. The result of the second study by Vermeer *et al* with an adherence of 58%, is more in line with our results, although it only included DMARD therapy change

and was limited to early RA.[13] Concerning therapy change, it should be mentioned that our centre has participated in a guideline adherence study before. In that particular study, therapy was changed in 33% of the visits with active disease, compared to 67% now.[22] This large improvement in guideline adherence is most likely caused by the introduction of the local RA guideline and the supportive actions afterwards.

Besides the assessment of guideline adherence we also looked if patient- and rheumatologist related determinants were associated with guideline adherence. Despite the low to moderate explained variance, some interesting observations can be made. For example, in qualitative studies factors like erosive disease, comorbidity and RF/aCCP status are often mentioned by rheumatologists as important reasons to intensify or not intensify treatment.[23] Therefore, we expected to find associations between these determinants and the guideline parameter on therapy change in case of active disease. Although we did not observe this association, we observed associations between number of treatment options, erosive disease, comorbidity and RF/aCCP status and the parameters on DMARD prescription, referral to a specialized nurse and nurse led DAS28 assessments. This implies that patient factors could, justly or unjustly, influence more decisions than treatment intensification only.

With regard to the rheumatologist related determinants, it is notable that only one association between a rheumatologist determinant (PhD) and a guideline adherence parameter (PA referral) was found. This was especially surprising as factors such as knowledge are frequently mentioned as a potential determinant of guideline adherence.[15, 24] This might imply that rheumatologist related determinants did not play a large role in our sample, but further studies on this subject are needed as guideline adherence is probably determined by a complex interplay of facilitators and barriers which makes it hard to capture.

Due to the explorative design of our study, replication of our results is warranted in other settings both inside and outside the Netherlands. However, the suggestion from our results that rheumatologists do not always practice what is preached, can be used more widely. It seems that despite the current focus on treat to target principles in RA literature, these principles are not automatically applied in daily practice.

This study provides an example for other centres to measure their quality of care and the determinants found in our sample might be reckoned with in future interventions. Recent developments around nationwide registries, such as the RISE (Rheumatology Informatics System for Effectiveness) registry, can facilitate measurements by providing real-time feedback on important aspects of quality of care.[25] Information gained from quality of care studies or registries can than serve as benchmark information for hospitals or individual physicians.[26] Furthermore, we would advocate for more attention of researchers and policy makers towards implementation of RA guidelines and quality of care. Besides replicating our results in larger studies, future research should focus on the identification of determinants

influencing adherence. This is crucial to gain insight into the most effective and feasible interventions to help rheumatologists adhere better to RA management guidelines and to improve patient outcomes in daily practice. Only then can patients benefit from the large body of evidence that already exists.

ACKNOWLEDGEMENTS

We would like to thank all rheumatologists that have participated in this study for their willingness to complete the questionnaires. Furthermore, we would like to thank Reinier Akkermans for his help with the multilevel analyses.

CONTRIBUTORS

NL, AdB and RvV contributed to the study design. NL collected the data. NL, AdB and MH participated in the data analysis. NL, AdB, MH and RvV assisted in the interpretation of the data. The manuscript was written by NL under the supervision of AdB, MH and RvV. All authors have revised the draft version of the manuscript, and read and approved the final version of the manuscript. NL is guarantor.

COMPETING INTEREST DECLARATION AND FUNDING

All authors have completed the Unified Competing Interest form at www.icmje.org/coi_disclosure.pdf (available on request from the corresponding author) and declare that RvV has received grants and personal fees from AbbVie, BMS, GSK, Pfizer, Roche, UCB, Biotest, Crescendo, Janssen, Lilly, Merck and Vertex outside the submitted work; all other authors reported no competing interests. This study received no external funding.

DATA SHARING STATEMENT

Patient level data, full dataset, technical appendix and statistical code are available at a reasonable request from the corresponding author. Consent from the patients was not obtained but the presented data are anonymized and risk of identification is low.

REFERENCES

1. Singh JA, Furst DE, Bharat A et al. 2012 update of the 2008 American College of Rheumatology recommendations for the use of disease-modifying antirheumatic drugs and biologic agents in the treatment of rheumatoid arthritis. *Arthritis Care Res* (Hoboken) 2012;64:625-39.
2. Smolen JS, Breedveld FC, Burmester GR et al. Treating rheumatoid arthritis to target: 2014 update of the recommendations of an international task force. *Ann Rheum Dis* 2015;12.
3. Smolen JS, Landewe R, Breedveld FC et al. EULAR recommendations for the management of rheumatoid arthritis with synthetic and biological disease-modifying antirheumatic drugs: 2013 update. *Ann Rheum Dis* 2014;73:492-509.
4. Schipper LG, van Hulst LT, Grol R, et al. Meta-analysis of tight control strategies in rheumatoid arthritis: protocolized treatment has additional value with respect to the clinical outcome. *Rheumatology (Oxford)* 2010;49:2154-64.
5. Schoels M, Knevel R, Aletaha D et al. Evidence for treating rheumatoid arthritis to target: results of a systematic literature search. *Ann Rheum Dis* 2010;69:638-43.
6. Stoffer MA, Schoels MM, Smolen JS et al. Evidence for treating rheumatoid arthritis to target: results of a systematic literature search update. *Ann Rheum Dis* 2015;19.
7. Kiely PD, Brown AK, Edwards CJ et al. Contemporary treatment principles for early rheumatoid arthritis: a consensus statement. *Rheumatology (Oxford) 2009*;48:765-72.
8. Fransen J, Laan RF, Van Der Laar MA, et al. Influence of guideline adherence on outcome in a randomised controlled trial on the efficacy of methotrexate with folate supplementation in rheumatoid arthritis. A*nn Rheum Dis* 2004;63:1222-6.
9. Fransen J, Moens HB, Speyer I, et al. Effectiveness of systematic monitoring of rheumatoid arthritis disease activity in daily practice: a multicentre, cluster randomised controlled trial. *Ann Rheum Dis* 2005;64:1294-8.
10. Benhamou M, Rincheval N, Roy C et al. The gap between practice and guidelines in the choice of first-line disease modifying antirheumatic drug in early rheumatoid arthritis: results from the ESPOIR cohort. *J Rheumatol* 2009;36:934-42.
11. Harrold LR, Harrington JT, Curtis JR et al. Prescribing practices in a US cohort of rheumatoid arthritis patients before and after publication of the American College of Rheumatology treatment recommendations. *Arthritis Rheum* 2012;64:630-8.
12. Escalas C, Dalichampt M, Combe B et al. Effect of adherence to European treatment recommendations on early arthritis outcome: data from the ESPOIR cohort. *Ann Rheum Dis* 2012;71:1803-8.
13. Vermeer M, Kuper HH, Bernelot Moens HJ et al. Adherence to a treat-to-target strategy in early rheumatoid arthritis: results of the DREAM remission induction cohort. *Arthritis Res Ther* 2012;14:R254.
14. Baker R, Camosso-Stefinovic J, Gillies C et al. Tailored interventions to address determinants of practice. *Cochrane Database Syst Rev* 2015;4:CD005470.
15. Flottorp SA, Oxman AD, Krause J et al. A checklist for identifying determinants of practice: a systematic review and synthesis of frameworks and taxonomies of factors that prevent or enable improvements in healthcare professional practice. *Implement Sci* 2013;8:35.
16. van Riel PL, Barendsen BC, van Croonenborg JJ, et al. Diagnostiek en behandeling van reumatoïde artrits. 2015. Available at http://www.diliguide.nl/document/3249/reumatoide-artritis.html
17. Hershberger PJ, Part HM, Markert RJ, et al. Development of a test of cognitive bias in medical decision making. *Acad Med* 1994;69:839-42.

18. Witteman C, Bercken van den J, Claes L, et al. Assessing rational and intuitive thinking styles. E*ur J Psychol Assess* 2009;25:39-47.

19. Denissen JJ, Geenen R, van Aken MA, et al. Development and validation of a Dutch translation of the Big Five Inventory (BFI). *J Pers Assess* 2008;90:152-7.

20. Cokely ET, Galesic M, Schulz E, et al. Measuring risk literacy: the Berlin Numeracy Test. *Judgement and Decision Making* 2012;7:25-47.

21. T.A.B.Snijders, R.J.Bosker. Multilevel analysis: an introduction to basic and advanced multilevel modeling. Thousand Oaks, CA: Sage, 1999.

22. van Hulst LT, Hulscher ME, van Riel PL. Achieving tight control in rheumatoid arthritis. *Rheumatology (Oxford)* 2011;50:1729-31.

23. van Hulst LT. Tight control in Rheumatoid Arthritis: bridging the gap between evidence and dailly practice [thesis]. 2011. Available at http://hdl.handle.net/2066/9646.

24. Cochrane LJ, Olson CA, Murray S, et al. Gaps between knowing and doing: understanding and assessing the barriers to optimal health care. *J Contin Educ Health Prof* 2007;27:94-102.

25. Yazdany J, Myslinski R, Francisco M et al. A national electronic health record-enabled registry in rheumatology: the ACR's Rheumatology Informatics System for Effectiveness (RISE) [abstract]. *Arthritis Rheumatol* 2015;67:18-12-0015.

26. Ivers N, Jamtvedt G, Flottorp S et al. Audit and feedback: effects on professional practice and healthcare outcomes. *Cochrane Database Syst Rev* 2012;6:CD000259.

27. Brewer NT, Chapman GB, Schwartz JA, et al. The influence of irrelevant anchors on the judgments and choices of doctors and patients. *Med Decis Making* 2007;27:203-11.

28. Ornstein S, Johnson A, Markert G, et al. Association between family medicine residents' personality and laboratory test-ordering for hypertensive patients. *J Med Educ* 1987;62: 603-5.

29. Ornstein SM, Markert GP, Johnson AH, et al. The effect of physician personality on laboratory test ordering for hypertensive patients. *Med Care* 1988;26:536-43.

30. Sladek RM, Bond MJ, Phillips PA. Why don't doctors wash their hands? A correlational study of thinking styles and hand hygiene. *Am J Infect Control* 2008;36:399-406.

31. Sladek RM, Bond MJ, Huynh LT, et al. Thinking styles and doctors' knowledge and behaviours relating to acute coronary syndromes guidelines. *Implement Sci* 2008;3:23.

32. Wegwarth O, Schwartz LM, Woloshin S, et al. Do physicians understand cancer screening statistics? A national survey of primary care physicians in the United States. *Ann Intern Med* 2012;156:340-9.

Online supplement 1: Guideline adherence parameters, corresponding treatment recommendations and level of measurement

Guideline adherence parameter	Dutch guideline recommendation	Local guideline recommendation	Level of measurement
Radiographs of hands, feet and thorax ordered within the first three visits, in patients with a disease duration ≤1 year	Radiographs of hand and feet at moment of diagnosis and one year thereafter (year 0 and 1).[16]	In case of clinical suspicion for RA, radiographs of hands, feet and thorax should be made.	Patients
Prescription of conventional and biological DMARDs in agreement with the local preferential sequence*	RA treatment should start with methotrexate. Combination therapy with another conventional DMARD or TNF-inhibitor should be given if disease activity remains high despite methotrexate in adequate dosage.[16]	All conventional and biological DMARDs should be prescribed in a pre-defined order which is stated in the local RA guideline.	Patients
Referral to a specialized nurse within the first three visits	Consultation with a specialized rheumatology nurse within one year after diagnosis.[16]]	All newly diagnosed RA patients should be referred to a specialized nurse at the moment of diagnosis in order to get more information on RA and discuss coping.	Patients
Referral to a PA or NP within the first year of treatment	Not applicable	All RA patients should be referred to a PA or NP within the first year of treatment. After referral patients are alternately seen by a PA/NP and rheumatologist, in order to share care between them.	Patients
Therapy change[Ŧ] in case of moderate to high disease activity[π]	Intensification of medication by a rheumatologist in case of a DAS28>3.2 and an adequate period of previous therapy. Adaptation of treatment based on DAS28 scores unless co-morbidity, extra-articular disease and/or side-effects prevent this.[16]	In case of active disease (DAS28 >3.2) therapy should be changed in order to reach low disease activity/remission again.	Visits
Regular outpatient clinic visits combined a nurse led DAS28 assessment	Frequent monitoring of disease activity, for example using the DAS28.[16]	All regular visits at the outpatient clinic for RA patients should be combined with a visit to a specialized nurse in order to assess the DAS28, which is available for the rheumatologist who sees the patient next	Visits

Correct intervals between regular outpatient clinic visits	Planned visit with a rheumatologist within 3 months of the last visit if DAS28 > 2.6. Planned visit with a rheumatologist within 6 months of the last visit if DAS28 < 2.6. Planned visit with a rheumatologist within one year after the last visit if DAS28 < 2.6.[16]	In the first year of treatment RA patients should be seen every three months, thereafter it depends on disease activity and medication use (3-montly visits in case of first year DMARD or biological use or active disease; otherwise 6-montly visits).	Visits

DMARD: Disease Modifying Anti-Rheumatic Drug; DAS28: Disease Activity Score in 28 joints; PA: physician assistant; NP: nurse practitioner. *The preferential order starts with methotrexate and hydroxychloquine combination therapy. After three months treatment is evaluated and if, despite adequate dosage, disease activity remains high (DAS28 >3.2) hydroxychloroquine will be exchanged for etanercept. Treatment effects will be evaluated every three months and in case of persistent high disease activity, the biological DMARD will be switched while the use of methotrexate remains stable. In case of intolerance of contraindications for methotrexate, other conventional DMARDs can be chosen. ‡Therapy change included the intensification of DMARD therapy (dosage increase, shortening of the interval, adding a new DMARD and/or biological, switching to another DMARD and/or biological), starting or increasing corticosteroids (dose), local corticosteroid injections. ¶DAS28 >3.2 or corresponding judgement from the rheumatologist if a DAS28 was not available.

***Online supplement 2**: Questionnaires used to measure rheumatologist-level determinants*

Determinant measured	Explanation of the determinant & relation with guideline adherence	Questionnaire used	Number of scales	Score range
Cognitive bias	Cognitive bias is an error in our thinking that occurs when people process and interpret information, trying to simplify information. Many different types of cognitive biases are known and they can influence medical decision making.[27]	Inventory for Cognitive Bias in Medicine (ICBM): 22 short case descriptions followed by a multiple choice question.[17]	1	0 to 22 (higher score indicates less cognitive bias)
Personality	Personality traits play an important part in our behaviour and can also influence medical decision making.[28, 29]	Big Five Inventory (BFI; Dutch version): 44 statements answered on a 5-point likert scale (completely disagree to completely agree).[19]	5 (extraversion, neuroticism, openness, conscientiousness and agreeableness)	1 to 5 on every subscale (a higher score indicating a more extravert, neurotic, open, conscience or agreeable personality)
Thinking styles	Thinking styles are relatively stable personality traits which address how people think. Two different thinking styles are distinguished: the rational ('deliberate' or 'conscious') and the experiential ('automatic' or 'unconscious') thinking style. These styles have been linked before to guideline adherence.[30, 31]	Rational Experiential Inventory (REI; Dutch version): 40 statements answered on a 5-point likert scale (completely disagree to completely agree).[18]	2 (rationality and experientality)	20 to 100 on every subscale (a higher score indicating a more rational/experiential thinking style)
Numeracy	Numeracy and risk literacy are important when interpreting and acting on risk information. These skills are important in many health decisions but physicians seem to struggle with numeracy.[32] As guideline adherence also involves medical decision making and risk interpretation, this measure was included.	Berlin Numeracy Test (BNT; Dutch version): seven open questions	1	0 to 7 (a higher score indicating a higher level of numeracy)
Knowledge	Guideline knowledge is often mentioned as an important determinant of guideline adherence.[15, 24]	A self-developed questionnaire with 2x10 multiple choice questions on general guideline knowledge and specific RA guideline knowledge	2 (general and specific knowledge)	General knowledge: 0 to 10 Specific knowledge: -5.2 to 10 (negative scores due to correction for guessing). On both scales a higher scores indicates more knowledge.

Guideline agreement	Guideline agreement is mentioned as an determinant of guideline adherence. [15, 24]	Two statements answered on 5-point likert scale (completely disagree to completely agree)	2 (general and specific agreement with the guideline)	1 to 5 on every subscale (a higher score indicating a higher level of agreement)
Outcome expectancy	Outcome expectancy is mentioned as an determinant of guideline adherence. [15, 24]	One statement answered 5-point likert scale (completely disagree to completely agree)	1	1 to 5 (a higher score indicating a higher level of outcome expectancy)


TREAT
TO
TARGET
Don't move... I'm sure I will
hit bull's eye this time!

9

Rheumatologists' guideline adherence in rheumatoid arthritis: a randomized controlled study on electronic decision support, education and feedback

Nienke Lesuis
Ronald F van Vollenhoven
Reinier P Akkermans
Lise M Verhoef
Marlies EJL Hulscher
Alfons A den Broeder

Submitted (Journal of Rheumatology)

ABSTRACT

Objectives

To assess the effects of education, feedback and a computerized decision support system (CDSS) versus education and feedback alone on rheumatologists' rheumatoid arthritis (RA) guideline adherence.

Methods

A single center, randomized controlled pilot study was performed among clinicians (rheumatologists, residents and physician assistants; n = 20) working at the study center, with a 1:1 randomization of included clinicians. A standardized sum score (SSS) on guideline adherence was used as the primary outcome (patient level). The SSS was calculated from 13 dichotomous indicators on quality of RA monitoring, treatment and follow-up.

The randomized controlled design was combined with a before-after design in the control group to assess the effect education and feedback alone.

Results

20 clinicians (mean age 44.3 ± 10.9 years; 55% female) and 990 patients (mean age 62 ± 13 years; 69% female; 72% rheumatoid factor and/or anti-CCP positive) were included. Addition of CDSS to education and feedback did not result in significant better quality of RA care than education and feedback alone (SSS difference 0.02; 95%-CI -0.04 to 0.08; p = 0.60). However, before/after comparison showed that education and feedback alone resulted in a significant increase in the SSS from 0.58 to 0.64 (difference 0.06; 95%-CI 0.02 to 0.11; p <0.01).

Conclusions

Our results suggest that CDSS did not have added value with regard to guideline adherence, whereas education and feedback can lead to a small but significant improvement of guideline adherence.

INTRODUCTION

Treatment of rheumatoid arthritis (RA) is based on the tight control principle in which disease activity monitoring and treatment changes if a pre-set target is not reached, is essential. Treating patients using a tight control strategy, especially in combination with a specified treatment protocol, results in lower disease activity and less functional damage compared to usual care.(1-5) In order to help practicing clinicians using a tight control strategy, many tight control based guidelines are available.(6-8) Next to the tight control based treatment guidelines, separate recommendations exist on topics such as shared care or risk management.(9;10)

Unfortunately, adherence to these guidelines is often suboptimal. For example, treatment is not changed on time in case of active disease or patients do not receive appropriate Disease Modifying Anti-Rheumatic Drug (DMARD) therapy.(11-13) For patients, the lack of adherence to tight control recommendations by rheumatologists can have severe consequences as non-adherence has been associated with more radiographic progression and functional impairment.(5)

Despite these observations on suboptimal guideline adherence and its consequences for patients, implementation research in rheumatology is scarce and almost no trials on improving rheumatologist guideline adherence exist (14). However, Cochrane reviews and two RA studies on often performed interventions (educational meetings, audit and feedback, and reminders) conclude that they all can improve care provided to patients.(15-17)

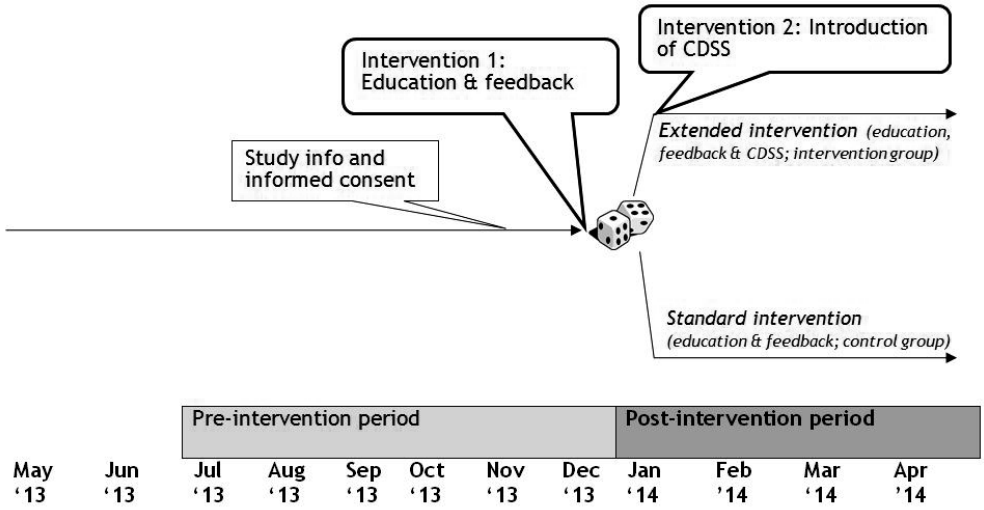
Based on the lack of intervention studies within rheumatology and the existence of effective interventions outside rheumatology, we aim to improve RA care by increasing rheumatologists' guideline adherence using education, feedback and Computerized Decision Support System (CDSS).

METHODS

Study design and participants

A single center, randomized controlled pilot study was performed to assess the effects of an extended intervention strategy including education, feedback and CDSS versus a standard strategy with education and feedback alone. In addition, the randomized controlled design was combined with a before-after design in the control group to assess the effect of the standard intervention strategy alone (figure 1).

Figure 1 *study design*



The study was conducted at the department of rheumatology at the Sint Maartenskliniek (specialized clinic in rheumatology, orthopedics and rehabilitation medicine, the Netherlands). All clinicians prescribing rheumatologic medication (rheumatologists, residents and physician assistants (PA)), working at this center between July 2013 and May 2014 were eligible. Only clinicians that were not willing to sign informed consent were excluded.

Although the interventions were aimed at clinicians, outcomes were measured in patients (provided care in accordance with the guideline 'yes' or 'no'). All adult patients with an ICD-9 code of RA (714.x) with a visit to an included clinician during the pre- or post-intervention period were eligible for inclusion. Participation in a biological DMARD (bDMARD) dose tapering trial, being held at the study center in the same period, was the only exclusion criterion as this trial could influence treatment decisions made during our study.

A random sample of all eligible patients was drawn both before- and after the intervention. This because approximately 2250 unique RA patients are treated at the study center and data collection for all those patients was deemed too labor intensive. Thus, patients were identified from two different time periods: July 2013 to December 2013 (pre-intervention) and January 2014 to April 2014 (post- intervention). Balancing precision and feasibility of data collection, we included 30 RA patients per clinician in both time periods. For those patients, only the first visit in the pre- or post-intervention period was used to assess guideline adherence, meaning that in this study the number of visits and patients is equal, and that the before after comparison is done between two unpaired groups.

Randomization

Included clinicians were randomized in a 1:1 ratio of intervention versus control group using two blocks (block size 10). A research physician allocated clinicians using a computer generated randomization list. After signing informed consent, the clinician received a sealed opaque envelope that contained the randomly assigned allocation. Due to the nature of the interventions, blinding of participants or researchers was not possible.

Interventions

The standard intervention was provided to all clinicians in the intervention- and control group, and comprised a one-hour group session combining education with feedback. The first part of the session focused on the importance of tight control and guideline adherence in RA patients (education). Next, feedback was given on group level and where possible on individual level (non-anonymous). The whole session was developed and provided by NL (PhD student) and AdB (rheumatologist), both working at the study center.

In the intervention group, CDSS was added to the previous intervention (extended intervention strategy). The CDSS was linked to the Electronic Health Record (EHR) used at the study center (EZIS version 5.2, Chipsoft). The CDSS was incorporated into the Computerized Physician Order Entry (CPOE) which was already integrated in the EHR and used by all clinicians. The CDSS worked with algorithms, using clinical information from the EHR, to automatically complete CPOE orders and to send reminders to the clinician about routine care. A week before the CDSS became available to the intervention group, they received 1.5 hour training and until the CDSS was released into the EHR, clinicians could practice with the CDSS in a special training version of the EHR. After implementation of the CDSS into the EHR, assistance from the developers was available for additional explanation of the system. The CDSS was designed in such a way that it could be specifically linked to the Chipsoft account of intervention group clinicians, making it impossible for control group clinicians to access the CDSS thereby preventing contamination between groups. The CDSS development was a close collaboration between clinicians (NL and AdB) and the Information Technology department of the study center.

During the intervention development, we took into account determinants of success as described in relevant reviews on this topic.(15;16;18-20) A more extensive explanation of both interventions can be found in supplement 1, with the PowerPoint slides used during the educational meeting provided in supplement 2.

Outcome measures

As no standard indicator set for quality of care in RA is available (21), we had to develop our own indicator set. We chose to base our set on the indicators stated in the Dutch national RA guideline (6) as this guideline most closely matches the situation at the study center.

As mentioned in the introduction, other recommendations besides tight control treatment are available to rheumatologists. Therefore, we chose to incorporate a broader set of indicators than in previous RA guideline adherence studies. This resulted in the selection of 13 indicators on treatment & monitoring, follow-up & shared care, and administration (table 1). In supplement 3 the development process is described in more detail.
Using the 13 indicators as separate primary outcomes would have resulted in multiple testing problems during the analysis. The primary outcome was therefore the mean difference in a standardized guideline adherence sum score (SSS) between the intervention and control group (i.e the primary analysis was conducted on pre-post intervention differences between education alone versus education and CDSS). The 13 separate indicators all had dichotomous outcomes (1 for 'yes' and 0 for 'no'). In this way the SSS could be calculated for every patient by totaling the score of the individual guideline adherence indicators and dividing this by the number of indicators that applied to this patient (score range 0 to 1, higher scores indicating more guideline adherence). Both the adherence percentages of the separate guideline indicators and the mean difference in SSS before and after the standard intervention in the control group are reported as secondary outcomes.

Data collection
As no real time feedback on all indicators was available at the study center and not all data could be automatically extracted from the EHR, we had to rely on manual EHR review for data collection. For every included patient, data from one visit was collected in either the pre- or post-intervention period. If the patient had visited the clinic more than once during the pre- or post-intervention period, only the first visit in this period was taken into account. Using pre-defined algorithms, the 13 guideline indicators could be calculated from visit data on demographics, disease characteristics, disease activity, functional status and current medication use.
Of note, during data collection we mainly relied on the CPOE orders done by included clinicians. For example, in case of indicator 11 (interval to the next visit) we looked at the corresponding CPOE order and noted the interval that the clinician had entered (i.e. three months, six months, etc.). In reality this follow-up visit could be planned a few weeks before or after the proposed interval due to organizational issues or patient factors. By using the CPOE orders, we were sure that the clinicians' decision had been noted and not organizational or other issues.
Blinded data collection was not possible as it could be directly seen from the EHR whether the patients' treating clinician used the CDSS (intervention group) or not (control group). However, double data extraction and entry was performed on two different random samples of patients in order to achieve high-quality data collection.

Table 1 *Guideline adherence indicators and the relation between indicators and interventions*

	Guideline adherence indicator		Topic covered during interventions	
			Education & feedback	*CDSS*
Treatment & monitoring	1.	DAS28 measurement performed during the outpatient clinic visit	✓	✓
	2.	Radiographs of hands, feet and thorax made at the moment of diagnosis and radiographs of hands and feet repeated 1 and 3 years thereafter	✓	
	3.	Yearly assessment of functional status using the HAQ	✓	✓
	4.	Prescription of conventional and biological DMARDs according to the preferential order[1] when initiating a new DMARD	✓	✓
	5.	Use or prescription of a concomitant conventional DMARD in case of biological use	✓	✓
	6.	Therapy change[2] in case of active disease as measured with the DAS28[3]	✓	
	7.	Dose reduction or interval lengthening (dose optimization) of biological DMARDs in case of low disease activity and stable biological use for the previous six months	✓	
Follow-up & shared care	8.	Referral of new RA patients to a specialized nurse within the two weeks after diagnosis[4]	✓	✓
	9.	Planned nurse led DAS28 assessment during the next regular outpatient clinic visit [5]	✓	✓
	10.	Referral to a PA[6]	✓	✓
	11.	Correct interval between the visit in the study period and the next planned regular outpatient clinic visit	✓	✓
Administration	12.	A letter to the general practitioner, sent within two weeks after diagnosis in case of a new RA patient (new patient letter)		✓
	13.	A letter to the general practitioner, sent once every 18 months (control patient letter)		✓

DAS28, Disease Activity Measurement in 28 joints; HAQ, Health Assessment Questionnaire; DMARD, Disease Modifying Anti-Rheumatic Drug; PA, Physician Assistant. [1]Preferred order in which conventional and biological DMARDs should be prescribed, the exact order is described in the local RA guideline used at the study center. [2]Therapy change included intensifying DMARD therapy, initiating or increasing the dose of oral corticosteroids and the use of corticosteroid injections (either intra-articular or intra-muscular).[3]Active disease as measured with the DAS28 depended on disease duration (disease duration ≤3 years active disease if DAS28 ≥2.6; disease duration >3 years active disease if DAS28 ≥3.2) [4]At the study center all patients should be seen by a specialized nurse in their first year of disease in order to receive more information on RA and discuss coping with the disease. [5]At the study center all RA patients should be seen by a nurse before their visit to the rheumatologist, in order to have the DAS28 measured. [6]At the study center rheumatologists and PA share care for RA patients, meaning that patients are alternatingly seen by a rheumatologist and PA. According to Dutch law PAs are allowed to prescribe DMARDs, but work always under supervision of a rheumatologist.

Ethical approval

This study was presented to the local ethics committee (CMO; Commissie Mensgebonden Onderzoek region Arnhem-Nijmegen), but according to Dutch Act on Medical Research Involving Human Subjects, the study did not need ethical approval (CMO reference number 2013/529). Written informed consent from all participating clinicians was obtained before study start.

All patient data were collected within the study hospital from the local EHR, after which the data was anonymized. As this data cannot be traced back to an individual patient, no written informed consent was needed from the patients according to Dutch Data Protection Act.

The study was registered with the Dutch trial register (www.trialregister.nl, NTR 4449). When reporting this study we followed the CONSORT and SQUIRE guidelines.(22;23)

Statistical analysis & reporting of results

All analyses were done using STATA version 13. Depending on the type of variable, descriptive statistics are presented as absolute numbers with the accompanying percentages, as means with standard deviations (SD) or as median with the interquartile range (IQR).

Based on an earlier retrospective study (24) we expected a mean SSS of 0.27±0.13 in both the intervention and control group before the intervention, increasing to 0.45 in the control group and 0.72 in the intervention group (mean SSS difference: 0.27). With one sided testing (α=0.05, 1−β=0.8) and a randomization ratio of 1:1, we calculated that 18 subjects would be needed for the before/after controlled design and 8 in the randomized controlled design. Potential clustering of patients within a clinician was already accounted for in the sample size calculation by taking the SSS as the primary outcome measure.

To assess our primary outcome, taking the hierarchical structure of our data into account (clustering of patients within clinicians), multilevel linear regression analysis was performed. In the regression model, the SSS was added as the dependent variable with study period, group allocation and the interaction between group allocation and study period as independent variables. By adding the interaction term we tested whether a baseline to post-treatment change in the dependent variable was greater for the intervention group than for the control group. The effect of the standard intervention alone was assessed with a multilevel linear regression model with study period as the independent variable, only using the data from the control group. Results from both multilevel regression analyses are reported as regression coefficients with the corresponding 95% confidence interval (95% CI) and p-value.

Secondary analyses were performed with the thirteen separate guideline indicators using multilevel logistic regression models assessing the added effect of the extended intervention and the separate effect of the standard intervention. Results from these analyses are

reported as odds ratios (OR) with the corresponding 95% confidence interval (95% CI) and p-value.
As not all 13 guideline adherence indicators could be covered with CDSS (table 1), a post-hoc sensitivity analysis was done in order to see if a SSS excluding the indicators not covered in the CDSS yielded different results than the SSS including all indicators.
A second post-hoc sensitivity analysis was performed to assess if SSS results were different when only rheumatologists were included. This was done because clinician randomization resulted in more PAs and residents being allocated to the control group.
Finally, the SSS was also calculated and analyzed for all three groups of indicators separately (treatment & monitoring, follow-up & shared care, administration) whether to see if this made a difference.

RESULTS

Participants
At study start 25 clinicians were assessed for eligibility and 20 fulfilled the inclusion criteria. All eligible clinicians signed informed consent and attended the allocated interventions. No loss to follow-up occurred. Table 2 shows the baseline clinician and patient characteristics. Altogether, 4648 unique adult patients with an ICD-9 code of RA visited the study clinic during the study period (pre- and post-intervention) and after drawing the random sample, 1102 of those patients were selected for participation. Of those, 60 had to be excluded due to participation in the dose tapering study. In addition, during the EHR review a small proportion of patients turned out not to fulfil the inclusion criteria (n= 52). For example, due to rescheduling of visits, no visit in the intervention period was available. This resulted in 990 patients being included in the final analysis (control group n= 508 patients; intervention group n= 482).

Intervention effects on the standardized sum score
Both the standard and extended intervention resulted in an increase of the SSS, with the mean SSS increasing from 0.58 to 0.64 for the standard intervention and from 0.55 to 0.63 for the extended intervention (mean SSS difference 0.02; 95%-CI -0.04 to 0.08; p = 0.60). In the before/after analysis in the control group, the increase in SSS after the standard intervention was statistically significant (mean difference 0.06; 95%-CI 0.02 to 0.11; p <0.01). All post-hoc sensitivity analyses yielded similar results to the primary analysis (table 3).

Table 2 *Clinician and patient characteristics at baseline*

	Control group	Intervention group
Clinician characteristics	***n= 10***	***n= 10***
Age, years (SD)	42.4 (11.1)	46.0 (11.0)
Female sex, n (%)	5 (50)	6 (60)
Rheumatologist, n (%)	9 (90)	6 (60)
Work experience, years (IQR)	5.0 (3.0 to 7.0)	8.0 (8.0 to 14.0)
Patient characteristics	***N= 508***	***N= 482***
Age, years (SD)	62.1 (12.5)	62.0 (12.6)
Female sex, n (%)	340 (66.9)	346 (71.8)
Disease duration, years (IQR)	8.0 (3.0 to 14.0)	7.0 (2.0 to 12.0)
Rheumatoid factor and/or anti-CCP positivity, n (%)	338 (76.5)	257 (67.5)
Erosive disease, n (%)	225 (47.3)	189 (44.0)

SD: standard deviation, IQR: interquartile range

Intervention effects on the individual indicators

The secondary analyses on the individual indicators yielded similar results to the primary analysis with no difference between the standard and extended intervention for any of the indicators. In the before/after comparison four out of thirteen indicators changed significantly after the standard intervention (table 4). Of those four, three improved after the intervention (DAS28 measurements, yearly HAQ assessment and PA referral) and one worsened (radiographs of hands, feet and thorax).

Table 3 *Results on the analyses performed*

Analysis performed	Mean SSS: standard intervention		Mean SSS: extended intervention		Intervention effects (difference; 95% CI (p-value))	
	Pre-intervention	*Post-intervention*	*Pre-intervention*	*Post-intervention*	*Standard vs extended*	*Before/after standard*
Primary analysis (all clinicians; all indicators)	0.58	0.64	0.55	0.63	0.02; -0.04 to 0.08 (0.60)	**0.06; 0.02 to 0.11 (<0.01)**
Post-hoc sensitivity analysis						
Only CDSS indicators*	0.59	0.67	0.58	0.67	0.01; -0.06 to 0.07 (0.85)	**0.08; 0.03 to 0.13 (<0.01)**
Only rheumatologists	0.53	0.58	0.56	0.63	0.02; -0.05 to 0.08 (0.65)	0.05;-0.002 to 0.10 (0.06)
Only indicators on monitoring & treatment	0.58	0.64	0.59	0.67	0.02;-0.7 to 0.12 (0.62)	0.06; -0.004 to 0.13 (0.07)
Only indicators on follow-up & referral	0.49	0.55	0.47	0.55	0.02; -0.06 to 0.10 (0.61)	**0.06; 0.002 to 0.12 (0.04)**
Only indicators on administration	0.73	0.75	0.73	0.72	-0.02; -0.14 to 0.08 (0.64)	0.02; -0.06 to 0.09 (0.67)

*SSS calculation from all indicators except indicator 2 (radiographs), 6 (therapy change) and 7 (biological dose optimization)

Table 4 *Guideline adherence to and intervention effects on the separate indicators*

Indicator	Adherence: standard intervention		Adherence: extended intervention		Odds ratio (95% CI)	
	Pre-intervention (n= 254)	*Post-intervention (n= 254)*	*Pre-intervention (n= 241)*	*Post-intervention (n= 241)*	*Standard vs extended*	*Before/after standard*
1.DAS28 measurement, % (n)	66.8 (144 / 216)	80.3 (183 / 228)	67.5 (139 / 206)	75.6 (164 / 217)	0.7 (0.4 to 1.4)	**2.0 (1.3 to 3.1)**
2.Radiographs, % (n)	48.4 (31 / 64)	13.2 (7 / 53)	35.3 (24 / 68)	23.0 (14 / 61)	3.2 (0.9 to 11.4)	**0.2 (0.1 to 0.4)**
3.Yearly HAQ, % (n)	68.5 (148 / 216)	79.4 (181 / 228)	68.1 (139 / 204)	83.0 (180 / 217)	1.4 (0.7 to 2.8)	**1.8 (1.1 to 2.7)**
4.Correct DMARD prescription, % (n)	50.0 (5 / 10)	46.2 (6 / 13)	78.6 (11 / 14)	65.0 (13 / 20)	1.1 (0.1 to 17.4)	0.4 (0.1 to 4.5)
5.Concomittant DMARD, % (n)	50.7 (36 / 71)	62.6 (62 / 99)	61.3 (49 /80)	59.5 (47 / 79)	0.6 (0.2 to 1.5)	1.6 (0.9 to 3.1)
6.Therapy change in case of active disease, % (n)	63.9 (23 / 36)	56.4 (31 / 55)	47.1 (16 / 34)	52.8 (28 / 53)	1.7 (0.5 to 5.9)	0.7 (0.3 to 1.7)
7.Biological dose optimization, % (n)	10.5 (4 / 38)	12.0 (6 / 50)	18.8 (6 / 32)	9.7 (3 / 31)	0.5 (0.1 to 4.0)	2.0 (0.2 to 4.2)
8.Referral to a specialized nurse, % (n)	30.8 (8 / 26)	30.8 (4 / 13)	15.8 (3 / 19)	23.1 (3 / 13)	6.2 (0.2 to 221.7)	0.9 (0.2 to 5.6)
9.Planned nurse-led DAS28 at next visit, % (n)	59.8 (143 / 239)	69.6 (156 / 224)	69.7 (156 / 223)	69.8 (148 / 212)	0.8 (0.4 to 1.5)	1.4 (0.9 to 2.1)
10.Referral to a PA, % (n)	22.1 (31 / 140)	39.2 (38 / 97)	14.6 (27 / 185)	34.3 (57 / 166)	1.3 (0.6 to 2.9)	**2.4 (1.3 to 4.4)**
11.Correct visit interval, % (n)	47.8 (108 / 226)	53.7 (116 / 216)	56.9 (119 / 209)	61.5 (126 / 205)	1.0 (0.6 to 1.7)	1.2 (0.8 to 1.8)
12.Timely new patient letter, % (n)	64.0 (16 / 25)	75.0 (9 / 12)	42.1(8 / 19)	53.9 (7 / 13)	1.0 (0.1 to 7.7)	1.7 (0.4 to 7.9)
13.Timely control patient letter, % (n)	73.2 (169 / 231)	74.9 (182 / 243)	75.7 (168 / 222)	72.4 (168 / 232)	0.8 (0.4 to 1.5)	1.1 (0.7 to 1.6)

Values in bold are significant (p-value <0.05). DAS28, Disease Activity Score in 28 joints; CDSS, Computerized Decision Support; HAQ, Health Assessment Questionnaire; DMARD, Disease Modifying Anti-Rheumatic Drug; PA, Physician Assistant.

DISCUSSION

To our knowledge this is one of the first randomized controlled trials within rheumatology trying to improve guideline adherence of clinicians. Our results show that CDSS has no added value in this context, whereas education and feedback did lead to a significant improvement in guideline adherence.

The strengths of this study are the use of a randomized design, a broad set of indicators, inclusion of different types of clinicians involved in RA care, inclusion of a large sample of both early and established RA patients reflecting daily clinical practice and the use of two different interventions.
However, this study has some limitations related to the internal validity and generalizability. Firstly, not all desired changes could be implemented in the CDSS. As a result, the SSS included indicators not covered with this intervention (table 1). This concerned indicator 2 (radiographs), 6 (therapy change) and 7 (biological dose optimization). Nevertheless, sensitivity analyses yielded no different results when excluding these indicators from the SSS calculation. Secondly, after randomization the control group included more rheumatologists than the intervention group (90% vs 60%), but this did not seem to have influenced our results as sensitivity analyses excluding non-rheumatologists gave similar results as the original analysis. Thirdly, due to our study design we are not able to infer a causal relation between the standard intervention and guideline adherence afterwards as other events in the same time period might have attributed to the observed results. However, we are not aware of any events during the study that could have influenced our results and during the study special attention was paid not to start other quality improvement projects. Fourthly, as this was a single center study the generalizability may be hampered due to differences on patient-, hospital-, or societal level. However, the RA population treated in the study center seems to represent a normal RA population, thus not hampering generalizability. Of course, the study center being a specialized clinic and the study only being performed in the Netherlands might have influenced our results, which stresses the need for replication of our study in other settings. Finally, the use of a broad set of indicators in combination with the sample size can also be seen as a disadvantage as not all patients could be included in all indicators. However, by using this set of indicators for the first time we were able to gain more insight into the broad concept of quality of care in rheumatology. However, future studies, preferably multi-centre, should use a larger sample in order to be able to confirm our results when using multiple indicators.

Being one of the first intervention studies to improve clinicians' RA guideline adherence also has a downside, as we cannot directly compare all of our results with other groups. With regard to our standard intervention we are aware of one other study using education and feedback to improve RA care (Metrix study). In this randomized controlled trial, rheumatologists receiving education and feedback (n= 10) collected more global assessments and HAQs than their colleagues not receiving these interventions (n=10). Furthermore, the researchers could calculate more composite scores in the intervention group (increase from 43% to 57%), whereas the control group did not show any change. Finally, the intervention group did change therapy in 57% of the patients with a high DAS compared to 38% of the rheumatologists in the control group.(25) Our results from the standard intervention on comparable topics are similar (therapy change in active disease) or better (HAQ and DAS28 measurement). However, the Metrix study did not measure if composite scores were actually calculated by the rheumatologists themselves and if they were used to guide treatment decisions. This makes our study probably more useful in judging the effect of education and feedback on the use of composite measures such as the DAS28 in daily practice.

With regard to CDSS, parallels can be drawn with other studies within rheumatology but of the four studies we are aware of, only one focuses on RA.(26-29) In this study a template, integrating information from different sources (i.e. physician itself, patients and/or EHR), was implemented. Following implementation of this system, a strong correlation was found between use of the system by rheumatologists and disease control, and more patients were in a state of low disease activity.(28) Although both the intervention and study population are not fully comparable with ours, these results imply that care for patients with rheumatic diseases could benefit from EHR changes.

Outside rheumatology far more studies have been performed on the effect of education and feedback. Two Cochrane reviews on this subject conclude that both educational meetings and feedback can improve clinical practice, although the effects are often small to moderate which resembles the effects found in this study.(15;16;30) Similarly, different reviews on CDSS have been performed outside rheumatology concluding that CDSS results are not always consistent but can improve practitioner performance. However, patient outcomes such as morbidity and mortality are at best moderately improved.(17;31-34)

Finally, it is interesting to notice that not all indicators did show an improvement after the intervention. For example, the indicators on ordering of radiographs and correct DMARD prescriptions worsened after the intervention in both the standard and extended intervention group. For both observations we do not have a good explanation. However, in the light of these results and the previously mentioned reviews, our results emphasize the need for better understanding why interventions work in one setting and not in another. Several reviews have addressed this issue and many factors could possibly influence successful uptake of the interventions. We have tried to take these factors into account

during the development and execution of our interventions, for example by making sure CDSS was integrated into the workflow and the messages were timely and relevant. Also, attendance during the educational meeting was high and feedback was provided by a direct colleague. However, it was not possible to incorporate all the potential factors for success, which might explain the small effects observed. In addition, guideline adherence might be classified as complex behavior due the many, and often interconnected, recommendations that have to be followed. This could have led to the small effect of education and feedback, with our CDSS not being adequate enough to fill in the gap between knowing about the recommendations and actually practicing them.

Despite the small effects observed, we feel that our study has important practical implications, especially within rheumatology. First of all, the results of this study confirm that improving guideline adherence is a challenge. However, the improvement resulting from our standard intervention is a first step in the right direction and again stresses the importance of more attention towards the implementation of guidelines. Secondly, this study probably could have benefited from a more formal barrier analysis before study start, in order to develop an even more targeted intervention. Although a Cochrane review on this subject is not conclusive, future studies should certainly consider such an approach.(35) Lastly, our study is an example of implementation research where we tried to bridge the gap between evidence and practice. So far, this type of research is scarce within rheumatology which was recently recognized by Buchbinder *et al*. We agree with these authors that only performing clinical research is not enough to improve care if no attention is given to the implementation of new findings in clinical practice.(14) Therefore, we would strongly advocate for more attention towards implementation science within rheumatology in order to let more patients benefit from optimal RA care.

ACKNOWLEDGEMENTS

We would like to thank all rheumatologists, residents and physician assistants who participated in this study for their willingness to attend the interventions. Furthermore, we want to thank Joost van Zadelhoff, Ilse Schasfoort and Marc Toonen for their invaluable help during the preparation, development and testing of the CDSS. Finally, we thank Lieke Nieboer for her help during data collection.

CONTRIBUTORS

All authors contributed to the study design. NL and AdB designed and executed the intervention. NL collected the data. NL, RA, AdB and MH participated in the data analysis. NL, AdB, MH and RvV assisted in the interpretation of the data. The manuscript was written by NL under the supervision of AdB, MH and RvV. All authors have revised the draft version of the manuscript, read and approved the final version of the manuscript. NL is guarantor.

COMPETING INTEREST DECLARATION AND FUNDING

All authors have completed the Unified Competing Interest form at www.icmje.org/coi_disclosure.pdf (available on request from the corresponding author) and declare that RvV has received grants and personal fees from AbbVie, Amgen, BMS, GSK, Pfizer, Roche, UCB, Biotest, Celgene, Crescendo, Janssen, Lilly, Merck and Vertex outside the submitted work; all other authors reported no competing interests. This study received no external funding.

DATA SHARING STATEMENT

Patient level data, full dataset, technical appendix and statistical code are available at a reasonable request from the corresponding author. Consent from the patients was not obtained but the presented data are anonymized and risk of identification is low.

REFERENCES

(1) Vermeer M, Kuper HH, Bernelot Moens HJ, Hoekstra M, Posthumus MD, van Riel PL, et al. Adherence to a treat-to-target strategy in early rheumatoid arthritis: results of the DREAM remission induction cohort. Arthritis Res Ther 2012;14:R254.

(2) Schipper LG, van Hulst LT, Grol R, van Riel PL, Hulscher ME, Fransen J. Meta-analysis of tight control strategies in rheumatoid arthritis: protocolized treatment has additional value with respect to the clinical outcome. Rheumatology (Oxford) 2010;49:2154-64.

(3) Fransen J, Moens HB, Speyer I, van Riel PL. Effectiveness of systematic monitoring of rheumatoid arthritis disease activity in daily practice: a multicentre, cluster randomised controlled trial. Ann Rheum Dis 2005;64:1294-8.

(4) Grigor C, Capell H, Stirling A, McMahon AD, Lock P, Vallance R, et al. Effect of a treatment strategy of tight control for rheumatoid arthritis (the TICORA study): a single-blind randomised controlled trial. Lancet 2004;364:263-9.

(5) Escalas C, Dalichampt M, Combe B, Fautrel B, Guillemin F, Durieux P, et al. Effect of adherence to European treatment recommendations on early arthritis outcome: data from the ESPOIR cohort. Ann Rheum Dis 2012;71:1803-8.

(6) van Riel PL, Barendsen BC, van Croonenborg JJ, et al. Diagnostiek en behandeling van reumatoïde artrits. 2015. Available at http://www.diliguide.nl/document/3249/reumatoide-artritis.html

(7) Smolen JS, Breedveld FC, Burmester GR, Bykerk V, Dougados M, Emery P, et al. Treating rheumatoid arthritis to target: 2014 update of the recommendations of an international task force. Ann Rheum Dis 2015;0:1-13.

(8) Smolen JS, Landewe R, Breedveld FC, Buch M, Burmester G, Dougados M, et al. EULAR recommendations for the management of rheumatoid arthritis with synthetic and biological disease-modifying antirheumatic drugs: 2013 update. Ann Rheum Dis 2014;73:492-509.

(9) Peters MJ, Symmons DP, McCarey D, Dijkmans BA, Nicola P, Kvien TK, et al. EULAR evidence-based recommendations for cardiovascular risk management in patients with rheumatoid arthritis and other forms of inflammatory arthritis. Ann Rheum Dis 2010;69:325-31.

(10) van Eijk-Hustings Y, van TA, Bostrom C, Braychenko E, Buss B, Felix J, et al. EULAR recommendations for the role of the nurse in the management of chronic inflammatory arthritis. Ann Rheum Dis 2012;71:13-9.

(11) Kiely P, Williams R, Walsh D, Young A. Contemporary patterns of care and disease activity outcome in early rheumatoid arthritis: the ERAN cohort. Rheumatology (Oxford) 2009;48:57-60.

(12) Harrold LR, Harrington JT, Curtis JR, Furst DE, Bentley MJ, Shan Y, et al. Prescribing practices in a US cohort of rheumatoid arthritis patients before and after publication of the American College of Rheumatology treatment recommendations. Arthritis Rheum 2012;64:630-8.

(13) Benhamou M, Rincheval N, Roy C, Foltz V, Rozenberg S, Sibilia J, et al. The gap between practice and guidelines in the choice of first-line disease modifying antirheumatic drug in early rheumatoid arthritis: results from the ESPOIR cohort. J Rheumatol 2009;36:934-42.

(14) Buchbinder R, Maher C, Harris IA. Setting the research agenda for improving health care in musculoskeletal disorders. Nat Rev Rheumatol 2015;11:597-605.

(15) Ivers N, Jamtvedt G, Flottorp S, Young JM, Odgaard-Jensen J, French SD, et al. Audit and feedback: effects on professional practice and healthcare outcomes. Cochrane Database Syst Rev 2012;6:CD000259.

(16) Forsetlund L, Bjorndal A, Rashidian A, Jamtvedt G, O'Brien MA, Wolf F, et al. Continuing education meetings and workshops: effects on professional practice and health care outcomes. Cochrane Database Syst Rev 2009;(2):CD003030.

(17) Shojania KG, Jennings A, Mayhew A, Ramsay CR, Eccles MP, Grimshaw J. The effects of on-screen, point of care computer reminders on processes and outcomes of care. Cochrane Database Syst Rev 2009;(3):CD001096.

(18) Roshanov PS, Fernandes N, Wilczynski JM, Hemens BJ, You JJ, Handler SM, et al. Features of effective computerised clinical decision support systems: meta-regression of 162 randomised trials. BMJ 2013;346:f657.

(19) Moxey A, Robertson J, Newby D, Hains I, Williamson M, Pearson SA. Computerized clinical decision support for prescribing: provision does not guarantee uptake. J Am Med Inform Assoc 2010;17:25-33.

(20) Kawamoto K, Houlihan CA, Balas EA, Lobach DF. Improving clinical practice using clinical decision support systems: a systematic review of trials to identify features critical to success. BMJ 2005;330:765.

(21) Mahmood S, Lesuis N, van Tuyl LH, van Riel PL, Landewe R. Quality in rheumatoid arthritis care. Best Pract Res Clin Rheumatol 2015;29:664-79.

(22) Schulz KF, Altman DG, Moher D. CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. BMJ 2010;340:c332.

(23) Ogrinc G, Davies L, Goodman D, Batalden P, Davidoff F, Stevens D. SQUIRE 2.0 (Standards for QUality Improvement Reporting Excellence): revised publication guidelines from a detailed consensus process. BMJ Qual Saf 2015;17:i13-i32.

(24) Lesuis N, Den Broeder AA, Hulscher ME, van Vollenhoven RF. Practice what you preach? An exploratory multilevel study on rheumatoid arthritis guideline adherence by rheumatologists. RMD Open 2016;2:e000195.

(25) Pope J, Thorne C, Cividino A, Lucas K. Effect of rheumatologist education on systematic measurements and treatment decisions in rheumatoid arthritis: the metrix study. J Rheumatol 2012;39:2247-52.

(26) Rolnick SJ, Jackson JM, Amundson JH. Development, implementation and evaluation of an electronic medical record prompt for bone density testing. Health Informatics J 2009;15:296-304.

(27) Ledwich LJ, Harrington TM, Ayoub WT, Sartorius JA, Newman ED. Improved influenza and pneumococcal vaccination in rheumatology patients taking immunosuppressants using an electronic health record best practice alert. Arthritis Rheum 2009;61:1505-10.

(28) Newman ED, Lerch V, Billet J, Berger A, Kirchner HL. Improving the quality of care of patients with rheumatic disease using patient-centric electronic redesign software. Arthritis Care Res (Hoboken) 2015;67:546-53.

(29) Edwards JJ, Jordan KP, Peat G, Bedson J, Croft PR, Hay EM, et al. Quality of care for OA: the effect of a point-of-care consultation recording template. Rheumatology (Oxford) 2015;54:844-53.

(30) Grimshaw JM, Thomas RE, MacLennan G, Fraser C, Ramsay CR, Vale L, et al. Effectiveness and efficiency of guideline dissemination and implementation strategies. Health Technol Assess 2004;8:iii-72.

(31) Moja L, Kwag KH, Lytras T, Bertizzolo L, Brandt L, Pecoraro V, et al. Effectiveness of computerized decision support systems linked to electronic health records: a systematic review and meta-analysis. Am J Public Health 2014;104:e12-e22.

(32) Roshanov PS, Misra S, Gerstein HC, Garg AX, Sebaldt RJ, Mackay JA, et al. Computerized clinical decision support systems for chronic disease management: a decision-maker-researcher partnership systematic review. Implement Sci 2011;6:92.

(33) Garg AX, Adhikari NK, McDonald H, Rosas-Arellano MP, Devereaux PJ, Beyene J, et al. Effects of computerized clinical decision support systems on practitioner performance and patient outcomes: a systematic review. JAMA 2005;293:1223-38.

(34) Bright TJ, Wong A, Dhurjati R, Bristow E, Bastian L, Coeytaux RR, et al. Effect of clinical decision-support systems: a systematic review. Ann Intern Med 2012;157:29-43.

(35) Baker R, Camosso-Stefinovic J, Gillies C, Shaw EJ, Cheater F, Flottorp S, et al. Tailored interventions to address determinants of practice. Cochrane Database Syst Rev 2015;4:CD005470.

(36) Cochrane LJ, Olson CA, Murray S, Dupuis M, Tooman T, Hayes S. Gaps between knowing and doing: understanding and assessing the barriers to optimal health care. J Contin Educ Health Prof 2007;27:94-102.

(37) Lawrence M, Olesen F. Indicators of Quality in Health Care. European Journal of General Practice 1997;3:103-8.

(38) Donabedian A. The quality of care. How can it be assessed? JAMA 1988;260:1743-8.

SUPPLEMENT 1: INTERVENTION DESCRIPTION

Intervention choice

The interventions used in this study were education combined with feedback, and a Computerized Decision Support System (CDSS). These specific interventions were chosen based on an informal barrier analyses at the study center, using the framework proposed by Cochrane *et al.*(36) This framework groups determinants of guideline adherence into different themes such as cognitive-behavioral barriers, barriers embedded in the guideline or evidence, and barriers related to support or resources. In the years before this study started, the study center already invested in improving knowledge and attitudes concerning RA treatment. As a result, amongst others, up to date RA treatment guidelines were available and a safe learning environment was created. Looking at the Cochrane framework, we concluded that the theme 'barriers related to support or resources' was not yet optimally covered at the study center. Therefore, we decided to develop our own CDSS to aid rheumatologists in their daily practice. As the latest updates from the local RA guideline (2013) had not been presented in an educational session before, we decided to also include education and feedback in our intervention strategy. This resulted in a standard intervention strategy (education and feedback) being tested against an extended intervention strategy (education, feedback and CDSS).

When developing both interventions we took into account existing reviews on factors of success for education, feedback and CDSS. For example, CDSS uptake was found to be more successful if adequate technical support and training were present, CDSS was integrated into the workflow and the messages were relevant and on time.(18-20) For education and feedback factors like attendance, the source of feedback and the complexity of the targeted behavior influence intervention effects.(15;16) Modifiable factors from these reviews were reckoned with during the development process.

Education & feedback

All included clinicians attended a one-hour group session, combining an educational meeting with feedback. During this meeting clinicians received background information on the effectiveness of tight control treatment strategies in RA treatment, the importance of guideline adherence for RA patients and the content of existing local RA treatment guidelines. With regard to the latter, extra attention was given to the local guideline on biological dose optimization, which was disseminated just before this study started. The session finished with feedback on current guideline adherence of the clinicians, using results from a previous study on guideline adherence in this center (manuscript under review) and existing feedback systems in the study hospital. Due to this combination we could give feedback on all but two indicators (concomitant cDMARD use and bDMARD dose reduction)

included in this study. The PowerPoint slides used during this intervention can be found in supplement 2.

Computerized Decision Support System

Background on the Computerized Physician Order Entry System used at the study clinic
In order to fully understand the CDSS used as an intervention in our study, it is first necessary to know how the EHR at the study clinic worked before the study. At the study center EZIS version 5.2 (Chipsoft) is used by all physicians and a Computerized Physician Order Entry (CPOE) system was already integrated in this EHR. As the CDSS intervention in this study focused at the CPOE, a description of the different CPOE categories before the intervention is given in table 1.

Description of the Computerized Decision Support System
The main aim of the changes made to the CPOE was to facilitate guideline adherence by clinicians working at the outpatient clinic of the rheumatology department. By reorganizing the CPOE and including CDSS it should be less easy for clinicians to accidently forget about important recommendations from the local RA guidelines. To achieve this goal, four changes were made to the CPOE.
The first of four changes included a reorganization of the CPOE system. As can be seen in table 1, the organization of the CPOE categories was mainly receiver-based and not very practical for clinicians. In the new version grouping of CPOE orders was done in a sender-based way, with four main categories: diagnostics, treatment, follow-up and administration. All orders from the old system were placed into the new categories and some orders were adapted to make the other changes possible.
With the next change we included hyperlinks to local guidelines in the CPOE system. For example, the order on DMARD initiation now included a hyperlink to the guidelines on the DMARD preferential order and DMARD toxicity follow-up.
The third change comprised the development of an algorithm which automatically completed some of the CPOE orders, based on patient-specific information from the EHR and the local RA guideline. An example of this change was the follow-up order with the algorithm using clinical information from the EHR, user login and local guidelines to complete the three main components of the order. These components were: 1) follow-up duration (3 or 6 months based on disease duration, disease activity and DMARD use), 2) preferred provider of care (PA or rheumatologist based on provider of the current visit), and 3) referral to a specialized nurse for routine DAS28 and HAQ assessments. Clinicians not agreeing with the suggestions done by the CDSS could always change the answers on all components of the order before sending the order away.

Table 1 *Main order categories within the pre-intervention Computerized Physician Order Entry system*

Order category	Suborders included into the category	Receiver
Clinical admissions	Clinical admissions and surgery	Surgery and clinical admission planner
Multidisciplinary treatments	No suborders were included in this category, although the choices following this order also included ordering of infusion therapies such as rituximab or infliximab.	Multidisciplinary treatment planner
Order to rheumatologist	No suborders were included in this category. This order only contained a text field in which a remark or question to another rheumatologist could be entered (limited number of characters).	Rheumatologist as chosen by the sender of the order
Order to front office	15 suborders were included in this category, ranging from orders on follow-up appointments to routine laboratory checks in DMARD users. This order could only be used if a patient was present at the outpatient clinic.	Nurses or supportive staff at the front office at the rheumatology outpatient clinic
Order to outpatient clinic nurse	15 suborders were included in this category, ranging from blood pressure measurement to the preparation of intra-articular injections.	Nurse at the rheumatology outpatient clinic
Order to secretary	Several suborders were included in this category such as retrieving patient information from other hospitals.	Secretary of the rheumatology department
Correspondence	No suborders were included in this category. This order only included a choice on what of letter should be made (new or control patient) and a few text fields were additional information on the receiver (general practitioner, other specialist) or letter (attachments) could be entered.	Secretary of the rheumatology department
Order to archive	No suborders were included in this category. This order could only be used to retrieve old paper chart from the archive.	Archive
Order to myself	No suborders were included in this category. This order only included a text field were a remark or question could be entered. This order often acted as a reminder for the sender to perform certain actions for the patient in question (for example calling the GP to discuss the patient).	Sender of the order
Order to back office	No suborders were included in this category. This order only included a text field were a remark or question could be entered. This order served as a substitute for the 'front office order' if a patient was not present at the outpatient clinic. For example, if after a telephone call with a patient, a follow-up visit should be planned, this order had to be used.	Nurses or supportive staff at the back office at the rheumatology outpatient clinic
Consulting other specialists	No suborders were included in this category. This order only included text fields in which clinical information about the patient and questions for the consulting specialist could be entered. This order only applied to clinically admitted patients.	Internal medicine or gerontology specialist
Laboratory tests	No suborders were included in this category. All available laboratory tests at the study center were included in this order and could be selected by clinicians.	Nurse at the rheumatology outpatient clinic

The final change included the development of a reminder system to assist clinicians in keeping their correspondence with the general physician up to date. With this system a reminder was created every time a new visit was entered into the EHR by a rheumatologist, PA or resident. This reminder consisted of an order stating that a letter to the GP should be created. However, this reminder only became visible to the clinician after a certain period of time (2 weeks or 18 months depending on the type of letter needed) if the clinician had not created a GP letter himself in the meantime. In this way we prevented unnecessary reminders for rheumatologists who did not need them.
In table 2 the relation between the CDSS changes and the guideline adherence indicators used in this study is stated.

Table 2 *CDSS changes in relation to the guideline adherence indicators*

Guideline adherence indicator	Topic covered with CDSS			
	Regrouping CPOE	*Hyperlink to guideline*	*Pre-fill orders*	*Reminders*
DAS28 measurement	✓		✓	
Radiographs of hands, feet and thorax	✓			
Yearly assessment of functional status using the HAQ	✓		✓	
Prescription of conventional and biological DMARDs according to the preferential order	✓	✓		
Concomitant conventional DMARD in case of biological use	✓	✓		
Therapy change in case of active disease as measured with the DAS28	✓	✓		
Dose reduction or interval lengthening (dose optimization) of biological DMARDs	✓	✓		
Referral of new RA patients to a specialized nurse	✓		✓	
Planned nurse led DAS28 assessment during the next regular outpatient clinic visit	✓		✓	
Referral to a PA	✓		✓	
Correct interval between the visit in the study period and the next planned regular outpatient clinic visit	✓		✓	
A letter to the general practitioner, sent within two weeks after diagnosis in case of a new RA patient	✓			✓
A letter to the general practitioner, sent once every 18 months (control patients)	✓			✓

SUPPLEMENT 2: POWERPOINT SLIDES USED DURING THE EDUCATIONAL MEETING

Supplement 2: PowerPoint slides used during the educational meeting

Treatment principles in rheumatoid arthritis

A translation of the PowerPoint slides used during the educational meeting

Authors: Nienke Lesuis (resident of rheumatology; MD) and Alfons A den Broeder (rheumatologist-epidemiologist; MD, PhD)
Sint Maartenskliniek, Nijmegen, the Netherlands

Content

- Background
- Treatment principles in RA
- Current situation at our hospital
- Summary

Content

- **Background**
- Treatment principles in RA
- Current situation at our hospital
- Summary

Background

We would rather prevent than treat these hands, but… treatment guidelines are not always followed optimally

Content

- Background
- **Treatment principles in RA**
- Current situation at our hospital
- Summary

Treatment principles in RA

- Many publications on effective RA treatments
- 'Treat to target', 'tight control' & 'hit hard, hit early' often mentioned

→ Cornerstone of ACR & EULAR RA guidelines

TREAT TO TARGET

Treatment principles in RA

Tight control: 'frequent assessment of disease activity combined with an objective structured protocol to make treatment changes that maintain low disease activity or remission at an agreed target' *(Kirly et al. Rheumatology 2009)*

Treat to target: 'the treatment aim was defined as remission with low disease activity being an alternative in patients with long-standing disease. Regular follow-up with appropriate therapeutic adaptation to reach the desired state within 3-6 months. Follow-up examinations ought to employ composite measures of disease activity which include joint counts. *(Smolen et al. Ann Rheum Dis 2010)*

Hit hard, hit early: 'early institution of DMARDs.' 'Window of opportunity'

Treatment principles in RA

Tight control: 'frequent assessment of disease activity combined with an objective structured protocol to make treatment changes that maintain low disease activity or remission at an agreed target' *(Grigor et al., Rheumatology 2006)*

Treat to target: 'the treatment aim was defined as remission with low disease activity being an alternative in patients with long-standing disease. Regular follow-up with appropriate therapeutic adaptation to reach the desired state within 3-6 months. Follow-up examinations ought to employ composite measures of disease activity which include joint counts. *(Smolen et al., Ann Rheum Dis 2010)*

Hit hard, hit early: 'early institution of DMARDs.' 'Window of opportunity'

Treatment principles in RA

"Yes, I agree with the general principle but…."

'*A number says nothing! I always look and listen to the patient, order an ESR and feel the joints*'

'*Patients don't want all those pills*'

'*My judgement is better than a composite measure, these numbers mean nothing to a patient*'

Treatment principles in RA

"Yes, I agree with the general principle but…."

DAS28 = 0.56 * sqrt(tender28) + 0.28 * sqrt(swollen28) + 0.70 * ln(ESR) + 0.014 * VAS *(http://www.das-score.nl/das28/nl/uitleg-das28/de-das28-score.html)*

In case of active disease, patients just want to get better. They only get worried about the number of pills later on. *(van Tuyl, Rheumatology 2008)*

A DAS28 <3.2 is associated with 50% less progression of radiographic damage and functional status (HAQ) is influenced by both active inflammation and radiographic damage. *(Fransen et al., Ann Rheum Dis 2005)*

Treatment principles in RA

Effect of a treatment strategy of tight control for rheumatoid arthritis (the TICORA study): a single-blind randomised controlled trial

Catriona Grigor, Hilary Capell, Anne Stirling, Alex D McMahon, Peter Lock, Rtsmay Vallance, Wilma Kincaid, Duncan Porter *Lancet 2004; 364: 263–69*

- RCT: routine vs intensive management
- Inclusion criteria: diagnosis of RA <5 years; DAS >2.4
- Study assessment 1x/3 months (DAS etc) by blinded assessor
- Primary endpoints:
  - Mean fall in disease activity
  - % patients with an EULAR good response

Treatment principles in RA

Effect of a treatment strategy of tight control for rheumatoid arthritis (the TICORA study): a single-blind randomised controlled trial

Catriona Grigor, Hilary Capell, Anne Stirling, Alex D McMahon, Peter Lock, Rtsmay Vallance, Wilma Kincaid, Duncan Porter *Lancet 2004; 364: 263–69*

- Intensive management:
  - Monthly visits including DAS measurement
  - Corticosteroid injection of any swollen joint or depomedrol 120mg i.m. if DAS >2.4
  - Strict medication protocol: dose increase or cDMARD switch every 1-3 month if DAS >2.4
- Routine care:
  - 3-montly visits
  - No routine DAS measurement or strict medication protocol

Treatment principles in RA

Effect of a treatment strategy of tight control for rheumatoid arthritis (the TICORA study): a single-blind randomised controlled trial

Catriona Grigor, Hilary Capell, Anne Stirling, Alex D McMahon, Peter Lock, Rtsmay Vallance, Wilma Kincaid, Duncan Porter *Lancet 2004; 364: 263–69*

	Intensive group (n=55)	Routine group (n=55)	Odds ratio (95% CI)	p*
EULAR good response	45 (82%)	24 (44%)	5·8 (2·4–13·9)*	<0·0001
EULAR remission	36 (65%)	9 (16%)	9·7 (3·9–23·9)*	<0·0001
ACR 20 response	50 (91%)	35 (64%)	5·7 (1·9–16·7)*	<0·0001
ACR 50 response	46 (84%)	22 (40%)	6·1 (2·5–14·9)*	<0·0001
ACR 70 response	39 (71%)	10 (18%)	11 (4·5–27)*	<0·0001

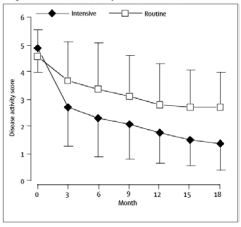
Intention-to-treat analysis of all patients randomised, including those who died or withdrew from the study. Analysis of patients completing the study is very similar (data not shown). *Mantel-Haenszel procedure used.

***Table 2:* Number of patients responding at 18-month assessment**

Treatment principles in RA

Effect of a treatment strategy of tight control for rheumatoid arthritis (the TICORA study): a single-blind randomised controlled trial

Catriona Grigor, Hilary Capell, Anne Stirling, Alex D McMahon, Peter Lock, Rtsmay Vallance, Wilma Kincaid, Duncan Porter *Lancet 2004; 364: 263–69*



Treatment principles in RA

Effect of a treatment strategy of tight control for rheumatoid arthritis (the TICORA study): a single-blind randomised controlled trial

Catriona Grigor, Hilary Capell, Anne Stirling, Alex D McMahon, Peter Lock, Rtsmay Vallance, Wilma Kincaid, Duncan Porter *Lancet 2004; 364: 263–69*

- Intensive management:
  - More corticosteroid injections (i.m./i.a.)
  - Higher doses of MTX
  - More frequent start of a new DMARD
  - Higher drug survival
  - Less medication side-effects

Conclusion: intensive treatment gives substantial improvement of disease activity

Treatment principles in RA

Effectiveness of systematic monitoring of rheumatoid arthritis disease activity in daily practice: a multicentre, cluster randomised controlled trial

J Fransen, H Bernelot Moens, I Speyer, P L C M van Riel

Ann Rheum Dis 2005;64:1294–1298. doi: 10.1136/ard.2004.030924

- Cluster RCT: monitoring DAS28 (12 centers) vs routine care (12 centers)
- DAS28 assessment by research nurse at 0 & 24 weeks
- Primary outcomes
  - % patients with DAS28 <3.2 (subgroup analysis due to organizational issues)
  - Changes in DMARD treatment (all patients)

Treatment principles in RA

Effectiveness of systematic monitoring of rheumatoid arthritis disease activity in daily practice: a multicentre, cluster randomised controlled trial

J Fransen, H Bernelot Moens, I Speyer, P L C M van Riel

Ann Rheum Dis 2005;64:1294–1298. doi: 10.1136/ard.2004.030924

- Intervention centers
  - DAS28 measurement for clinical use at 0,4,12 & 24 weeks by treating rheumatologist
  - Study advice: change medication if DAS28 >3.2
- Control centers
  - Visit at week 0,4,12 & 24
  - No systematic monitoring or treatment advices

Treatment principles in RA

Effectiveness of systematic monitoring of rheumatoid arthritis disease activity in daily practice: a multicentre, cluster randomised controlled trial

J Fransen, H Bernelot Moens, I Speyer, P L C M van Riel

Ann Rheum Dis 2005;64:1294–1298. doi: 10.1136/ard.2004.030924

- Baseline data
  - ± 70% women; 58 years; ± 80% RF positivity; disease duration 6 years
  - DAS28 4.5; 13% low disease activity

Of note, DAS28 only measured by research nurse in 142 patients (61 intervention group; 81 control group)

Treatment principles in RA

Effectiveness of systematic monitoring of rheumatoid arthritis disease activity in daily practice: a multicentre, cluster randomised controlled trial

J Fransen, H Bernelot Moens, I Speyer, P L C M van Riel

Ann Rheum Dis 2005;64:1294–1298. doi: 10.1136/ard.2004.030924

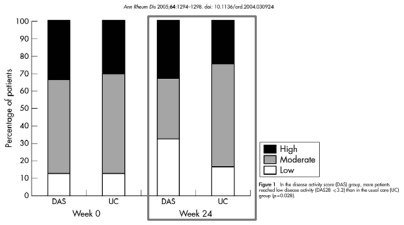


Figure 1 In the disease activity score (DAS) group, more patients reached low disease activity (DAS28 <3.2) than in the usual care (UC) group (p<0.05).

Treatment principles in RA

Effectiveness of systematic monitoring of rheumatoid arthritis disease activity in daily practice: a multicentre, cluster randomised controlled trial

J Fransen, H Bernelot Moens, I Speyer, P L C M van Riel

Ann Rheum Dis 2005;64:1294–1298. doi: 10.1136/ard.2004.030924

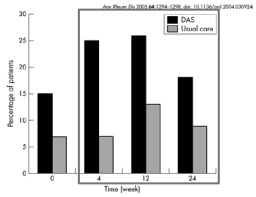


Figure 2 In the disease activity score (DAS) group, more changes in disease modifying antirheumatic drug (DMARD) treatment occurred during the course of the study (p=0.013).

Treatment principles in RA

Effectiveness of systematic monitoring of rheumatoid arthritis disease activity in daily practice: a multicentre, cluster randomised controlled trial

J Fransen, H Bernelot Moens, I Speyer, P L C M van Riel

Ann Rheum Dis 2005;64:1294–1298. doi: 10.1136/ard.2004.030924

- No significant differences in mean MTX, sasp and prednisone dose
- No differences in side-effects
- Intervention group
  - DAS28 measured in 99% of the visits
  - 98% of patients in which medication was changed had a DAS28 >3.2
  - 20% of patients with a DAS28 >3.2 had their medication changed

Conclusion: standard monitoring of disease activity in daily practice, can lead to more DMARD-changes compared to usual care

Treatment principles in RA

Meta-analysis of tight control strategies in rheumatoid arthritis: protocolized treatment has additional value with respect to the clinical outcome

Lydia G. Schipper[1], Laura T. C. van Hulst[1], Richard Grol[2], Piet L. C. M. van Riel[1], Marlies E. J. L. Hulscher[2] and Jaap Fransen[1]

Rheumatology 2010;49:2154–2164

- PubMed & Cochrane library 1995 – 2009: monitoring of disease activity combined with treatment protocols vs monitoring alone
- Inclusion: studies on routine care vs tight control
- Primary outcome
  - Mean change in DAS28 (year 0 vs 1)

Treatment principles in RA

Meta-analysis of tight control strategies in rheumatoid arthritis: protocolized treatment has additional value with respect to the clinical outcome

Lydia G. Schipper[1], Laura T. C. van Hulst[1], Richard Grol[2], Piet L. C. M. van Riel[1], Marlies E. J. L. Hulscher[2] and Jaap Fransen[1]

Rheumatology 2010;49:2154–2164

- Included studies (n= 6)
  - 4 RCT, 2 CCT; study duration between 12 and 24 months
  - 3 studies monitoring + treatment protocol, 3 studies monitoring alone
  - 2 studies in early, DMARD-naive RA; others in early & late RA
- Baseline data
  - 110 to 435 pt per study; 60% to 70% female; 42% to 80% RF positive; DAS28 >3.2

Treatment principles in RA

Meta-analysis of tight control strategies in rheumatoid arthritis: protocolized treatment has additional value with respect to the clinical outcome

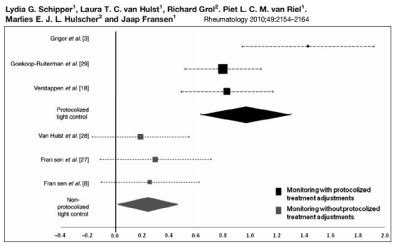
Lydia G. Schipper[1], Laura T. C. van Hulst[1], Richard Grol[2]. Piet L. C. M. van Riel[1]. Marlies E. J. L. Hulscher[2] and Jaap Fransen[1] Rheumatology 2010;49:2154-2164

- Results clinical effectiveness tight control
  - In 5 studies tight control better than routine care
  - More medication changes, better physical functioning and less radiographic damage with tight control
  - Toxicity similar
- Results meta-analysis
  - Tight control vs usual care: tight control is more effective, 0.6 DAS28-point more decrease in DAS28
  - Within tight control studies, monitoring + protocol is more effective than monitoring alone: 0.66 DAS28-point more decrease in DAS28

Treatment principles in RA

Meta-analysis of tight control strategies in rheumatoid arthritis: protocolized treatment has additional value with respect to the clinical outcome

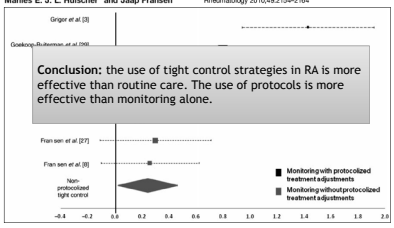
Lydia G. Schipper[1], Laura T. C. van Hulst[1], Richard Grol[2]. Piet L. C. M. van Riel[1]. Marlies E. J. L. Hulscher[2] and Jaap Fransen[1] Rheumatology 2010;49:2154-2164



Treatment principles in RA

Meta-analysis of tight control strategies in rheumatoid arthritis: protocolized treatment has additional value with respect to the clinical outcome

Lydia G. Schipper[1], Laura T. C. van Hulst[1], Richard Grol[2]. Piet L. C. M. van Riel[1]. Marlies E. J. L. Hulscher[2] and Jaap Fransen[1] Rheumatology 2010;49:2154-2164



Treatment principles in RA

"Yes, but these are clinical trials. The effect in daily practice has not been proven."

Treatment principles in RA

"Yes, but these are clinical trials. The effect in daily practice has not been proven."

Effect of adherence to European treatment recommendations on early arthritis outcome: data from the ESPOIR cohort

Cécile Escalas, Marie Dalichampt, Bernard Combe, et al.
Ann Rheum Dis 2012 71: 1803-1808

Influence of guideline adherence on outcome in a randomised controlled trial on the efficacy of methotrexate with folate supplementation in rheumatoid arthritis

J Fransen, R F J M Laan, M A F J van der Laar, T W J Huizinga, P L C M van Riel

Ann Rheum Dis 2004;63:1222–1226. doi: 10.1136/ard.2003.018861

Treatment principles in RA

"Yes, but these are clinical trials. The effect in daily practice has not been proven."

ESPOIR study: less radiographic damage after 1 year and functional deterioration after 2 years (early RA).

MTX study: larger decrease in DAS28 after 48 weeks (established RA)

Influence of guideline adherence on outcome in a randomised controlled trial on the efficacy of methotrexate with folate supplementation in rheumatoid arthritis

J Fransen, R F J M Laan, M A F J van der Laar, T W J Huizinga, P L C M van Riel

Ann Rheum Dis 2004;63:1222–1226. doi: 10.1136/ard.2003.018861

Treatment principles in RA

Summary:

Tight control gives better results and guideline adherence makes a difference to patients. However…

Treatment principles in RA

Cohort	Recommendation	Guideline adherence
ESPOIR (2002-2005)	First DMARD in early RA (EULAR)	54%
ERAN (2002-2007)	DMARD in early RA	Median time to DMARD start: 8 months (97% DMARD; 67% after 3 years DAS28 >3,2)
North-America (2002-2009)	DMARD in case of active RA (ACR)	25-50% (publication of updated ACR guideline no difference)
DREAM remission induction (2006)	DMARD in case of active RA	70% (98% DAS28 available)

Treatment principles in RA

Cohort	Recommendation	Guideline adherence
ESPOIR (2002-20		
ERAN (2002-200		e to DMARD nths 7% after 3 years
North-America (2002-2009)	DMARD in case of active RA (ACR)	25-50% (publication of updated ACR guideline no difference)
DREAM remission induction (2006)	DMARD in case of active RA	70% (98% DAS28 available)

Guideline adherence is not always optimal…

But how are things in our hospital?

Content

- Background
- Treatment principles in RA
- **Current situation at our hospital**
- Summary

Current situation at our hospital

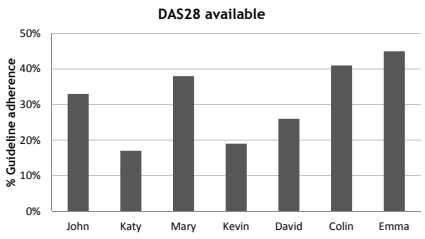
- Local RA treatment guideline available
  - Based on tight control: measuring DAS28, target based on disease duration, changing treatment if target is not reached, adequate follow-up
- Also in the guideline:
  - Shared care (nurses, PA)
  - Monitoring of functional damage
  - Other treatment modalities (physical therapy)
  - Risk management

Current situation at our hospital

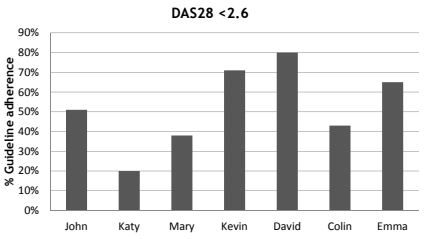
- Other guidelines on related themes also available
  - Preferential order of c/bDMARDs and NSAIDs
  - bDMARD dose optimization

Optimal RA care is more than tight control alone

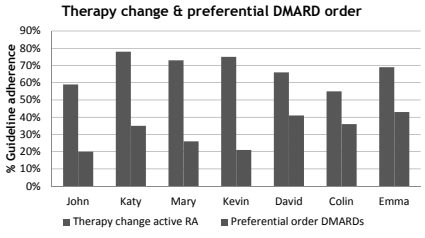
Current situation at our hospital



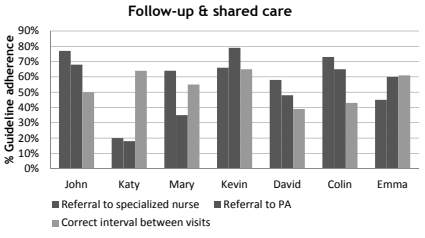
Current situation at our hospital



Current situation at our hospital



Current situation at our hospital



Current situation at our hospital

- Guideline adherence not always optimal
- Guideline adherence does not need to be 100%: deviations are allowed, but explain them!

Content

- Background
- Treatment principles in RA
- Current situation at our hospital
- **Summary**

Summary

- Treatment principles in RA
  - Tight control: important strategy that benefits our patients
  - Application in daily practice not yet optimal, but seems feasible
- RA treatment in this hospital
  - Improvement possible on many indicators

SUPPLEMENT 3: DEVELOPMENT OF GUIDELINE INDICATORS

Knowing how to describe quality of care is a prerequisite for its measurement. Often quality indicators are used to assess quality of care. A quality indicator is 'a measurable element of practice performance for which there is evidence or consensus that it can be used to assess the quality, and hence change the quality of care provided'.(37) Quality indicators are often grouped into structure-, process- and outcome indicators. Outcome indicators reflect the result of the care that was provided by the healthcare provider, while process indicators reflect the actual care given to patients ('what is done'). Structure indicators, on the other hand, describe organizational aspects ('what is available').(38)

Different groups and organizations have developed indicator sets and for the purpose of this study we used the process indicators covering monitoring, drug treatment, follow-up and documentation from the Dutch national RA guideline (CBO indicators).(6) The Dutch set is selected because it bests reflects care provided at the study center. Structure indicators are not taken into account because this will be a single center study and as a consequence all structure indicators will be the same for all clinicians. The CBO indicators are not always very specific, therefore some indicators are modified as to better reflect the recommendations from the local RA guideline at the study center. An overview of all indicators is given in table 1.

Besides adaptations of existing CBO indicators, extra indicators were added to the set used in our study. This concerned two indicators in the follow-up & referral domain (PA referral and nurse-led DAS28 assessments) and the two indicators in the domain on administration (new and control patient letters to the general physician).

The indicators on PA referral and nurse-led DAS28 assessments were added to cover specific shared care practices at the study center. In this center all RA patients should be seen by a nurse prior to the visit with the rheumatologist (nurse led DAS28 assessment). During this visit the DAS28 is done by a specialized rheumatology nurse, together with routine laboratory tests and assessment of current medication use. All information is provided to the rheumatologist. Furthermore RA patients can be treated by both a physician assistant (PA) and rheumatologist, with alternating visits between them. The PA is allowed to prescribe rheumatologic medication and make treatment decisions, but the final responsibility always lies with the rheumatologist.

Finally, a new group of indicators concerning administration was added. In the Dutch RA guideline no recommendations were given on correspondence with other relevant clinicians, especially the general physician (GP). We chose to add these indicators as it is of crucial importance that the GP knows if a patient uses DMARD or biological therapy because of potentially severe side effects or interactions with other commonly prescribed medication.

Table 1 *Indicators in national Dutch guideline and study*

Group of indicators	Indicators	
	Dutch national guideline (CBO)(6)	**Study**
Treatment & monitoring	*Monitoring of disease activity*	
	DAS28 measurement done at every visit by either the rheumatologist or another health care professional (HCP)	DAS28 measurement done at every outpatient clinic visit by either the rheumatologist or another health care professional
	Monitoring of structural damage	
	X-rays of hand and feet at moment of diagnosis and one year thereafter (year 0 and 1)	X-rays of hands, feet and thorax, done at the year 0,1 and 3.
	X-rays of hand and feet done after a period of high disease activity (DAS28 >3.2 at two consecutive visits), if not done in the year before	
	Yearly assessment of functional status (Health Assessment Questionnaire) by either the rheumatologist of another health care professional	Yearly assessment of functional status (Health Assessment Questionnaire) by either the rheumatologist of another health care professional
	Treatment	
	Use of biological therapy	Prescription of conventional and biological DMARDs according to the preferential order* when initiating a new DMARD
	Use of methotrexate >20 mg/week	
	Use of prednisone >5 mg/day	N/A
	Intensification of medication by a rheumatologist in case of a DAS28>3.2 and an adequate period of previous therapy	Change in therapy† in case of active disease based on DAS28 score
	Adaptation of treatment based on DAS28 scores unless co-morbidity, extra-articular disease and/or side-effects prevent this	
	N/A	Use or prescription of a concomitant conventional DMARD in case of biological use
	N/A	Biological dose optimization in case of low disease activity and stable biological use for the previous six months

Follow-up & referral	*Shared care*	
	Consultation with a specialized rheumatology nurse within one year after diagnosis	Referral to a specialized rheumatology nurse within two weeks after diagnosis
	N/A	Planned nurse led DAS28 assessment during the next regular outpatient clinic visit‡
	N/A	Referral to a physician assistant
	Follow-up	
	Planned visit with a rheumatologist within 3 months of the last visit if DAS28 > 2.6	Correct intervals between consecutive visits, based on disease activity and medication use as stated in the local RA guideline¶
	Planned visit with a rheumatologist or other relevant HCP within 6 months of the last visit if DAS28 < 2.6	
	Planned visit with a rheumatologist within one year after the last visit if DAS28 < 2.6	
Administration	*Correspondence*	
	N/A	A letter to the general physician, send within two weeks after diagnosis in case of a new RA patient (new patient letter)
		A letter to the general physician, send once every 18 months (control patient letter)

DAS28: Disease Activity Score based on 28 joints; DMARD: Disease Modifying Anti-Rheumatic Drug
*Preferred order in which DMARDs and biological should be prescribed, stated in the local RA guideline at the study center. †The following options are seen as changes in therapy: intensifying DMARD therapy (dose increase, adding a new DMARD, switching to another DMARD and/or biological), starting or increasing prednisone dose and/or local corticosteroid injections. ‡Nurse led center were RA patients are seen prior to the visit with the rheumatologist or physician assistant (PA) with measurement of the DAS28, routine laboratory assessments and asking about current medication use. All information is entered in the electronic patient record and therefore directly available to the rheumatologist or PA who's sees the patient next.
¶Correct follow-up visit schedule is as follows: week 0 and 6; every 3 months during the first year of DMARD/biological therapy and once every six months thereafter. In case of DAS28 >3.2 a visit every 3 months, independent of medication use.


TREAT
TO
TARGET
Don't move... I'm sure I will
hit bull's eye this time!

10

Implementation of protocolized tight control and biological dose optimization in daily clinical practice: results of a pilot study

Nienke Lesuis
Lise M Verhoef
Lieke M Nieboer
George A Bruyn
Paul Baudoin
Ronald F van Vollenhoven
Marlies EJL Hulscher
Frank HJ van den Hoogen
Alfons A den Broeder

Accepted for publication in the Scandinavian Journal of Rheumatology

ABSTRACT

Objectives

Assess the effects of education, guideline development and individualised treatment advices on rheumatologist adherence to tight control based treatment and biological dose optimization in rheumatoid arthritis (RA), psoriatic arthritis (PsA) and spondyloarthropathy (SpA) patients.

Methods

This pilot study, among two rheumatologists and two specialized nurses in a general hospital, combined education, feedback, local guideline development and individualized treatment advices. Outcomes (baseline and 1 year post-intervention) were the percentage of patients with a Disease Activity Score in 28 joints (DAS28) or Bath Ankylosing Spondylitis Disease Activity Index (BASDAI) measured during the visit, mean DAS28/ BASDAI and the percentage of patients using a reduced biological dose. DAS28 outcomes only apply to RA and PsA patients, BASDAI outcomes only to SpA patients, and outcomes on biological dose apply to all patients.

Results

232 patients were included (67% RA, 58% female, mean age 56 ± 15 years). The percentage of DAS28 and BASDAI measurements performed increased after the intervention (DAS28: 15% to 51%, OR 3.3, 95% CI 2.1 to 5.5; BASDAI 23% to 50%, OR 2.2, 95% CI 1.0 to 5.5), with mean DAS28 and BASDAI scores remaining similar (DAS28: mean difference 0.1, 95% CI -0.3 to 0.5; BASDAI: mean difference 0.03, 95% CI -1.8 to 1.9). Use of reduced biological dose increased from 10% to 61% (OR 3.9, 95% CI 2.4 to 6.5).

Conclusions

A multi-component intervention strategy aimed at rheumatologists can lead to improved adherence to tight control based treatment and a reduction in biological use in RA, SpA and PsA patients.

INTRODUCTION

Treatment of rheumatoid arthritis (RA) is based on tight control principles: setting a target, frequent assessment of disease activity and a structured protocol to make treatment changes. This strategy leads to lower disease activity and less functional damage compared to usual care.[1, 2] Unfortunately, the dissemination of tight control based guidelines does not seem to influence the daily practice of rheumatologists enough.[3] For example, treatment is not always changed on time in case of active disease or patients are not receiving correct Disease Modifying Anti-Rheumatic Drugs (DMARDs).[3]

According to the 2014 treat-to-target recommendations "maintenance of the treatment target does not in itself imply maintenance of treatment"; this refers towards dose reduction or stopping of biological DMARDs (bDMARDs; bDMARD dose optimization).[4] In recent years, several studies have shown that this strategy can be successful in patients while preserving low disease activity.[5, 6] The high costs and dose-related side effects of bDMARDs make dose optimization a desirable goal for implementation in daily practice. Despite this, actual adherence seems to be not optimal.[7]

So far tight control based treatment and biological dose optimization have mainly been studied in RA. However, slowly evidence is gained that for PsA and SpA the same principles might apply [8, 9] and treat to target recommendations have been recently published.[10] Combing all the existing evidence, we conducted a pilot study to improve RA, PsA and SpA tight control treatment and bDMARD optimization using a multi-component intervention strategy.

METHODS

This pilot study was conducted in a general hospital in the Netherlands with two rheumatologists and two specialized nurses between May and October 2014. Although the intervention was aimed at clinicians, outcomes were measured in patients. All adult patients with an ICD-9 code of RA (714.0), psoriatic arthritis (PsA; 696) and axial spondyloarthropathy (SpA; 720), using a bDMARD at study start and having visited their rheumatologist during the pre- and post-intervention period were eligible for inclusion.

The intervention strategy consisted of: 1) an educational meeting combined with feedback and local guideline development (bDMARD dose optimization and tight control based treatment of RA, PsA and SpA), 2) individualized treatment advices in all bDMARDs users, written in their electronic health record (EHR), and 3) feedback after three and six months. An example of the PowerPoint slides used during the educational meeting can be found in supplement 1. This strategy was developed and provided by a rheumatologist-

epidemiologist, a rheumatology PhD student and an administrative assistant (AdB, NL and LN) from the Sint Maartenskliniek, a specialized rheumatology clinic in the Netherlands, with experience in using tight control based guidelines and dose optimization. The choice of the different steps was based on Cochrane reviews on effective interventions [11, 12] and previous experience of the authors. The different steps of the intervention took place between May and October 2014.

The outcome measures used in this study are stated in table 1. All outcomes were compared between the pre- and post-intervention period. As the intervention took place between May and October 2014, the patients' visits most closely situated before and after this time period were used as pre and post-intervention visits respectively. For all patients a single visit per period was used for data collection (data as recorded in the EHR).

Table 1 *Outcome measures*

Outcome measure	Patient population
Percentage of patients with a disease activity measure	
DAS28	RA and PsA patients
BASDAI	SpA patients
Mean score of the disease activity measures	
DAS28	RA and PsA patients
BASDAI	SpA patients
Percentage of patients using a reduced dose of their bDMARD	All patients using a bDMARD
Percentage of patients using a concomitant cDMARD	All RA patients using a bDMARD

DAS28, Disease Activity Score in 28 joints; BASDAI, Bath Ankylosing Spondylitis Disease Activity Index; bDMARD, biological Disease Modifying Anti-Rheumatic Drug; cDMARD, conventional Disease Modifying Anti-Rheumatic Drug.

The study hospital approved this study; the rheumatologists were informed beforehand about this study and asked if they would participate. As this was a quality assessment performed in the hospital where two authors of this study worked, no written informed consent was asked from the patients. In addition, data collection was done within the study hospital and directly afterwards all patient data was anonymized.

Depending on the type of variable, descriptive statistics are presented as percentages with the accompanying absolute numbers or as means with standard deviations. Outcome comparison between the two time periods was done using appropriate statistics (t-test or McNemar; two-sided, $\alpha = 0.05$). As bDMARD dose optimization, according to the local guideline, should only be done in patients with low disease activity (DAS28 <3.2/BASDAI <4 or, if not available, the judgement of the rheumatologist) and bDMARD use of minimally 6 months, a post-hoc sensitivity analysis was done on the percentage of patients using a reduced bDMARD dose, only including patients fulfilling both criteria. All analyses were done using STATA version 13.

RESULTS

All rheumatologists and nurses (n= 4) participated in this study and in all intervention steps. At study start, 258 RA, SpA and PsA patients were using a bDMARD. Of those patients, 24 were lost to follow-up (n= 8: moving to another city; n= 16: not visited the study clinic during the post-intervention period). The remaining 232 patients were included in the final analysis (table 2).

Table 2 *Description of the patient population*

Variable	Result (n= 232)
Females, % (n)	58% (131)
Age, in years	56 ± 16
Disease duration, in years	9 ± 8
Diagnosis, % (n)	
RA	67% (153)
PsA	15% (34)
SpA	18% (40)
Type of bDMARD, % (n)	
Adalimumab	42% (98)
Etanercept	22% (50)
Tocilizumab	13% (30)
Other	23% (54)

After the intervention more disease activity measurements (DAS28 and BASDAI) were done and more patients used a concomitant DMARD (table 3). Similarly, after the intervention more patients used a reduced biological dose while disease remained stable (table 3 and supplemental file 2).

Of note, after the intervention 20 of the 232 included patients (9%) stopped their bDMARD. Nine of them stopped for other reasons than dose optimization (pregnancy, infection, ineffectiveness), leaving 11 patients who successfully stopped their bDMARD after dose optimization (5%).

Finally, the sensitivity analysis on reduced bDMARD use, including only patients with data on disease activity and duration of bDMARD use (71% available), yielded no different results from the primary analysis.

Table 3 *Outcomes on DAS28, BASDAI and bDMARD use*

Outcome	Pre-intervention *(n= 232)*	Post-intervention *(n= 232)*	Mean difference (MD) or odds ratio (OR) (95% confidence interval)	P-value
DAS28 performed*, % (n)	15% (29)	51% (97)	OR 3.3 (2.1 to 5.2)	<0.01
BASDAI performed‡, % (n)	23% (9)	50% (20)	OR 2.2 (1.0 to 5.5)	0.04
Mean DAS28*	2.2 ± 0.9	2.1 ± 0.9	MD 0.1 (-0.3 to 0.5)	0.51
Mean BASDAI‡	4.4 ± 2.5	4.3 ± 2.1	MD 0.03 (-1.8 to 1.9)	0.97
Patients using a reduced bDMARD dose†, % (n)	10% (21)	61% (124)	OR 3.9 (2.4 to 6.5)	<0.01
Patients using a concomitant cDMARD¶, % (n)	42% (63)	52% (72)	OR 1.1 (0.7 to 1.5)	0.79

*Outcome only assessed in RA and PsA patients. ‡Outcome only assessed in SpA patients. †Outcome assessed in all bDMARD users. ¶Outcome only assessed in RA patients using a bDMARD.

DISCUSSION

To our knowledge this is one of the first studies within rheumatology describing an improvement strategy on tight control based bDMARD dose optimization combining education, feedback and individualized treatment advices. Our results suggest that implementation of relatively new treatment principles in daily practice is feasible, resulting in increased adherence to tight control based treatment and a sizable reduction in bDMARD use.

The main strengths of this study are the short time between the publication of positive trial results on bDMARD dose reduction and the conduct of this pilot study, the combined focus on tight control and dose optimization (with tight control being a necessary prerequisite for safe and patient friendly tapering), and the inclusion of RA, PsA and SpA patients to aid generalizability. On the other hand, the main limitation of our study is its small scale, stressing the need to replicate our trial in a larger sample of rheumatologists. Also, due to our uncontrolled study design we are not able to comment on which part of our strategy was most effective or to infer a definite causal relation between the intervention strategy and the results afterwards as other events in the same time period might have attributed to the observed results. However, we are not aware of any external factors during the study that could have influenced our results.

Within rheumatology not many comparable intervention studies on tight control implementation exist. However, one Canadian study also used education and feedback to improve daily practice.[13] In this study, education and feedback resulted in more disease activity measures being collected by the rheumatologists (DAS28 measurement from 43% to 57%). These results are somewhat in line with our study, although the increase of DAS28 measurements was higher in our study and mean DAS28 lower (2.1 versus 3.05).[13]

Interestingly, with our intervention strategy we were able to replicate the results from the only two randomized controlled trials on bDMARD dose optimization (DRESS and STRASS study).[5, 14] For example, in the DRESS study 43% of the RA patient could taper their adalimumab or etanercept dose and 20% could stop their bDMARD.[5] In our pilot study even more patients used a reduced bDMARD dose (61%), however fewer patients completely stopped their bDMARD (9%). This might be explained by a shorter follow up in our study (12 versus 18 months) and inclusion of SpA patients, in whom stopping is probably less successful.[15] Nevertheless, our study shows that replication of trial results in daily practice is possible, if, however enough attention is given to optimal implementation of the required changes.

In other settings the separate components of our strategy have shown to be effective before [11, 12], but we cannot discriminate between the effects of the different components of our intervention strategy. In order to gain some insight in this topic, a short interview with the participating rheumatologists and nurses was done after the study. During this evaluation it was suggested that the individual treatment advices in the EHR of included patients were of crucial importance because they acted as a reminder. In addition, the educational session and development of local guidelines were seen as necessary prerequisites to change behavior. Finally, the feedback acted as a trigger to improve their practice and the close contact with the research team was positively evaluated.

Despite the use of our strategy in only one centre, we feel that our study has important practical implications as it shows that implementation of tight control and bDMARD dose optimization in daily practice is feasible. The enthusiasm of the rheumatologists and nurses at the study centre has strengthened our view that rheumatologists are able to apply new treatment strategies if they are assisted in doing so. In our opinion, this study stresses the fact that implementation research is of crucial importance for the field of rheumatology in order to bridge the gap between theory and practice. Therefore, we are planning a randomized controlled trial to assess the effectiveness of our intervention strategy in a multi-centre study aimed at tight control based bDMARD dose optimization (Rheumatoid Arthritis ImplemeNtation of Biological dose Optimization in real World; RAINBOW).

ACKNOWLEGDEMENTS

We would like to thank Marian de Waal and Ria de Kort for their enthusiastic participation during the implementation of all changes, hospitality and invaluable practical help with the data collection.

FUNDING

This study received no external funding.

CONTRIBUTORS

All authors contributed to the study design. NL and AdB designed and executed the intervention. NL, LV and LN participated in data collection. NL, AdB and LV participated in the data analysis. NL and AdB interpreted the data, with help of GB, PB and MH. The manuscript was written by NL under the supervision of AdB. All authors have revised the draft version of the manuscript, and read and approved the final version of the manuscript. NL is guarantor.

COMPETING INTEREST DECLARATION

All authors have completed the Unified Competing Interest form at www.icmje.org/coi_disclosure.pdf (available on request from the corresponding author) and declare that RvV has received grants and personal fees from AbbVie, BMS, GSK, Pfizer, Roche, UCB, Biotest, Crescendo, Janssen, Lilly, Merck and Vertex outside the submitted work; all other authors reported no competing interests.

DATA SHARING STATEMENT

Patient level data, full dataset, technical appendix and statistical code are available at a reasonable request from the corresponding author. Consent from the patients was not obtained but the presented data are anonymized and risk of identification is low.

REFERENCES

(1) Schipper LG, van Hulst LT, Grol R, van Riel PL, Hulscher ME, Fransen J. Meta-analysis of tight control strategies in rheumatoid arthritis: protocolized treatment has additional value with respect to the clinical outcome. Rheumatology (Oxford) 2010;49:2154-64.

(2) Escalas C, Dalichampt M, Combe B, Fautrel B, Guillemin F, Durieux P, et al. Effect of adherence to European treatment recommendations on early arthritis outcome: data from the ESPOIR cohort. Ann Rheum Dis 2012;71:1803-8.

(3) Harrold LR, Harrington JT, Curtis JR, Furst DE, Bentley MJ, Shan Y, et al. Prescribing practices in a US cohort of rheumatoid arthritis patients before and after publication of the American College of Rheumatology treatment recommendations. Arthritis Rheum 2012;64:630-8.

(4) Smolen JS, Breedveld FC, Burmester GR, Bykerk V, Dougados M, Emery P, et al. Treating rheumatoid arthritis to target: 2014 update of the recommendations of an international task force. Ann Rheum Dis 2015;0:1-13.

(5) van Herwaarden N, van der Maas A, Minten MJ, van den Hoogen FH, Kievit W, van Vollenhoven RF, et al. Disease activity guided dose reduction and withdrawal of adalimumab or etanercept compared with usual care in rheumatoid arthritis: open label, randomised controlled, non-inferiority trial. BMJ 2015;350:h1389.

(6) van Herwaarden N, Den Broeder AA, Jacobs W, van der Maas A, Bijlsma JW, van Vollenhoven RF, et al. Down-titration and discontinuation strategies of tumor necrosis factor-blocking agents for rheumatoid arthritis in patients with low disease activity. Cochrane Database Syst Rev 2014;9:CD010455.

(7) Zhang J, Xie F, Delzell E, Yun H, Lewis JD, Haynes K, et al. Impact of biologic agents with and without concomitant methotrexate and at reduced doses in older rheumatoid arthritis patients. Arthritis Care Res (Hoboken) 2015;67:624-32.

(8) Schoels MM, Braun J, Dougados M, Emery P, Fitzgerald O, Kavanaugh A, et al. Treating axial and peripheral spondyloarthritis, including psoriatic arthritis, to target: results of a systematic literature search to support an international treat-to-target recommendation in spondyloarthritis. Ann Rheum Dis 2014;73:238-42.

(9) Coates LC, Moverley AR, McParland L, Brown S, Navarro-Coy N, O'Dwyer JL, et al. Effect of tight control of inflammation in early psoriatic arthritis (TICOPA): a UK multicentre, open-label, randomised controlled trial. Lancet. Published online 30th September 2015. doi: 10.1016/S0140-6736(15)00347-5.

(10) Smolen JS, Braun J, Dougados M, Emery P, Fitzgerald O, Helliwell P, et al. Treating spondyloarthritis, including ankylosing spondylitis and psoriatic arthritis, to target: recommendations of an international task force. Ann Rheum Dis 2014;73:6-16.

(11) Pope J, Thorne C, Cividino A, Lucas K. Effect of rheumatologist education on systematic measurements and treatment decisions in rheumatoid arthritis: the metrix study. J Rheumatol 2012;39:2247-52.

(12) Fautrel B, Pham T, Alfaiate T, Gandjbakhch F, Foltz V, Morel J, et al. Step-down strategy of spacing TNF-blocker injections for established rheumatoid arthritis in remission: results of the multicentre non-inferiority randomised open-label controlled trial (STRASS: Spacing of TNF-blocker injections in Rheumatoid ArthritiS Study). Ann Rheum Dis. Published online 23th June 2015. doi: 10.1136/annrheumdis-2014-206696.

(13) Song IH, Haibel H, Poddubnyy D, Braun J, Sieper J. Withdrawal of biologic therapy in axial spondyloarthritis: the experience in early disease. Clin Exp Rheumatol 2013;31:S37-S42.

(14) Ivers N, Jamtvedt G, Flottorp S, Young JM, Odgaard-Jensen J, French SD, et al. Audit and feedback: effects on professional practice and healthcare outcomes. Cochrane Database Syst Rev 2012;6:CD000259.

(15) O'Brien MA, Rogers S, Jamtvedt G, Oxman AD, Odgaard-Jensen J, Kristoffersen DT, et al. Educational outreach visits: effects on professional practice and health care outcomes. Cochrane Database Syst Rev 2007;CD000409.

SUPPLEMENTAL FILE 1: POWERPOINT SLIDES USED DURING THE EDUCATIONAL MEETING

Supplement 1: PowerPoint slides used during the educational meetings

Treatment principles in rheumatoid arthritis

A translation of the PowerPoint slides used during the first educational meeting

Authors: Nienke Lesuis (resident of rheumatology; MD) and Alfons A den Broeder (rheumatologist-epidemiologist; MD, PhD)
Sint Maartenskliniek, Nijmegen, the Netherlands

Goals

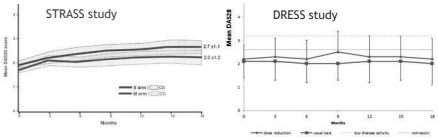
To optimize biological use of rheumatologists in this hospital

What do we need for rational use of biologicals?

Goals

To optimize biological use of rheumatologists in this hospital

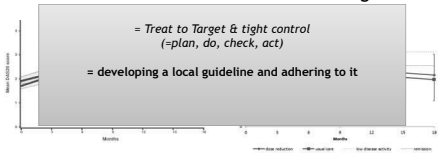
What do we need for rational use of biologicals?



Goals

To optimize biological use of rheumatologists in this hospital

What do we need for rational use of biologicals?



Content

- Background
- Treatment principles in RA
- Current situation at this hospital
- Summary
- How to continue?

Content

- **Background**
- Treatment principles in RA
- Current situation at this hospital
- Summary
- How to continue?

Background

We would rather prevent than treat these hands, but… treatment guidelines are not always followed optimally

Content

- Background
- **Treatment principles in RA**
- Current situation at our hospital
- Summary
- How to continue?

Treatment principles in RA

- Many publications on effective RA treatments
- 'Treat to target', 'tight control' & 'hit hard, hit early' often mentioned

→ Cornerstone of ACR & EULAR RA guidelines

Treatment principles in RA

Tight control: 'frequent assessment of disease activity combined with an objective structured protocol to make treatment changes that maintain low disease activity or remission at an agreed target' *(Kiely et al., Rheumatology 2009)*

Treat to target: 'the treatment aim was defined as remission with low disease activity being an alternative in patients with long-standing disease. Regular follow-up with appropriate therapeutic adaptation to reach the desired state within 3-6 months. Follow-up examinations ought to employ composite measures of disease activity which include joint counts. *(Grigor et al., Ann Rheum Dis 2010)*

Hit hard, hit early: 'early institution of DMARDs.' 'Window of opportunity'

Treatment principles in RA

Tight control: 'frequent assessment of disease activity combined with an objective structured protocol to make treatment changes that maintain low disease activity or remission at an agreed target' *(Kiely et al., Rheumatology 2009)*

Treat to target: 'the treatment aim was defined as remission with low disease activity being an alternative in patients with long-standing disease. Regular follow-up with appropriate therapeutic adaptation to reach the desired state within 3-6 months. Follow-up examinations ought to employ composite measures of disease activity which include joint counts. *(Grigor et al., Ann Rheum Dis 2010)*

Hit hard, hit early: 'early institution of DMARDs.' 'Window of opportunity'

Treatment principles in RA

"Yes, I agree with the general principle but…."

'*A number says nothing! I always look and listen to the patient, order an ESR and feel the joints*'

'*Patients don't want all those pills*'

'*My judgement is better than a composite measure, these numbers mean nothing to a patient*'

Treatment principles in RA

"Yes, I agree with the general principle but…."

DAS28 = 0.56 * sqrt(tender28) + 0.28 * sqrt(swollen28) + 0.70 * ln(ESR) + 0.014 * VAS *(http://www.das-score.nl/das28/nl/uitleg-das28/de-das28-score.html)*

In case of active disease, patients just want to get better. They only get worried about the number of pills later on. *(van Tuyl, Rheumatology 2008)*

A DAS28 <3.2 is associated with 50% less progression of radiographic damage and functional status (HAQ) is influenced by both active inflammation and radiographic damage. *(Fransen et al., Ann Rheum Dis 2005)*

Treatment principles in RA

Effect of a treatment strategy of tight control for rheumatoid arthritis (the TICORA study): a single-blind randomised controlled trial

Catriona Grigor, Hilary Capell, Anne Stirling, Alex D McMahon, Peter I Lock, Romnay Vallance, Wilma Kincaid, Duncan Porter *Lancet 2004; 364: 263–69*

- RCT: routine vs intensive management
- Inclusion criteria: diagnosis of RA <5 years; DAS >2.4
- Study assessment 1x/3 months (DAS etc) by blinded assessor
- Primary endpoints:
  - Mean fall in disease activity
  - % patients with an EULAR good response

Treatment principles in RA

Effect of a treatment strategy of tight control for rheumatoid arthritis (the TICORA study): a single-blind randomised controlled trial

Catriona Grigor, Hilary Capell, Anne Stirling, Alex D McMahon, Peter I Lock, Romnay Vallance, Wilma Kincaid, Duncan Porter *Lancet 2004; 364: 263–69*

- Intensive management:
  - Monthly visits including DAS measurement
  - Corticosteroid injection of any swollen joint or depomedrol 120mg i.m. if DAS >2.4
  - Strict medication protocol: dose increase or cDMARD switch every 1-3 month if DAS >2.4
- Routine care:
  - 3-montly visits
  - No routine DAS measurement or strict medication protocol

Treatment principles in RA

Effect of a treatment strategy of tight control for rheumatoid arthritis (the TICORA study): a single-blind randomised controlled trial

Catriona Grigor, Hilary Capell, Anne Stirling, Alex D McMahon, Peter Lock, Ramsay Wallace, Wilma Kincaid, Duncan Porter Lancet 2004; 364: 263–69

	Intensive group (n=55)	Routine group (n=55)	Odds ratio (95% CI)	p*
EULAR good response	45 (82%)	24 (44%)	5·8 (2·4–13·9)*	<0·0001
EULAR remission	36 (65%)	9 (16%)	9·7 (3·9–23·9)*	<0·0001
ACR 20 response	50 (91%)	35 (64%)	5·7 (1·9–16·7)*	<0·0001
ACR 50 response	46 (84%)	22 (40%)	6·1 (2·5–14·9)*	<0·0001
ACR 70 response	39 (71%)	10 (18%)	11 (4·5–27)*	<0·0001

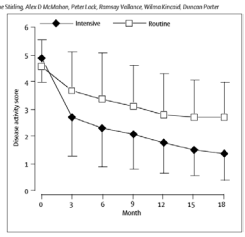
Intention-to-treat analysis of all patients randomised, including those who died or withdrew from the study. Analysis of patients completing the study is very similar (data not shown). *Mantel-Haenszel procedure used.

Table 2: **Number of patients responding at 18-month assessment**

Treatment principles in RA

Effect of a treatment strategy of tight control for rheumatoid arthritis (the TICORA study): a single-blind randomised controlled trial

Catriona Grigor, Hilary Capell, Anne Stirling, Alex D McMahon, Peter Lock, Ramsay Wallace, Wilma Kincaid, Duncan Porter Lancet 2004; 364: 263–69



Treatment principles in RA

Effect of a treatment strategy of tight control for rheumatoid arthritis (the TICORA study): a single-blind randomised controlled trial

Catriona Grigor, Hilary Capell, Anne Stirling, Alex D McMahon, Peter Lock, Ramsay Wallace, Wilma Kincaid, Duncan Porter Lancet 2004; 364: 263–69

- Intensive management:
  - More corticosteroid injections (i.m./i.a.)
  - Higher doses of MTX
  - More frequent start of a new DMARD
  - Higher drug survival
  - Less medication side-effects

Conclusion: intensive treatment gives substantial improvement of disease activity

Treatment principles in RA

Effectiveness of systematic monitoring of rheumatoid arthritis disease activity in daily practice: a multicentre, cluster randomised controlled trial

J Fransen, H Bernelot Moens, I Speyer, P L C M van Riel
Ann Rheum Dis 2005;64:1294–1298. doi: 10.1136/ard.2004.030924

- Cluster RCT: monitoring DAS28 (12 centers) vs routine care (12 centers)
- DAS28 assessment by research nurse at 0 & 24 weeks
- Primary outcomes
  - % patients with DAS28 <3.2 (subgroup analysis due to organizational issues)
  - Changes in DMARD treatment (all patients)

Treatment principles in RA

Effectiveness of systematic monitoring of rheumatoid arthritis disease activity in daily practice: a multicentre, cluster randomised controlled trial

J Fransen, H Bernelot Moens, I Speyer, P L C M van Riel
Ann Rheum Dis 2005;64:1294–1298. doi: 10.1136/ard.2004.030924

- Intervention centers
  - DAS28 measurement for clinical use at 0,4,12 & 24 weeks by treating rheumatologist
  - Study advice: change medication if DAS28 >3.2
- Control centers
  - Visit at week 0,4,12 & 24
  - No systematic monitoring or treatment advices

Treatment principles in RA

Effectiveness of systematic monitoring of rheumatoid arthritis disease activity in daily practice: a multicentre, cluster randomised controlled trial

J Fransen, H Bernelot Moens, I Speyer, P L C M van Riel
Ann Rheum Dis 2005;64:1294–1298. doi: 10.1136/ard.2004.030924

- Baseline data
  - ± 70% women; 58 years; ± 80% RF positivity; disease duration 6 years
  - DAS28 4.5; 13% low disease activity

Of note, DAS28 only measured by research nurse in 142 patients (61 intervention group; 81 control group)

10

Treatment principles in RA

Effectiveness of systematic monitoring of rheumatoid arthritis disease activity in daily practice: a multicentre, cluster randomised controlled trial

J Fransen, H Bernelot Moens, I Speyer, P L C M van Riel
Ann Rheum Dis 2005;64:1294–1298. doi: 10.1136/ard.2004.030924

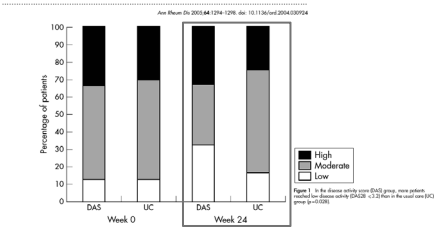


Figure 1. In the disease activity score (DAS) group, more patients reached low disease activity (DAS28: <3.2) than in the usual care (UC) group (p=0.028).

Treatment principles in RA

Effectiveness of systematic monitoring of rheumatoid arthritis disease activity in daily practice: a multicentre, cluster randomised controlled trial

J Fransen, H Bernelot Moens, I Speyer, P L C M van Riel
Ann Rheum Dis 2005;64:1294–1298. doi: 10.1136/ard.2004.030924

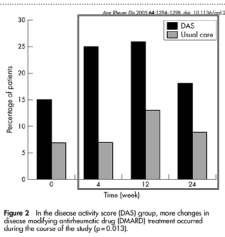


Figure 2 In the disease activity score (DAS) group, more changes in disease modifying antirheumatic drug (DMARD) treatment occurred during the course of the study (p=0.013).

Treatment principles in RA

Effectiveness of systematic monitoring of rheumatoid arthritis disease activity in daily practice: a multicentre, cluster randomised controlled trial

J Fransen, H Bernelot Moens, I Speyer, P L C M van Riel

Ann Rheum Dis 2005;64:1274–1276. doi: 10.1136/ard.2004.030924

- No significant differences in mean MTX, sasp and prednisone dose
- No differences in side-effects
- Intervention group
  - DAS28 measured in 99% of the visits
  - 98% of patients in which medication was changed had a DAS28 >3.2
  - 20% of patients with a DAS28 >3.2 had their medication changed

Conclusion: standard monitoring of disease activity in daily practice, can lead to more DMARD-changes compared to usual care

Treatment principles in RA

Meta-analysis of tight control strategies in rheumatoid arthritis: protocolized treatment has additional value with respect to the clinical outcome

Lydia G. Schipper[1], Laura T. C. van Hulst[1], Richard Grol[2], Piet L. C. M. van Riel[1], Marlies E. J. L. Hulscher[2] and Jaap Fransen[1]

Rheumatology 2010;49:2154–2164

- PubMed & Cochrane library 1995 – 2009: monitoring of disease activity combined with treatment protocols vs monitoring alone
- Inclusion: studies on routine care vs tight control
- Primary outcome
  - Mean change in DAS28 (year 0 vs 1)

Treatment principles in RA

Meta-analysis of tight control strategies in rheumatoid arthritis: protocolized treatment has additional value with respect to the clinical outcome

Lydia G. Schipper[1], Laura T. C. van Hulst[1], Richard Grol[2], Piet L. C. M. van Riel[1], Marlies E. J. L. Hulscher[2] and Jaap Fransen[1]

Rheumatology 2010;49:2154–2164

- Included studies (n= 6)
  - 4 RCT, 2 CCT; study duration between 12 and 24 months
  - 3 studies monitoring + treatment protocol, 3 studies monitoring alone
  - 2 studies in early, DMARD-naive RA; others in early & late RA
- Baseline data
  - 110 to 435 pt per study; 60% to 70% female; 42% to 80% RF positive; DAS28 >3.2

Treatment principles in RA

Meta-analysis of tight control strategies in rheumatoid arthritis: protocolized treatment has additional value with respect to the clinical outcome

Lydia G. Schipper[1], Laura T. C. van Hulst[1], Richard Grol[2], Piet L. C. M. van Riel[1], Marlies E. J. L. Hulscher[2] and Jaap Fransen[1]

Rheumatology 2010;49:2154–2164

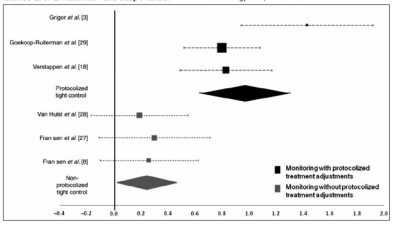
- Results clinical effectiveness tight control
  - In 5 studies tight control better than routine care
  - More medication changes, better physical functioning and less radiographic damage with tight control
  - Toxicity similar
- Results meta-analysis
  - Tight control vs usual care: tight control is more effective, 0.6 DAS28-point more decrease in DAS28
  - Within tight control studies, monitoring + protocol is more effective than monitoring alone: 0.66 DAS28-point more decrease in DAS28

Treatment principles in RA

Meta-analysis of tight control strategies in rheumatoid arthritis: protocolized treatment has additional value with respect to the clinical outcome

Lydia G. Schipper[1], Laura T. C. van Hulst[1], Richard Grol[2], Piet L. C. M. van Riel[1], Marlies E. J. L. Hulscher[2] and Jaap Fransen[1]

Rheumatology 2010;49:2154–2164

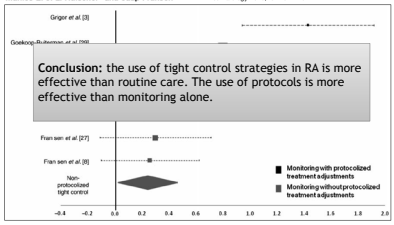


Treatment principles in RA

Meta-analysis of tight control strategies in rheumatoid arthritis: protocolized treatment has additional value with respect to the clinical outcome

Lydia G. Schipper[1], Laura T. C. van Hulst[1], Richard Grol[2], Piet L. C. M. van Riel[1], Marlies E. J. L. Hulscher[2] and Jaap Fransen[1]

Rheumatology 2010;49:2154–2164



Conclusion: the use of tight control strategies in RA is more effective than routine care. The use of protocols is more effective than monitoring alone.

Treatment principles in RA

"Yes, but these are clinical trials. The effect in daily practice has not been proven."

Treatment principles in RA

"Yes, but these are clinical trials. The effect in daily practice has not been proven."

Effect of adherence to European treatment recommendations on early arthritis outcome: data from the ESPOIR cohort

Cécile Escalas, Marie Dalichampt, Bernard Combe, et al.
Ann Rheum Dis 2012 71: 1803-1808

Influence of guideline adherence on outcome in a randomised controlled trial on the efficacy of methotrexate with folate supplementation in rheumatoid arthritis

J Fransen, R F J M Laan, M A F J van der Laar, T W J Huizinga, P L C M van Riel

Ann Rheum Dis 2004;63:1222–1226. doi: 10.1136/ard.2003.018861

Treatment principles in RA

"Yes, but these are clinical trials. The effect in daily practice has not been proven."

ESPOIR study: less radiographic damage after 1 year and functional deterioration after 2 years (early RA).

MTX study: larger decrease in DAS28 after 48 weeks (established RA)

Influence of guideline adherence on outcome in a randomised controlled trial on the efficacy of methotrexate with folate supplementation in rheumatoid arthritis

J Fransen, R F J M Laan, M A F J van der Laar, T W J Huizinga, P L C M van Riel

Ann Rheum Dis 2004;**63**:1222–1226. doi: 10.1136/ard.2003.018861

Treatment principles in RA

Summary:

Tight control gives better results and guideline adherence makes a difference to patients. However…

Treatment principles in RA

Cohort	Recommendation	Guideline adherence
ESPOIR (2002-2005)	First DMARD in early RA (EULAR)	54%
ERAN (2002-2007)	DMARD in early RA	Median time to DMARD start: 8 months (97% DMARD; 67% after 3 years DAS28 >3,2)
North-America (2002-2009)	DMARD in case of active RA (ACR)	25-50% (publication of updated ACR guideline no difference)
DREAM remission induction (2006)	DMARD in case of active RA	70% (98% DAS28 available)

Treatment principles in RA

Cohort	Recommendation	Guideline adherence
ESPOIR (2002-20		
ERAN (2002-200		e to DMARD nths 7% after 3 years
North-America (2002-2009)	DMARD in case of active RA (ACR)	25-50% (publication of updated ACR guideline no difference)
DREAM remission induction (2006)	DMARD in case of active RA	70% (98% DAS28 available)

Guideline adherence is not always optimal…

But how are things in our hospital?

Content

- Background
- Treatment principles in RA
- **Current situation at this hospital**
- Summary
- How to continue?

Current situation at this hospital

		This hospital	Reference hospital
%DAS28 measured *(all RA)*		?	84%
Mean DAS28 *(all RA)*		?	2,58
%DAS28 <3.2 *(all RA)*		?	79%
%DAS28 <2.6 *(all RA)*		?	66%
%Biological users		RA (35%) SpA (37%) PsA (12%)	RA (28%) SpA (59%) PsA (3%)
Diagnosis amon biological users(%)	1. 2. 3.	RA (57%) SpA (20%) PsA (16%)	RA (66%) SpA (30%) PsA (3%)
Most used biological	1. 2. 3.	Adalimumab (96%) Etanercept (86%) Tocilizumab (102%)	Adalimumab (87%) Etanercept (92%) Rituximab (43%)

10

Content

- Background
- Treatment principles in RA
- Current situation at this hospital
- **Summary**
- How to continue?

Summary

- Treatment principles in RA
  - Tight control: important strategy that benefits our patients
  - Application in daily practice not yet optimal, but seems feasible
- RA treatment in this hospital
  - Improvement possible on many indicators

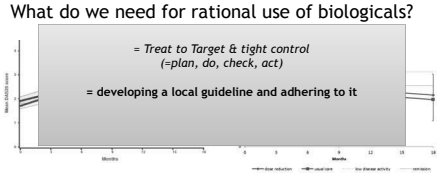
Content

- Background
- Treatment principles in RA
- Current situation at this hospital
- Summary
- **How to continue?**

Goals

To optimize biological use of rheumatologists in this hospital

What do we need for rational use of biologicals?



How to continue?

1. Development of local guidelines
2. Optimizing work flow where necessary
   (for example implementation of standard monitoring of disease activity)
3. Individualized treatment advices in the medical charts of all biological users
4. Feedback after 3 & 6 months

Treatment principles in rheumatoid arthritis & feedback

A translation of the PowerPoint slides used during the second educational meeting

Authors: Nienke Lesuis (resident of rheumatology; MD) and Alfons A den Broeder (rheumatologist-epidemiologist; MD, PhD)
Sint Maartenskliniek, Nijmegen, the Netherlands

Goals

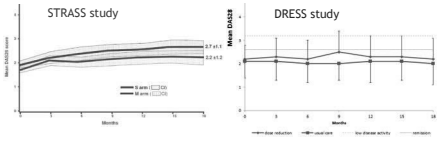
To optimize biological use of rheumatologists in this hospital

What do we need for rational use of biologicals?

Goals

To optimize biological use of rheumatologists in this hospital

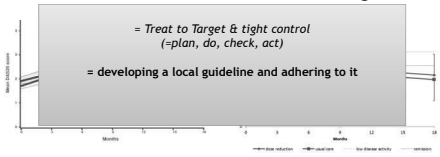
What do we need for rational use of biologicals?



Goals

To optimize biological use of rheumatologists in this hospital

What do we need for rational use of biologicals?



Content

- Background
- Treatment principles in RA
- Current situation at this hospital
- Summary
- How to continue?

Content

- **Background**
- Treatment principles in RA
- Current situation at this hospital
- Summary
- How to continue?

Background

We would rather prevent than treat these hands, but… treatment guidelines are not always followed optimally

Content

- Background
- **Treatment principles in RA**
- Current situation at our hospital
- Summary
- How to continue?

Treatment principles in RA

- Many publications on effective RA treatments
- 'Treat to target', 'tight control' & 'hit hard, hit early' often mentioned

→ Cornerstone of ACR & EULAR RA guidelines

Treatment principles in RA

Tight control: 'frequent assessment of disease activity combined with an objective structured protocol to make treatment changes that maintain low disease activity or remission at an agreed target' *(Kiely et al, Rheumatology 2009)*

Treat to target: 'the treatment aim was defined as remission with low disease activity being an alternative in patients with long-standing disease. Regular follow-up with appropriate therapeutic adaptation to reach the desired state within 3-6 months. Follow-up examinations ought to employ composite measures of disease activity which include joint counts. *(Smolen et al, Ann Rheum Dis 2010)*

Hit hard, hit early: 'early institution of DMARDs.' 'Window of opportunity'

Treatment principles in RA

Tight control: 'frequent assessment of disease activity combined with an objective structured protocol to make treatment changes that maintain low disease activity or remission at an agreed target' *(Kiely et al, Rheumatology 2009)*

Treat to target: 'the treatment aim was defined as remission with low disease activity being an alternative in patients with long-standing disease. Regular follow-up with appropriate therapeutic adaptation to reach the desired state within 3-6 months. Follow-up examinations ought to employ composite measures of disease activity which include joint counts. *(Smolen et al, Ann Rheum Dis 2010)*

Hit hard, hit early: 'early institution of DMARDs.' 'Window of opportunity'

Treatment principles in RA

"Yes, I agree with the general principle but…."

'*A number says nothing! I always look and listen to the patient, order an ESR and feel the joints*'

'*Patients don't want all those pills*'

'*My judgement is better than a composite measure, these numbers mean nothing to a patient*'

Treatment principles in RA

"Yes, I agree with the general principle but…."

DAS28 = 0.56 * sqrt(tender28) + 0.28 * sqrt(swollen28) + 0.70 * ln(ESR) + 0.014 * VAS *(http://www.das-score.nl/das28/nl/uitleg-das28/de-das28-score.html)*

In case of active disease, patients just want to get better. They only get worried about the number of pills later on. *(van Tuyl, Rheumatology 2008)*

A DAS28 <3.2 is associated with 50% less progression of radiographic damage and functional status (HAQ) is influenced by both active inflammation and radiographic damage. *(Fransen et al, Ann Rheum Dis 2005)*

Treatment principles in RA

Effect of a treatment strategy of tight control for rheumatoid arthritis (the TICORA study): a single-blind randomised controlled trial

Catriona Grigor, Hilary Capell, Anne Stirling, Alex D McMahon, Peter Lock, Rramsay Vallance, Willma Kincaid, Duncan Porter Lancet 2004; 364: 263–69

- RCT: routine vs intensive management
- Inclusion criteria: diagnosis of RA <5 years; DAS >2.4
- Study assessment 1x/3 months (DAS etc) by blinded assessor
- Primary endpoints:
  - Mean fall in disease activity
  - % patients with an EULAR good response

Treatment principles in RA

Effect of a treatment strategy of tight control for rheumatoid arthritis (the TICORA study): a single-blind randomised controlled trial

Catriona Grigor, Hilary Capell, Anne Stirling, Alex D McMahon, Peter Lock, Rramsay Vallance, Willma Kincaid, Duncan Porter Lancet 2004; 364: 263–69

- Intensive management:
  - Monthly visits including DAS measurement
  - Corticosteroid injection of any swollen joint or depomedrol 120mg i.m. if DAS >2.4
  - Strict medication protocol: dose increase or cDMARD switch every 1-3 month if DAS >2.4
- Routine care:
  - 3-montly visits
  - No routine DAS measurement or strict medication protocol

Treatment principles in RA

Effect of a treatment strategy of tight control for rheumatoid arthritis (the TICORA study): a single-blind randomised controlled trial

Catriona Grigor, Hilary Capell, Anne Stirling, Alex D McMahon, Peter Lock, Rramsay Vallance, Willma Kincaid, Duncan Porter Lancet 2004; 364: 263–69

	Intensive group (n=55)	Routine group (n=55)	Odds ratio (95% CI)	p*
EULAR good response	45 (82%)	24 (44%)	5·8 (2·4–13·9)*	<0·0001
EULAR remission	36 (65%)	9 (16%)	9·7 (3·9–23·9)*	<0·0001
ACR 20 response	50 (91%)	35 (64%)	5·7 (1·9–16·7)*	<0·0001
ACR 50 response	46 (84%)	22 (40%)	6·1 (2·5–14·9)*	<0·0001
ACR 70 response	39 (71%)	10 (18%)	11 (4·5–27)*	<0·0001

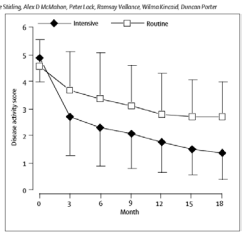
Intention-to-treat analysis of all patients randomised, including those who died or withdrew from the study. Analysis of patients completing the study is very similar (data not shown). *Mantel-Haenszel procedure used.

Table 2: **Number of patients responding at 18-month assessment**

Treatment principles in RA

Effect of a treatment strategy of tight control for rheumatoid arthritis (the TICORA study): a single-blind randomised controlled trial

Catriona Grigor, Hilary Capell, Anne Stirling, Alex D McMahon, Peter Lock, Rramsay Vallance, Willma Kincaid, Duncan Porter Lancet 2004; 364: 263–69



Treatment principles in RA

Effect of a treatment strategy of tight control for rheumatoid arthritis (the TICORA study): a single-blind randomised controlled trial

Catriona Grigor, Hilary Capell, Anne Stirling, Alex D McMahon, Peter Lock, Rramsay Vallance, Willma Kincaid, Duncan Porter Lancet 2004; 364: 263–69

- Intensive management:
  - More corticosteroid injections (i.m./i.a.)
  - Higher doses of MTX
  - More frequent start of a new DMARD
  - Higher drug survival
  - Less medication side-effects

Conclusion: intensive treatment gives substantial improvement of disease activity

Treatment principles in RA

Effectiveness of systematic monitoring of rheumatoid arthritis disease activity in daily practice: a multicentre, cluster randomised controlled trial

J Fransen, H Bernelot Moens, I Speyer, P L C M van Riel

Ann Rheum Dis 2005;64:1294–1298. doi: 10.1136/ard.2004.030924

- Cluster RCT: monitoring DAS28 (12 centers) vs routine care (12 centers)
- DAS28 assessment by research nurse at 0 & 24 weeks
- Primary outcomes
  - % patients with DAS28 <3.2 (subgroup analysis due to organizational issues)
  - Changes in DMARD treatment (all patients)

Treatment principles in RA

Effectiveness of systematic monitoring of rheumatoid arthritis disease activity in daily practice: a multicentre, cluster randomised controlled trial

J Fransen, H Bernelot Moens, I Speyer, P L C M van Riel

Ann Rheum Dis 2005;64:1294–1298. doi: 10.1136/ard.2004.030924

- Intervention centers
  - DAS28 measurement for clinical use at 0,4,12 & 24 weeks by treating rheumatologist
  - Study advice: change medication if DAS28 >3.2
- Control centers
  - Visit at week 0,4,12 & 24
  - No systematic monitoring or treatment advices

Treatment principles in RA

Effectiveness of systematic monitoring of rheumatoid arthritis disease activity in daily practice: a multicentre, cluster randomised controlled trial

J Fransen, H Bernelot Moens, I Speyer, P L C M van Riel

Ann Rheum Dis 2005;64:1294–1298. doi: 10.1136/ard.2004.030924

- Baseline data
  - ± 70% women; 58 years; ± 80% RF positivity; disease duration 6 years
  - DAS28 4.5; 13% low disease activity

Of note, DAS28 only measured by research nurse in 142 patients (61 intervention group; 81 control group)

Treatment principles in RA

Effectiveness of systematic monitoring of rheumatoid arthritis disease activity in daily practice: a multicentre, cluster randomised controlled trial

J Fransen, H Bernelot Moens, I Speyer, P L C M van Riel

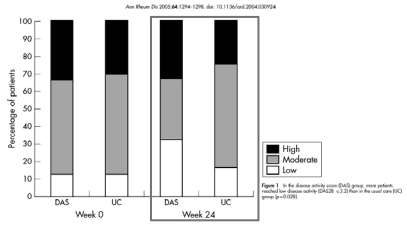


Figure 1 In the disease activity score (DAS) group, more patients reached low disease activity (DAS28 <3.2) than in the usual care (UC) group (p=0.028).

Treatment principles in RA

Effectiveness of systematic monitoring of rheumatoid arthritis disease activity in daily practice: a multicentre, cluster randomised controlled trial

J Fransen, H Bernelot Moens, I Speyer, P L C M van Riel

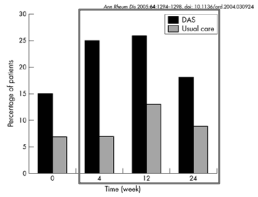


Figure 2 In the disease activity score (DAS) group, more changes in disease modifying antirheumatic drug (DMARD) treatment occurred during the course of the study (p=0.013).

Treatment principles in RA

Effectiveness of systematic monitoring of rheumatoid arthritis disease activity in daily practice: a multicentre, cluster randomised controlled trial

J Fransen, H Bernelot Moens, I Speyer, P L C M van Riel

- No significant differences in mean MTX, sasp and prednisone dose
- No differences in side-effects
- Intervention group
  - DAS28 measured in 99% of the visits
  - 98% of patients in which medication was changed had a DAS28 >3.2
  - 20% of patients with a DAS28 >3.2 had their medication changed

Conclusion: standard monitoring of disease activity in daily practice, can lead to more DMARD-changes compared to usual care

Treatment principles in RA

Meta-analysis of tight control strategies in rheumatoid arthritis: protocolized treatment has additional value with respect to the clinical outcome

Lydia G. Schipper[1], Laura T. C. van Hulst[1], Richard Grol[2]. Piet L. C. M. van Riel[1]. Marlies E. J. L. Hulscher[2] and Jaap Fransen[1] Rheumatology 2010;49:2154-2164

- PubMed & Cochrane library 1995 – 2009: monitoring of disease activity combined with treatment protocols vs monitoring alone
- Inclusion: studies on routine care vs tight control
- Primary outcome
  - Mean change in DAS28 (year 0 vs 1)

Treatment principles in RA

Meta-analysis of tight control strategies in rheumatoid arthritis: protocolized treatment has additional value with respect to the clinical outcome

Lydia G. Schipper[1], Laura T. C. van Hulst[1], Richard Grol[2]. Piet L. C. M. van Riel[1]. Marlies E. J. L. Hulscher[2] and Jaap Fransen[1] Rheumatology 2010;49:2154-2164

- Included studies (n= 6)
  - 4 RCT, 2 CCT; study duration between 12 and 24 months
  - 3 studies monitoring + treatment protocol, 3 studies monitoring alone
  - 2 studies in early, DMARD-naive RA; others in early & late RA
- Baseline data
  - 110 to 435 pt per study; 60% to 70% female; 42% to 80% RF positive; DAS28 >3.2

Treatment principles in RA

Meta-analysis of tight control strategies in rheumatoid arthritis: protocolized treatment has additional value with respect to the clinical outcome

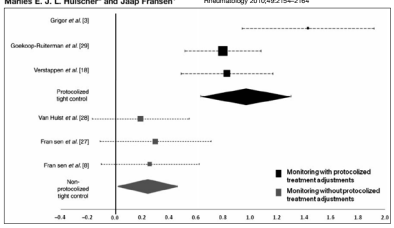
Lydia G. Schipper[1], Laura T. C. van Hulst[1], Richard Grol[2]. Piet L. C. M. van Riel[1]. Marlies E. J. L. Hulscher[2] and Jaap Fransen[1] Rheumatology 2010;49:2154-2164

- Results clinical effectiveness tight control
  - In 5 studies tight control better than routine care
  - More medication changes, better physical functioning and less radiographic damage with tight control
  - Toxicity similar
- Results meta-analysis
  - Tight control vs usual care: tight control is more effective, 0.6 DAS28-point more decrease in DAS28
  - Within tight control studies, monitoring + protocol is more effective than monitoring alone: 0.66 DAS28-point more decrease in DAS28

Treatment principles in RA

Meta-analysis of tight control strategies in rheumatoid arthritis: protocolized treatment has additional value with respect to the clinical outcome

Lydia G. Schipper[1], Laura T. C. van Hulst[1], Richard Grol[2]. Piet L. C. M. van Riel[1]. Marlies E. J. L. Hulscher[2] and Jaap Fransen[1] Rheumatology 2010;49:2154-2164

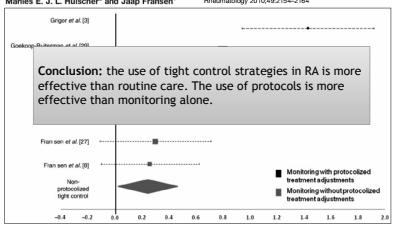


Treatment principles in RA

Meta-analysis of tight control strategies in rheumatoid arthritis: protocolized treatment has additional value with respect to the clinical outcome

Lydia G. Schipper[1], Laura T. C. van Hulst[1], Richard Grol[2]. Piet L. C. M. van Riel[1]. Marlies E. J. L. Hulscher[2] and Jaap Fransen[1] Rheumatology 2010;49:2154-2164

Conclusion: the use of tight control strategies in RA is more effective than routine care. The use of protocols is more effective than monitoring alone.



Treatment principles in RA

"Yes, but these are clinical trials. The effect in daily practice has not been proven."

Treatment principles in RA

"Yes, but these are clinical trials. The effect in daily practice has not been proven."

Effect of adherence to European treatment recommendations on early arthritis outcome: data from the ESPOIR cohort

Cécile Escalas, Marie Dalichampt, Bernard Combe, et al.
Ann Rheum Dis 2012 71: 1803-1808

Influence of guideline adherence on outcome in a randomised controlled trial on the efficacy of methotrexate with folate supplementation in rheumatoid arthritis

J Fransen, R F J M Laan, M A F J van der Laar, T W J Huizinga, P L C M van Riel

Ann Rheum Dis 2004;**63**:1222–1226. doi: 10.1136/ard.2003.018861

Treatment principles in RA

"Yes, but these are clinical trials. The effect in daily practice has not been proven."

ESPOIR study: less radiographic damage after 1 year and functional deterioration after 2 years (early RA).

MTX study: larger decrease in DAS28 after 48 weeks (established RA)

Influence of guideline adherence on outcome in a randomised controlled trial on the efficacy of methotrexate with folate supplementation in rheumatoid arthritis

J Fransen, R F J M Laan, M A F J van der Laar, T W J Huizinga, P L C M van Riel

Ann Rheum Dis 2004;**63**:1222–1226. doi: 10.1136/ard.2003.018861

Treatment principles in RA

Summary:

Tight control gives better results and guideline adherence makes a difference to patients. However…

Treatment principles in RA

Cohort	Recommendation	Guideline adherence
ESPOIR (2002-2005)	First DMARD in early RA (EULAR)	54%
ERAN (2002-2007)	DMARD in early RA	Median time to DMARD start: 8 months (97% DMARD; 67% after 3 years DAS28 >3,2)
North-America (2002-2009)	DMARD in case of active RA (ACR)	25-50% (publication of updated ACR guideline no difference)
DREAM remission induction (2006)	DMARD in case of active RA	70% (98% DAS28 available)

Treatment principles in RA

Cohort	Recommendation	Guideline adherence
ESPOIR (2002-20		
ERAN (2002-200		e to DMARD nths 7% after 3 years
North-America (2002-2009)	DMARD in case of active RA (ACR)	25-50% (publication of updated ACR guideline no difference)
DREAM remission induction (2006)	DMARD in case of active RA	70% (98% DAS28 available)

Guideline adherence is not always optimal…

But how are things in our hospital?

Content

- Background
- Treatment principles in RA
- **Current situation at this hospital**
- Summary
- How to continue?

Current situation at this hospital

		This hospital	Reference hospital
%DAS28 measured *(all RA)*		?	84%
Mean DAS28 *(all RA)*		?	2,58
%DAS28 <3.2 *(all RA)*		?	79%
%DAS28 <2.6 *(all RA)*		?	66%
%Biological users		RA (35%) SpA (37%) PsA (12%)	RA (28%) SpA (59%) PsA (3%)
Diagnosis among biological users (%)	1. 2. 3.	RA (57%) SpA (20%) PsA (16%)	RA (66%) SpA (30%) PsA (3%)
Most used biological	1. 2. 3.	Adalimumab (96%) Etanercept (86%) Tocilizumab (102%)	Adalimumab (87%) Etanercept (92%) Rituximab (43%)

Current situation at this hospital

		This hospital	Reference hospital
%DAS28 measured *(all RA)*		?	84%
Mean DAS28 *(all RA)*		?	2,58
%DAS28 <3.2 *(all RA)*		?	79%
%DAS28 <2.6 *(all RA)*		?	66%
%Biological users		?	RA (28%) SpA (59%) PsA (3%)
Diagnosis among biological users (%)	1. RA (57%) 2. SpA (20%) 3. PsA (16%)		RA (66%) SpA (30%) PsA (3%)
Most used biological	1. Adalimumab (96%) 2. Etanercept (86%) 3. Tocilizumab (102%)		Adalimumab (87%) Etanercept (92%) Rituximab (43%)

But this was six months ago...

Current situation at this hospital

- 344 biological users (275 pre-intervention, 69 post-intervention)
- 57% female, mean age of 57 years
- 65% RA, 15% PsA, 15% SpA
- 37% adalimumab, 21% etanercept, 13% tociluzimab

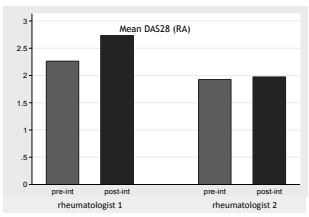
Current situation at this hospital

	Pre-intervention	Post-intervention
Disease activity (measured)		
- **% DAS28** *(all RA)*	40%	
- **% BASDAI** *(alle SpA)*	20%	
Disease activity (mean)		
- **DAS28** *(all RA)*	2,1	
- **BASDAI** *(all SpA)*	3,6	
Disease activity (% low disease activity)		
- **%DAS28 <3.2** *(all RA)*	86%	
- **%BASDAI <4** *(all SpA)*	50%	

Current situation at this hospital

	Pre-intervention	Post-intervention
Disease activity (measured)		
- **% DAS28** *(all RA)*	40%	63%
- **% BASDAI** *(alle SpA)*	20%	38%
Disease activity (mean)		
- **DAS28** *(all RA)*	2.1	2.4
- **BASDAI** *(all SpA)*	3.6	5.0
Disease activity (% low disease activity)		
- **%DAS28 <3.2** *(all RA)*	86%	87%
- **%BASDAI <4** *(all SpA)*	50%	33%

Current situation at this hospital



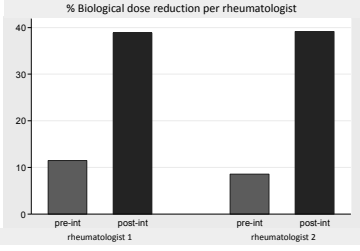
Current situation at this hospital

	Pre-intevention	Post-intevention
Biological dosage (all biologicals)		
- **Startdose**	82%	
- **Dose reduction step 1**	4%	
- **Dose reduction step 2**	5%	
- **Dose reduction step 3**	1%	
Biological dosage (adalimumab)		
- **Startdose**	87%	
- **Dose reduction step 1**	8%	
- **Dose reduction step 2**	3%	
- **Dose reduction step 3**	2%	
Biological dosage (etanercept)		
- **Startdose**	93%	
- **Dose reduction step 1**	0%	
- **Dose reduction step 2**	5%	
- **Dose reduction step 3**	0%	
Biological dosage (tocilizumab)		
- **Startdose**	72%	
- **Dose reduction step 1**	4%	
- **Dose reduction step 2**	17%	
- **Dose reduction step 3**	0%	

Current situation at this hospital

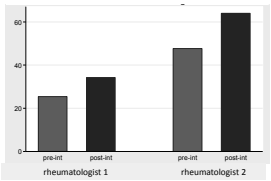
	Pre-intevention	Post-intervention
Biological dosage (all biologicals)		
- **Startdose**	82%	58%
- **Dose reduction step 1**	4%	31%
- **Dose reduction step 2**	5%	7%
- **Dose reduction step 3**	1%	0%
Biological dosage (adalimumab)		
- **Startdose**	87%	62%
- **Dose reduction step 1**	8%	38%
- **Dose reduction step 2**	3%	0%
- **Dose reduction step 3**	2%	0%
Biological dosage (etanercept)		
- **Startdose**	93%	50%
- **Dose reduction step 1**	0%	33%
- **Dose reduction step 2**	5%	17%
- **Dose reduction step 3**	0%	0%
Biological dosage (tocilizumab)		
- **Startdose**	72%	56%
- **Dose reduction step 1**	4%	22%
- **Dose reduction step 2**	17%	0%
- **Dose reduction step 3**	0%	0%

Current situation at this hospital



Current situation at this hospital

	Pre-intervention	Post-intervention
Concomitant DMARD use (RA)	44%	58%



Summary

- More DAS28 and BASDAI measurements done
- More patients using a reduced biological dose
- Mean disease activity unchanged
- Slight improvement in concomitant DMARD use

However...

- Individual differences in results on disease activity measurements and concomitant DMARD use

Keep up the good work!

SUPPLEMENTAL FILE 2: DETAILED OUTCOMES ON bDMARD USE

Table 1 *Outcomes on bDMARD use per patient category*

Outcome	Pre-intervention *(n= 232)*	Post-intervention *(n= 232)*	Odds ratio (95% confidence interval)	P-value
Patients using a reduced bDMARD dose†, % (n)				
RA	12% (18)	67% (89)	2.4 (1.3 to 4.4)	<0.01
PsA	5% (2)	50% (16)	8.0 (1.9 to 71.7)	<0.01
SpA	2% (1)	42% (15)	21 (3.4 to 868.5)	<0.01

Table 2 *Outcomes on bDMARD use per bDMARD*

Outcome	Pre-intervention *(n= 232)*	Post-intervention *(n= 232)*	Odds ratio (95% confidence interval)	P-value
Patients using a reduced bDMARD dose†, % (n)				
Abatacept	0% (0)	25% (1)	-	-
Adalimumab	11% (11)	78% (66)	1.7 (0.8 to 4.0)	0.20
Certolizumab	0% (0)	40% (2)	-	-
Etanercept	8% (4)	60% (28)	4.8 (1.6 to 19.2)	<0.01
Golimumab	0% (0)	15% (2)	-	-
Infliximab	0% (0)	28% (5)	-	-
Rituximab	0% (0)	0% (0)	-	-
Tociluzimab	25% (6)	77% (20)	1.0 (0.3 to 3.7)	1.00
Ustekinumab	-	0% (0)	-	-

"The lab test results are back — now it's time to roll the dice."

11

Summary and general discussion

11.1 SUMMARY

GENERAL INTRODUCTION

Quality of care is the overarching theme of this thesis and can be described as 'the degree to which health services for individuals and populations increase the likelihood of desired health outcomes and are consistent with current professional knowledge'. In this thesis we look at quality of care within rheumatology, focusing on two different themes: the use of diagnostic laboratory tests in rheumatic diseases and guideline adherence of rheumatologists in the treatment of rheumatoid arthritis. These topics are not extensively studied within rheumatology. This thesis aims to fill this gap by exploring the following questions: '*Do rheumatologists provide evidence-based care in their daily practice?*', '*What factors influence whether evidence-based care is provided?*' and '*How can the provision of evidence-based care be improved?*'.

THEME 1: USE OF DIAGNOSTIC LABORATORY TESTS IN RHEUMATIC DISEASES

Chapter 1: Introduction

In this chapter an introduction to the first theme (use of diagnostic laboratory tests) is provided. The topic of this theme was chosen because ordering laboratory tests is daily routine for many rheumatologists. This chapter starts with describing the relation between sensitivity, specificity, pre-test probability on disease and the positive- and negative predictive value (PPV and NPV respectively) of a test. For clinicians it is of crucial importance to realise that even a good test (high sensitivity and specificity) can have a low PPV or NPV if the pre-test probability on disease is either too low or too high. Despite this knowledge on optimal test use, many laboratory tests are ordered inappropriately, leading to overuse of these tests. Consequently, this leads to a higher rate of false-positives, a higher patient burden due to uncertainty and additional testing, and higher costs. This makes overuse of laboratory test an important problem in medicine and **chapter 2** to **6** will explore this topic within rheumatology.

Chapter 2: Use of ANA testing

Here we focus on Antinuclear Antibody (ANA) testing since this test is often used by rheumatologists as a screening test for rheumatic diseases although this practice is not supported by evidence. In this study we assessed ANA use by rheumatologists in three

different hospitals before and after a targeted intervention. This intervention consisted of education and feedback, incorporating a short guideline on when (not) to order an ANA test. Before the intervention, ANA use was high (ANA/new patient ratio (APR) 0.37) but it decreased significantly afterwards (APR 0.11; odds ratio 0.19, 95% confidence interval (95%-CI) 0.17 to 0.22, p <0.01). Furthermore, the percentage of repeated ANA tests and the variation between rheumatologists also decreased significantly. Only the percentages of positive ANA tests and the percentage of patients with an ANA-associated disease did not change after the intervention. Based on these results we concluded that it is possible to decrease ANA overuse by rheumatologists through a relatively simple intervention.

Chapter 3: Determinants of ANA testing

This study is an extension of the previous chapter and explores determinants of ANA overuse both before and after the intervention. In this study we found associations between rheumatologist gender, years of work experience, personality and several ANA outcomes before and after the intervention. For example, rheumatologist with more work experience and a less extravert personality ordered more ANA tests before the intervention (β 0.01, 95%-CI 0.003 to 0.02, p= 0.01; β -0.11, -0.21 to -0.01, p= 0.04 respectively; R^2 47%). After the intervention, female rheumatologists changed less than their male colleagues with regard to the number of ordered ANA tests (β 0.15, 95%-CI 0.03 to 0.26, p= 0.02; R^2 25%). Besides these quantitative analyses we also conducted focus group meetings with rheumatologists after the intervention. Eight rheumatologists participated and they identified seven themes that influenced improvement in ANA overuse: determinants related to the intervention and study; individual health professionals; patients; professional interactions; incentives and resources; capacity for organizational change; and social, political and legal factors.

Chapter 4: Choosing Wisely

In this chapter we describe a 'Top 5 list of things rheumatologists and patients should question', which is part of the Dutch Choosing Wisely campaign. This campaign originates from the United States and aims to 'advance a national dialogue on avoiding wasteful or unnecessary medical tests, treatments and procedures' by publishing top five lists of 'things that physicians and patients should question'. On behalf of the Dutch Society for Rheumatology such a list was developed and the final version includes five statements on diagnostic tests and treatment that should not be done routinely. One of these statements concerns ANA testing and advices only to use ANA testing in patients with a reasonable pre-test probability on ANA-associated diseases such as systemic sclerosis.

Chapter 5: CK and TSH testing in fibromyalgia

Chapter 5 again concerns tests much used in daily rheumatology practice but for which the evidence of this use is scarce. Both creatine kinase (CK) and thyroid stimulating hormone (TSH) are frequently used tests in the routine diagnostic work-up of fibromyalgia (FMS), due to the presumed similarities between FMS and myopathies or hypothyroidism. As the diagnostic value of CK and TSH testing in the context of FMS has not been assessed before, we aimed to study this in 373 patients with suspected FMS. In those patients only 0.5% (95%-CI 0.2% to 1.9%) had an abnormal CK test and none of them was diagnosed with a disease related to elevated CK. For TSH similar results were found: 3.5% (95%-CI 2.1% to 5.9%) of the patients had an elevated TSH and 1.4% (95-CI 0.6% to 3.1%) a lowered TSH but the final diagnosis was FMS in all these patients. Based on these results we concluded that abnormal CK and TSH values are rare in patients with suspected FMS and do not result in an alternative diagnosis. Therefore it seems that routine testing of CK and TSH levels, in patients with suspected FMS referred to secondary care, is not useful.

Chapter 6: Effect of computer reminders on laboratory testing

In the final chapter of theme 1 we examined the effects of two different interventions on the use of rarely indicated tests in rheumatology (complement, immunoglobulins, cryoglobulins and M-protein). Again, these tests were relatively frequently ordered by rheumatologists at the study centre, despite negative evidence for this practice. In this trial, using interrupted time series analysis, we assessed whether trends in the number of laboratory tests ordered by rheumatologists changed following an educational meeting and introduction of computer reminders. The analyses were done for the set of tests on which both interventions had focused (intervention tests; complement, cryoglobulins, immunoglobins, M protein) and a set of tests unrelated to the interventions. After the educational meeting both the level and trend of ordered intervention and control tests did not change significantly. After implementation of the reminders the level of ordered intervention tests decreased with 85.0 tests (95%-CI -133.3 to -36.8, $p < 0.01$). The level of control tests did not change following the introduction of reminders. In this study we concluded that the educational meeting alone was not effective in decreasing the number of ordered intervention tests but in combination with computer reminders it did result in a large decrease of those tests.

THEME 2: GUIDELINE ADHERENCE IN THE TREATMENT OF RHEUMATOID ARTHRITIS

Chapter 7: Introduction

The second theme of this thesis, guideline adherence of rheumatologists in the treatment of rheumatoid arthritis (RA), is introduced in this chapter. RA is a chronic inflammation of the synovial joints, leading to pain, swelling and stiffness, giving limited function of the affected joints. Furthermore, ongoing inflammation can give bone and cartilage destruction, leading to progressive joint damage and consequently loss of function. This irreversible damage can be prevented by adequate and timely treatment with Disease Modifying Anti-Rheumatic Drugs (DMARDs). Besides this 'hit hard, hit early' principle, it is of equal importance to use a 'tight control' based treatment strategy. This means that disease activity should be measured regularly and treatment changed accordingly. Use of tight control strategies leads to lower disease activity, better functional status and less radiographic damage when compared to usual care. As adherence of rheumatologists to tight control principles is not yet optimal, the chapters in this theme all relate to rheumatologists' guideline adherence in the treatment of RA.

Chapter 8: Guideline adherence of rheumatologists

Here we describe the results of an observational study on rheumatologists' adherence to seven different recommendations from the local RA guideline. These recommendations concerned RA diagnostics, treatment and follow-up. Guideline adherence was assessed in 994 visits from the first year of treatment in 137 RA patients. Guideline adherence to the different recommendations varied between 21% and 72%, with referral to the physician assistant as the lowest scoring parameter and referral to a specialized nurse as the highest scoring one. Variation in guideline adherence between rheumatologists was also present, with the parameter on ordering of radiographs showing the highest variation (adherence between 29% and 100%). Furthermore, patient sex, the number of DMARD options, presence of erosions, comorbidity, RF/aCCP positivity, type of patient and the rheumatologists' scientific education status were associated with adherence to one or more guideline parameters. Based on these results we concluded that guideline adherence varied between recommendations and rheumatologists, showing that there is room for improvement.

Chapter 9: Improving rheumatologists guideline adherence

Next we conducted a randomised controlled trial to assess the effects of two different interventions on rheumatologists RA guideline adherence. In this study 20 clinicians (rheumatologists, residents and physician assistants) were randomized in two groups, with the control group receiving the standard intervention (education and feedback)

and the intervention group receiving an extended intervention (education, feedback and a Computerized Decision Support System (CDSS)). Guideline adherence to 13 different indicators (RA treatment, follow-up and administration) was assessed in 990 RA patients visiting the study clinic. A standardized sum score (SSS) on guideline adherence was used as the primary outcome (patient level). The SSS was calculated from the 13 dichotomous indicators. Addition of CDSS to education and feedback did not result in significant better quality of RA care than education and feedback alone (SSS difference 0.02; 95%-CI -0.04 to 0.08; p = 0.60). However, a before/after comparison showed that education and feedback alone resulted in a significant increase in the SSS from 0.58 to 0.64 (difference 0.06; 95%-CI 0.02 to 0.11; p <0.01). So, our results suggest that CDSS did not have an added value with regard to guideline adherence, whereas education and feedback can lead to a small but significant improvement of guideline adherence.

Chapter 10: Implementation of tight control care and biological dose optimisation in a general hospital

The final chapter of this thesis again describes an intervention study to improve RA treatment provided by rheumatologists. At the Sint Maartenskliniek much attention is given to the translation of evidence into practice, and the majority of the studies described in this thesis are conducted at the Sint Maartenskliniek. However, other rheumatology practices are also experiencing that RA treatment is not always optimally executed. Therefore, we conducted a pilot study in another hospital than the Sint Maartenskliniek to improve tight control based care and biological DMARD (bDMARD) dose optimization using a multi-component intervention strategy. As bDMARD dose optimization was an important aspect of this study, not only RA patients were included but also patients with psoriatic arthritis (PsA) and spondyloarthropathy (SpA). Rheumatologists and specialized nurses (n=4) working at the study centre received three consecutive interventions, being: 1) education, feedback and local guideline development, 2) individualised treatment advices written down in the medical charts of patients using a bDMARD, and 3) repeated feedback after three and six months. After these interventions we observed an increase in the percentage of disease activity measures (DAS28: 15% to 51%, OR 3.3, 95% CI 2.1 to 5.5; BASDAI 23% to 50%, OR 2.2, 95% CI 1.0 to 5.5). Furthermore, the percentage of patients using a reduced dose of their bDMARD increased from 10% before the interventions to 61% afterwards (OR 3.9, 95% CI 2.4 to 6.5) while disease activity remained the same (DAS28: mean difference 0.1, 95% CI -0.3 to 0.5; BASDAI: mean difference 0.03, 95% CI -1.8 to 1.9). Based on these results it seems possible to improve tight control based treatment and bDMARD dose optimization in daily practice.

"The lab test results are back — now it's time to roll the dice."

CHAPTER 11.2 GENERAL DISCUSSION

In this chapter we discuss the main findings of this thesis in more detail, followed by some methodological considerations. This chapter finishes with practical implications for rheumatologists and with recommendations for future research.

1. MAIN FINDINGS

The common thread running through this thesis is the behaviour of rheumatologists with regard to daily patient care. In the introduction three questions were formulated that reappeared in both subthemes described in this thesis: 'do rheumatologists provide evidence-based care in daily practice?', 'what factors influence whether evidence-based care is provided?' and 'how can the provision of evidence-based care be improved?'. These questions are used to group the main findings of this thesis and to compare results across studies.

Do rheumatologists provide evidence-based care in daily practice?

1. Overuse of ANA, complement, cryoglobuline, gammaglobuline and M-protein testing by rheumatologists is present in daily practice as compared to optimal test use for various diagnostic tests within rheumatology (**chapter 2** and **6**)
2. Adherence to existing local rheumatoid arthritis (RA) treatment guidelines by rheumatologists is often suboptimal and varies between recommendations and between rheumatologists (**chapter 8, 9** and **10**)

What factors influence whether evidence-based care is provided?

1. Variation in outcomes between rheumatologists could be better explained for ANA overuse than for RA guideline adherence, as observed by the explained variance of the different models (**chapter 3** and **8**)
2. Determinant analysis using quantitative methods mainly yielded relations between outcomes and non-modifiable determinants (e.g. personality, scientific education status), whereas qualitative methods also revealed potential modifiable determinants (e.g. knowledge, awareness) (**chapter 3** and **8**)
3. Frequently mentioned or promising potential determinants of behaviour or behavioural change were not associated with any outcome on ANA use or RA guideline adherence (**chapter 3** and **8**)

How can the provision of evidence-based care be improved?

1. Both improvement studies on laboratory tests overuse showed the same pattern: the number of tests was vastly reduced, whereas the percentage of abnormal lab results did not change (**chapter 2** and **6**)

2. The interventions aimed to reduce laboratory test overuse all resulted in a significant reduction of overuse, whereas the interventions aimed to improve guideline adherence gave less consistent results (**chapter 2, 6, 9** and **10**)

2. DISCUSSION OF MAIN FINDINGS

2.1 Do rheumatologists provide evidence-based care in their daily practice?

1. *Overuse of ANA, complement, cryoglobuline, gammaglobuline and M-protein testing by rheumatologists is present in daily practice as compared to optimal test use for various diagnostic tests within rheumatology (**chapter 2** and **6**)*
2. *Adherence to existing local rheumatoid arthritis (RA) treatment guidelines by rheumatologists is often suboptimal and varies between recommendations and between rheumatologists (**chapter 8**, **9** and **10**)*

Both main findings are not surprising as previous literature on laboratory test use and guideline adherence have found similar results[1-3]. Although these conclusions seem relatively straightforward, it is important to note that no absolute norm on optimal adherence to evidence-based practices exists. However, other methods can be used to judge quality of care.

For example, in the ANA intervention study (**chapter 2**) Bayes theorem was used to show that the number of ANA tests ordered was substantially higher than could be expected based on the population at risk of an ANA associated disease. The observation that the percentages of positive ANA tests were low and repeated ANA tests were relatively high supported this conclusion. So, in case of laboratory test use it is possible to detect overuse on group level without the use of an absolute norm on the correct number of laboratory tests that should be done.

With regard to guideline adherence it is important to note that indicators used to measure adherence are derived from specific guideline recommendations. This implies that adherence to these indicators should be as high as possible while taking for example patient co morbidity or medication side-effects into account. No absolute norms on optimal guideline adherence exist as we do not know what the cut-off value for 'optimal' should be. For that reason, relative norms are often used to judge optimal guideline adherence. Benchmarking, i.e. comparing individual data with adherence data from one or more peers (e.g. similar hospitals or professionals), is an example of a using such a relative norm. This comparison shows whether and what improvements are necessary. In addition a 'best practice' is identified which can serve as an example for other hospitals and individuals. The latter implies that the norm on optimal guideline adherence is subject to change as best practices might keep on improving.

Inter-rheumatologists variation in the degree of ANA overuse and RA guideline adherence was observed in this thesis. Again this has been described before, similar to variation in adherence between recommendations as observed in **chapter 8**[4]. As these observations closely relate to the question '*What factors influence whether evidence-based care is provided*?' this topic will be described in more detail in the section on this question.

Overall, our results on laboratory test use and guideline adherence are in line with the literature. It is striking, however, that in this thesis laboratory tests are mainly associated with overuse and not underuse, whereas the pattern is the other way around in guideline adherence. Making a comparison to the governance of states, a physician can be seen as a state were the legislative, executive and judiciary functions are fused into one system. Similar to state governance, this system carries the risk of conflicts of interest between the different functions. It could be argued that in guideline adherence a physician needs to combine all functions (legislative: indicating treatment; executive: prescribing medication; judiciary: evaluating treatment effects), whereas ordering laboratory tests only includes two functions (legislative: indicating a test; judiciary: interpreting test result). Consequently, this could lead to different forms of conflicting interests between the two behaviours. In case of guideline adherence the workload of performing the indicated actions directly comes to the physician. This probably results in, consciously or unconsciously, not performing some of the guideline recommendations to prevent further workload (conflict between executive and legislative function). In contrast, ordering a laboratory test does not have this conflict as the test is performed by laboratory personal. When the combination of legislative, executive and evaluating functions would be reversed between ordering tests and prescribing treatments, it is possible that then laboratory tests would be underused and treatments overused.

2.2 What factors influence whether evidence-based care is provided?

1. *Variation in outcomes between rheumatologists could be better explained for ANA overuse than for guideline adherence, as observed by the explained variance of the different models (**chapter 3** and **8**)*

Several explanations can be given for this observation. Firstly, the literature on laboratory test use shows that some determinants -such as work experience- are not consistently related with overuse of laboratory tests[5;6]. The same may hold true for the relative impact of those determinants on either test overuse or guideline adherence, although no comparisons exist between overuse and guideline adherence determinants. In addition, laboratory test use and guideline adherence are different sorts of behaviours. Consequently, different determinants might be associated with both behaviours, or similar determinants might not have the same impact on rheumatologists' behaviour. Finally, the sample size in the

ANA study was somewhat higher than in the guideline adherence study (20 vs 14 included clinicians), which might also have influenced our results.

2. *Determinant analysis using quantitative methods mainly yielded relations between outcomes and non-modifiable determinants (e.g. personality, scientific education status), whereas qualitative methods also revealed potential modifiable determinants (e.g. knowledge, awareness) (**chapter 3** and **8**)*

This main finding relates to differences in determinants resulting from the quantitative and qualitative analyses, with the first only finding non-modifiable determinants (gender, rheumatoid factor positivity, etc), whereas qualitative analysis in **chapter 3** also yielded many modifiable determinants such as knowledge. Although the self-reported determinants generated from the qualitative analysis might provide new clues to the selection of effective interventions, we have to keep in mind that physicians are not always capable of judging their own behaviour[7;8]. In addition, the association of these determinants with actual behaviour is not proven by quantitative analysis. Furthermore, determinants such as knowledge might be hard to catch with questionnaires, making that no association was found with behaviour in the quantitative analysis. This issue is described in more detail in the next paragraph.

3. *Frequently mentioned or promising potential determinants of behaviour or behavioural change were not associated with any outcome on ANA use or RA guideline adherence (**chapter 3** and **8**)*

In this thesis we did not observe any relation between frequently mentioned determinants (e.g. knowledge) or promising determinants (e.g. cognitive bias) and outcomes on both ANA overuse and guideline adherence, despite inter-clinician variation in determinants and outcomes.

Firstly, the absence of an association between knowledge and outcomes in this thesis is surprising as knowledge is frequently mentioned in reviews, is a target for many interventions, and was also extensively discussed during the focus group meetings where it seemed to be an important driver of behaviour change. This discrepancy has been described more often[5;6], but in our case a suboptimal measurement instrument might also have attributed to the discrepancy with previous literature. For example, in our ANA study we wanted to assess if knowledge about ANA testing was related to ANA overuse but no ANA knowledge questionnaire was already available. Therefore we developed such a questionnaire ourselves to at least have some form of objective measurement of ANA knowledge. As this was a non-validated questionnaire it might not have been adequate enough to capture ANA knowledge.

Secondly, we also noticed that promising determinants selected from the literature were not always related to our outcomes. In this thesis, this issue was predominantly present for the determinant 'cognitive bias'. Multiple studies suggest that cognitive bias influences medical decision making, but we could not confirm this relationship [9]. If our findings are true, the difference with the literature might have to do with differences in study design. Unfortunately no systematic review is available, but when looking at individual studies they either merely describe potential biases[9-12] or they use fictive scenarios to assess the influence of cognitive biases on decision making[13-16]. By linking cognitive bias to actual behaviour we are probably the first group to have studied cognitive bias in daily practice, explaining the discrepancy.
Besides being a true finding, we could also have missed associations due to our measurement method of cognitive bias. We used the validated Inventory of Cognitive Biases in Medicine (ICBM) questionnaire, which is specifically aimed at physicians and addresses multiple types of cognitive biases[17]. However, ANA overuse and RA guideline adherence might be linked to specific forms of cognitive bias only, making the ICBM not suitable to detect these associations. Therefore, it might be useful to assess cognitive bias in a more indirect way as no other measurement instrument exists. For example, if availability bias (judging things to be more likely if they readily come to mind) plays a role in ANA testing, one would expect rheumatologists to order more ANA tests directly after they have encountered a patient with a positive ANA and a related disease. Unfortunately, practical constraints prevented us from doing this analysis with our ANA data, but it might be an interesting future direction.

In summary, based on the findings related to the question '*What factors influence whether evidence-based care is provided?*' we conclude that (1) rheumatologists behaviour is probably determined by a complex interplay of different factors and (2) self-reported determinants or determinants described in the literature are not always linked to actual behaviour in daily practice.

2.3 How can the provision of evidence-based care be improved?

1. *Both improvement studies on laboratory tests overuse showed the same pattern: the number of tests was vastly reduced, whereas the percentage of abnormal lab results did not change (**chapter 2** and **6**)*

This observation contradicts Bayes theorem as one would expect a rise in the percentage of related diagnoses and positive test results if test overuse decreases. This discrepancy can be explained by the presumption that rheumatologists have not sufficiently applied Bayesian reasoning yet, resulting in test orders in patients in whom pre-test probability is too low for the test to be helpful. In other words: although test overuse has decreased, the decrease was not large enough yet to overcome the drowning of true positive test results into all false positive results.

Linked to this topic might be the degree of numeracy of the rheumatologists, as it has been shown before in other settings that low numeracy might hamper rational decision making on for example the value of cancer screening[18]. However, the rheumatologists in our studies were all highly numerate (95% scored a 6 or 7 out of 7 on the numeracy questionnaire; **chapter 2**), which makes numeracy an unlikely explanation for suboptimal Bayesian reasoning. On the other hand, Bayesian reasoning is more than understanding numbers alone, so numeracy might not capture the broader concept of Bayesian reasoning. Moreover, the numeracy questionnaire used was not specifically developed for physicians or medical decisions, so the questionnaire might not be suited for use in a medical setting.

2. *The interventions aimed to reduce laboratory test overuse all resulted in a significant reduction of overuse, whereas the interventions aimed to improve guideline adherence gave less consistent results (**chapter 2, 6, 9** and **10**)*

In this thesis the four intervention studies (**chapter 2, 6, 9** and **10**) used four different types of interventions (education, feedback, computerized decision support (CDSS) and tailored treatment advices) in various combinations (table 1).

***Table 1** Interventions used in this thesis*

Study topic Intervention	Laboratory test overuse		Guideline adherence	
	Chapter 2	Chapter 6	Chapter 9	Chapter 10
Education		×		
Education & feedback	✓		±	
Education & CDSS		✓	×	
Education, feedback & tailored treatment advices				✓

×: no (additional) effect, ±: small positive effect, ✓: positive effect

As can be seen in table 1, similar interventions gave different results across studies. Of the four studies mentioned in table 1 three were performed in the same hospital. This rules out differences in setting as an explanation for differences in results, leaving the intervention itself or the targeted behaviour as potential explanations. These factors will be described below in more detail, followed by a short notice on the differences in setting for **chapter 2, 6** and **9** versus **chapter 10**.

Firstly, it is known from various Cochrane reviews on the improvement interventions used in this thesis that in general these are effective, but the amount of improvement differs between studies[19-22]. Probably, one of the main reasons for this observation is the notion that interventions should be as closely linked as possible to relevant determinants of the targeted behaviour. Such tailored interventions are more effective than non-tailored interventions;

the most effective method of tailoring is, however, not yet clear[23]. During the preparation of the intervention studies in this thesis, we discussed potential barriers and facilitators of the targeted behaviour within the study team, taking into account our own experiences and literature on the specific topics. For example, before the ANA intervention study we noticed that rheumatologists were not aware of the correct indications for ANA testing and did not know whether they ordered few or many ANA tests. To address these barriers, the combination of education and feedback was used as the intervention in this study. In the case of ANA testing, this informal barrier analysis resulted in the selection of an effective intervention. However, the same method did not result in an effective intervention in our intervention study on RA guideline adherence (**chapter 9**). Here education and feedback only had a small positive effect on the final outcome and CDSS was not effective at all. In hindsight, our studies might have benefitted from a systematic diagnostic analysis assessing the specific barriers and facilitators to the targeted behaviour.

Secondly, similar to education and feedback, education and CDSS were effective in the overuse study (**chapter 6**), but not in the guideline adherence study (**chapter 9**). With regard to the studies including CDSS as an intervention (**chapter 6** and **9**), the complexity of the intervention might have played an additional role besides the aforementioned suboptimal barrier analysis at study start. In the overuse study CDSS consisted of one simple reminder stating that a particular test should not be done, after which a reason had to be given if the rheumatologists wanted to continue the order. So, the CDSS reminder was only a small addition to an existing system. In contrast, CDSS in the guideline adherence study had multiple different components and included a complete make-over of the existing computerized ordering system. This difference in complexity of the intervention might have contributed to the differences in intervention effects between the studies.

Thirdly, differences in targeted behaviour might explain the differences in results in the overuse and guideline adherence studies. For example, decreasing one's ANA overuse can be seen as a less complex behavioural change than improving one's guideline adherence. As complexity of the targeted behaviour is associated with the success of an intervention[4;24], this might partly explain the difference in effectiveness between the studies.
In addition, the complexity of a guideline itself is also known to influence guideline adherence[25]. Again, following one relatively simple recommendation ('no ANA unless reasonable clinical suspicion on a related disease is present') is probably less complicated than following 13 different recommendations which also interact with each other and might all have specific exceptions. Furthermore, the ANA advice was a negative advice stating that something should not be done ("don't"), while in contrast the RA recommendations all stated that something should be done in a certain situation ("do"; for example, change

therapy if RA is active). There is some evidence from a previous study that "do's" are less frequently adhered to than "don'ts"[4], being in line with our observation.

Fourthly, although not directly supported by other studies, it might also be that the targeted behaviour and the interventions were viewed differently by the rheumatologists in the various studies, especially with regard to professional autonomy. It might be that in the guideline adherence study –as compared to the laboratory overuse study- the intervention was felt as a larger breach of professional autonomy as application of the new behaviour would influence a greater part of daily patient care than the application of the ANA advice. Similarly, the complex CDSS used in the RA guideline adherence study might have felt as a larger breach on professional autonomy than the simple CDSS (single reminder) used in the overuse study.

Finally, the study described in **chapter 10** is the only study in this thesis not executed at the Sint Maartenskliniek. As the studies in **chapter 9** and **chapter 10** both aimed to improve RA guideline adherence, but within a different setting we want to comment on the difference in effects between the two studies (small effect in **chapter 9**, substantial effect in **chapter 10**). Firstly, the intervention strategies differed between the studies with education and feedback being the only overlap. Secondly, the baseline values were different in both studies with much lower values in chapter 10 compared to chapter 9, giving the rheumatologists in chapter 10 more room for improvement. Finally, the study in **chapter 10** was performed in another hospital than most of the studies in this thesis. Therefore setting and organizational issues might have played a role. But probably more important was the fact that the rheumatologists in **chapter 10** had an additional financial incentive to improve their practice: if they did not reduce their biological costs, insurance companies would not fully pay for these costs.

All in all it can be concluded that changing rheumatologists' behaviour is a challenge that can be achieved. Various determinants of success probably play a role, including hospital setting, complexity of the intervention, targeted behaviour and recommendations. Incorporating all these factors in a formal diagnostic analysis before selecting the interventions probably would have enhanced the effects of our interventions.

3. METHODOLOGICAL CONSIDERATIONS

In an ideal situation clinical trials include a control group, are double-blinded, and have enough follow-up time and subjects to adequately measure relevant and validated outcomes. When performing implementation studies with clinicians as the main study population, adherence to the standard of clinical trials can be difficult as implementation studies have additional methodological issues. These methodological issues were also encountered during the execution of the studies described in this thesis and will therefore be discussed in more detail below.

3.1 Inclusion of participants in implementation studies

Inclusion of sufficient participants can be a challenge in any type of trial, but implementation trials pose some specific problems as participants are often clinicians instead of patients. Firstly, by default less clinicians than patients are available, leading to smaller target populations in implementation trials. Secondly, inclusion in implementation trials often needs to be done on hospital or ward level as it would otherwise lead to problems with contamination between groups or patient clustering. This results in a group of physicians needing to be included as a whole instead of including physicians on individual basis, meaning that group processes are very likely to influence successful inclusion. In addition, in a group of physicians the motivation to change behaviour -or not- might differ between individuals. Many different psychological models describe different steps of behaviour change, for example distinguishing an orientation, insight, acceptance, change and preservation phase[26]. These phases of behaviour change influence the willingness of physicians to take part in implementation studies and can therefore aid or hamper inclusion.

In this thesis we also struggled with the inclusion of participants in our studies, especially during our multi-centre trial (**chapter 2**, intervention to decrease ANA overuse). Only 3 out of 7 hospitals participated and in hindsight this might have been caused by differences in behaviour change phase the rheumatologists in those hospitals were in. Therefore, our study could have benefitted from a more extensive target group analysis aiming to gain more insight into which phase of behaviour change rheumatologists were in before study start and to adapt our inclusion strategy accordingly[26].

3.2 Use of a control group and blinding of participants in implementation studies

In a classic medication trial, the effectiveness of a medicine is tested using a control and intervention group with participants and researchers blinded for group allocation. Using this design it is relatively easy to assess the effectiveness of the new medicine, but with implementation trials two problems arise.

Firstly, contamination between intervention- and control group poses a real problem if the trial is a single-centre study. In such a trial clinicians might talk about the intervention with their control-group colleagues thus influencing those clinicians as well. Furthermore, patients can be treated by more than one clinician which further increases the risk of contamination between groups as outcomes are often measured on patient level. Randomizing on hospital or ward level rather than on clinician level (cluster randomization) prevents this problem but also requires multi-centre rather than single-centre studies. This again poses an extra challenge to the researchers regarding the inclusion of multiple sites. In this thesis the control group issue was encountered multiple times and for pragmatic reasons we often chose not to use a randomized controlled study design. However, were possible we tried to find other solutions to overcome this problem. For example, we decided to use the number of routine laboratory tests as a control group in combination with an interrupted time series design (**chapter 6**) or to randomize only one of two interventions (**chapter 10**).

Secondly, in an implementation trial behaviour change of physicians is almost always the primary aim. Reaching this aim often requires multi-component interventions such as the combination of education and feedback. This can make it difficult to separate intervention effects of the individual components of a multi-modal intervention strategy. This problem was for example encountered in **chapter 10** (pilot study on education, feedback and tailored treatment advices) of this thesis, warranting the need for replication of our results in a randomized controlled setting.

Finally, due to the type of interventions performed in implementation trials it is virtually impossible to blind participants for group allocation. For example, providing a 'placebo educational meeting' with non-relevant or even incorrect information to a control group would most probably be immediately noticed by the participants and if not, could lead to non-desired behaviour. On the other hand, blinding of researchers is possible but in our randomized controlled trial this could not be done as the researcher performing data analysis was also the researcher performing the interventions.

3.3 The selection of outcome measures in implementation studies

The selection of validated outcome measures can be another major issue in implementation trials. In general, quality of care is measured using quality indicators. These indicators can be further divided into structure (organizational aspects of care), process (provided care to a patient), and outcome (result of provided care) indicators[27]. Compared to outcomes used in clinical trials, use of indicators poses several specific problems.

Firstly, quality of care is a broad concept and often a set of multiple indicators is used to measure several important aspects of care. Especially in combination with a small number of study participants this method can easily lead to multiple testing problems as the number of outcome measures is too high for the number of study subjects. To overcome this issue a sum score can be calculated from individual indicators if those indicators are dichotomous variables. This method has disadvantages as well. For example, if a sum score is used, the scores for indicators not directly related to each other end up in one score. In addition, the relative impact of the intervention on the separate indicators cannot be seen and comparison with other studies might me hampered if different indicators or sum scores are used. In our guideline adherence intervention trial (**chapter 9**) we encountered the multiple-testing issue, and weighing this against the potential disadvantages of sum scores, we chose to use an indicator sum score as the primary outcome measure, reporting the separate indicator scores as secondary outcomes.

Secondly, in an ideal situation indicator sets used in implementation studies should be validated and widely accepted, similar to outcome measures in clinical trials. As for the last criterion: widely accepted indicator sets are virtually non-existing within rheumatology. For example, during the literature search for our RA quality review (**chapter 7**) we encountered seven different RA indicator sets. Although their content overlapped, the number and type (structure, process or outcome) of indicators included in the sets varied[28-34]. None of these sets is internationally accepted, hampering comparison of quality of care studies within RA. In clinical trials the quality of outcome measures used can be assessed using the COSMIN checklist, which comprises aspects on reliability, responsiveness and validity of outcome measure[35]. Although some of those aspects (internal consistency, face validity) can be applied to quality indicators, others are more difficult as for example often no golden standard is available for quality indicators (criterion validity). However, a recent review from the Guidelines International Network describes nine reporting standards for guideline indicators including standards such as measure specification and measure testing/ validating[36]. In addition to this review, other tools are available to judge the quality of indicators. The Appraisal of Indicators through Research and Evaluation (AIRE) is such a tool using 20 statements on four themes to assess the quality of an indicator[37]. The issue of indicator selection was encountered during the preparation of the studies in this thesis, especially those on guideline adherence. In these studies we chose to be pragmatic by selecting indicators from the set stated in the Dutch national RA guideline, as this most closely resembled the local situation at the study hospital.

Thirdly, outcomes in implementation trials are often process indicators, thus reflecting care that is provided to a patient but not the outcomes of that care. Thus, process indicators are essentially surrogate markers for the real outcome in patients. The same can also be seen in clinical trials were for example high cholesterol is a surrogate marker for risk of subsequent myocardial infarction. Often such surrogate markers are easier to measure than the real outcome. However, they carry the risk that results on surrogate markers do not translate into results for hard outcomes. In contrast, for quality of care research process indicators can be used as real outcomes and indeed should be considered the desired outcome measure, when in other types of studies a clear causal relation between provided care and patient outcome has been established. For example, different trials have proven that strict use of tight control strategies benefits RA patient outcomes such as disease activity and functional status[38;39]. Therefore, process indicators on the provision of tight control care by rheumatologists could be used as an outcome measure to judge quality and quality improvement of RA care. In this thesis we mainly used process indicators to assess quality of care, but we are aware of the fact that not for all indicators used a strong relation between process and outcomes of care is established, leaving an issue for future research.

3.4 Measuring outcomes in implementation studies

When measuring outcomes in any type of study it is important to have sufficient study participants to ensure adequate precision of the results. However, due to aforementioned difficulties with participant inclusion this can be a challenge in implementation studies. In **chapter 8** and **9** (RA guideline adherence) of this thesis this issue was encountered during study preparation. In both studies the number of clinicians that could be included was relatively low (14 and 20 respectively), so we needed to include as many patients per clinician as possible in order to ensure adequate precision. However, in **chapter 8** the number of eligible patients was limited due to our in- and exclusion criteria (first year of treatment, second opinion partly excluded). This might have hampered the precision of our results. In contrast, in **chapter 9** our patient inclusion criteria were less strict, leading to large numbers of eligible patients. However, due to manual data collection (chart review) we had to balance precision with feasibility of data collection.

Besides the number of participants, results can also be influenced by other events that take place during follow-up. Especially in uncontrolled studies embedded in daily practice this can be a real issue. Our overuse intervention studies (**chapter 2** and **6**) are examples of these kinds of studies. However, in none of those studies we are aware of any event that could have given the same effect size. Nevertheless, we did learn from the ANA intervention study (**chapter 2**) and in our next overuse intervention study (**chapter 6**) we choose to strengthen the study design by using a control group and an interrupted time series design.

Secondly, the measurement of outcomes in implementation studies can also be complicated by multilevel issues. Interventions in those studies are often aimed at clinicians, leading to behaviour change with regard to care provided to their patients. This means that the intervention and outcome measurement takes place on different levels (clinician vs patient), and in case of multi-centre studies a third level is added (hospital). This calls for more advanced statistical techniques and less power when analysing study results.

Finally, when measuring quality indicators it is also important to consider how they are measured, especially if this concerns process indicators on guideline adherence. In general the two most used methods are: self-reported guideline adherence by the study population or registered behaviour, as measured by medical chart review of patients treated by the study population. The latter method is more labour intensive, but judged to be more objective. In the only review on this subject, comparing self-reported adherence with medical chart review, self-reported adherence exceeded more objective methods in almost all instances and a median over-estimation of adherence of 27% was observed[7]. Therefore, we chose to only measure guideline adherence using medical chart review (**chapter 9** to **11**). This still had some practical issues as there can be a difference between what a rheumatologists decided to do and what is actually done (visit interval of 3 months as decided by the rheumatologist vs actual interval of 4 months due to holiday of the patient). As we were mainly interested in the decisions made by rheumatologists, we chose to rely on the written instructions of the rheumatologists as entered in the chart or orders.

3.5 Generalizability of results

As a final part of this section on methodological issues we comment on the generalizability of results from implementation studies. In order to be able to generalize results outside the specific study setting, it is important to include a sample representative of the target population in daily practice. One way to do this in implementation studies is by running a multi-centre trial including different types of centres, preferably in different settings. In this thesis only one multi-centre trial was performed and even within this trial (ANA intervention study, **chapter 2**) generalizability might be hampered as all participating centres were located in the Netherlands and only 3 out of 7 hospitals participated. However, we did include various types of hospital (general, academic and specialized), aiding generalisation of our results. For our guideline adherence studies, this issue is probably more profound as all were single-centre studies mainly conducted in one specialized rheumatology clinic. Nevertheless, due to the scarcity of quality of care research in rheumatology, our studies provide a first valuable insight into daily rheumatology practice. However, multi-centre studies should be the goal of future research.

Secondly, generalizability of results from implementation studies might be hampered by the fact that interventions are often not similar between studies. Using the intervention taxonomy proposed by the EPOC group[40] can counter this to some extent by at least using the same terminology to describe interventions.

All in all, the preparation and execution of implementation trials has some major challenges compared to clinical trials, especially regarding participant inclusion and outcome measures. Nevertheless, the standards of clinical trials also apply to implementation research.

4. IMPLICATIONS OF THIS THESIS FOR PRACTICING RHEUMATOLOGISTS

With this thesis we have tried to provide more insight into different aspects of quality of rheumatologic care, but above all we hope that our studies will encourage other rheumatologists to assess and improve their quality of care. Therefore, the practical implication of this thesis described below specifically focus on rheumatologists, providing a stepwise approach to improve quality of care.

4.1 How to choose a topic for improvement?

As 'quality of rheumatic care' is a far too broad topic to cover with a single improvement project, specific topics must be chosen. The two most common ways are to act on signals coming either from daily practice (e.g. the observation of large differences in the number of laboratory tests ordered between rheumatologists) or from evidence (e.g. publication of a new RA treatment guideline). To identify areas with room for improvement, a more systematic and continuous approach to quality improvement can also be used by applying a risk matrix often used in safety procedures. These matrixes use the frequency and severity of the consequences of an event to assess the risk level. For example, applying this to the treatment of patients with a rheumatic disease, one can chose to focus on RA patients with high disease activity as this constitutes a large patient population (frequency) and high disease activity has a potentially large impact on patients live (consequences). Depending on the level on which this matrix is applied (patient, hospital, society) factors such as costs could also be taken into account when selecting a topic for improvement. A similar method is proposed by Flottorp *et al*. In their manuscript they provide useful worksheets to facilitate the selection of a topic for improvement[41].

4.2 How to engage rheumatologists in implementation research?

As described in the previous section inclusion of participants in quality or implementation projects can be difficult. Moreover, due to differences in behaviour change motivation, it

might be necessary to apply different strategies in different groups of rheumatologists. For those reasons, we recommend rheumatologists who aim to improve quality of care to first estimate the level of motivation of the targeted colleagues. If they are already aware of the quality of care being suboptimal and are in principle willing to change, the focus can be directed towards the intervention and how this can help them improve the quality of care. However, if the target population is still in an early phase of behaviour change (orientation or insight), this method will not work. In this situation it is first necessary to create awareness about the proposed change and provide insight into the current situation. Most probably the targeted rheumatologists do not exactly know their own performance on the selected topic, so the first step would be to offer rheumatologists a non-committal assessment of their own quality of care. When quality of care proves to be suboptimal after the analysis, the focus can change to the proposed intervention and how this can improve quality of care. In addition, the rheumatologists probably feel bad about their suboptimal quality of care, aiding internal motivation and increasing the likelihood that they accept the intervention as a 'way out' of this feeling.

4.3 How to select outcomes measure for improvement?

As often no widely accepted indicator set is available, the choice of quality measurements used in daily practice should be pragmatic. This means that if indicator sets for the chosen topic are available, the set best reflecting care for that specific setting should be chosen. Such an indicator set can slightly be adapted, for example when the indicators are not defined specific enough (i.e. changing 'regular DAS28 measurements' to 'DAS28 measurements every 3 months'). If no indicator set is available, own indicators should be developed. When developing new indicators it is of crucial importance to be very specific and careful about the choice of denominator and numerator as this greatly influences interpretation of the result. During the process of indicator selection and/or development it is also valuable to use tools such as the AIRE instrument[37].

This topic also relates to another aspect of quality improvement projects: goal setting. In order to easily judge current quality of care or improvement after an intervention it can be helpful to set a goal ('70% of all RA patients should have a DAS28 measured every 3 months'). Furthermore, there are suggestions from previous literature that goal setting aids guideline adherence[19]. However, as mentioned before it is often not known what such a goal should be. Another option is to choose a relative instead of an absolute goal. For example, if after a first measurement of the indicator adherence percentages vary between 20% and 40%, a reasonable goal in this setting could be 50% adherence.

4.4 How to choose an intervention?

Many different types of interventions are available and, as also shown in this thesis, results are not consistent across different studies and different settings, making it impossible to know exactly when to use what intervention. However, in general interventions selected or developed to address barriers specific to the target population are more successful than interventions which are not targeted to barriers[23;26;42]. Therefore, the choice of an intervention starts with a 'diagnostic analysis' to assess which determinants (both barriers and facilitators) drive current practices. Based on this analysis, an intervention can be chosen that specifically targets the observed barriers. During this process worksheets developed by Flottorp *et al*. can be helpful as they provide help to prioritize determinants and select interventions[41]. In addition, the reviews from the Cochrane Effective Practice and Organization of Care (EPOC) group provide up to date evidence for the effectiveness of various types of interventions (http://epoc.cochrane.org).

4.5 How to retain change?

If after the selected interventions rheumatologists' behaviour has improved, it is important that results are sustained. Additional actions are probably necessary to prevent that the effect wears off. An important element in this seems to be the continuation of regular monitoring of the desired behaviour. So, the chosen set of indicators could be used to regularly measure adherence. This information can then be used to guide additional interventions to sustain or improve achieved results. In addition, topics generated from the first, or a new, barrier analysis not yet covered could also be targeted in this phase. Regular monitoring and a new barrier analysis are also important if the intervention did not give the desired result, further stressing the need for additional actions to reach the goal. Of note, in a busy day to day practice attention for these kind of projects may also wear off after a certain time, so it might be helpful to make somebody responsible for quality of care (not necessarily being a rheumatologist) in order to ensure long-term attention to the topic.

5. IMPLICATIONS OF THIS THESIS FOR FURTHER RESEARCH

Research almost always raises more questions than it answers. This thesis also raised several interesting directions for further research. Based on our results we propose that it is important to:

- Assess whether incorrect use of diagnostic tests is also present in other areas than laboratory tests. Based on the frequency of use, imaging techniques such as plain X-ray's or Magnetic Resonance Imaging could be relevant targets for rheumatology

- Assess the added value of general history taking and/or physical examinations in the diagnostic work-up of rheumatic diseases. Both are used to base further clinical decisions upon and incorrect or non-relevant findings may cause incorrect use of diagnostic modalities later on in the diagnostic process
- Develop a set of internationally accepted indicators for the diagnosis, treatment and follow-up of RA patients. Such a set should preferably be developed and endorsed by both the American College of Rheumatology (ACR) and the European League Against Rheumatism (EULAR)
- Establish strong causal relations between proposed process indicators and relevant outcome indicators
- Assess the relation between complexity and workload of guideline recommendations and targeted behaviour in the setting of rheumatology
- Assess the influence of cognitive bias and/or numeracy on actual physician behaviour using indirect measures of bias or better suited questionnaires
- Make a head-to-head comparison of different multi-modal intervention strategies in order to identify the most (cost) effective combination in a certain setting
- Explore the link between rheumatologists' behaviour and what patients do with the advices and treatments provided to them, for example the influence of the way rheumatologists communicate about risks of treatment or rheumatologists' own beliefs about medication
- Systematically explore the reasons for non-participation in implementation trials and consequently develop strategies to enhance participation

REFERENCES

(1) Zhi M, Ding EL, Theisen-Toupal J, Whelan J, Arnaout R. The landscape of inappropriate laboratory testing: a 15-year meta-analysis. *PLoS One* 2013; 8(11):e78962.

(2) Cochrane LJ, Olson CA, Murray S, Dupuis M, Tooman T, Hayes S. Gaps between knowing and doing: understanding and assessing the barriers to optimal health care. *J Contin Educ Health Prof* 2007; 27(2):94-102.

(3) MacLean CH, Louie R, Leake B, McCaffrey DF, Paulus HE, Brook RH et al. Quality of care for patients with rheumatoid arthritis. *JAMA* 2000; 284(8):984-992.

(4) Van den Berg MJ, de Bakker DH, Spreeuwenberg P, Westert GP, Braspenning JC, van der Zee J et al. Labour intensity of guidelines may have a greater effect on adherence than GPs' workload. *BMC Fam Pract* 2009; 10:74.

(5) Whiting P, Toerien M, de S, I, Sterne JA, Dieppe P, Egger M et al. A review identifies and classifies reasons for ordering diagnostic tests. *J Clin Epidemiol* 2007; 60(10):981-989.

(6) Sood R, Sood A, Ghosh AK. Non-evidence-based variables affecting physicians' test-ordering tendencies: a systematic review. *Neth J Med* 2007; 65(5):167-177.

(7) Adams AS, Soumerai SB, Lomas J, Ross-Degnan D. Evidence of self-report bias in assessing adherence to guidelines. *Int J Qual Health Care* 1999; 11(3):187-192.

(8) Davis DA, Mazmanian PE, Fordis M, Van HR, Thorpe KE, Perrier L. Accuracy of physician self-assessment compared with observed measures of competence: a systematic review. *JAMA* 2006; 296(9):1094-1102.

(9) Bornstein BH, Emler AC. Rationality in medical decision making: a review of the literature on doctors' decision-making biases. *J Eval Clin Pract* 2001; 7(2):97-107.

(10) Ely JW, Graber ML, Croskerry P. Checklists to reduce diagnostic errors. *Acad Med* 2011; 86(3):307-313.

(11) Croskerry P. The importance of cognitive errors in diagnosis and strategies to minimize them. *Acad Med* 2003; 78(8):775-780.

(12) Pines JM. Profiles in patient safety: confirmation bias in emergency medicine. *Acad Emerg Med* 2006; 13(1):90-94.

(13) Sorum PC, Mullet E, Shim J, Bonnin-Scaon S, Chasseigne G, Cogneau J. Avoidance of anticipated regret: the ordering of prostate-specific antigen tests. *Med Decis Making* 2004; 24(2):149-159.

(14) Brewer NT, Chapman GB, Schwartz JA, Bergus GR. The influence of irrelevant anchors on the judgments and choices of doctors and patients. *Med Decis Making* 2007; 27(2):203-211.

(15) Schwartz J, Chapman G, Brewer N, Bergus G. The effects of accountability on bias in physician decision making: going from bad to worse. *Psychon Bull Rev* 2004; 11(1):173-178.

(16) Roswarski TE, Murray MD. Supervision of students may protect academic physicians from cognitive bias: a study of decision making and multiple treatment alternatives in medicine. *Med Decis Making* 2006; 26(2):154-161.

(17) Hershberger PJ, Part HM, Markert RJ, Cohen SM, Finger WW. Development of a test of cognitive bias in medical decision making. *Acad Med* 1994; 69(10):839-842.

(18) Wegwarth O, Schwartz LM, Woloshin S, Gaissmaier W, Gigerenzer G. Do physicians understand cancer screening statistics? A national survey of primary care physicians in the United States. *Ann Intern Med* 2012; 156(5):340-349.

(19) Ivers N, Jamtvedt G, Flottorp S, Young JM, Odgaard-Jensen J, French SD et al. Audit and feedback: effects on professional practice and healthcare outcomes. *Cochrane Database Syst Rev* 2012; 6:CD000259.

(20) Forsetlund L, Bjorndal A, Rashidian A, Jamtvedt G, O'Brien MA, Wolf F et al. Continuing education meetings and workshops: effects on professional practice and health care outcomes. *Cochrane Database Syst Rev* 2009;(2):CD003030.

(21) O'Brien MA, Rogers S, Jamtvedt G, Oxman AD, Odgaard-Jensen J, Kristoffersen DT et al. Educational outreach visits: effects on professional practice and health care outcomes. *Cochrane Database Syst Rev* 2007;(4):CD000409.

(22) Shojania KG, Jennings A, Mayhew A, Ramsay CR, Eccles MP, Grimshaw J. The effects of on-screen, point of care computer reminders on processes and outcomes of care. *Cochrane Database Syst Rev* 2009;(3):CD001096.

(23) Baker R, Camosso-Stefinovic J, Gillies C, Shaw EJ, Cheater F, Flottorp S et al. Tailored interventions to address determinants of practice. *Cochrane Database Syst Rev* 2015; 4:CD005470.

(24) Grilli R, Lomas J. Evaluating the message: the relationship between compliance rate and the subject of a practice guideline. *Med Care* 1994; 32(3):202-213.

(25) Francke AL, Smit MC, de Veer AJ, Mistiaen P. Factors influencing the implementation of clinical guidelines for health care professionals: a systematic meta-review. *BMC Med Inform Decis Mak* 2008; 8:38.

(26) Wensing M, Bosch M, Grol R. Beïnvloedende factoren bij implementatie. In: Grol R, Wensing M, editors. Implementatie. Effectieve verbetering van de patiëntenzorg. 4 ed. Amsterdam: Reed Business; 2011. 231-253.

(27) Braspenning J, Hermens R, Calsbeek H, Grol R. Kwaliteit en veiligheid in de zorg: de rol van indicatoren. In: Grol R, Wensing M, editors. Implementatie. Effectieve verbetering van de patiëntenzorg. 4 ed. Amsterdam: Reed Business; 2011. 195-229.

(28) van Hulst LT, Fransen J, Den Broeder AA, Grol R, van Riel PL, Hulscher ME. Development of quality indicators for monitoring of the disease course in rheumatoid arthritis. *Ann Rheum Dis* 2009; 68(12):1805-1810.

(29) Petersson IF, Strombeck B, Andersen L, Cimmino M, Greiff R, Loza E et al. Development of healthcare quality indicators for rheumatoid arthritis in Europe: the eumusc.net project. *Ann Rheum Dis* 2014; 73(5):906-908.

(30) MacLean CH, Saag KG, Solomon DH, Morton SC, Sampsel S, Klippel JH. Measuring quality in arthritis care: methods for developing the Arthritis Foundation's quality indicator set. *Arthritis Rheum* 2004; 51(2):193-202.

(31) Navarro-Compan V, Smolen JS, Huizinga TW, Landewe R, Ferraccioli G, da Silva JA et al. Quality indicators in rheumatoid arthritis: results from the METEOR database. *Rheumatology (Oxford)* 2015; 54(9):1630-1639.

(32) NHS Employers, General Practitioners Committee, British Medical Association. 2014/15 General medical services contract quality and outcomes framework: guidance for GMS contract 2014/15.

(33) Australian Institute of Health and Welfare. National indicators for monitoring osteoarthritis, rheumatoid arthritis and osteoporosis. AIHW cat. no. PHE 77. 2006. Canberra, AIHW.

(34) Kwoh C, MacLean CH, Anderson LG, Dodge C, Fitterman N, Grunwald M et al. Rheumatoid arthritis physician performance measurement set. 2008. American Medical Association and National Committee for Quality Assurance.

(35) Mokkink LB, Terwee CB, Patrick DL, Alonso J, Stratford PW, Knol DL et al. The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. *J Clin Epidemiol* 2010; 63(7):737-745.

(36) Nothacker M, Stokes T, Shaw B, Lindsay P, Sipilä R, Follmann M et al. Reporting standards for guideline-based performance measures. *Implement Sci* 2016; 11(6).

(37) de Koning J, Smulders A, Klazinga N. Appraisal of Indicators through Research and Evaluation (AIRE). Version 2.0. 2007. Available at: http://www.google.nl/ url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&cad=rja&uact=8&ved=0ahUKEwj58oz2iuDJAhXIqg 4KHQ63BjYQFggdMAA&url=http%3A%2F%2Fold.cbo.nl%2FDownloads%2F407% 2FAIRE%2520Instrument %2520(VERSIE%25202.0).pdf&usg=AFQjCNF6TxHzG7_-eHz6aDDmAT1bgnhgAg

(38) Schipper LG, van Hulst LT, Grol R, van Riel PL, Hulscher ME, Fransen J. Meta-analysis of tight control strategies in rheumatoid arthritis: protocolized treatment has additional value with respect to the clinical outcome. *Rheumatology (Oxford)* 2010; 49(11):2154-2164.

(39) Escalas C, Dalichampt M, Combe B, Fautrel B, Guillemin F, Durieux P et al. Effect of adherence to European treatment recommendations on early arthritis outcome: data from the ESPOIR cohort. *Ann Rheum Dis* 2012; 71(11):1803-1808.

(40) Effective Practice and Organisation of Care (EPOC). EPOC Taxonomy 2015. Oslo: Norwegian Knowledge Centre for the Health Services. Available at: http://epoc.cochrane.org/epoc-taxonomy

(41) Flottorp SA, Oxman AD, Krause J, Musila NR, Wensing M, Godycki-Cwirko M et al. A checklist for identifying determinants of practice: a systematic review and synthesis of frameworks and taxonomies of factors that prevent or enable improvements in healthcare professional practice. *Implement Sci* 2013; 8:35.

(42) Grimshaw JM, Thomas RE, MacLennan G, Fraser C, Ramsay CR, Vale L et al. Effectiveness and efficiency of guideline dissemination and implementation strategies. *Health Technol Assess* 2004; 8(6):iii-72.

11

"The lab test results are back — now it's time to roll the dice."

12

Nederlandse samenvatting

INLEIDING

Kwaliteit van zorg is een breed begrip. Een veel gebruikte Engelse definitie is 'the degree to which health services for individuals and populations increase the likelihood of desired health outcomes and are consistent with current professional knowledge'. Vrij vertaald naar het Nederlands is kwaliteit van zorg 'de mate waarin de zorg voor individuen en populaties de kans vergroot dat de gewenste gezondheidsresultaten behaald worden en de mate waarin de zorg consistent is met de huidige kennis van de professionals'. In dit proefschrift staat kwaliteit van zorg binnen de reumatologie centraal waarbij de focus ligt op twee thema's: het gebruik van diagnostische testen en de behandeling van reumatoïde artritis.

THEMA 1: HET GEBRUIK VAN DIAGNOSTISCHE TESTEN BIJ REUMATISCHE AANDOENINGEN

Hoofdstuk 1: Inleiding

Voor het eerste thema is gekozen omdat diagnostische testen veel gebruikt worden door reumatologen om te kijken of een patiënt een bepaalde ziekte wel (aantonen van ziekte) of juist niet (uitsluiten van ziekte) heeft. Dit wordt vaak gedaan met behulp van bloedtesten. Helaas is geen enkele test perfect en worden mensen soms onterecht als ziek of gezond aangemerkt. Hoe goed of slecht een test dit onderscheid kan maken wordt aangegeven met de sensitiviteit en specificiteit van een test (de kans op een positieve test bij zieken en de kans op een negatieve test bij gezonden). Dit zijn belangrijke kenmerken van een test, maar voor een reumatoloog is de omgekeerde vraag veel belangrijker: wat is de kans dat mijn patiënt ziek is nu de testuitslag positief is of juist gezond nu de test negatief is? Wat de kansen hierop zijn, wordt uitgedrukt met de positief en negatief voorspellende waarde. Met deze twee kansen is wel iets bijzonders aan de hand: ze hangen niet alleen af van hoe goed de test is, maar ook van het vóórkomen van de ziekte. In de praktijk betekent dit dat een bijna perfecte test voor een zeldzame ziekte nog steeds een lage positief voorspellende waarde kan hebben. Hoe lager de kans op de ziekte is en hoe vaker een dergelijke test wordt aangevraagd, hoe sterker dit effect is. Reumatologen die voor hun patiënten een test aanvragen moeten dus van te voren bedenken hoe groot de kans op een bepaalde ziekte is zodat ze achteraf de testuitslag kunnen vertrouwen. In de dagelijkse praktijk blijkt dit echter lastig te zijn en worden er vaak teveel testen aangevraagd (ook wel overdiagnostiek genoemd). In **hoofdstuk 2** tot **6** brengen we voor een aantal veel gebruikte testen deze overdiagnostiek in kaart en beschrijven we hoe we het gebruik hebben proberen terug te dringen.

Hoofdstuk 2 en 3: Het gebruik van de ANA-test

In deze twee hoofdstukken richten we ons op een veelgebruikte test binnen de reumatologie: de Antinuclear Antibody (ANA) test. Deze test is eigenlijk bedoeld voor een aantal zeldzame reumatische ziekten en zou daarom ook niet vaak aangevraagd moeten worden. Desondanks wordt deze test toch veel gebruikt door reumatologen. Daarom hebben we, door middel van een training, geprobeerd 29 reumatologen in drie verschillende Nederlandse ziekenhuizen minder ANA-testen te laten aanvragen. Tijdens deze training kregen de artsen informatie over de ANA-test en hoe deze te gebruiken. Ook kregen ze een spiegel voorgehouden over het aantal testen dat ze zelf aanvroegen, waarbij ze zichzelf met directe collega's konden vergelijken. Na een jaar bleek deze interventie zeer succesvol te zijn geweest aangezien het aantal aangevraagde ANA-testen in alle deelnemende ziekenhuizen sterk gedaald was (ANA/nieuwe patiëntratio van 0,37 naar 0,11; odds ratio 0,19, 95%-betrouwbaarheidsinterval 0,17 tot 0,22, p-waarde <0,001). Dit onderzoek staat beschreven in **hoofdstuk 2**. Het volgende hoofdstuk, **hoofdstuk 3**, is een verdieping van het ANA-onderzoek en probeert te achterhalen welke factoren bij de reumatologen hebben bijgedragen aan het succes van de interventie. Uiteindelijk bleken artseigenschappen zoals geslacht, werkervaring en persoonlijkheid gerelateerd te zijn aan het aantal ANA-aanvragen voor de interventie en de verandering in het aantal aanvragen daarna. Zo bleek bijvoorbeeld dat reumatologen met meer werkervaring en een minder extraverte persoonlijkheid meer ANA-testen aanvroegen (respectievelijk regressiecoëfficiënten van 0,01 en -0,11 met 95%-BI van 0,003 tot 0,02 en -0,21 tot -0,01). Verder werden door de reumatologen zelf nog vele andere factoren genoemd die hebben bijgedragen aan het succes, zoals de kwaliteit van de interventie en het enthousiasme van het onderzoeksteam.

Hoofdstuk 4: Verstandig Kiezen

In dit hoofdstuk beschrijven we een 'Top 5 van Verstandige Keuzen', zoals opgesteld door de Nederlandse Vereniging voor Reumatologie (NVR) en wat onderdeel uitmaakt van de Nederlandse campagne 'Verstandig Kiezen'. Dit project is geïnspireerd op de Amerikaanse campagne 'Choosing Wisely' en heeft als doel om medische specialisten en patiënten te ondersteunen bij beslissingen over gepaste zorg. De 'Top 5 van Verstandige Keuzen' is bedoeld als praktisch handvat hierbij. Namens de NVR hebben wij deze top 5 voor de reumatologie opgesteld. Deze bevat stellingen over diagnostische tests en behandelingen die niet routinematig gedaan zouden moeten worden. Een van deze stellingen is: 'Vraag niet standaard een ANA aan, maar alleen bij patiënten die op basis van anamnese en lichamelijk onderzoek een redelijke kans hebben op een zeldzame reumatische ziekte'.

Hoofdstuk 5: Schildklier- en spiertesten bij fibromyalgie

In dit hoofdstuk gaan we in op twee andere veelgebruikte testen binnen de reumatologie: de testen op het schildklierhormoon TSH en op het spierenzym CK. Deze laboratoriumtesten worden regelmatig gebruikt in de diagnostiek rondom het pijnsyndroom fibromyalgie om schildklier- of spierziekten uit te sluiten. Dit wordt gedaan omdat de symptomen van fibromyalgie soms kunnen lijken op die van schildklier- of spierziekten. Maar aan het nut van deze testen kan sterk getwijfeld worden en bovendien is dit nooit goed uitgezocht. Daarom hebben wij bij 373 patiënten met fibromyalgiesymptomen het TSH en CK geprikt en gekeken hoeveel van de testuitslagen afwijkend waren. Vervolgens is er ook gekeken of er bij mensen met een dergelijke afwijkende testuitslag uiteindelijk ook een schildklier- of spierziekte vastgesteld werd. Uit dit onderzoek bleek dat maar 5% van de mensen met mogelijke fibromyalgie een afwijkend TSH had en bij CK was dit zelfs nog lager (0,5%). Na verder onderzoek bleek er bij geen van deze mensen een schildklier- of spierziekte aanwezig te zijn en een dergelijke ziekte was dus ook niet de verklaring van de fibromyalgieklachten. Op basis van deze resultaten lijkt het er dus op dat het niet zinvol is om bij alle mensen met een mogelijk fibromyalgiesyndroom het TSH en CK te testen.

Hoofdstuk 6: Het effect van pop-ups op het aanvraaggedrag van reumatologen

Naast de ANA-test zijn er binnen de reumatologie nog meer testen die, ondanks de zeldzaamheid van gerelateerde ziekten, toch regelmatig worden gebruikt. Ook voor deze testen (complement, gammaglobulines, cryoglobulines, M-proteïne) wilden we het gebruik terugdringen. Net als bij het ANA-onderzoek kregen de reumatologen scholing over het gebruik van deze testen (deze keer zonder feedback over hoe vaak ze zelf de testen aanvroegen). Ook werd hen geadviseerd om de testen minder vaak aan te vragen. Acht maanden na deze scholing werd er nog een aanvullende maatregel genomen: wanneer een reumatoloog een van de betreffende testen wilde aanvragen in het Elektronisch Patiënten Dossier (EPD), verscheen er eerst een mededeling (‘pop-up’). Deze bestond uit één zin over de (on)zin van de test binnen reumatologie en het advies om nog een keer goed over de aanvraag na te denken. Wilde de reumatoloog de test toch aanvragen, dan moest er een reden ingevuld worden voor de aanvraag. Uit de resultaten van dit onderzoek bleek dat de scholing weinig effect had op het aantal aanvragen, maar dat na de invoering van de pop-ups het aantal aanvragen flink daalde (-85,0 testen; 95%-BI -133,3 tot -36,8, p <0.01). Ter controle is er gekeken naar een groep andere testen waarover geen scholing was gegeven en waarbij geen pop-up verscheen tijdens de aanvraag. Bij deze testen veranderde er in de gehele studieperiode van drie jaar weinig aan het aantal aangevraagde testen. Hierdoor kunnen we dus concluderen dat het instellen van een pop-up kan leiden tot een forse daling in het aantal aangevraagde laboratoriumtesten.

THEMA 2: DE BEHANDELING VAN REUMATOIDE ARTRITIS - PROTOCOLADHERENTIE

Hoofdstuk 7: Inleiding

In het tweede thema gaan we in op de behandeling van een van de meest voorkomende vormen van ontstekingsreuma: reumatoïde artritis (RA). Patiënten met RA hebben meestal een symmetrische ontsteking van de gewrichten aan de handen en voeten. Deze ontstekingen geven pijn, stijfheid en zwelling waardoor deze gewrichten minder goed functioneren. Uiteindelijk kan door de ontsteking ook onherstelbare schade aan de gewrichten ontstaan. Reumatologen proberen deze schade te voorkomen door al snel na het begin van de RA te starten met reumaremmers (Disease Modifying Anti-Rheumatic Drugs, DMARDs). Naast het snel starten van DMARDs is het volgens de laatste wetenschappelijke inzichten ook belangrijk om patiënten te behandelen volgens het 'tight control'-principe. Dit betekent dat regelmatig de ziekteactiviteit gemeten moet worden, er een behandeldoel gesteld moet worden (bijvoorbeeld het volledig rustig zijn van de RA) en dat de behandeling aangepast moet worden als dit doel niet gehaald wordt. Patiënten die strikt volgens dit principe behandeld worden hebben uiteindelijk een lagere ziekteactiviteit, functioneren beter en lijken ook minder onherstelbare gewrichtsschade te hebben. Vanwege deze gunstige resultaten zijn (inter)nationale behandelrichtlijnen voor RA gebaseerd op het tight control-principe. Helaas is dit nog niet genoeg om ervoor te zorgen dat alle reumatologen zich altijd aan dit principe, en dus de richtlijnen, houden (ook wel protocoladherentie genoemd). Daarom richten de **hoofdstukken 8**, **9** en **10** zich op protocoladherentie van reumatologen bij de behandeling van RA-patiënten.

Hoofdstuk 8: Protocoladherentie van reumatologen

In het eerste onderzoek van dit thema hebben we gekeken naar de protocoladherentie van reumatologen in de Sint Maartenskliniek. Hiervoor hebben we in de medische status van 137 RA-patiënten gekeken die net begonnen waren met hun behandeling in de Sint Maartenskliniek. Bij al deze patiënten hebben we bij alle visites in het eerste jaar van hun behandeling gekeken of de reumatoloog zich aan zeven belangrijke adviezen uit het RA-protocol hield. Zo keken we bijvoorbeeld of de ziekteactiviteit werd gemeten, of de medicatie op tijd werd aangepast en of patiënten gezien werden door een reumaverpleegkundige. Het bleek dat de ene aanbeveling beter werd gevolgd dan de andere, met protocoladherentie die varieerde tussen de 21% en 72%. Daarnaast verschilde de protocoladherentie ook nog tussen de reumatologen onderling. In het onderzoek hebben we ook gekeken naar redenen voor deze variatie. Het bleek dat verschillende kenmerken van patiënten geassocieerd waren met de protocoladherentie van artsen. Zo werd er bijvoorbeeld gevonden dat patiënten met gewrichtsschade en bijkomende ziekten (co-morbiditeit) minder vaak naar

de reumaverpleegkundige verwezen werden (respectievelijk odds ratio 0,68, 95%-BI 0,16 tot 0,93 en odds ratio 0,68, 95%-BI 0,13 tot 1,00). Daarnaast verwezen gepromoveerde artsen hun patiënten vaker naar de physician assistant dan niet gepromoveerde artsen (odds ratio 4,14, 95%-BI 1,33 tot 12,86). Al met al concludeerden wij uit deze resultaten dat de protocoladherentie van reumatologen nog niet optimaal is en dat verschillende arts- en patiëntkenmerken geassocieerd kunnen zijn met protocoladherentie.

Hoofdstuk 9: Interventie ter verbetering van protocoladherentie reumatologen
Uit de vorige studie kwam naar voren dat protocoladherentie van reumatologen in de Sint Maartenskliniek nog niet op alle onderdelen optimaal was. Daarom is er een vervolgonderzoek gedaan waarbij we geprobeerd hebben deze protocoladherentie te verbeteren. Hiervoor werden de 20 deelnemers (reumatologen, arts-assistenten en physician assistants) ingedeeld in een controle- en interventiegroep. Vervolgens kregen alle deelnemers onderwijs over optimale RA-zorg en het belang van protocoladherentie. Ook kregen ze de resultaten van de vorige studie te horen (feedback). Daarna kreeg de interventiegroep nog een extra maatregel: bij hen werd het EPD zo aangepast dat het makkelijker moest worden om het RA-protocol op te volgen. Zo deed het EPD bijvoorbeeld al een voorstel over het juiste aantal maanden tot de volgende afspraak met de patiënt. Om het effect van beide interventies te testen is er gekeken naar de adherentie van reumatologen aan 13 verschillende aanbevelingen, waarbij er zowel voor als na de interventie bijna 500 patiënten gescoord zijn. Na vergelijking tussen de interventie- en controlegroep bleek dat de gemiddelde protocoladherentie over die 13 punten gelijk was (gemiddelde verschil tussen de groepen: 0.02, 95%-betrouwbaarheidsinterval -0,04 tot 0,08, p-waarde = 0,60). Wel bleek dat het onderwijs voor een kleine, maar significante, stijging van de protocoladherentie had gezorgd (gemiddelde verschil: 0,06, 95%-BI 0,02 tot 0,11, p-waarde <0,01). Deze resultaten betekenen dat in dit geval de computeraanpassingen geen extra effect hebben gehad bovenop het onderwijs, maar dat het onderwijs wel effectief is geweest in het verbeteren van de protocoladherentie van reumatologen.

Hoofdstuk 10: Invoering van optimale reumazorg in een praktijk buiten de Sint Maartenskliniek
Binnen de Sint Maartenskliniek is er veel ervaring opgedaan met het uitvoeren van tightcontrol-reumazorg. Omdat dit principe in andere ziekenhuizen nog niet altijd optimaal gevolgd wordt, hebben we er in dit onderzoek voor gekozen om te proberen of we het tight control-principe ook konden invoeren in een ander ziekenhuis. Daarnaast werd er in dezelfde periode bekend dat het mogelijk was om dure reumaremmers (‘biologicals’) – mits patiënten goed werden gecontroleerd - ook in lagere doseringen aan patiënten voor te schrijven zonder nadelige gevolgen voor patiënten. Het ziekenhuis in Lelystad was

bereid om mee te werken aan dit project en in verschillende stappen werden zowel tight contol-reumazorg als biological dosisoptimalisatie ingevoerd. Als eerste kregen de twee reumatologen en twee reumaverpleegkundigen onderwijs over optimale reumazorg en werd er feedback gegeven op de huidige situatie. Daarna werden er protocollen opgesteld die, onder andere, de behandeling van RA en biological dosisoptimalisatie beschreven. Vervolgens heeft een team van de Sint Maartenskliniek de medische statussen bekeken van alle reumapatiënten die een biological gebruikten. Indien deze patiënten nog niet volledig volgens het nieuwe protocol behandeld werden, werd in de status opgeschreven wat er nog beter kon (bijvoorbeeld: 'denk aan het doen van regelmatige ziekteactiviteitsmetingen'). Tot slot werd er na drie en zes maanden feedback gegeven aan de reumatologen over de huidige situatie. Dit alles resulteerde in een sterke toename van het aantal uitgevoerde ziektemetingen (DAS28 van 15% naar 51%; oddsratio 3,3, 95%-BI 2,1 tot 5,5), toename van het gebruik van een gereduceerde biological dosis (van 10% naar 61%; oddsratio 3,9, 95%-BI 2,4 tot 6,5). Ondanks het verlagen van de biological dosering bleef de ziekteactiviteit gelijk (gemiddelde verschil DAS28 0,1, 95%-BI -0.3 tot 0.5). Uit de resultaten van deze pilot studie blijkt dat het mogelijk is om ook in andere ziekenhuizen de reumazorg te verbeteren.

Hoofdstuk 11: Algemene discussie

In dit hoofdstuk worden de bevindingen uit dit proefschrift samengevat en in meer detail besproken. De belangrijkste bevindingen zijn als volgt:

- Zowel wat betreft het aanvragen van laboratoriumtesten als het volgen van behandelprotocollen is er een gat tussen theorie en praktijk (**hoofdstuk 2**, **6**, **8**, **9** en **10**)
- Variatie tussen reumatologen kan beter verklaard worden voor ANA-overdiagnostiek dan voor RA-protocoladherentie (**hoofdstuk 3** en **8**)
- De determinanten gevonden in de studies verschillen op een aantal punten van andere bestaande studies en hangen daarnaast af van de gekozen analysemethode (**hoofdstuk 3** en **8**)
- De verbetering in het aanvraaggedrag van artsen vertoont hetzelfde patroon in beide overdiagnostiekstudies: het aantal testen daalt sterk terwijl het percentage abnormale testresultaten niet verandert (**hoofdstuk 2** en **6**)
- Gelijksoortige interventies geven niet altijd dezelfde resultaten (**hoofdstuk 2**, **6**, **9** en **10**)

In het vervolg van **hoofdstuk 11** wordt er vooral ingegaan op de tegenstellingen binnen en tussen de onderzoeken. Het is bijvoorbeeld opvallend dat gelijksoortige interventies niet altijd dezelfde effecten geven. Zo is de combinatie van onderwijs en feedback gebruikt in twee studies (ANA-overdiagnostiek in **hoofdstuk 2** en verbeteren protocoladherentie in **hoofdstuk 9**), maar in de ANA-studie was het effect van deze interventie veel groter dan

in de protocoladherentie-studie. Dit verschil zou mogelijk te maken kunnen hebben met verschillen in het aangepakte gedrag tussen beide studies. Zo ging het in de ANA-studie bijvoorbeeld maar om één opdracht die opgevolgd moest worden ('minder ANA's aanvragen'), terwijl bij de protocoladherentie studie zeven verschillende adviezen opgevolgd moesten worden. Hierdoor zou het voor de reumatologen makkelijker kunnen zijn om het ANA-advies op te volgen, wat heeft geleid tot de betere interventie-effecten in **hoofdstuk 2** vergeleken met **hoofdstuk 9**. Naast dit soort verschillen wordt ook de praktische betekenis van de resultaten uit dit proefschrift besproken en worden er aanbevelingen gedaan voor vervolgstudies. Onder andere de volgende aanbevelingen worden gedaan:

- Nagaan of het incorrect gebruiken van diagnostische testen ook voorkomt bij bijvoorbeeld het aanvragen van röntgenfoto's of scans
- Ontwikkeling van een algemeen geaccepteerde set van indicatoren voor de diagnostiek, behandeling en follow-up van reumatoïde artritis. Bij voorkeur wordt zo'n set door de American College of Rheumatology (ACR) en de European League Against Rheumatism (EULAR) opgesteld
- Uitgebreidere analyse van factoren die mogelijk het gedrag van artsen beïnvloeden
- Het direct vergelijken van verschillende multimodale interventiestrategieën zodat de meest effectieve combinatie van interventies vastgesteld kan worden
- Nagaan wat de redenen zijn voor ziekenhuizen om niet deel te nemen aan implementatieonderzoek en deze informatie gebruiken om effectieve strategieën te ontwikkelen die deelname bevorderen


calvin and Hobbes
WATTERSON ©1995
WOW, IT REALLY SNOWED LAST NIGHT! ISN'T IT WONDERFUL?
EVERYTHING FAMILIAR HAS DISAPPEARED! THE WORLD LOOKS BRAND-NEW!
A NEW YEAR... A FRESH, CLEAN START!
IT'S LIKE HAVING A BIG WHITE SHEET OF PAPER TO DRAW ON!
A DAY FULL OF POSSIBILITIES!
IT'S A MAGICAL WORLD, HOBBES, OL' BUDDY...
...LET'S GO EXPLORING!

List of publications
(including earlier theses of the Sint Maartenskliniek)

Curriculum Vitae

Dankwoord

LIST OF PUBLICATIONS

International publications

Within thesis

Lesuis N, Hulscher ME, Piek E, Demirel H, van der Laan-Baalbergen N, Meek I, van Vollenhoven RF, den Broeder AA. Choosing Wisely in daily practice: An intervention study on Antinuclear Antibody testing by rheumatologists. Arthritis Care Res (Hoboken) 2016;68:562-9. Doi 10.1002/acr.22725.

Lesuis N, den Broeder AA, van Vollenhoven RF, Vriezekolk J, Hulscher MEJL. Choosing Wisely in daily practice: a mixed methods study on determinants of Antinuclear Antibody testing by rheumatologists. [manuscript accepted for publication in the Scandinavian Journal of Rheumatology]

Lesuis N, van Vliet J, Boers N, den Broeder N, Hulscher MEJL, Cats H, Verrips A, den Broeder AA. The value of routine creatine kinase and thyroid stimulating hormone testing in patients with suspected fibromyalgia: a cross-sectional study. Rheumatology (Oxford) 2016;55:1273-6. Doi 10.1093/rheumatolog/kew046.

Lesuis N, den Broeder N, Boers N, Piek E, Hulscher MEJL, van Vollenhoven RF, den Broeder AA. The effects of an educational meeting and subsequent computer reminders on the ordering of laboratory tests by rheumatologists. [Accepted for publication in Clinical and Experimental Rheumatology]

Lesuis N, den Broeder AA, Hulscher MEJL, van Vollenhoven RF. Practice what you preach? An explanatory multilevel study on rheumatoid arthritis guideline adherence by rheumatologists. RMD Open 2016;2:e000195. Doi 10.1136/rmdopen-2015-000195.

Lesuis N, van Vollenhoven RF, Hulscher MEJL, den Broeder AA. Rheumatologists' guideline adherence in rheumatoid arthritis: a randomized controlled study on electronic decision support, education and feedback. [manuscript submitted for publication]

Lesuis N, Verhoef LM, Nieboer LM, Bruyn GM, Baudoin P, van Vollenhoven RF, Hulscher MEJL, van den Hoogen FHJ, den Broeder AA. Implementation of protocolized tight control and biological dose optimization in daily clinical practice: results of a pilot study. [Accepted for publication in the Scandinavian Journal of Rheumatology]

Mahmood S, Lesuis N, van Tuyl L, van Riel P, Landewé R. Quality in rheumatoid arthritis care. Best Practice & Research Clinical Rheumatology 2015;29: 664-679.

Not within thesis
Lesuis N, Befrits R, Nyberg F, van Vollenhoven RF. Gender and the treatment of immune-mediated chronic inflammatory diseases: rheumatoid arthritis, inflammatory bowel disease and psoriasis: an observational study. BMC Medicine. 2012;10:82. doi: 10.1186/1741-7015-10-82.

Lesuis N, Zwikker HE, de Bruin M, van Dulmen S, van Vollenhoven RF, den Broeder AA, van den Bemt BJ. What do rheumatologists do to improve medication adherence in patients with rheumatoid arthritis? [manuscript submitted for publication]

Zwikker HE, Lesuis N, den Broeder AA, van den Ende CH, Horne R, van Dulmen S, van den Bemt BJ. Rheumatologists' beliefs about medication barely differ from patients' medication beliefs. [manuscript submitted for publication]

Lesuis N, van Vollenhoven RF. Rheumatoid Arthritis. Handbook of clinical gender medicine. 2012. Karger, Germany.

Other publications

Within thesis
Lesuis N, den Broeder AA. Verstandig kiezen. [Dutch Society for Rheumatology - Choosing Wisely].
Nederlands Tijdschrift voor Reumatologie 2015;2:56-7.

Not within thesis
Lesuis N, van Vollenhoven RF, Hulscher MEJL, den Broeder AA. Te veel onnodig onderzoek in de tweede lijn. [Too much unnecessary testing in secondary care]. Medisch Contact. 2013;49: 2596-97.

Straten VHHP, Lesuis N, den Broeder AA. Naar een eenduidige uitvoering van de DAS28: de 'DAS28-battle'.[Towards uniform execution of DAS28 measurements: the 'DAS28-battle']. Nederlands Tijdschrift voor Reumatologie. 2012;3:10-3.

Lesuis N, Baudoin P, Bruin G, den Broeder A. Op biologicals valt flink te besparen. [Relevant savings can be made on biological costs]. Medisch Contact. 2015;22: 1086-89.

Lesuis N, Bevers K. Van prikpieper naar prikbokaal: de effecten op het aantal gewrichtspuncties uitgevoerd door arts-assistenten.[From injection-telephone to injection-trophy: the results on joint injections done by rheumatologists in training]. Nederlands Tijdschrift voor Reumatologie. 2015;4:6-8.

Conference abstracts

Lesuis N, Befrits R, Nyberg F, van Vollenhoven RF. Gender differences in disease characteristics in rheumatoid arthritis, inflammatory bowel disease and psoriasis: evidence for a greater symptomatic burden in females. Poster presentation at: ACR annual meeting 2011, Chicago, United States of America.

Lesuis N, Piek E, Demirel H, van Vollenhoven RF, den Broeder AA. Requesting behavior of rheumatologists regarding Anti-Nuclear Antibodies (ANA) in patients with rheumatic complaints. Poster presentation at: NVR Najaarsdagen 2012, Arnhem, the Netherlands.

Lesuis N, Piek E, Demirel H, van Vollenhoven RF, den Broeder AA. Less is more? A targeted intervention results in substantial improvement in rheumatologists' ordering of Anti-Nuclear Antibodies (ANA) in patients with rheumatic complaints. Poster presentation at: EULAR conference 2013, Madrid, Spain; NVR Najaarsdagen 2013, Arnhem, the Netherlands.
Lesuis N, Piek E, Demirel H, van Vollenhoven RF, den Broeder AA. Requesting Behavior of Rheumatologists Regarding Anti-Nuclear Antibodies in Patients with Rheumatic Complaints. Poster presentation at: ACR annual meeting 2013, San Diego, United States of America.

Lesuis N, Zwikker H, de Bruin M, van Vollenhoven RF, van Dulmen S, den Broeder AA, van den Bemt B. What do rheumatologists do to improve medication adherence of their rheumatoid arthritis patients? Poster presentation at: NVR Najaarsdagen 2014, Arnhem, the Netherlands.

Lesuis N, van Vollenhoven RF, Hulscher M, den Broeder AA. Practice what you preach? Suboptimal guideline adherence by rheumatologists in patients with rheumatoid arthritis. Poster presentation at: NVR Najaarsdagen 2014, Arnhem, the Netherlands; ACR annual meeting 2014, Boston, United States of America.

Lesuis N, van Vollenhoven RF, Hulscher M, den Broeder AA. Effects of two interventions on rheumatologists' adherence to optimal care recommendations in rheumatoid arthritis: randomized controlled pilot trial. Poster presentation at: NVR Najaarsdagen 2015, Arnhem, the Netherlands; ACR annual meeting 2015, San Francisco, United States of America.

Boers N, den Broeder N, van Vliet J, van Vollenhoven RF, Lesuis N, den Broeder AA. The value of routine use of creatine kinase and thyroid stimulating hormone tests in patients with suspected fibromyalgia: results of a cross-sectional study. Poster presentation at: NVR Najaarsdagen 2015, Arnhem, the Netherlands; ACR annual meeting 2015, San Francisco, United States of America.

Lesuis N, Bruyn GAW, Baudoin P, Nieboer L, den Broeder AA. Impact of protocolized tight control and biological dose optimization in daily clinical practice: results of a pilot implementation study. Oral presentation at: NVR Najaarsdagen 2015, Arnhem, the Netherlands. Poster presentation at: ACR annual meeting 2015, San Francisco, United States of America.

Earlier theses of the Sint Maartenskliniek

De Rooij, D. (1988). Clinical and serological studies in the connective tissue diseases. University of Nijmegen, Nijmegen, The Netherlands.

Geurts, A. (1992). Central adaptation of postural organization to peripheral sensorimotor impairments. University of Nijmegen, Nijmegen, The Netherlands.

Van Lankveld, W. (1993). Coping with chronic stressors of rheumatoid arthritis. University of Nijmegen, Nijmegen, The Netherlands.

Tromp, E. (1995). Neglect in action: a neuropsychological exploration of some behavioural aspects of neglect. University of Nijmegen, Nijmegen, The Netherlands.

Van Balen, H. (1997). A disability-oriented approach to long-term sequelae following traumatic brain injury. Neuropsychological assessment for post-acute rehabilitation. University of Nijmegen, Nijmegen, The Netherlands.

De Kleuver, M. (1998). Triple osteotomy of the pelvis.An anatomical, biomechanical and clinical study. University of Nijmegen, Nijmegen, The Netherlands.

Hochstenbach, J. (1999). The cognitive, emotional, and behavioural consequenses of stroke. University of Nijmegen, The Netherlands.

Donker, S. (2002). Flexibility of human walking: a study on interlimb coordination. Groningen University, Groningen, The Netherlands.

Hendricks, H. (2003). Motor evoked potentials in predicting motor and functional outcome after stroke. University of Nijmegen, Nijmegen, The Netherlands.

Hosman, A. J. F. (2003). Idiopathic thoracic spinal deformities and compensatory mechanisms. University of Nijmegen, Nijmegen, The Netherlands.

Jongerius, P. (2004). Botulinum toxin type-A to treat drooling. A study in children with cerebral palsy. Radboud University, Nijmegen, The Netherlands.

Van de Crommert, H. (2004). Sensory control of gait and its relation to locomotion after a spinal cord injury. Radboud University, Nijmegen, The Netherlands.

Van der Linde, H. (2004). Prosthetic prescription in lower limb amputation. Development of a clinical guideline in the Netherlands. Groningen University, Groningen, The Netherlands.

De Haart, M. (2005). Recovery of standing balance in patients with a supratentorial stroke. Radboud University, Nijmegen, The Netherlands.

Den Otter, R. (2005). The control of gait after stroke: an electromyographic approach to functional recovery. Groningen University, Groningen, The Netherlands.

Weerdesteyn, V. (2005). From the mechanisms of obstacle avoidance towards the prevention of falls. Radboud University, Nijmegen, The Netherlands.

Baken, B. (2007). Reflexion on reflexes. Modulation during gait. Radboud University, Nijmegen, The Netherlands.

Gaasbeek, R. (2007). High tibial osteotomy. Treatment of varus osteoarthritis of the knee. Radboud University, Nijmegen, The Netherlands.

Koëter, S. (2007). Patellar instability. Diagnosis and treatment. Radboud University, Nijmegen, The Netherlands.

Langeloo, D. (2007). Monitoring the spinal cord during corrective spinal surgery: a clinical study of TES-MEP. Radboud University, Nijmegen, The Netherlands.

Ruiter, M. (2008). Speaking in ellipses. The effect of a compensatory style of speech on functional communication in chronic agrammatism. Radboud University, Nijmegen, The Netherlands.

Van den Bemt, B. (2009). Optimizing pharmacotherapy in patients with rheumatoid arthritis: an individualized appraoch. Radboud University, Nijmegen, The Netherlands.

Van Nes, I. (2009). Balance recovery after supratentorial stroke. Influence of hemineglect and the effects of somatosensory stimulation. Radboud University, Nijmegen, The Netherlands.

Aarts, P. (2010). Modified constraint-induced movement therapy for children with unilateral spastic cerebral palsy: the Pirate group intervention. Radboud University, Nijmegen, The Netherlands.

Groen, B. (2010). Martial arts techniques to reduce fall severity. Radboud University, Nijmegen, The Netherlands.

Van Koulil, S. (2010). Tailored cognitive behavioral therapy in fibromyalgia. Radboud University, Nijmegen, The Netherlands.

Boelen, D. (2011). Order out of chaos? Assessment and treatment of executive disorders in brain-injured patients. Radboud University, Nijmegen, The Netherlands.

Heesterbeek, P. (2011). Mind the gaps! Clinical and technical aspects of PCL-retaining total knee replacement with the balanced gap technique. Radboud University, Nijmegen, The Netherlands.

Hegeman, J. (2011). Fall risk and medication. New methods for the assessment of risk factors in commonly used medicines. Radboud University, Nijmegen, The Netherlands.

Smulders, E. (2011). Falls in rheumatic diseases. Risk factors and preventive strategies in osteoporosis and rheumatoid arthritis. Radboud University, Nijmegen, The Netherlands.

Snijders, G. (2011). Improving conservative treatment of knee and hip osteoarthritis. Radboud University, Nijmegen, The Netherlands.

Vriezekolk, J. (2011). Targeting distress in rheumatic diseases. Utrecht University, Utrecht, The Netherlands.

Willems, P. (2011). Decision making in surgical treatment of chronic low back pain. The performance of prognostic tests to select patients for lumbar spinal fusion. Maastricht University, Maastricht, The Netherlands.

Beijer, L. (2012). E-learning based speech therapy (EST). Exploring the potentials of E-health for dysarthric speakers. Radboud University, Nijmegen, The Netherlands.

Hoogeboom, T. (2012). Tailoring conservative care in osteoarthritis. Maastricht University, Maastricht, The Netherlands.

Brinkman, M. (2013). Fixation stability and new surgical concepts of osteotomies around the knee. Utrecht University, Utrecht, The Netherlands.

Kwakkenbos, L. (2013). Psychological well-being in systemic sclerosis: Moving forward in assessment and treatment. Radboud University, Nijmegen, The Netherlands.

Severens, M. (2013). Towards clinical BCI applications: assistive technology and gait rehabilitation. Radboud University, Nijmegen, The Netherlands.

Stukstette, M. (2013). Understanding and treating hand osteoarthritis: a challenge. Utrecht University, Utrecht, The Netherlands.

Van der Maas, A. (2013). Dose reduction of TNF blockers in Rheumatoid Arthritis: clinical and pharmacological aspects. Radboud University, Nijmegen, The Netherlands.

Zedlitz, A. (2013). Brittle brain power. Post-stroke fatigue, explorations into assessment and treatment. Radboud University, Nijmegen, The Netherlands.

Koenraadt, K. (2014). Shedding light on cortical control of movement. Radboud University, Nijmegen, The Netherlands.

Smink, A. (2014). Beating Osteoarthritis. Implementation of a stepped care strategy to manage hip or knee osteoarthritis in clinical practice. VU University Amsterdam, Amsterdam, The Netherlands.

Stolwijk, N. (2014). Feet 4 feet. Plantar pressure and kinematics of the healthy and painful foot. Radboud University, Nijmegen, The Netherlands.

Van Kessel, M. (2014). Nothing left? How to keep on the right track. Spatial and non-spatial attention processes in neglect after stroke. Radboud University, Nijmegen, The Netherlands.

Altmann, V. (2015). Impact of trunk impairment on activity limitation with a focus on wheelchair rugby. Leuven University, Leuven, Belgium.

Bevers, K. (2015). Pathophysiologic and prognostic value of ultrasonography in knee osteoartrhitis. Utrecht University, Utrecht, The Netherlands.

Cuperus, N. (2015). Strategies to improve non-pharmacological care in generalized osteoarthritis. Radboud University, Nijmegen, The Netherlands.

Kilkens, A. (2015). De ontwikkeling en evaluatie van het Communicatie Assessment & Interventie Systeem (CAIS) voor het aanleren van (proto-)imperatief gedrag aan kinderen met complexe ontwikkelingsproblemen. Radboud University, Nijmegen, The Netherlands.

Penning, L. (2015). The effectiveness of injections in cuffdisorders and improvement of diagnostics. Maastricht University, Maastricht, The Netherlands.

Stegeman, M. (2015). Fusion of the tarsal joints: outcome, diagnostics and management of patient expectations. Utrecht University, Utrecht, The Netherlands.

Van Herwaarden, N. (2015). Individualised biological treatment in rheumatoid arthritis. Utrecht University, Utrecht, The Netherlands.

Wiegant, K. (2015). Uitstel kunstknie door kniedistractie. Utrecht University, Utrecht, The Netherlands.

Willems, L. (2015). Non-pharmacological care for patients with systemic sclerosis. Radboud University, Nijmegen, The Netherlands.

Witteveen, A. (2015). The conservative treatment of ankle osteoarthritis. University of Amsterdam, Amsterdam, The Netherlands.

Zwikker, H. (2015). All about beliefs. Exploring and intervening on beliefs about medication to improve adherence in patients with rheumatoid arthritis. Radboud University, Nijmegen, The Netherlands.

Luites, J. (2016). Innovations in femoral tunnel positioning for anatomical ACL reconstruction. Radboud University, Nijmegen, The Netherlands.

Pakvis, D. (2016). Survival, primary stability adn bone remodeling assessment of cementless sockets. An appraisal of Wolff's law in the acetabulum. Radboud University, Nijmegen, The Netherlands.

CURRICULUM VITAE

Nienke Lesuis werd op 21 september 1986 geboren in Oosterhout (Noord-Brabant). Zij behaalde haar VWO diploma in 2004 op het Candea College te Duiven. In datzelfde jaar begon ze ook aan de studie Geneeskunde in Nijmegen (Radboud Universiteit). Aan het einde van haar studie deed zij in het kader van de wetenschappelijke stage onderzoek naar sekseverschillen bij de behandeling van reumatoïde artritis. Deze stage vond plaats in het Karolinska Instituut in Stockholm (Zweden) onder begeleiding van prof. Ronald van Vollenhoven. Begin 2011 behaalde zij haar artsexamen (bene meritum).

Na haar afstuderen begon Nienke als arts-assistent niet in opleiding op de klinische afdeling reumatologie van de Sint Maartenskliniek. Deze werkzaamheden voerde ze uit onder supervisie van dr. Maurice Jeurissen, dr. Henk Martens en dr. Marcel Franssen. Een jaar later werd er gestart met het promotietraject waarbij kwaliteit van zorg het centrale thema was. Dit traject werd begeleid door dr. Alfons den Broeder, prof. Ronald van Vollenhoven en prof. Marlies Hulscher. De resultaten hiervan zijn beschreven in dit proefschrift en gepresenteerd tijdens verschillende (inter)nationale conferenties.

Per 1 januari 2016 is zij begonnen met haar vooropleiding interne geneeskunde in het ziekenhuis Rijnstate te Arnhem (opleider dr. L.J.M. Reichert) in het kader van de opleiding tot reumatoloog (opleider dr. A.E. van Ede).

DANKWOORD

Na alle wetenschappelijke inhoud van dit boekje is het nu tijd voor een wat luchtiger, maar niet minder belangrijk hoofdstuk: het dankwoord. Dit boekje was er niet geweest zonder de praktische, maar vooral ook morele steun van heel veel mensen.

Als eerste wil ik mijn co-promotor dr. A.A. den Broeder bedanken. Beste Alfons, zonder één bepaald diner onder aan de berg in Tante Koosje had dit boekje er waarschijnlijk niet gelegen en had ik misschien niet eens gekozen voor de reumatologie. Met je enthousiasme, maar ook begrip voor mijn twijfels, wist je me toen toch om te praten. Ook daarna heb je dit goed weten vol te houden en heb ik een fantastische onderzoekstijd gehad. Dank voor het delen van je bijna oneindige hoeveelheid (onderzoeks)kennis en optimisme over het doen van onderzoek, ook als het met mede-auteurs even niet opschoot of een artikel wéér werd afgewezen. Maar ik wil je vooral bedanken voor alles buiten het promoveren: de gezellige theemomenten, goede gespreken over het leven, maffe (metaaldetector)-verhalen, interessante boekentips en natuurlijk de gesprekken over onze gedeelde liefde: cognitieve bias en alles wat daarmee maar enigszins samenhangt!
Prof. M.E.J.L. Hulscher, beste Marlies, het duurde even voordat je we jou als eerste promotor hadden gevonden. En ondanks dat je in een al (vrij hard?) rijdende trein stapte, heb ik ontzettend veel aan je begeleiding gehad. Jij zorgde voor de broodnodige 'kwaliteit van zorg' input en zorgde ervoor dat ik & Alfons niet met te wilde ideeën aan kwamen zetten. Daarnaast heb ik altijd heel veel plezier gehad tijdens onze overleggen, zelfs als je mijn artikelen weer eens helemaal had omgeschreven (en ik al je pijlen en commentaar op commentaar moest ontcijferen). Dank voor je enthousiasme, positieve houding en vertrouwen in een goede afloop van de studies in dit boekje.
Prof. R.F. van Vollenhoven, mijn tweede promotor, beste Ronald, ook jij hebt letterlijk aan de wieg gestaan van dit boekje. Als begeleider van mijn onderzoeksstage als student geneeskunde liet je me voor het eerst zien dat onderzoek doen in de eerste plaats leuk is en pas in de tweede plaats (soms) ook lastig. Tijdens mijn promotie was je opeens een begeleider op afstand, maar desondanks kon je ook zo prima de grote lijn bewaken en mij met ogenschijnlijk simpele commentaren weer op het juiste pad brengen. Dank ook voor de gastvrijheid tijdens onze bezoeken aan Stockholm (een leuke bonus van een Zweedse promotor), maar ook voor je bezoekjes aan ons plekje op de berg.

De Sint Maartenskliniek is tijdens mijn promotie mijn tweede 'thuis' geweest, niet in de laatste plaats vanwege de erg fijne werksfeer op de afdeling reumatologie. Daarom wil ik ook graag prof. F.H.J. van den Hoogen bedanken. Beste Frank, dank voor je interesse in mijn projecten en het mogelijk maken van mijn promotietraject.

Geen enkel onderzoek is mogelijk zonder de deelname en inzet van de proefpersonen, daarom een speciaal woord van dank aan alle reumatologen, arts-assistenten, physician assistants en verpleegkundigen die op een of andere manier hebben deelgenomen aan een (of meerdere) van de onderzoeken uit dit boekje. Jullie waren een geweldige groep om mee te werken en ondanks de soms 'spannende' onderwerpen van de studies lieten jullie je niet kennen en deden jullie met veel enthousiasme mee. Heel veel dank daarvoor!

Achter alle onderzoeken uit dit boekje gaat een grote hoeveelheid werk schuil wat mij niet was gelukt zonder de hulp van allerlei mensen. Michiel, in je eerste periode hier had je nog tijd om Acces databases voor anderen te bouwen en daarvan heb ik dankbaar gebruik gemaakt. Verder wil ik je bedanken voor je hulp als ik er met STATA net niet helemaal uit kwam.
Joke, naast de directe bijdrage aan hoofdstuk 3 uit dit boekje heb je ook heel veel praktische & morele steun geboden toen ik als neventaak 'even' de JAK-trial erbij moest doen, dank hiervoor.
Niet alle onderzoeken uit de afgelopen jaren zijn in dit boekje gekomen, maar desondanks wil ik jou, Hanneke, bedanken voor de ontzettend leuke & bij vlagen hilarische samenwerking tijdens het artsenbeliefsproject.
Van alle onderzoeken in dit boekje heeft de voorbereiding van hoofdstuk 9 het meeste tijd & moeite gekost. Ilse, onder jouw strakke leiding liep 'project checklist aka ITS BETTER' op rolletjes en heb ik ook nog eens wat geleerd over projectmanagement (die nobelprijs is helaas nog niet gelukt, maar leuk was het wel!). Ik hoop in de toekomst opnieuw met je te kunnen samenwerken. Marc, zonder jou was de hele interventie er niet geweest en ik wil je ontzettend bedanken voor je inzet, geduld en enthousiasme tijdens dit project. De vele uren bij jou op de kamer maakten me duidelijk dat ik het goede beroep heb gekozen maar dat een inkijk achter de EPD-schermen wel erg nuttig is, bedankt hiervoor. Isabelle & Margo, ook jullie wil ik even noemen en bedanken voor alle input om de EPD wijzigingen ook voor de poli in goede banen te leiden. Lieke Nieboer, ook jij was onmisbaar tijdens dit project: bedankt voor de berg werk die je hebt verzet door van 1000 patiënten de status door te nemen en de info in te voeren. Na dit project ging je gelukkig nog niet weg en konden we samen op pad naar Lelystad om daar de data voor hoofdstuk 10 te verzamelen. Toen ook Lise hierbij aansloot werd het helemaal leuk, dank dames voor de hulp & gezellige autoritjes van en naar de polder. Lise, je bent qua onderwerp toch wel een beetje mijn 'opvolgster', dus bij deze wil ik je heel veel succes wensen met je promotie en ik hoop dat ik in de toekomst toch nog verder kan helpen met de RAINBOW.
Nathan & Nadine, jullie kwamen in het laatste jaar van mijn promotie om de hoek kijken en jullie hebben er, met veel enthousiasme, voor gezorgd dat er opeens nog 2 hoofdstukken in het boekje bij konden. Dankjewel hiervoor.

Reinier, jou wil ik bedanken voor al je hulp bij de statistische kant van veel van de onderzoeken in dit proefschrift. Je uitleg en adviezen waren altijd de moeite van het heen en weer fietsen waard.
Tot slot mogen ook Majella en de andere dames van het secretariaat hier niet ontbreken. Bedankt dames voor alle administratieve ondersteuning en de gezellige praatjes tussen de bedrijven door.

Het 'thuisgevoel' bij de Sint Maartenskliniek heeft alles te maken met de ontzettend fijne (ex-)collega's terplekke. Daarom, Alfons, Karen, Henk, Joost van Zadelhoff, Aatke, Hans, Marcel Flendrie, Marcel Franssen, Maartje, Annemiek, Hatice, Agnes, Frank, Elien, Maurice, Esther, Joost Huijs, Vincent, Susan, Regina, Iris, Anne Wennemers, Calin, Gijs, Anne Cremers, Sandra, Delia & Fleur: ontzettend bedankt voor de enorm gezellige, fijne & leerzame tijd die ik bij/met jullie heb gehad.
Ook alle andere reumaonderzoekers ('uit het W-gebouw') wil ik bedanken voor de afgelopen tijd. Buiten de onderzoeksoverleggen zagen we elkaar niet altijd veel, maar dankzij de jaarlijkse schrijfdagen is er toch een band gesmeed. De 'Ski-radio' en andere maffe uitspraken tijdens 30 seconds (of andere spelletjes) zal ik niet snel vergeten.
In de rijtjes hierboven missen bewust 3 mensen, want mijn (ex-)kamergenootjes Noortje, Chantal & Lieke wil ik graag speciaal bedanken. Als eerstje Noortje: vanaf het eerste moment dat ik als co-assistent bij de reuma begon, was jij altijd in de buurt. Naast de medisch/onderzoeks-inhoudelijke kant konden we op persoonlijk gebied ook veel met elkaar delen. Ik vond het ontzettend fijn om jou zo lang als collega te hebben gehad en gelukkig hebben we tijdens de opleiding tot reumatoloog nog genoeg momenten waarop we elkaar weer tegen kunnen komen. Chantal, ook met jou heb ik bijna vanaf het begin samen gewerkt. Vanaf het begin konden we het goed met elkaar vinden en hebben we ondertussen heel wat af gekletst. Dit laatste kwam ook goed van pas tijdens de vele 4Daagse trainingen als we uren over saaie wegen door de regen moesten lopen, gelukkig bezorgden vissers langs de route ons altijd genoeg afleiding. Ik vind het dan ook erg leuk dat ik je straks als directe collega weer in Rijnstate tegen kom. Lieke, je theebeker was vaak nog halfvol als we alweer voor de volgende ronde gingen, maar hierdoor hadden we wel weer even tijd om te kletsen over vanalles en nog wat. Ook met jou heb ik een hele gezellige tijd gehad en totdat ook jij in opleiding gaat, hoop ik je nog op andere momenten te treffen.

Gelukkig hebben er in de afgelopen jaren ook mensen gezorgd voor afleiding buiten het werk om. Annemiek, sinds groep 3 ben je een constante factor in mijn leven en hoewel je mijn promotie vooral vanuit andere landen dan Nederland hebt meegekregen wil ik je toch bedanken voor de gezellige, fijne en mooie gesprekken via alle mogelijke communicatiemanieren. Lieke, ook onze geschiedenis gaat terug tot op de basisschool en in

de afgelopen jaren heb ik je oprechte interesse in mij enorm gewaardeerd. Wat mij betreft mag je nog wel even in Arnhem blijven wonen zodat we nog vaak kunnen bijkletsen bij een kopje thee.
Wienke, Maaike, Joep, Ramon, Lyvonne, Lianne, Monique, Nicky en Yara, ook jullie mogen in dit dankwoord niet ontbreken. Na een zeer geslaagde gezamenlijke studietijd ben ik heel blij dat jullie in de afgelopen jaren nog steeds voor de broodnodige afleiding zorgen met alle gezellige eetafspraken, musicalbezoeken, schaatstrainingen, spelletjesavonden en natuurlijk het jaarlijkse weekendje-weg.
Sybilla, als ex-co-groepgenootje hebben we tijdens de laatste fase van onze studie veel lief & leed gedeeld. Ik ben erg blij dat het contact ook hierna is gebleven en de bezoekjes aan jou en Rick in het Zuid-Limburgse Mechelen zijn altijd een feest. Bedankt voor deze bijzondere vriendschap en ik loop graag nog een keer het Pieterpad met jullie mee.
Katerina, a special word of thanks is for you. I would never have thought that I would get a Greek friend living in Sweden, who I can meet most easily at international conferences in Chicago, Boston or San Francisco. I hope we can keep this up, but for now you deserve a big 'thank you' for the cartoon on the front of this thesis. Tack så mycket, dankjewel, σας ευχαριστώ!

Met familie verspreid over bijna het hele land is afspreken niet altijd makkelijk, maar toch wil ik Oma Lesuis en alle ooms, tantes, neven & nichten Lesuis en Olthof bedanken voor de gezellige familiebijeenkomsten en de altijd aanwezige interesse in mijn doen & laten. In 2013 werd hier ook nog een schoonfamilie aan toegevoegd: Jan, Janneke & Karen, dankjewel voor het warme welkom in jullie gezin. Ditzelfde geldt ook voor de hele familie Ligtenberg, en al zal ik nooit een echte Toeteboer worden, Roode Klif heb ik door jullie toch maar mooi gezien!

Lieve papa, mama en Jasper, de goede basis die ik bij jullie heb meegekregen heeft er zeker voor gezorgd dat dit boekje nu voor jullie ligt. Mama, ondanks al je eigen drukte rondom de naailes & tig andere bezigheden wil je altijd heel graag weten hoe het mij gaat. De hele wetenschappelijke inhoud van mijn promotie is volgens mij wat meer langs je heen gegaan, maar ik ben je enorm dankbaar voor alle gezellige momenten in de afgelopen jaren én natuurlijk je praktische hulp (potjes jam aanleveren, naaivragen oplossen, verhuishulp…). Papa, tijdens onze gesprekken over ons beider werk kwam ik er al snel achter dat het weinig uitmaakt of je nu fabrieksmedewerkers betere buizen wilt laten maken of artsen betere zorg wil laten leveren. Dit maakte dat ik jou interesse in mij & dit boekje steeds meer ben gaan waarderen, en net als mama ben je nooit te beroerd om ook op andere gebieden hulp te leveren (taxi spelen, helpen in de tuin, rommel naar de stort brengen…). Dankjewel hiervoor! Jasper, als mijn 'kleine grote' broer neem je voor mij een bijzondere plek in.

Niemand kan met zoveel oprechte interesse de meest directe & persoonlijke vragen stellen op een moment dat je dit totaal niet verwacht. Ik ben heel blij dat we nog steeds aardig dicht bij elkaar wonen want onze (eet)afspraken wil ik zeker niet missen!

Lieve, lieve Rob, dit keer is het laatste woord niet aan jou maar voor jou. Je bent het liefste, leukste en grappigste vriendje dat ik me kan wensen en vanaf het eerste begin ben je een grote steun voor mij geweest. Als nuchtere Twentenaar weet je me altijd weer tot rust te brengen als ik te snel wil of vrolijk je me met maffe woordgrapjes weer op als ik iets teveel stress heb. En nu we ook echt samen in 1 huis wonen weet ik het zeker: ik wil je nooit meer kwijt!

Sint Maartenskliniek