Maastricht University

A general practitioner's approach to lower urinary tract symptoms

Citation for published version (APA):

Norg, R. J. C. (2008). A general practitioner's approach to lower urinary tract symptoms. [Doctoral Thesis, Maastricht University]. Datawyse / Universitaire Pers Maastricht. https://doi.org/10.26481/dis.20080125rn

Document status and date:
Published: 01/01/2008

DOI:
10.26481/dis.20080125rn

Document Version:
Publisher's PDF, also known as Version of record

Please check the document version of this publication:

• A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
• The final author version and the galley proof are versions of the publication after peer review.
• The final published version features the final layout of the paper including the volume, issue and page numbers.

Link to publication

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

• Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
• You may not further distribute the material or use it for any profit-making activity or commercial gain
• You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

www.umlib.nl/taverne-license

Take down policy
If you believe that this document breaches copyright please contact us at:

repository@maastrichtuniversity.nl

providing details and we will investigate your claim.

Download date: 20 Apr. 2024

A general practitioner's approach to lower urinary tract symptoms

Proefschrift

ter verkrijging van de graad van doctor aan de Universiteit Maastricht
op gezag van de Rector Magnificus, prof.mr. G.P.M.F. Mols,
volgens het besluit van het college van Decanen
in het openbaar te verdedigen
op vrijdag 25 januari 2008 om 14.00 uur
door

Roelf Johannes Cornelis Norg

UM
UNIVERSITAIRE PERS MAASTRICHT

Promotoren:

prof.dr. J.A. Knottnerus

prof.dr. C.P. van Schayck

Co-promotor:

dr. P.J.M. Portegijs

Beoordelingscommissie:

prof.dr. Ph.E.V. van Kerrebroeck (voorzitter)

dr. A.D.M. Kester

prof.dr. M. Prins

prof.dr. F.G. Schellevis (Nivel / Vrije Universiteit Amsterdam)

dr. G.G.M.C. Wolfs

A general practitioner's approach to lower urinary tract symptoms

The studies presented in this theses were conducted at the Care and Public Health Research Institute (Caphri). Caphri is part of the Universiteit Maastricht and participates in the Netherlands School of Primary Care Research (CaRe). CaRe is acknowledged in 1995 by the Royal Dutch Academy of Science (KNAW).

The studies were financially supported by The Netherlands Organisation for Health Research and Development (ZonMw), the Health Care Insurance Board (CVZ), and Astellas B.V. Astellas is producer of tamsulosin.

© Roelf J.C. Norg, The Netherlands. All rights are reserved.
ISBN 978 90 5278 693 3

Ὁ ποιήσας τὸ ἔλεος μετ᾽αὐτοῦ.

List of abbreviations

AUR	acute urinary retention
BPH	benign prostatic hyperplasia
Dan-PSS	Danish prostate symptom score
DRE	digital rectal examination
GP	general practitioner
IPSS	international prostate symptoms score
IPSS-QoI	quality of life score (question 8 of the IPSS)
LUTS	lower urinary tract symptoms
PSA	prostate specific antigen
Qmax	maximum urinary flow rate
Qave	average urinary flow rate
QoI	quality of life
UTI	urinary tract infection

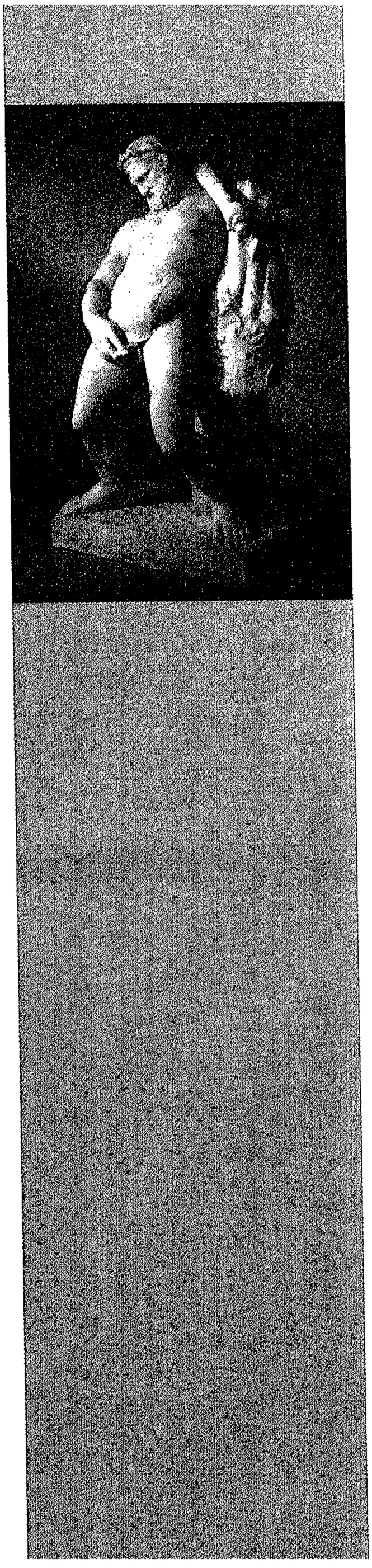
Contents

	Woord vooraf	9
Chapter 1	General introduction	11
Chapter 2	A decision aid for general practitioners for the treatment of elderly male patients with lower urinary tract symptoms	21
Chapter 3	The effectiveness of a treatment protocol for male lower urinary tract symptoms in general practice - a practical randomised controlled trial	37
Chapter 4	Please be seated? Position-related differences in voiding in men with lower urinary tract symptoms - a 20-fold replicated single case experiment	55
Chapter 5	Objective reading of uroflowmetric measurements - increased reliability and precision	73
Chapter 6	Formal quality criteria for websites do not always reflect the quality of their content - the example of benign prostatic hyperplasia	87
Chapter 7	General practitioners, their patients and the Internet - a qualitative study of opinions and attitudes, and their interrelation with the doctor-patient relationship	99
Chapter 8	General discussion	115
	Summary	143
	Samenvatting	153
	Woord van dank	163
	Curriculum vitae	169

Woord vooraf

Meer dan andere soorten curatieve artsen, brengt de huisarts de zorg dichtbij de patiënt. Niet alleen fysiek dichtbij, door 'om de hoek' te praktiseren of een huisbezoek af te leggen, maar ook figuurlijk dichtbij, door samen de bij de patiënt passende vertaling van de klacht in een medisch probleem te maken en vervolgens een aanpak van de klacht te kiezen. Daarbij is het nodig om een probleem van verschillende kanten te benaderen. Verschillende patiënten vereisen een verschillende aanpak.

Zo heb ik ook het onderwerp van dit proefschrift benaderd. Van verschillende kanten kijkend en met verschillende manieren van aanpak onderzoekend. Ik hoop u mee te kunnen nemen in deze verschillende invalshoeken, waarbij uiteindelijk ook de 'helicopterview' gekozen wordt. Ik wens u veel leesplezier toe, en vergeet uw kritische blik niet!



CHAPTER 1

General introduction

General introduction

outline of the study domain

For most people who have grown out of their early childhood diapers micturition and urinary continence have become self-evident phenomena. This is especially the case for men. Women may experience the vulnerability of the mechanisms of urine storage and voiding during or after a pregnancy. Men however, are generally less acquainted with micturition problems during the first decades of their lives. Instead, they are more likely to contest with each other on the ability to urinate furthest, both literally and metaphorically.

During life the quality of urine storage and voiding often changes. The muscular bladder wall becomes less flexible due to an increase in connective tissue. This leads to a need for an increased intravesicular pressure to stretch the bladder wall. The prostate grows under the influence of testosterone and its derivatives. The resulting central enlargement of the prostate may narrow the prostatic part of the urethra, and increase the urethral resistance (i.e. obstruction). All these changes take place very, very slowly. The evolution from 'health' to 'disease' - if we may call it a disease at all - takes place very gradually over decades.

Like other urogenital ailments, micturition problems often are a taboo. Most people hesitate to talk about it with others, be they friends, relatives or health care professionals. Micturition is a private matter, and a solitary one. People lack a proper reference standard for self-assessment of the normality of one's micturition behaviour. One may compare his micturition pattern with that of an earlier age, but many men consider this not correct. Some alteration of the quality of voiding due to or corresponding with age is considered normal. Therefore they may also compare it with an (assumed) pattern 'corrected for age'. In other words, elderly male are also comparing their voiding with the way they suppose others at their age will void. This latter supposed pattern is considered to be 'normal' and consequently the reference standard. The thus constructed reference standard is not validated, lacks precision, and is often biased. It is highly subjective, and often influenced by some sort of wishful thinking, be it for the better or the worse.

During one of my contacts with the participants of the randomised controlled trial, one subject told me that he always used a small PVC-pipe when voiding. He used this to reduce the volume that was spilled on the floor when voiding. When travelling, this small pipe was not useful: it became dirty, and could then not be stored in a convenient way. Therefore he used old toilet paper rolls as a disposable alternative. He suggested I would recommend this to other patients and gave me a test sample.

Impaired micturition quality may become bothersome or be dealt with in sometimes very creative ways (see box).

The bother caused by the symptoms may be very situational. A certain micturition pattern may be difficult to handle in some circumstances and remain unnoticed in other situations. Indoors, at one's own home, an altered micturition pattern often has few consequences. Elsewhere, this may be different. It is very unpleasant not to be able to see a play in the theatre without thinking of the shortest way to the toilet during the third act. The consequences may reflect on the subject himself, but on the partner as well. She or he may be the one who complaints of yellow spots in trousers and dirty toilets. She or he can urge him to void sitting for reasons of hygiene.

With regard to micturition the optimal state of health is clear. However, the state of 'disease' is not defined at all, except for the sudden inability to void at all (acute urinary retention), or frequent recurrent painful infections (urinary tract infections or prostatitis). A static anatomical model may define the state of disease as an enlarged prostate (benign prostatic hyperplasia). In a dynamic physiologic model the 'disease' is considered to be the state of low urinary flow rate due to insufficient bladder contraction, increased urethral resistance due to compression (infravesicular outflow obstruction), or a combination of both. Psychological and cultural models will explain the 'disease' from their consequences for daily life (bother) or the relation with others (interpersonal consequences).

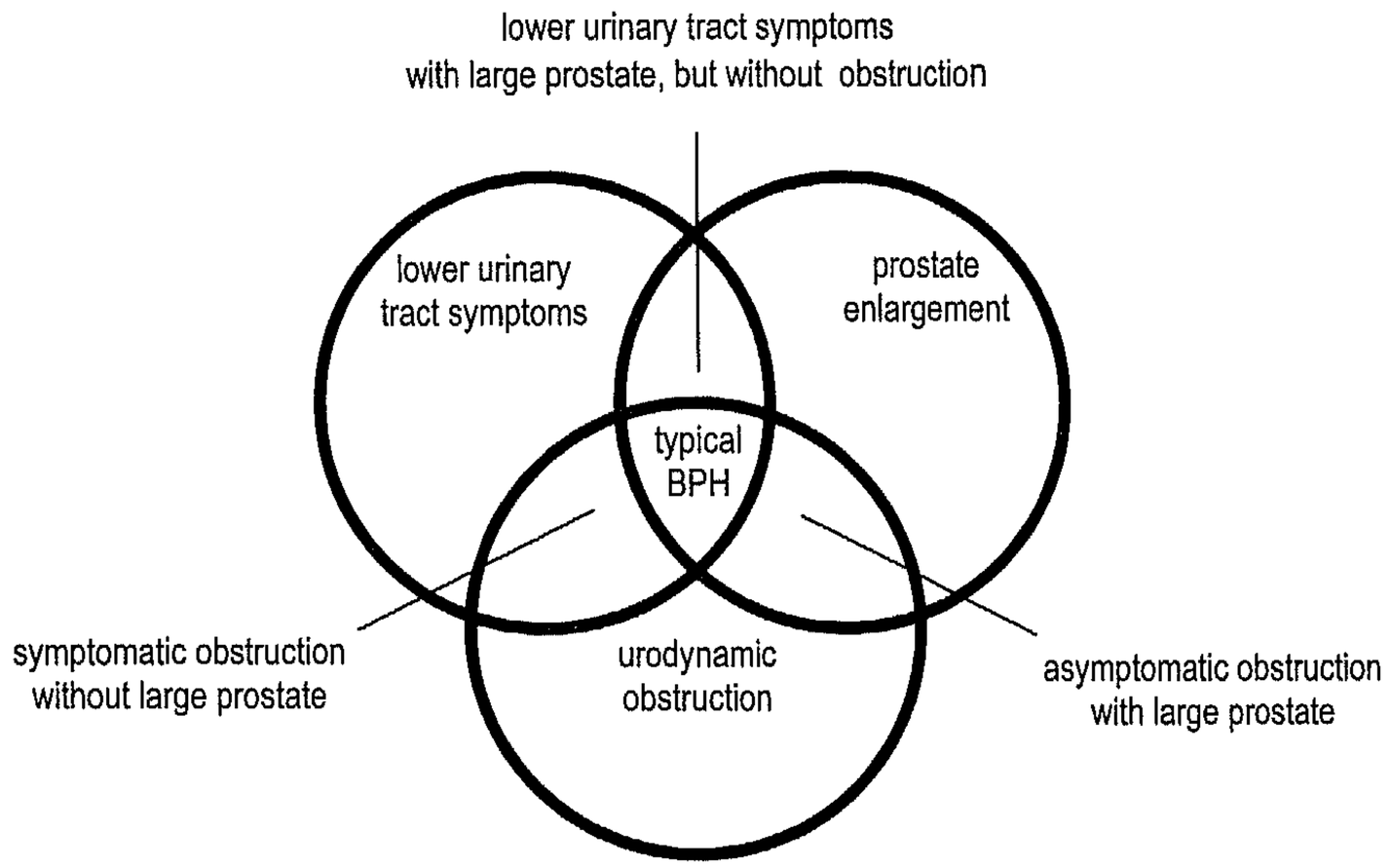
the concept of lower urinary tract symptoms

Medical models require physical and/or mental targets for intervention by doctors as explanatory mechanism. In former days, doctors spoke of the 'disease' as benign prostatic hyperplasia (BPH) with corresponding symptoms of prostatism. Later, when there appeared to be only a poor to moderate correlation between symptoms and 'objective' parameters measured with urinary flow measurement, urodynamics and ultrasound, the medical profession started to use a first abstraction of the anatomic concept: *clinical* BPH. The relation and discrepancies between the main aspects of this concept of clinical BPH were visualized in three overlapping circles (figure 1.1).

At present, in the predominant medical model the 'disease' is called 'lower urinary tract symptoms', which is a second step in the abstraction from the physical target of medical intervention to the conceptualisation of the 'disease'. This term immediately shows that this concept consists of a combination of symptoms rather than a well-defined disease. The consequence of this term is that the translation of this 'disease' concept to anatomic and physiologic targets for diagnosis and treatment has become obscure. One way or the other however, doctors need such a translation. Their job requires them to go from the symptoms of the patient to a certain state or process in his body that should be altered by therapy in order to reduce these symptoms. One could say that the

abstracted 'disease' concept has led to a situation where the concrete (anatomic or physiologic) diagnosis has been left and replaced by a working hypothesis. Consequently the target of the diagnostic work-up has returned from the general debate in medical science to the consulting room where the physician meets with the individual patient.

figure 1.1: Interrelated aspects of (the function of) the male urogenital system, leading to different concepts of 'disease' (after Hald[1])



'Typical BPH' is defined here as: presence of lower urinary tract symptoms, urodynamic obstruction and prostate enlargement. Large-scale general population studies to reveal the 'real sizes' of the three overlapping circles are difficult to perform due to the invasive nature of urodynamic measurements. However, using urinary flow measurement as the 'best (ethical and practical) proxy' one may conclude that the circles differ very much in size and are age[2] and population-specific.

the context of the present studies: general practice

General practitioners in The Netherlands are specialists in integrated medicine. One of the key qualities of a general practitioner is the ability not only to combine, but to integrate information of different natures from different sources. A personal contact in or nearby the natural habitat of the patient facilitates this. This may be used for diagnostic, prognostic and treatment purposes. For instance, chronic diseases may not be cured, and symptoms not reduced, but their bother may be diminished by non-physical interventions.

General practitioners are also specialists in uncertainty medicine. With limited technical means they try to come from the reason for encounter of the patient to a reasonable working strategy. This strategy is based on the most likely hypothesis ((differential) 'diagnosis') in the absence of serious (e.g. life-

threatening) other possible explanations of the condition of the patient. This most likely hypothesis may be a sound or probable diagnosis, only a possibility, or just a certain physical state that justifies a certain action. This latter situation often is the case when clear symptoms can be caused by several diseases, but require the same (initial) treatment. Uncertainty about the diagnosis then is no problem, provided the symptomatic treatment is correct and follow-up secured.
General practitioners are gatekeepers. This does not mean that their task is to secure the gate and to keep people out of the health care system, but to show them the right door leading to health improvement. This door may be one of the many doors of a hospital, but also that of other health care professionals, the GP's own door, or - very often - the door of the patient himself. The gatekeeping function leads to a selection of patients with similar diseases and similar stages of the disease visiting the specialist.
Finally, general practitioners are 'life-time' doctors. They are supposed to offer continuity. Although this has somewhat eroded during the past decades, individual general practitioners often are recognized as 'their doctor' by patients. The fact that the GP almost always is the first health care professional that a patient contacts for health problems, whatever their nature, contributes to this. Continuity also offers the option of 'watchful waiting', a strategy in which the bottom-line is: wait until the stage of the disease becomes serious enough to do what at present is not (yet) indicated. Such a watchful waiting strategy may or may not include provisional interventions (e.g. lifestyle advices). It requires long-time follow-up sometimes. This aspect of continuity requires some structural improvement in general practice, and a change of attitude of general practitioners. They often leave the initiative to the patient and are not very used to make pre-planned appointments for the follow-up of a condition.

current trends

In the past decades three main trends have influenced the management of lower urinary tract symptoms in general practice. One was mentioned above: the redefinition of the concept of lower urinary tract symptoms. A second is the introduction of medication for the treatment of lower urinary tract symptoms, especially alpha-adrenergic receptor blocking agents (α-blockers). As a result lower urinary tract symptoms have become a medical problem not only for secondary, but also for primary care. The third is an increased awareness of the problem among both elderly men and primary care physicians. As a result the average man who consults his doctor for lower urinary tract symptoms is younger, and less symptomatic. In conclusion: GPs treat new groups of patients with new stages of the 'disease' with new medication in a new setting.

guidelines

In many countries primary care physicians and urologists have tried to formulate evidence-based guidelines for the management of lower urinary tract symptoms. These guidelines vary somewhat due to cultural differences between nations or different health care systems in the various countries. Most guidelines however, share - sometimes implicitly, sometimes explicitly - three elements:

1. A urinary tract infection or prostate cancer should be excluded;
2. Treatment should focus on the reduction of (bother due to) symptoms;
3. Treatment options should be explained; the choice for a specific treatment should be made by the patient.

treatment options for lower urinary tract symptoms

Lower urinary tract symptoms can be treated with surgery, medication and non-pharmacological interventions (lifestyle advices). The most well-known surgical technique is the trans-urethral resection of the prostate. During this procedure the surgeon enters the urethra with a small dissecting loop to slice pieces of tissue from the prostate. This way, the prostatic urethra is widened, and an obstruction removed. The main side effects are incontinence and erectile dysfunction. Restenosis of the urethra due to scar formation may lead to a reobstruction and consequently to a recurrence of voiding complaints. Medication consists of 5-α-reductase inhibitors (finasteride, dutasteride) and α-blockers (alfuzosine, doxazosine, prazosine, tamsulosine en terazosine). Apart from these medicines several other agents are used to 'improve prostate functioning' such as serenoa repens and bèta-sitosterol. These are not officially registered in The Netherlands, but available in over-the-counter preparations, often in combinations with other agents. There is no inventory of the non-pharmacological interventions performed by Dutch doctors, but a British study showed that many different interventions exist. The Dutch College of General Practitioners (NHG) recommends the following: void regularly, and do not postpone voiding; take enough time to void; empty your urethra completely ('urethral milking'); take care with alcohol; void sitting on a (conventional) toilet. Most of these recommendations are not evidence-based. Most of them have hardly been studied at all. (And although this was acknowledged by the last revision of the NHG-guideline these recommendations can still (November 2007) be found in the NHG-patient information leaflets.)

research questions and outline of the thesis

The guidelines for lower urinary tract symptoms are patient-centred, and promote 'shared decision making'. Patients, however, can only really take part in the decision making if they are provided with clear information requiring their subjective judgement. Up till now, the scientific body of knowledge fails to provide both doctors and patients with important pieces of information to make real, evidence-based choices. Determinants of prognosis are still under investigation, the indications for the different treatment options are not well defined, and it is unclear to what extent efficacy in randomised controlled trials leads to effectiveness of treatment in daily primary care. Non-pharmacological interventions are often used, but have hardly been investigated.

In conclusion: although much information is available, it is difficult to find useful - that is: true and applicable – information, especially in the primary care situation. This thesis consists of 'applied research' studies, answering research questions emerging from, and with consequences for, daily care for elderly male patients with lower urinary tract symptoms in general practice.

Chapter 2 describes the development of a tool for general practitioners to improve their treatment strategy for elderly patients with lower urinary tract symptoms.

Chapter 3 presents a pragmatic randomised controlled trial of the effectiveness of a treatment protocol for male lower urinary tract symptoms in a primary care setting.

In Chapter 4 I present the results of a 20-fold replicated randomised single case experiment, and its meta-analysis. This experiment was designed to investigate the best position for voiding in male patients with moderate lower urinary tract symptoms.

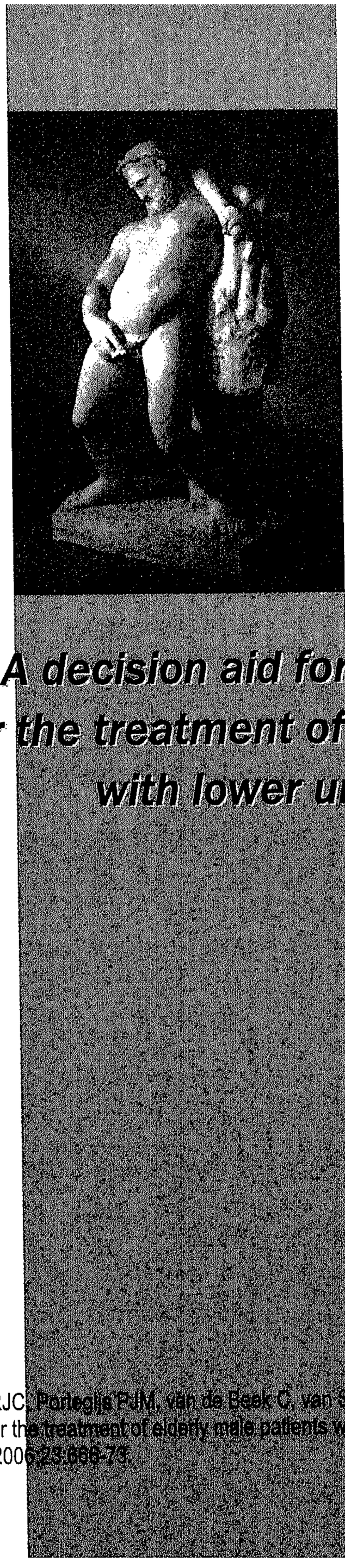
The next chapter, chapter 5, forms a spin-off of our single case experiments. We studied the need and possibilities to enhance the precision of urinary flow measurements.

Finally, Chapters 6 and 7 deal with the value and influence of health information on the Internet for daily practice. In chapter 6 we investigate the relation between formal quality criteria for websites regarding male lower urinary tract symptoms and the quality of their content. Chapter 7 describes a qualitative study into the perceptions and experiences of both patients and general practitioners with regard to Internet in general and its influence on the doctor-patient relationship in particular.

The general discussion of this thesis reviews the results of these studies and their consequences for daily care in general practice, and can be found in chapter 8.

References

1. Hald T. Urodynamics in benign prostatic hyperplasia: a survey. Prostate (Suppl.) 1989;2:69-77.
2. Blanker MH, Groeneveld FP, Prins A, Bernsen RM, Bohnen AM, Bosch JL. Strong effects of definition and nonresponse bias on prevalence rates of clinical benign prostatic hyperplasia: the Krimpen study of male urogenital tract problems and general health status. BJU Int 2000;85(6):665-71.



CHAPTER 2

A decision aid for general practitioners for the treatment of elderly male patients with lower urinary tract symptoms

Norg RJC, Portegijs PJM, van de Beek C, van Schayck CP, Knottnerus JA. A decision aid for GPs for the treatment of elderly male patients with lower urinary tract symptoms (LUTS). *Fam Pract* 2006;23:666-73.

Abstract

background:

General practitioners have four main treatment options for lower urinary tract symptoms (LUTS): watchful waiting, α-blockers, 5-α-reductase inhibitors, or (referral for) surgery. Guidelines do not provide clear cut-off values for (combinations of) symptoms and physical examination results to decide which treatment is best.

objective:

1. To develop a decision aid ('checklist') for general practitioners for the treatment of patients with LUTS.

2. To assess its value for use in a primary care population.

methods and materials:

Population-based cross-sectional study. Included were subjects with uncomplicated LUTS for whom treatment in primary care may be appropriate. (International Prostate Symptom Score (IPSS) ≥ 8, no prior prostate surgery, PSA value < 10 ng/ml). For each subject the appropriateness of surgery and α-blocker treatment was determined using a previously validated formalised international expert panel judgement. Regression models using data available in primary care were constructed to predict the panel judgement. Subsequently these models were transformed into simple checklists. Finally, the efficiency of these checklists was calculated.

results:

The best checklists consisted of age, symptoms severity, type of symptoms, a quality of life score, and PSA-value. Assuming one would like to provide at least 95% of the subjects for whom a certain treatment is appropriate with this treatment (i.e. 'sensitivity' of the checklist ≥ 95%), one can reach a positive predictive value of 50-60%.

conclusion:

Simple checklists based on the judgement of experts regarding the most appropriate therapy can help general practitioners to advise their patients a rational treatment strategy.

Introduction

Lower urinary tract symptoms (LUTS) are becoming an increasing health problem in the elderly male. Ageing and more awareness and recognition among both patients and doctors have led to an increase in the demand for and costs of diagnosis and treatment of LUTS.[1,2] Almost half of the new patients can be treated in primary care.[3] New patients tend to be younger and present themselves with less severe symptoms than before.[4] Therefore, the involvement of general practitioners (GPs) is expected to grow in the years to come.

Primary care guidelines for LUTS often start with excluding conditions that can be easily diagnosed in primary care (e.g. urinary tract infections (UTI) or acute urinary retention (AUR)), or should be referred to secondary care for further diagnosis because of the seriousness of the disease (e.g. prostate cancer). The rest of the patients are regarded to have LUTS 'suggestive' for benign prostatic hyperplasia, also referred to as uncomplicated LUTS. These patients may be treated with watchful waiting (WW), α-blockers (AB), 5-α-reductase inhibitors (5ARI), or surgery. Current guidelines offer little help to primary care physicians in deciding which of these treatments to choose. They leave much room for interpretation and shared decision making by the GP and the patient.[2,3,5]

In the process of shared decision making, the GP would be helped if he could fall back on a 'reasonable' treatment strategy. Such a strategy would be to anticipate the expected treatment strategy of the specialist. This means that the GP asks two questions. One is: *Should I refer?* A second is: *Should I myself treat?* In case of LUTS these questions are specified to: *Would surgery (probably) be appropriate?*, and: *Should I prescribe an α-blocker?*, or - alternatively - *Should I maintain a watchful waiting strategy?* (5-α-reductase inhibitors have limited value in primary care.[5]) The order of these questions may vary. A general practitioner may first determine whether or not to refer and subsequently determine the appropriateness of α-blocker treatment; or he may also first prescribe an α-blocker if appropriate, and (if it should fail) refer for surgery if indicated. Some general practitioners may prefer a specific strategy; others may leave the choice to the patient, who may either have a preference for a quick and definitive solution if indicated, or may want a less invasive approach if possible. Both treatment strategies therefore are defendable, depending on the circumstances.

The question whether surgery or an α-blocker would be appropriate was answered in an international study that used a formalised expert opinion procedure.[6,7] In this study, 39 urologists considered vignettes of men ≥ 50 years of age, referred by their GP for the first time with -according to the first general diagnostic evaluation- uncomplicated LUTS, and no influencing co-morbidity or medication. Each case had a different combination of a total of eight diagnostic

variables. For each case the experts judged a treatment as 'appropriate', 'inappropriate', or 'uncertain' compared to watchful waiting.
Such a study revealing the general opinion of an international expert panel could be helpful to general practitioners to determine their treatment strategy. However, application in the primary care setting is problematic. Usually only five of these diagnostic variables are available in primary care (i.e. age, IPSS score, type of symptoms, bother, and history of urinary tract infections); two require specialised diagnostic procedures that are usually not available in primary care (maximum urinary flow (Q_{max}) and post-void residual volume (PVR)); one is measured differently in primary care (prostate size via digital rectal examination vs. ultrasound). Furthermore, the case-mix in primary care differs from that in secondary care. Most patients in primary care have only mild to moderate LUTS. Consequently, a modification and validation of this expert opinion for the primary care setting is needed, preferably into a simple checklist with (weighted) key symptoms and investigations, and clear cut-off values determining the best treatment option available.
We developed such a checklist in a population-based sample of elderly males with LUTS. In this article we report on its development, and provide an estimate of its value for the primary care situation.

Methods and materials

study population

Our source population consisted of all men aged 55 or older registered with 14 general practices. They were invited by mail to participate in a cohort study. During a qualification period the participants repeatedly filled in questionnaires (International Prostate Symptom Score (IPSS), Danish Prostate Symptom Score (Dan-PSS), micturition behaviour), underwent urinary flow measurement, and physical examination. Also, prostate specific antigen (PSA) and creatinin were measured, and the subjects kept a 3-day micturition diary and underwent ultrasound measurement of PVR and prostate size. Included in this analysis were subjects with untreated uncomplicated LUTS, i.e. an IPSS ≥ 8 points, no prior prostate surgery or current use of medication for LUTS, and a PSA value < 10 ng/ml. Consulting for LUTS was not an inclusion criterion.

analysis

We used the combined opinion of the expert panel as published in the studies of Stoevelaar[6] and McDonnell[7] to determine the appropriateness of α-blockers and surgery for each member of our study sample. Subsequently, we constructed logistic regression models that predicted whether or not a certain treatment would be appropriate or not (that is: 'appropriate' versus 'inappropriate' or

'uncertain'). We used a backward stepwise approach with -2 log likelihood ($p < 0.05$) as criterion. Only data available in primary care were used in the primary analyses, although we also constructed similar models including Q_{max} to investigate the added value of urinary flow measurement. We checked for collinearity between the included variables. We also compared the Receiver Operating Characteristic curves (ROC) of the models, and calculated the area under these curves (AUC). ROC curves indicate the sensitivity and the corresponding specificity at various cut-off points of a test (i.e. our regression models and our checklist). The AUC indicates the diagnostic accuracy of the checklists. An AUC of 0.50 indicates that the checklist cannot discriminate between appropriate and inappropriate treatment. An AUC of 1.0 would imply perfect discriminative power. In a second step, the final models were transformed into simpler checklists, preferably based only on simple dichotomous items.

During the analysis we followed the two possible strategies described in the introduction. According to strategy 1 (referral first strategy) we first constructed regression models to predict the appropriateness of surgery. Subsequently, for those for whom surgery (and therefore referral) was not indicated, we constructed regression models to predict the appropriateness of α-blocker treatment.

In strategy 2 the sequence was reversed (medication first strategy). First we constructed regression models to select those for whom an α-blocker would be appropriate. Within this group we constructed models to predict whether surgery would be appropriate in case of failure of α-blocker treatment.

Having translated the regression models into simpler checklists we - as a preliminary investigation - calculated the sensitivity, specificity and positive predictive value. In this specific setting the sensitivity represents - within the group of subjects for whom a referral (or prescription of medication) would be appropriate - the percentage who - using the checklist - would indeed be referred (or be prescribed medication). The specificity represents - within the group for whom therapy is considered not appropriate - the percentage of subjects that would not receive a treatment. The positive predictive value represents the percentage of referred subjects among those for whom surgery would be appropriate.

Our aim was to create checklists that would have a high sensitivity, whereas the specificity and positive predictive value of the models were less important. Indeed, a general practitioner must select and refer those for whom surgery is appropriate, accepting that a certain percentage of those referred will not be operated at all. This is not problematic, because the urologist makes a second selection using more extensive diagnostic tools (ultrasound, urinary flow measurement, urodynamics). This second selection will guarantee that no

unnecessary operations are executed. (Of course there is a trade-off here: the best way to secure that no patient is withheld possibly beneficial therapy is being not selective at all. That however would mean that a lot of patients would be subjected to unpleasant, costly and unnecessary investigations with the possibility of iatrogenic damage or side-effects.) The general practitioner must however not be too selective. A too selective GP might withhold patients for whom surgery would be appropriate possibly beneficial therapy. A similar reasoning holds for α-blocker treatment. Current guidelines state that -when in doubt- a trial of treatment is appropriate. Consequently, the GP should aim to select those for whom α-blocker treatment will be useful, accepting that some of those for whom these are less appropriate will also receive such a prescription. Overall, in the case of LUTS, the general practitioner in his role of *first-selector out of two* can let sensitivity prevail over specificity and positive predictive value.

Results

Included in the analysis were 512 subjects. Table 2.1 (see page 27) shows the baseline characteristics and their association with the appropriateness of surgery and α-blocker treatment. 432 Subjects (86%) had moderate symptoms. The median age was 66 years. α-Blockers would be appropriate in 86 subjects (17%); surgery in 39 subjects (8%). There was much overlap between these groups: for 37 subjects both surgery and α-blockers would be considered appropriate.

analysis of the treatment strategies

The results of the analysis of the treatment strategies are presented in table 2.2 (see page 28).

1. referral first strategy

The regression model that indicates for whom surgery would be appropriate consists of age, IPSS score ≥ 20, IPSS-QoI score ≥ 4, irritative symptoms, and PSA-value. Inclusion of Q_{max} resulted in a significantly better regression model ($p < 0.001$), which, however, was hardly visible in the ROC curve. For those for whom surgery was not indicated (N=473), a similar model without PSA-value fitted best to predict the appropriateness of α-blocker treatment. Again, inclusion of Q_{max} resulted in a significantly better regression model ($p < 0.001$) with a comparable area under the ROC curve.

2. medication first strategy

The best fitting models included age, IPSS score ≥ 20, IPSS-QoI score ≥ 4, and irritative symptoms for α-blocker treatment, and the same variables plus PSA-value and prostate size for surgery. Inclusion of Q_{max} resulted in a significantly better regression models ($p < 0.001$), with comparable AUCs.

table 2.1. Baseline characteristics of the study population, and results of bivariate analysis between determinants and appropriateness for surgery and α-blocker treatment.

variable	surgery			α-blocker			total study population
	appropriate (N=39)	inappropriate / uncertain (N=473)		appropriate (N=86)	inappropriate / uncertain (N=426)		(N=512)
	median [iqr]	median [iqr]	p-value *	median [iqr]	median [iqr]	p-value *	median [iqr]
age (years)	**64** [60 – 68]	**66** [62 – 72]	0.06	**64** [60 - 71]	**66** [62 - 72]	0.86	**66** [62 – 72]
IPSS	**23** [21 – 27]	**12** [10 – 15]	< 0.001	**22** [20 - 25]	11 [9 - 14]	< 0.001	**12** [10 – 17]
Dan-PSS	**24** [17 – 38]	**8** [4 – 15]	< 0.001	**20** [12 - 31]	**8** [4 - 14]	< 0.001	9 [4 – 16]
prostate size (ultrasound, ml)	**30** [24 - 47]	**26** [20 - 35]	0.02	**28** [23 - 38]	**26** [20 - 36]	0.13	**26** [20 - 36]
PVR (ultrasound, ml)	**63** [17 - 166]	**29** [13 - 62]	0.006	**41** [17 - 104]	**29** [12 - 61]	0.02	**32** [13 - 68]
PSA (ng/ml)	**2.4** [0.8 – 3.4]	**1.3** [0.7 – 2.8]	0.1	**1.8** [1.0 - 3.1]	**1.3** [0.7 - 2.8]	0.12	**1.4** [0.7 – 2.8]
Qmax (ml/s)	**9.6** [6.3 – 13.9]	**12.1** [8.6 – 16.2]	0.01	**10.1** [7.1 - 15.3]	**12.3** [8.8 - 16.4]	0.02	**11.9** [8.5 – 16.0]
	N [%]	N [%]		N [%]	N [%]		N [%]
prostate size #							
small	**5** [14%]	**64** [14%]	0.005 **	**10** [12%]	**59** [14%]	0.05 **	**69** [14%]
normal	**7** [19%]	**209** [46%]		**24** [29%]	192 [46%]		**216** [44%]
enlarged	**19** [51%]	**152** [33%]		**37** [43%]	**134** [33%]		**171** [35%]
very large	**6** [16%]	**31** [7%]		**10** [12%]	27 [7%]		**37** [8%]
age < 70 years	**32** [82%]	**320** [68%]	0.06	**62** [72%]	290 [68%]	0.46	352 [69%]
IPSS ≥20	**36** [93%]	**34** [7%]	< 0.001	**67** [78%]	**3** [1%]	< 0.001	**70** [14%]
dominating irritative symptoms	**30** [77%]	**67** [14%]	< 0.001	**58** [67%]	**39** [9%]	< 0.001	97 [19%]
IPSS-QoI ≥ 4	**12** [31%]	**6** [1%]	< 0.001	**16** [19%]	2 [0%]	< 0.001	**18** [4%]
Qmax (ml/s)							
< 10	**7** [18%]	**151** [32%]	0.006 **	**22** [26%]	**136** [32%]	0.01 **	**158** [31%]
10-15	**10** [26%]	**164** [35%]		**21** [24%]	**153** [36%]		**174** [34%]
> 15	**22** [56%]	**158** [33%]		**43** [50%]	**137** [32%]		**180** [35%]

iqr: inter quartile range; * Mann Whitney U test; ** Kruskal Wallis test; # according to the judgement of the general practitioner using digital rectal examination.
IPSS: International Prostate Symptom Score; a score ≥ 20 represents severe symptoms; Dan-PSS: Danish Prostate Symptom Score; PSA: prostate specific antigen; Q_{max}: maximum urinary flow rate; IPSS-QoI: IPSS-Quality of Life score; a score ≥ 4 represents a qualification 'mostly dissatisfied', or worse.

table 2.2. Results of the multiple logistic regression analyses.

	referral first strategy				medication first strategy			
	appropriateness of surgery		appropriateness of α-blocker *		appropriateness of α-blocker		appropriateness of surgery **	
	without Q_{max}	with Q_{max}	without Q_{max}	with Q_{max}	without Q_{max}	with Q_{max}	without Q_{max}	with Q_{max}
variable	regression coefficient (p-value)	regression coefficient (p-value)	regression coefficient (p-value)	regression coefficient (p-value)	regression coefficient (p-value)	regression coefficient (p-value)	regression coefficient (p-value)	regression coefficient (p-value)
age (years)	-0.104 (0.02)	-0.137 (0.006)	-0.092 (0.03)	-0.149 (0.003)	-0.076 (0.06)	-0.142 (0.003)	-0.115 (0.01)	-0.158 (0.003)
IPSS ≥ 20	4.452 (< 0.001)	4.674 (< 0.001)	5.871 (< 0.001)	7.302 (<0.001)	6.142 (< 0.001)	7.514 (< 0.001)	4.912 (< 0.001)	5.190 (<0.001)
dominating irritative symptoms	1.739 (0.003)	1.794 (0.004)	3.200 (< 0.001)	3.813 (<0.001)	3.099 (< 0.001)	3.784 (< 0.001)	1.206 (0.04)	1.265 (0.05)
IPSS-QoI ≥ 4	2.143 (0.01)	2.702 (0.002)	2.676 (0.035)	3.683 (0.02)	3.816 (0.001)	4.697 (< 0.001)	2.245 (0.01)	2.707 (0.003)
PSA (ng/ml)	0.378 (0.015)	0.365 (0.02)					0.389 (0.04)	0.343 (0.04)
prostate size 0 = normal; 1 = enlarged							1.298 (0.03)	1.278 (0.05)
Q_{max} 0 = normal (>15 ml/s) 1 = intermediate (10-15) 2 = low (<10))		1.164 (0.001)		1.595 (<0.001)		1.718 (0.001)		1.284 (0.001)
constant	0.162	0.679	1.661	3.045	0.718	2.390	0.150	1.207
AUC (95% CI):	**0.96** (0.92-1.00)	**0.98** (0.96-1.00)	**0.94** (0.90-0.99)	**0.97** (0.95-1.00)	**0.97** (0.95-1.00)	**0.98** (0.97-1.00)	**0.94** (0.89-0.98)	**0.95** (0.91-0.98)

referral first strategy: first, determination of the appropriateness of surgery; second, of the appropriateness of α-blocker treatment; medication first strategy: first, determination of the appropriateness of α-blocker treatment; second, of the appropriateness of surgery.
* for those not referred for surgery; ** for those in whom α -blocker treatment failed.
Q_{max} = maximum urinary flow rate; PSA = prostate specific antigen; AUC = Area under the Receiver Operating Characteristic curve. IPSS: International Prostate Symptom Score; a score ≥ 20 represents severe symptoms; IPSS-QoI: IPSS-Quality of Life score; a score ≥ 4 represents a qualification 'mostly dissatisfied', or worse.

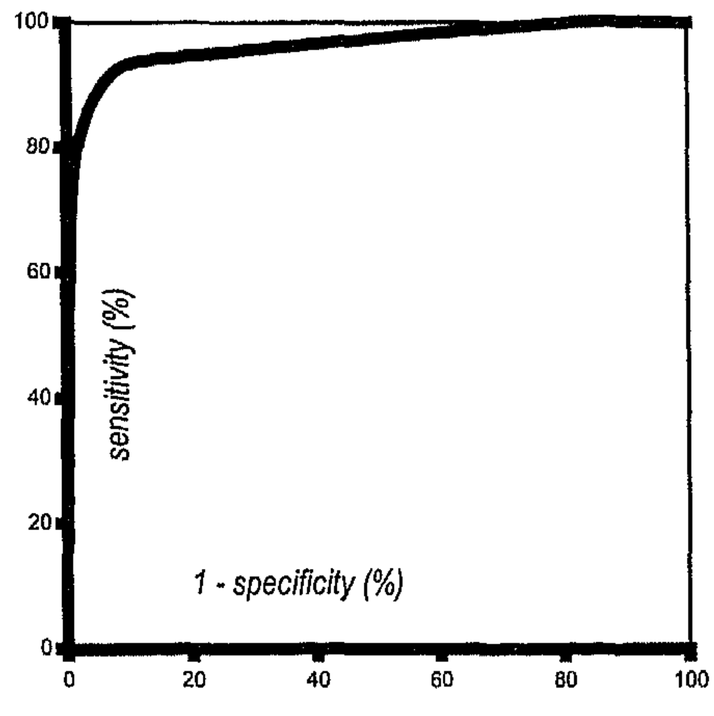
evaluation of the checklists

The most feasible situation for daily practice would be if all regression equations would consist of the same variables, making it possible to transform them into one single checklist with different cut-off values for the different situations. However, this was not possible without considerable loss of predictive value of the regression equations. Our trade-off between feasibility and predictive power is shown in table 2.3 (see page 30). Figure 2.1 shows the corresponding ROC curves.

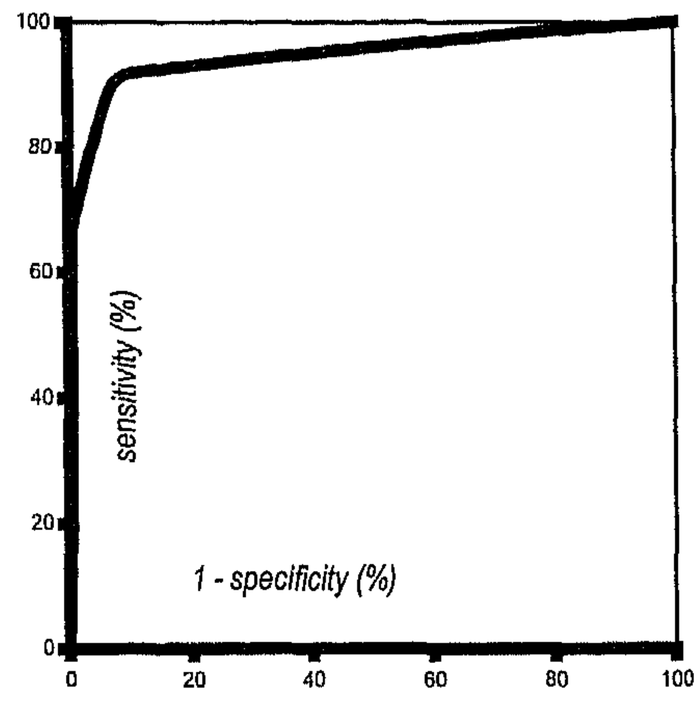
figure 2.1. Receiver Operating Characteristic (ROC) curves of the treatment strategies.

1. ROC curves of the referral first strategy.

checklist 1: surgery appropriate?

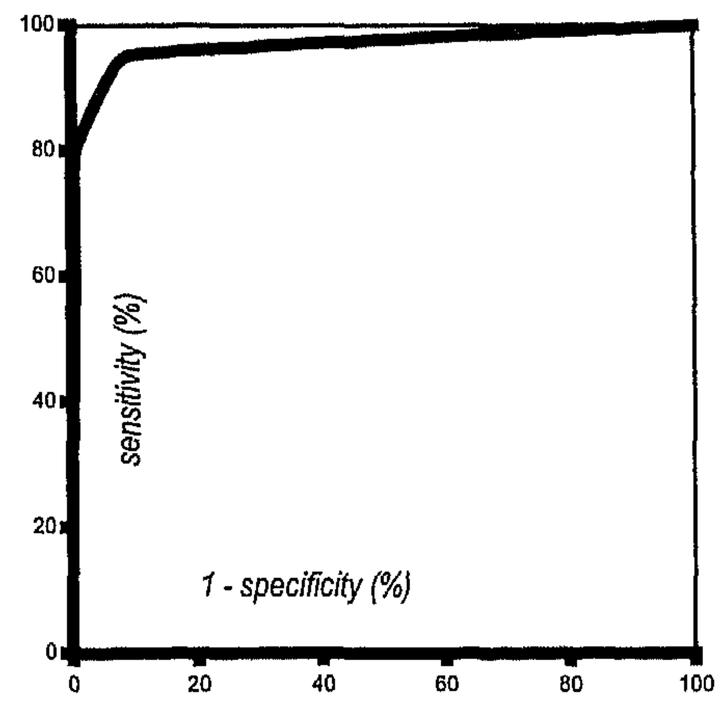


(for those not referred for surgery)
checklist 2: medication appropriate?



2. ROC curves of the medication first strategy.

checklist 3: medication appropriate?



(for those in whom an α-blocker failed)
checklist 4: surgery appropriate?

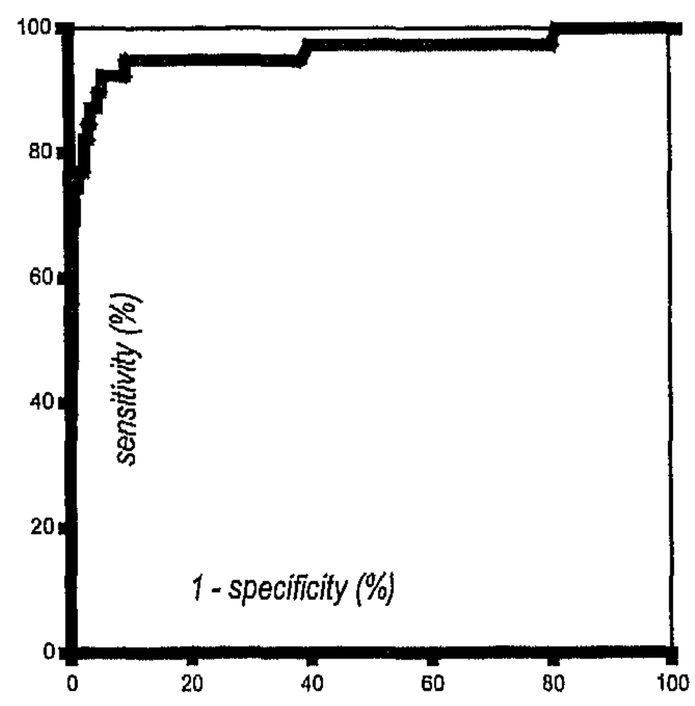


table 2.3. checklists for the referral for surgery and prescription of α-blockers.

	referral first strategy		medication first strategy	
checklist	1	2	3	4
	appropriateness of surgery	appropriateness of α-blockers *	appropriateness of α-blockers	appropriateness of surgery **
variable	weight	weight	weight	weight
age < 70 years	1	1	1	3
IPSS ≥ 20	2	6	6	6
mainly irritative symptoms	1	3	3	2
IPSS-QoI ≥ 4	1	3	3	2
PSA ≥4 ng/ml	1			
PSA ng/ml				0.5
prostate enlarged				1
total range:	*0 - 6*	*0 - 13*	*0 - 13*	*0 - 19*
proposed cut-off value for prescribing treatment:	*3*	*3*	*3*	*7*

The checklists in table 2.3 are simplifications of the regression models presented in table 2.2.

referral first strategy: first, determination of the appropriateness of surgery; secondly of the appropriateness of α-blocker treatment; medication first strategy: first, determination of the appropriateness of α-blocker treatment; secondly of the appropriateness of surgery.
* for those not referred for surgery; ** for those in whom α-blocker treatment failed.
Q_{max} = maximum urinary flow rate; PSA = prostate specific antigen; AUC = Area under the Receiver Operating Characteristic curve. IPSS: International Prostate Symptom Score; a score ≥ 20 represents severe symptoms; IPSS-QoI: IPSS-Quality of Life score; a score ≥ 4 represents a qualification 'mostly dissatisfied', or worse.

example (medication first strategy):
If a patient with uncomplicated LUTS visits his GP, the GP may use checklist 3 to determine whether the patient should be prescribed an α-blocker; if the total score would be high (≥ 3 points), the GP would do so. If the α-blocker would fail, the GP may use checklist 4 to decide whether to refer (high score, ≥ 7 points) or to follow-up without active treatment (watchful waiting) in case of a low score.

1. checklists of referral first strategy

Within the referral first strategy, checklist 1 can be used to select patients for referral. In our population referral of anyone with ≥ 2 points would have a sensitivity of 95% and a specificity of 78% (figure 2.1). This means that in 26% (37/141) of the patients referred for (determination of the eligibility and necessity of) surgery an operation would indeed be appropriate. Using ≥ 3 points as the cut-off these values would be 92%, 92% and 49% (36/73) respectively.

Checklist 2 can be used to decide whom to prescribe medication. In our population prescription of an α-blocker for anyone with ≥ 3 points would have a sensitivity of 92% and a specificity of 90%. In 52% (45/87) of the patients who received a prescription of an α-blocker, this would be appropriate; the others (48%, 42/87) were classified by the expert panel as 'uncertain'; none of them was classified as inappropriate (according to the expert panel). The cut-off of 1 point would almost unselective: even all subjects for whom an α-blocker was considered inappropriate by the expert panel would receive this medication.

2. checklists of medication first strategy

Checklist 3 can be used to select those for whom an α-blocker would be appropriate. It is identical to the situation in which already a selection of those for whom surgery would be indicated had taken place. In our population prescription of an α-blocker for anyone with ≥ 3 points would have a sensitivity of 95% and a specificity of 90%, and a positive predictive value of 66% (82/124). Again, the others (34%, 42/124) consisted only of people for whom the appropriateness was uncertain.

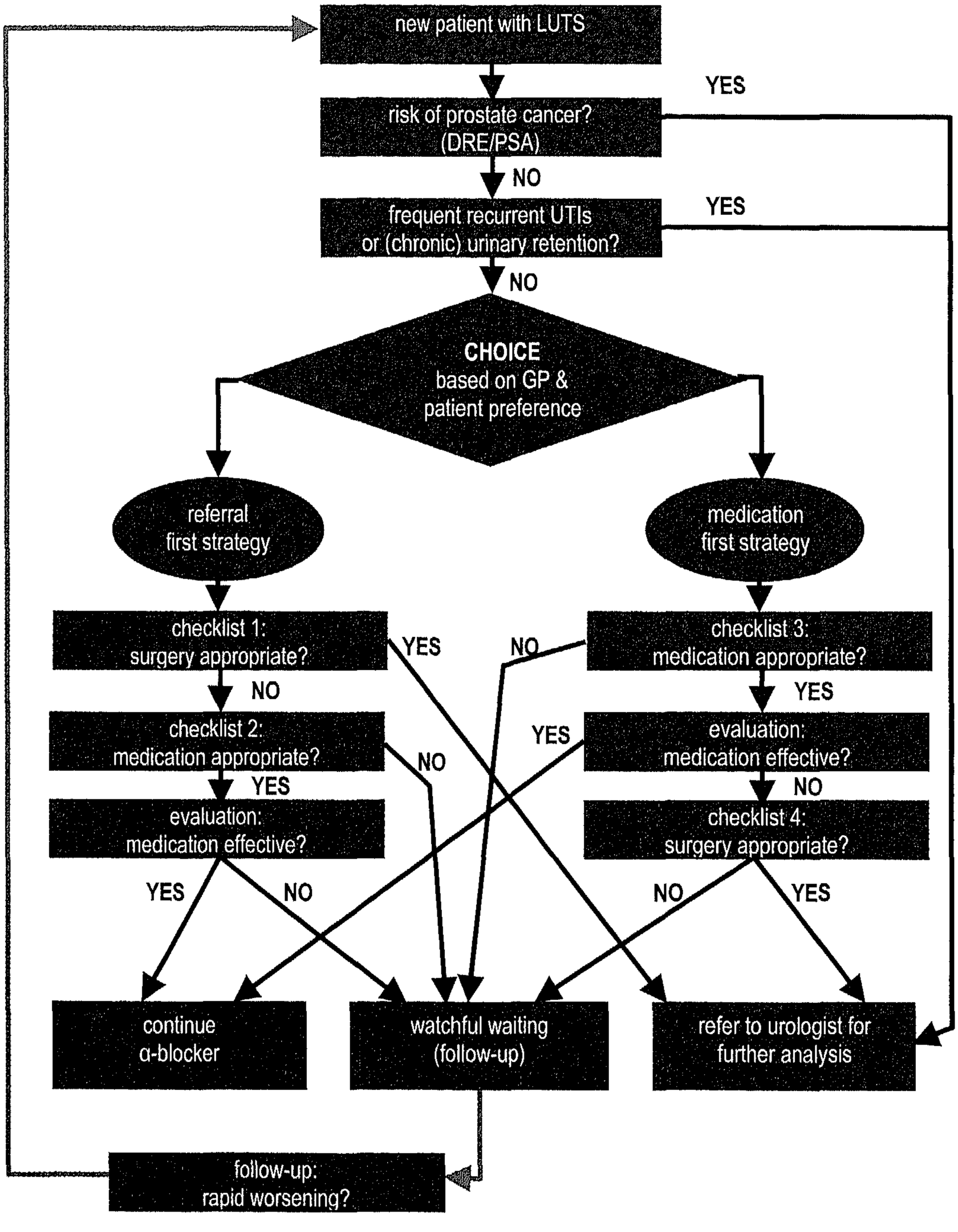
Since surgery would only be considered in those situations that an α-blocker has failed, we considered the sensitivity and specificity in the 124 subjects that according to the checklist 3 would have received an α-blocker. Using ≥ 7 points on checklist 4 as the cut-off, there would be a calculated sensitivity of 95%, and a corresponding specificity of 90%.

decision aid

All this can be summarised in the following scheme (figure 2.2, page 32).

figure 2.2. Decision aid for a rational treatment strategy for lower urinary tract symptoms in primary care.

Depending on the preferences of both general practitioner (GP) and patient, a referral first strategy (left) or medication first strategy (right) may be followed:



DRE = digital rectal examination; PSA = prostate specific antigen; UTI = urinary tract infection; GP = general practitioner; LUTS = lower urinary tract symptoms

Discussion

main results

We developed a decision aid using an approach in which general practitioners anticipate on the probable treatment of the urologist (figure 2.2, page 32). Using simple checklists general practitioners can follow a rational treatment strategy without gross overprescription or undertreatment of patients.

strengths and limitations

Any research using consensus procedures or panel judgements to define the presence or absence of a health status or - like in this study - the appropriateness of a certain treatment, is subject to changes due to new insights and developments in medicine. However, there has not been a recent major breakthrough.

Our study population was based on the general population. This is very similar to actual everyday patients in primary care where new patients with LUTS tend to be less symptomatic than in former days.

Our study was based on the results of a study with a Dutch expert panel, which was subsequently validated with an international panel of urologists. Our results can therefore be generalised beyond the Dutch situation. Off course, a local urologist may have different opinions than the expert panel. Opinions and daily practice have been shown to differ sometimes substantially. Even then, however, the strategies proposed here are defendable due to the multinational composition of the expert panel.

We developed the checklists and subsequently calculated the effects of their use with regard to the number of adequate referrals and prescriptions. These figures were not calculated in a separate 'test' sample, but in the same sample in which the models were constructed. They therefore must be regarded as (somewhat optimistic) indications of the real sensitivity, specificity and predictive value. We did not try to quantify the included optimism. The checklists should be validated in other primary care populations, ideally with real decisions of an expert panel as outcome parameter. Real patients rather than plain paper vignettes are necessary to be able to calculate proper predictive values. These require knowledge of the distribution of (combinations of) different variables in the studied population. Such studies however, are difficult to perform and probably expensive.

The expert panel classified the cases into three categories: appropriate, uncertain and inappropriate. In a primary care situation, and for this ailment, it is reasonable to treat the 'uncertain' group as (for the time being) 'not appropriate'. In case of referred patients (secondary care) the primary care physician has already decided not to follow a watchful waiting strategy. It is

therefore logical that urologists who take their primary care colleagues seriously want to further investigate whether there are specific circumstances that may justify a certain treatment in an individual patient. In a primary care situation however, it is often justified to start with a watchful waiting strategy, provided this includes a proper follow-up policy. The in- and exclusion criteria of our study signify that our study does not apply to patients with mild symptoms (IPSS < 8), nor to situations of acute urinary retention or active urinary tract infections, which require different therapies. It furthermore presupposes that the GP has checked for (a high risk for) prostate cancer first (using digital rectal examination and PSA measurement). It is however applicable in the situation in which an elevated PSA (between 2.5 and 10 ng/ml) requires follow-up ('watchful waiting') for prostate cancer (but not necessarily also for LUTS).

what this study adds

In a process of shared decision making, the patient should make a personal decision, weighing the information provided by the health professional. Preferably, this information is based on evidence from randomised controlled trials with hard clinical endpoints performed in a similar population and under similar circumstances. This means: effectiveness studies in primary care. Lacking such trials we think our study is the next best option to create a rational and transparent treatment strategy. Rationality and transparency are important conditions for adequate patient information. Our study shows that such an opinion-based guideline can be translated into a simple and practical tool for primary care. It can be used in the overwhelming majority of patients consulting with LUTS. We consider our checklist feasible for daily practice since the IPSS and IPSS-QoI consist of only a total of 8 questions. These are designed to be filled out by patients themselves. The regular use of the IPSS/IPSS-QoI has been recommended in primary care guidelines.[3,4,8] The time required is less than one minute. The length of a regular consultation will therefore not increase considerably.

Our checklists predict the treatment policy of urologists well. The added value of urinary flow measurements is too small to promote its use for improving decision making in general practice.

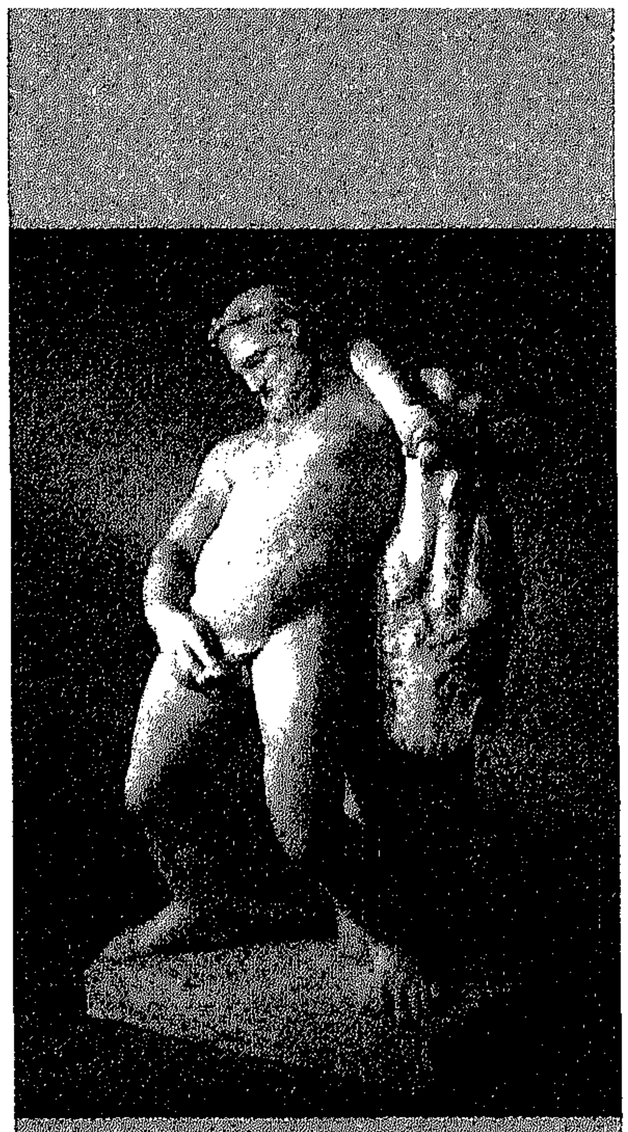
We analysed two different strategies, which in principle are equally defendable. However, our results suggest that the medication first strategy is the most sensible of the two. This strategy guarantees the most efficient trial of α-blockers as recommended by the current guidelines.[3]

conclusion

Using short checklists general practitioners can follow a reasonable strategy for the treatment of LUTS. However, our decision aid should be challenged and validated in other primary care settings and populations.

References

1. Department of Health. Prescription Cost Analysis, England 2003. http://www.dh.gov.uk/PublicationsAndStatistics/Publications/PublicationsStatistics/PublicationStatisticsArticle/fs/en?CONTENT_ID=4081720&chk=kV0up3. (Accessed Nov 1 2006)
2. De la Rosette JJ, Madersbacher S, Alizivatos G, Rioja Sanz C, Emberton M, Nordling J. *Guideline on benign prostatic hyperplasia*: European Association of Urology, Dec 2004.
3. Speakman MJ, Kirby RS, Joyce A, Abrams P, Pocock R. Guideline for the primary care management of male lower urinary tract symptoms. *BJU Int* 2004;93(7):985-90.
4. Bruskewitz R. Management of symptomatic BPH in the US: who is treated and how? *Eur Urol* 1999;36 Suppl 3:7-13.
5. Wolters RJ, Spigt MG, van Reedt Dorlant PF, Gercama AJ, Klomp ML, Romeijnders AC, et al. NHG-Standaard Bemoeilijkte mictie bij oudere mannen. *Huisarts en Wetenschap* 2004;47(12):571-86.
6. Stoevelaar HJ, Van de Beek C, Casparie AF, McDonnell J, Nijs HG. Treatment choice for benign prostatic hyperplasia: a matter of urologist preference? *J Urol* 1999;161(1):133-8.
7. McDonnell J, Stoevelaar HJ, Bosch JL, Kahan JP. The appropriateness of treatment of benign prostatic hyperplasia: a comparison of Dutch and multinational criteria. *Health Policy* 2001;57(1):45-56.
8. American Urological Association. Guideline on the management of Benign Prostatic Hyperplasia (BPH), 2003 (accessible via www.auanet.org).



CHAPTER 3

The effectiveness of a treatment protocol for male lower urinary tract symptoms in general practice - a practical randomised controlled trial

trial registration number: NCT00111592

Norg RJC, van de Beek C, Portegijs PJM, van Schayck CP, Knottnerus JA. The effectiveness of a treatment protocol for male lower urinary tract symptoms in general practice: a practical randomised controlled trial. *Br J Gen Pract* 2006;56:938-44.

Abstract

background:

Randomised controlled trials have shown the efficacy of several treatment modalities for lower urinary tract symptoms (LUTS) in selected populations. The effectiveness in daily practice has hardly been investigated, especially in primary care. It is dependent on choices between all possible treatment options and is best investigated in a comprehensive study, including all treatment modalities (watchful waiting, α-blockers, 5-α-reductase inhibitors, and surgery).

objective:

Assessment of the effectiveness of a comprehensive treatment protocol for LUTS in primary care.

methods and materials:

Randomised controlled trial. Intervention: treatment protocol, based on a formalised expert opinion. Control condition: usual care. Study population: 208 subjects with moderate to severe LUTS (IPSS ≥ 8, median 13). Outcome measures: symptom severity (IPSS), bother score (Dan-PSS), and maximum urinary flow (Q_{max}); incidence of acute urinary retention and urinary tract infections.

results:

In the intervention group markedly more subjects used an α-blocker at end of follow-up than in the usual care group (24% vs. 6%). No significant differences were found between intervention and control group in IPSS, Q_{max} or Dan-PSS.

conclusion:

Alpha-blockers and watchful waiting are the most frequent treatment modalities for LUTS in primary care. Our study showed no evidence that a protocol using well-defined indications for all possible treatment modalities based on a formalised expert opinion procedure has added value. Based on our results, we cannot recommend a broadening of the indication for α-blockers, which, however, seems to be the current trend.

Introduction

The attention for male lower urinary tract symptoms (LUTS) has increased during the past decade. Ageing and the availability of medication have led to an increase in elderly men being diagnosed and treated for LUTS. Consequently, costs are increasing.[1] Furthermore, patients visiting both general practitioners (GPs) and urologists nowadays are younger and have less severe symptoms.[2]

Patients with lower urinary tract symptoms may be offered four main treatment options: watchful waiting, α-blockers, 5-α-reductase inhibitors, and surgery. Alpha-blockers provide moderate improvement of symptoms and urinary flow measures.[3,4] 5-α-reductase inhibitors can reduce symptoms and the incidence of acute urinary retention in patients with large prostates.[5-7] Surgery is most powerful in reducing symptoms at the cost of a higher risk of irreversible side effects.[8,9] The efficacy of these treatments has been investigated in randomised clinical trials with often well defined, selected, 'homogeneous' populations from secondary care settings. Inclusion and exclusion criteria required investigations like urinary flow measurements or urodynamics that are not directly available to general practitioners.[3,4,8]

Therefore, the efficacy figures of these trials cannot easily be generalised to the heterogeneous situation in primary care with its generally less developed clinical stages.[10,11] As a result guidelines fail to provide specific diagnostic criteria and sharp indications for therapy. They leave much room for interpretation.[9,12-14] Consequently, daily care in primary care varies widely.[2,15-19] General practitioners must treat patients on a likely (symptoms of impaired micturition) rather than a confirmed diagnosis (proven infravesicular outflow obstruction). In situations like this, studies into the effectiveness of treatments in daily practice may be of great practical value if they enrol patients on presenting symptoms rather than definitive test results.[20,21]

Such 'practical' or 'pragmatic' randomised controlled trials may be used to investigate individual treatment options or, alternatively, comprehensive treatment regimens. A comprehensive approach - including various treatment options - may better resemble actual practice: in daily care the decision to start or stop a treatment depends on the physician's assessment of not only an individual treatment option, but of the relative effectiveness, and risk of side effects and complications of all available treatments. We present such a study comparing a comprehensive treatment protocol with usual care.

Lacking an evidence-based protocol with sharp indications, we chose a feasible and useful alternative for the development of a treatment protocol for primary care. The adequate treatment policy for the included patients could be determined using a previously validated algorithm.[14,22] This algorithm reflects the treatment choices of an international panel of urologists. Such a formalised

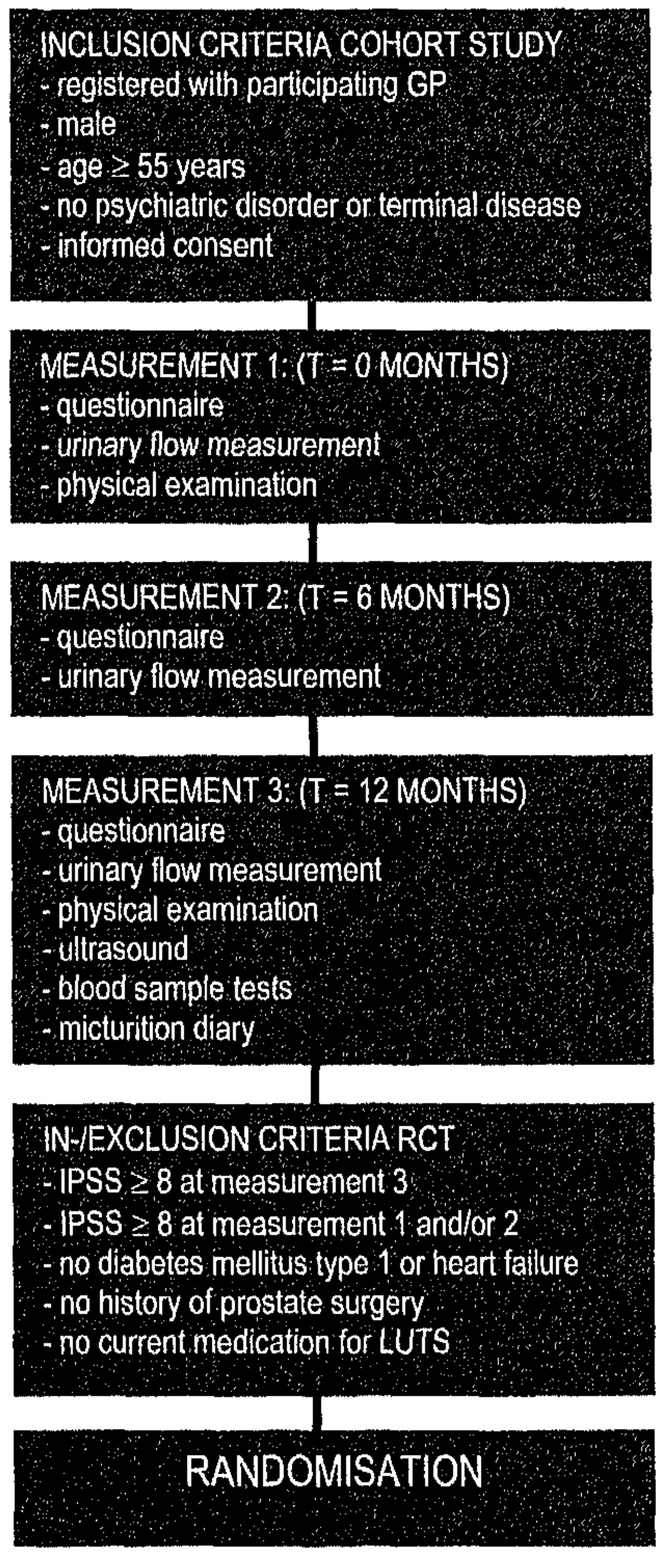
panel judgement rationalises the individual management choices, and thereby may improve both the efficiency and patient outcome.[12] It also resembles daily practice, where - lacking clear guidelines - general practitioners are likely to anticipate on the (expected) treatment of specialists. If a general practitioner expects the urologist to prescribe medication, he may do so himself. If he expects the urologist to perform surgery, he probably will refer his patient for (assessment of the eligibility for) surgery.

Methods and materials

study population

The participants were recruited from a 12-months cohort study to which all men aged 55 or older registered in 14 general practices (26 GPs) were invited to participate. During this cohort study participants kept a 3-day micturition diary and filled in questionnaires with general questions (e.g. demographics, micturition behaviour), and questions that assessed the symptoms (International Prostate Symptom Score (IPSS)[23], referred to as 'symptom score') and the bother caused by the symptoms (Danish Prostate Symptom Score (Dan-PSS)[24], referred to as 'bother score'). They underwent physical examination, and measurement of urinary flow, PSA, creatinin, post-void residual volume, and prostate size (figure 3.1).

figure 3.1. Recruitment of the study population



Eligible for inclusion in this trial were all subjects with moderate or severe lower urinary tract symptoms (IPSS ≥8) at the time of recruitment, and at least one earlier occasion (to minimise spontaneous symptom improvement due to regression to the mean). Exclusion criteria were known or increased risk of prostate cancer (PSA > 10 ng/ml, or PSA >2.5 ng/ml and a Free to Total (F/T) ratio <25%); prior prostate surgery; and diabetes mellitus type 1 or heart failure (to prevent disturbances of symptom assessment due to neuropathy or irregular use of diuretics). Since these exclusion criteria are easily available from the patient record the study population is easy to recognise in a primary care setting.

informed consent and randomisation

Eligible subjects received written information with neutral information regarding the details of intervention and control condition. They returned signed informed consent forms to the research nurse. She gave unique patient identification numbers - including practice identifier - to RJCN, who performed the randomisation blind for patient data (block randomisation, block size 4, stratified per general practice).

intervention

For the intervention group, the treatment protocol followed a stepwise approach (see figure 3.2, pages 44 and 45). All participants were offered an α-blocker (tamsulosin 0.4 mg once daily). After three months this was evaluated. If the α-blocker was well tolerated and the symptom score had improved (≥3 points) the medication was prolonged for 6 months; then the medication was stopped for three months, and re-evaluated. If symptoms had worsened again ≥3 points (indicating a 'real' effect of the medication), the α-blocker was restarted for one year (the rest of the study period). For those who had not improved after three months we established the indication for other treatment options using the algorithm of the formalised expert opinion (see appendix 3.1, page 53)[14,22]. If an indication for a 5-α-reductase inhibitor existed, finasteride (5 mg once daily) was prescribed. In case surgery was indicated patients were offered referral to the urologist for this purpose. The intervention was performed by RJCN. Subjects were instructed to report and discuss any intermittent health questions related to lower urinary tract symptoms to RJCN.

control condition

The control group received 'usual care'. They were not actively offered a certain treatment proposal. Any treatment was coordinated and prescribed by their general practitioner. To prevent contamination of their usual care, the general practitioners were blinded to the content of the protocol. GPs were explained why they should not have any knowledge of the study protocol, and were instructed not to ask their patients about the study or their treatment. They received no special training on diagnosis or treatment of LUTS.

data-collection

After 12 and 24 months of follow-up all subjects completed a questionnaire (symptom and bother scores, any contact with doctors for micturition problems), and underwent urinary flow measurements. Details of consultations, referrals, and prescriptions were obtained from the patient records in the general practice, if necessary. Flow measurements were performed in a private room in the practice using a Da Capo Homeflowmeter®. Flow measurements with a voided volume below 100 ml. were repeated.

outcome parameters and sample size

The primary outcome parameters were improvement of symptom score and maximum urinary flow rate. The sample size calculation was based on a relevant difference of 2.5 points on the symptom score (with a supposed standard deviation (SD) of 7 points)[25] and 2 ml/sec on the maximum urinary flow (supposed SD 4.9 ml/sec)[26]. Consequently, 240 subjects should be included in the analysis ($\alpha = 0.05$, $\beta = 0.80$, equal group sizes, highest needed number).

Secondary outcome parameter was improvement on the bother score. Furthermore, we registered the number of acute urinary retentions and urinary tract infections.

statistical analysis

The difference between intervention and control after 24 months was tested using the Mann Whitney U test (IPSS, Qmax, Dan-PSS), or the χ^2 test (acute urinary retention and urinary tract infections), based on the intention-to-treat principle.

We analysed the baseline characteristics of the subjects in the different treatment subgroups at end of follow-up. This was done to explore possible explanations for differences between the intervention and control condition in outcome and the number of patients receiving a certain treatment.

Furthermore, we acknowledged that the treatment protocol in some ways is a 'black box'. Its overall effectiveness depends on the specific efficacies of the various treatments options and the quality of the decision which treatment to offer an individual patient. Therefore, we calculated the differences in outcome measures at end of follow-up to check whether the individual treatments were effective in the subgroups identified by the treatment protocol.

Results

Between August 2000 and June 2002 208 subjects were randomised to the intervention (N=104) and control group (N=104); baseline characteristics are shown in table 3.1 (see page 43). Figure 3.2 presents the flow of the subjects through the study (see pages 44 and 45).

intervention group

After randomisation four subjects were excluded: three had started using an α-blocker, one had been referred for suspected prostate cancer. The remaining 100 subjects were offered tamsulosin. 13 Of them did not want medication; 11 stopped because of side effects (dizziness (N=4), sexual problems (N=1), rash (N=1), interaction with co-medication (N=1), increasing LUTS (N=3), leg-cramps (N=1)); one stopped after abdominal surgery for nephrogenic haematuria.

Consequently, 75 completed the first three months period of α-blocker treatment according to the treatment protocol. Evaluation showed a mean change in symptom score of -2.7 points, with (sufficient) improvement in 40 subjects (mean -6.8 points), and no improvement in 35 subjects (mean +2.1 points).

The 40 'improved' subjects continued the protocol (6 months α-blocker/3 months stop). Re-evaluation at t=12 months showed that 26 had worsened after withdrawal (they therefore restarted the α-blocker); 10 subjects showed no worsening (they consequently changed therapy (watchful waiting, N=9; 5-α-reductase inhibitors, N=1)). Four subjects stopped the medication because of 'second thoughts' (N=1), dizziness (N=1), sexual problems (N=1) and angina pectoris (based on the advice of his cardiologist, N=1).

Of the 35 'unimproved' subjects 31 subjects changed to watchful waiting; the other 4 were referred for surgery. Two of these were indeed operated, but the other two did not risk the side effects and complications of surgery, and continued tamsulosin treatment (N=1), and no further treatment (N=1) respectively. One subject of the watchful waiting group died before the end of the study.

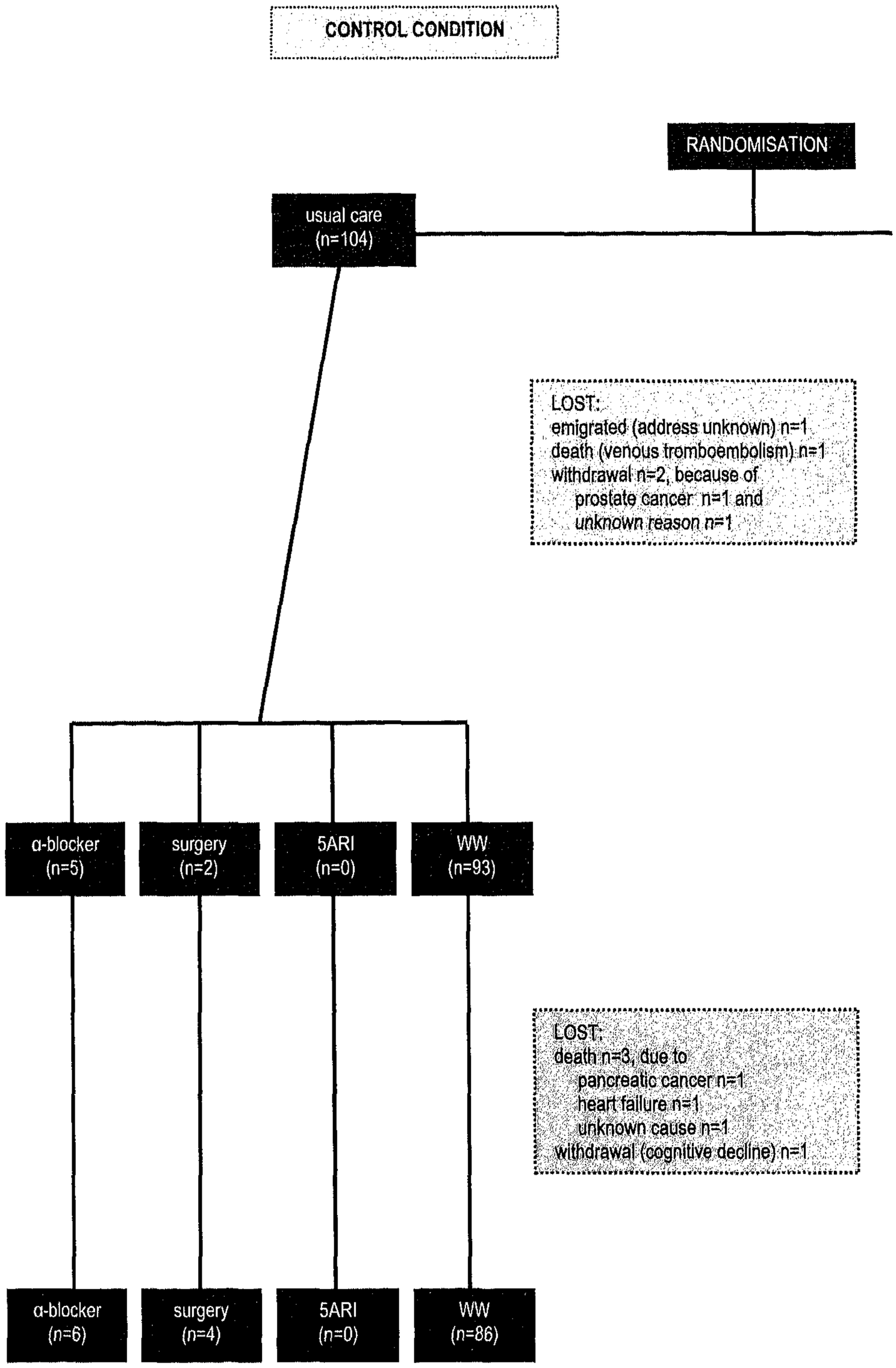
Finally, 67 subjects (64%) completed the study according to the protocol; complete data were available from 103 subjects.

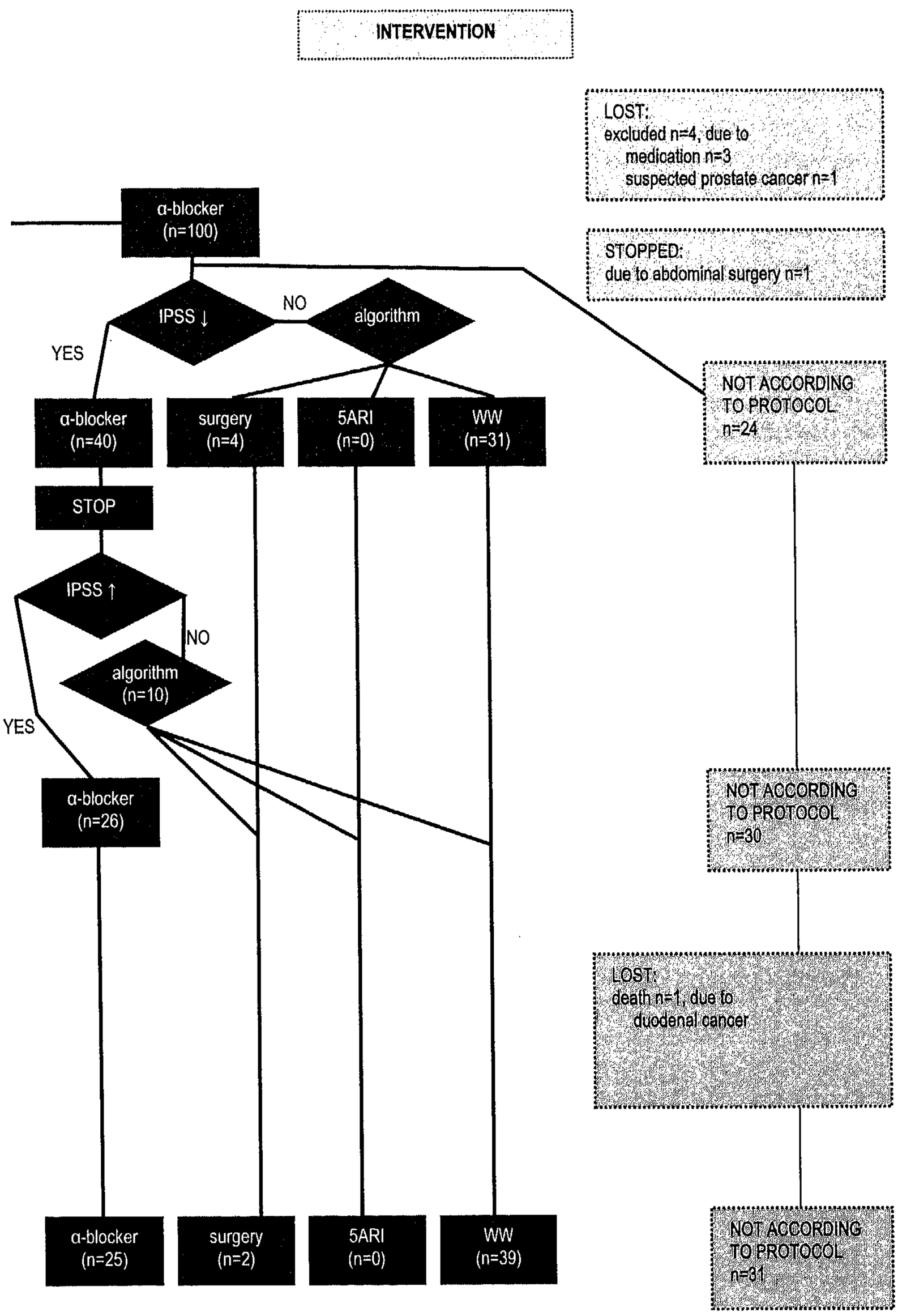
table 3.1. Study population, baseline characteristics of intervention (N=104) and control group (N=104).

	intervention		control		
	median	[iqr]	median	[iqr]	
IPSS	**13**	[10 - 17]	**13**	[10 - 17]	p=0.90
Q_{max} (ml/sec)	**11.4**	[7.6 - 16.6]	**12.8**	[8.3 - 16.5]	p=0.81
Dan-PSS	**10**	[5 - 18]	**8**	[5 - 19]	p=0.91
age (years)	**64**	[60 - 70]	**65**	[60 - 71]	p=0.56
prostate volume (ml)	**25**	[19 - 32]	**24**	[20 - 33]	p=0.35
post-void residual volume (ml)	**34**	[15 - 65]	**38**	[19 - 61]	p=0.68
PSA (ng/ml)	**1.0**	[0.7 - 2.0]	**1.3**	[0.7 - 2.0]	p=0.27
	N	[%]	N	[%]	
post-void residual volume > 100 ml.	**18**	[17%]	**20**	[19%]	p=0.70

IPSS = International Prostate Symptom Score; Q_{max} = maximum urinary flow; Dan-PSS = Danish Prostate Symptom Score; PSA = Prostate Specific Antigen.

figure 3.2. Flow of participants through the study.





IPSS = International Prostate Symptom Score; 5ARI = 5-α-reductase inhibitor; WW = watchful waiting

control group

In the first year five subjects received an α-blocker and two were referred for surgery. One subject died and three were lost to follow-up (emigration, withdrawal of informed consent). During the second year three participants were prescribed an α-blocker; one subject changed from α-blocker to surgery, and one from watchful waiting to surgery. Three subjects died. Two were lost to follow-up (withdrawal of informed consent, emigration (only returned a questionnaire)). As a result 96 subjects completed the study, of which six used an α-blocker, four had undergone surgery and 86 were on watchful waiting. No participant used a 5-α-reductase inhibitor.

outcome

Significant changes between baseline and end of follow-up in symptom score, bother score, and maximum urinary flow were not found, neither in the intervention, nor in the control group (table 3.2). A difference between intervention and control group in change scores was not found either. During the first year of follow-up two subjects suffered from an acute urinary retention (intervention (N=1); control (N=1)); another three suffered from a urinary tract infection (intervention (N=1); control (N=2)). During the second year four subjects were treated for a urinary tract infection (intervention (N=1), control (N=3)), and one for an acute urinary retention (intervention).

table 3.2. Comparison of symptom score, maximum urinary flow and bother score, and the incidence of acute urinary retention and urinary tract infections between baseline and end of follow-up (t=24 months).

	intervention (N=104)				control (N=104)				
	baseline		outcome		baseline		outcome		
	median [iqr]		median [iqr]	p(B-O) *	median [iqr]		median [iqr]	p(B-O) *	p(I-C) **
IPSS	13	[10 -17]	14	[10 -18] p=0.79	13	[10-17]	14	[10-17] p=0.80	p=0.98
Q_{max}	11.4	[7.6-16.6]	11.3	[8.1-15.8] p=0.85	12.8	[8.3-16.5]	11.8	[8.3-16.0] p=0.76	p=0.45
Dan-PSS	9	[4 -17]	9	[4-21] p=0.92	8	[5-19]	10	[5-19] p=0.25	p=0.98
AUR			2	[1,9%]			1	[1,0%]	p=0.62
UTI			2	[1,9%]			5	[4.8%]	p=0.45

IPSS = International Prostate Symptom Score; Q_{max} = maximum urinary flow; Dan-PSS = Danish Prostate Symptom Score; AUR = Acute Urinary Retention; UTI = Urinary Tract Infection.
* Wilcoxon signed rank test; ** Mann Whitney U test or χ^2 test (due to small numbers the Fisher Exact test was used), where applicable.

Our exploratory analysis did not show significant differences in baseline characteristics within the treatment subgroups between intervention and control condition (table 3.3).

table 3.3. Baseline characteristics of the study population; subgroups are defined by the treatment status at end of follow-up.

	treatment subgroups at end of follow-up							
	α-blocker		5-α-reductase inhibitors		surgery		watchful waiting	
	I (N=25)	C (N=6)	I (N=1)	C (N=0)	I (N=2)	C (N=4)	I (N=39)	C (N=86)
age (years)	63 [60 - 68]	69 [63 - 73]	68	-	76	70 [60 - 74]	64 [61 - 71]	64 [60 - 70]
IPSS	13 [11 - 18]	14 [11 - 21]	32	-	14 [13 - 15]	12 [9 - 20]	11 [9 - 14]	13 [10 - 17]
Q_{max} (ml/s)	9.2 [7.0 - 16.2]	14.1 [8.9 - 23.8]	8.1	-	4.3 [4.3 - 4.3]	10.9 [3.9 - 16.1]	11.6 [7.6 - 15]	12.8 [8.2 - 16.6]
Dan-PSS	8 [4 - 15]	19 [11 - 24]	71	-	27 [22 - 32]	9 [3 - 23]	8 [2 - 12]	8 [5 - 18]
prostate size (ml)	25 [18 - 32]	25 [21 - 36]	87	-	27	31 [23 - 69]	23 [17 - 29]	24 [20 - 31]
PVR (ml)	51 [34 - 82]	33 [17 - 97]	308	-	21	42 [40 - 55]	23 [12 - 58]	39 [19 - 72]
PSA (ng/ml)	1.0 [0.7 - 1.7]	1.5 [0.8 - 3.7]	3.4	-	2.2	3.0 [1.9 - 5.8]	1.0 [0.6 - 1.7]	1.2 [0.7 - 1.8]

numbers represent medians [inter quartile range]

	treatment subgroups at end of follow-up		
	not according to protocol		'watchful waiting' /\ *
	I (N=31)	C (N=0)	I (N=69)
age (years)	66 [60 - 71]	-	64 [60 - 71]
IPSS	15 [11 - 19]	-	12 [10 - 16]
Q_{max} (ml/s)	12.0 [7.6 - 18.1]	-	11.6 [8.0 - 17.5]
Dan-PSS	12 [5 - 20]	-	9 [4 - 17]
prostate size (ml)	26 [19 - 32]	-	24 [18 - 32]
PVR (ml)	31 [20 - 52]	-	28 [13 - 56]
PSA (ng/ml)	1.3 [0.7 - 2.3]	-	1.1 [0.7 - 2.0]

I = intervention
C = control

* watchful waiting (I) = pooled data from all who were on watchful waiting at end of follow-up, i.e. 39 subjects from the watchful waiting subgroup according to the treatment protocol, and 30 (of 31) subjects from the not-according-to-protocol subgroup.

All significance values of the differences between I and C group within a treatment subgroup were $p \geq 0.05$.

IPSS = International Prostate Symptom Score; Q_{max} = maximum urinary flow rate; Dan-PSS = Danish Prostate Symptom Score; PSA = prostate specific antigen.

The analysis of the 'black box' showed that in the intervention condition all active treatment subgroups symptom and bother scores, and Qmax had improved, while the watchful waiting group had worsened. This indicates that the protocol was internally consistent. Noteworthy, in the control group the majority of the users of α-blockers at end of follow-up had an increased symptom and bother score, and a decreased maximum urinary flow, leading to an overall (though not statistically significant) worse outcome for this subgroup (table 3.4).

table 3.4. Analysis of the 'black box': comparison of symptom score, maximum urinary flow and bother score between baseline and end of follow-up (t=24 months); subgroups are defined by the treatment status at end of follow-up.

		treatment subgroups at end of follow-up							
		α-blocker		5-α-reductase inhibitors		surgery		watchful waiting	
		I (N=25)	C (N=6)	I (N=1)	C (N=0)	I (N=2)	C (N=4)	I (N=39)	C (N=86)
IPSS	baseline	**13** [11 - 18]	**14** [11 - 21]	**32**	-	**14** [13 - 15]	**12** [9 - 20]	**11** [9 - 14]	**13** [10 - 17]
	outcome	**12** [10 - 16]	**19** [13 - 24]	**22**	-	**7** [7 - 7]	**5** [4 - 8]	**14** [9 - 18]	**14** [10 - 17]
Q_{max}	baseline	**9.2** [7.0 - 16.2]	**14.1** [8.9 - 23.8]	**8.1**	-	**4.3** [4.3 - 4.3]	**10.9** [3.9 - 16.1]	**11.6** [7.6 - 15]	**12.8** [8.2 - 16.6]
	outcome	**11.0** [8.1 - 17.6]	**11.8** [8.5 - 20.5]	**8.8**	-	**15.3** [15.0-15.6]	**20.2** [11.0-32.0]	**12.2** [7.7 - 17.6]	**11.8** [8.2 - 15.8]
Dan-PSS	baseline	**8** [4 - 15]	**19** [11 - 24]	**71**	-	**27** [22 - 32]	**9** [3 - 23]	**8** [2 - 12]	**8** [5 - 18]
	outcome	**7** [1 - 11]	**22** [9 - 35]	**43**	-	**11** [10 - 12]	**3** [1 - 7]	**9** [4 - 22]	**10** [6 - 19]

numbers represent medians [inter quartile range]

		treatment subgroups at *end of follow-up*		
		not according to protocol		'watchful waiting' /\ *
		I (N=31)	C (N=0)	I (N=69)
IPSS	baseline	**15** [11 - 19]	-	**12** [10 - 16]
	outcome	**15** [12 - 19]	-	**14** [10 - 19]
Q_{max}	baseline	**12.0** [7.6 - 18.1]	-	**11.6** [8.0 - 17.5]
	outcome	**10.7** [7.5 - 14.6]	-	**11.1** [7.8 - 16.8]
Dan-PSS	baseline	**12** [5 - 20]	-	**9** [4 - 17]
	outcome	**13** [7 - 23]	-	**11** [5 - 22]

Note!
The treatment subgroups in this analysis were defined post-hoc, not beforehand.
Due to confounding by indication the results are not suitable for statistical testing of the differences between two treatments or the treatment arms.

* watchful waiting (i) = pooled data from all intervention subjects who were on watchful waiting at end of follow-up, i.e. 39 subjects from the watchful waiting subgroup according to the treatment protocol, and 30 (of 31) subjects from the not-according-to-protocol subgroup.

I = intervention; C = control
IPSS = International Prostate Symptom Score; Q_{max} = maximum urinary flow rate; Dan-PSS = Danish Prostate Symptom Score.

Discussion

main results

The structured approach of our treatment protocol did not result in a relevant and statistically significant improvement in outcome as compared to current usual care.

strengths and limitations

Several considerations concerning the generalisability of our results need to be acknowledged. Our study population reflects the heterogeneous population in primary care, which increases the external validity. However, we studied the pool of symptomatic patients from which the consulting population self-selects. Characteristics of consulting patients will vary in time or place, because of cultural aspects, health education or promotional activities, but in general they will be more symptomatic than the subjects in our study. Our results indicate that a more active treatment policy in consulting patients will not improve overall outcome.

Several considerations may help to explain the absence of relevant and significant results in this study. The increased attention for lower urinary tract symptoms in general practice may have reduced the room for improvement. Although the general practitioners were not acquainted with the intervention nor involved in its execution, they knew that a study on LUTS was performed. They may have been (or become) more interested in lower urinary tract symptoms, which may have influenced their behaviour. They maybe already had adopted the guidelines, either explicitly or implicitly because the guideline confirmed their knowledge or clinical experience. Theoretically, any trial with a usual care control group does not provide a fair estimate of the value of an intervention if the usual care has already started to resemble the intervention studied (i.e. the guideline). This may be the case with regard to restricted use of 5-α-reductase inhibitors. For α-blockers current treatment guidelines themselves leave a lot of room for interpretation.[27] A wider use of these medicines is advocated. But we noticed that they are prescribed less frequently in the usual care situation without deleterious effect.

Although our intervention was based on the opinion of a large international sample of urologists, one may prefer a different interpretation, and consequently argue that some patients should have been treated differently. However, it is unlikely that individual changes in treatments would have changed overall results.

We included 208 subjects and could analyse the data of 199 subjects, and used non-parametric tests. This may have lead to loss of statistical power. However, the variance of the IPSS was lower than presumed (SD 5.2 instead of 7) which sufficiently compensates this (power was actually about 90%).

this study and the literature

Our study primarily aimed at the reduction of symptom severity. A focus on the prevention of complications (like acute urinary retentions) could have led to a somewhat more prominent place for 5-α-reductase inhibitors.[12] However, these are most effective in patients with large prostates. In our sample 16 (7.7%) participants had a prostate size ≥50 gr. Only 1 subject (of 3) who developed an acute urinary retention had such a prostate size. Therefore, a more rigorous prescription of 5-α-reductase inhibitors to this group probably would not have had effect on our primary outcome parameters and conclusions.

Little consensus exists on the appropriateness of α-blocker treatment.[14] Recently published guidelines acknowledge this, stating that 'a trial of treatment is appropriate'.[12,13] We tried to overcome this problem and select 'real' responders using a trial of treatment combined with a similar trial of withdrawal of treatment. In this way, the effect cannot be attributed to the natural course of the disease, or 'regression to the mean', although it does not fully protect against the placebo effect. Despite this, the subjects who were treated with tamsulosin at end of follow-up had a relatively small improvement of both symptoms and maximum urinary flow compared to what could be expected if the initial effect after the first three months would have been sustained throughout the whole study period.

At end of follow-up only six subjects in the control condition were treated with α-blockers. Most of them had worse outcome parameters at end of follow-up compared to baseline. Probably general practitioners reserve medication for those who experience a rapid deterioration. This may be too restrictive: some untreated patients in the control group may have been better off using medication. However, our study results showed no evidence that a protocol starting with an α-blocker has added value. The larger group in the intervention condition who had been prescribed an α-blocker did not have a more favourable outcome than the controls who had not received this treatment.

conclusion

A primary care treatment regimen based on the consensus among urologists does not improve effectiveness. Especially for the large 'grey area' in which α-blockers are considered, our study does not support a broadening of the indication for α-blockers, which, however, seems to be the current trend. We should continue to look for sharper criteria for treatment, resulting in adequate therapy for those who may benefit and prevention of overuse of medication for those who may expect little effect.

References

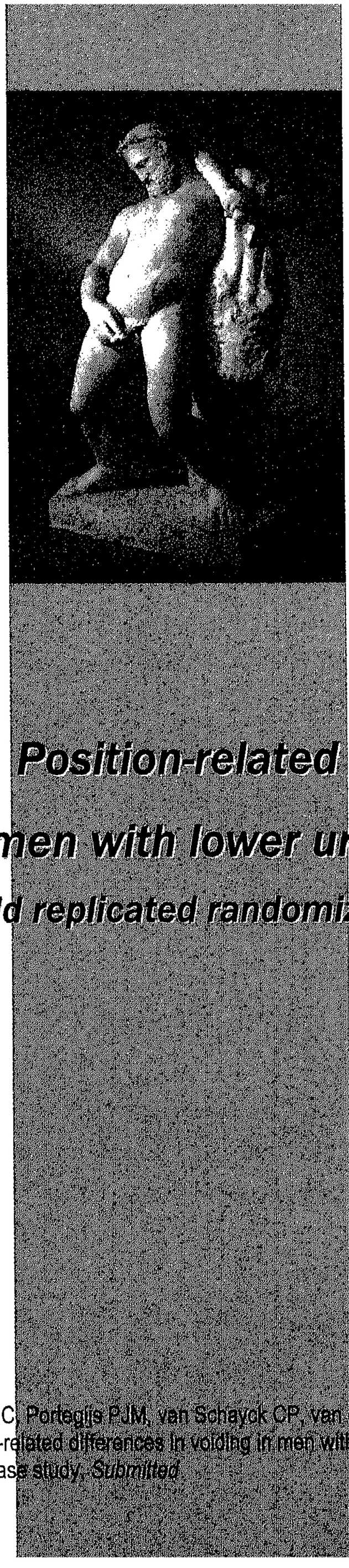
1. Department of Health. Prescription Cost Analysis. England 2003. http://www.dh.gov.uk/PublicationsAndStatistics/Publications/PublicationsStatistics/PublicationStatisticsArticle/fs/en?CONTENT_ID=4081720&chk=kVOup3. (Accessed 1 Nov 2006)
2. Bruskewitz R. Management of symptomatic BPH in the US: who is treated and how? *Eur Urol* 1999;36 Suppl 3:7-13.
3. Wilt TJ, Howe RW, Rutks IR, MacDonald R. Terazosin for benign prostatic hyperplasia. *Cochrane Database Syst Rev* 2002(4):CD003851.
4. Wilt TJ, Mac Donald R, Rutks I. Tamsulosin for benign prostatic hyperplasia. *Cochrane Database Syst Rev* 2003(1):CD002081.
5. Roehrborn CG, Marks LS, Fenter T, Freedman S, Tuttle J, Gittleman M, et al. Efficacy and safety of dutasteride in the four-year treatment of men with benign prostatic hyperplasia. *Urology* 2004;63(4):709-15.
6. Abrams P, Schafer W, Tammela TL, Barrett DM, Hedlund H, Rollema HJ, et al. Improvement of pressure flow parameters with finasteride is greater in men with large prostates. Finasteride Urodynamics Study Group. *J Urol* 1999;161(5):1513-7.
7. Marberger MJ. Long-term effects of finasteride in patients with benign prostatic hyperplasia: a double-blind, placebo-controlled, multicenter study. PROWESS Study Group. *Urology* 1998;51(5):677-86.
8. Speakman MJ. Who should be treated and how? Evidence-based medicine in symptomatic BPH. *Eur Urol* 1999;36 Suppl 3:40-51.
9. McConnell JD, Barry MJ, Bruskewitz RC, et al. *Benign Prostatic Hyperplasia: Diagnosis and Treatment. Clinical Practice Guideline, Number 8. AHCPR Publication No. 94-0582.* Rockville, MD: Agency for Health Care Policy and Research, Public Health Service, U.S. Department of Health and Human Services., 1994.
10. Masumori N, Tanaka Y, Takahashi A, Itoh N, Ogura H, Furuya S, et al. Lower urinary tract symptoms of men seeking medical care--comparison of symptoms found in the clinical setting and in a community study. *Urology* 2003;62(2):266-72.
11. Knottnerus JA. Medical decision making by general practitioners and specialists. *Fam Pract* 1991;8(4):305-7.
12. Speakman MJ, Kirby RS, Joyce A, Abrams P, Pocock R. Guideline for the primary care management of male lower urinary tract symptoms. *BJU Int* 2004;93(7):985-90.
13. Wolters RJ, Spigt MG, van Reedt Dortant PF, Gercama AJ, Klomp ML, Romeijnders AC, et al. NHG-Standaard Bemoeilijkte mictie bij oudere mannen. *Huisarts en Wetenschap* 2004;47(12):571-86.
14. McDonnell J, Stoevelaar HJ, Bosch JL, Kahan JP. The appropriateness of treatment of benign prostatic hyperplasia: a comparison of Dutch and multinational criteria. *Health Policy* 2001;57(1):45-56.
15. Plawker MW, Fleisher JM, Nitti VW, Macchia RJ. Primary care practitioners: an analysis of their perceptions of voiding dysfunction and prostate cancer. *J Urol* 1996;155(2):601-4.
16. Collins MM, Barry MJ, Bin L, Roberts RG, Oesterling JE, Fowler FJ. Diagnosis and treatment of benign prostatic hyperplasia. Practice patterns of primary care physicians. *J Gen Intern Med* 1997;12(4):224-9.
17. Fawzy A, Fontenot C, Guthrie R, Baudier MM. Practice patterns among primary care physicians in benign prostatic hyperplasia and prostate cancer. *Fam Med* 1997;29(5):321-5.
18. Chapple CR. BHP Disease Management. Introduction and concluding remarks. *Eur Urol* 1999;36 Suppl 3:1-6.
19. McNicholas TA. Management of symptomatic BPH in the UK: who is treated and how? *Eur Urol* 1999;36 Suppl 3:33-9.
20. Schwartz D, Lellouch J. Explanatory and pragmatic attitudes in therapeutical trials. *J Chronic Dis* 1967;20(8):637-48.
21. Tunis SR, Stryer DB, Clancy CM. Practical clinical trials: increasing the value of clinical research or decision making in clinical and health policy. *Jama* 2003;290(12):1624-32.
22. Stoevelaar HJ, Van de Beek C, Casparie AF, McDonnell J, Nijs HG. Treatment choice for benign prostatic hyperplasia: a matter of urologist preference? *J Urol* 1999;161(1):133-8.
23. Barry MJ, Fowler FJ, Jr., O'Leary MP, Bruskewitz RC, Holtgrewe HL, Mebust WK, et al. The American Urological Association symptom index for benign prostatic hyperplasia. The Measurement Committee of the American Urological Association. *J Urol* 1992;148(5):1549-57; discussion 1564.
24. Hansen BJ, Flyger H, Brasso K, Schou J, Nordling J, Thorup Andersen J, et al. Validation of the self-administered Danish Prostatic Symptom Score (DAN-PSS-1) system for use in benign prostatic hyperplasia. *Br J Urol* 1995;76(4):451-8.

25. Bosch JL, Hop WC, Kirkels WJ, Schroder FH. The International Prostate Symptom Score in a community-based sample of men between 55 and 74 years of age: prevalence and correlation of symptoms with age, prostate volume, flow rate and residual urine volume. *Br J Urol* 1995;75(5):622-30.
26. Rosier PF, de Wildt MJ, Wijkstra H, Debruyne FF, de la Rosette JJ. Clinical diagnosis of bladder outlet obstruction in patients with benign prostatic enlargement and lower urinary tract symptoms: development and urodynamic validation of a clinical prostate score for the objective diagnosis of bladder outlet obstruction. *J Urol* 1996;155(5):1649-54.
27. Stoevelaar HJ, van de Beek C, Casparie AF, Nijs HG, McDonnell J, Janknegt RA. [Variation in the diagnosis and treatment of benign prostatic hyperplasia in urological practice]. *Ned Tijdschr Geneeskd* 1996;140(15):837-42.

appendix 3.1. The algorithm based on the formalised expert opinion procedure.[14,22] Variables and coefficients for the used algorithm to determine the appropriateness of referral for surgery or prescription of a 5-α-reductase inhibitor by general practitioners.

				surgery	5-α-reductase inhibitor
variable		categories		coefficient	coefficient
	constant			β_0 = -15.70	β_0 = -26.40
X_1 =	age	0 1	< 70 ≥ 70	β_1 = 0.53	β_1 = 1.19
X_2 =	total symptom score (IPSS)	0 1	moderate severe	β_2 = 5.32	β_2 = 10.38
X_3 =	symptoms	0 1	non-specific irritative	β_3 = 1.01	β_3 = -2.16
X_4 =	bother	0 1	acceptable not acceptable	β_4 = 5.71	β_4 = 2.15
X_5 =	history of urinary tract infections	0 1	< 2 ≥ 2	β_5 = 3.76	β_5 = 1.19
X_6 =	maximum urinary flow rate	0 1	> 15 ml/sec $10 \leq X_6 \leq 15$ ml/sec	β_6 = 4.63	β_6 = -0.55
		0 1	> 15 ml/sec < 10 ml/sec	β_6 = 7.53	β_6 = -0.55
X_7 =	prostate volume	0 1	< 30 g $30 \leq X_7 \leq 50$ g	β_7 = 2.95	β_7 = -0.01
		0 1	< 30 g > 50 g	β_7 = 5.19	β_7 = 10.86
X_8 =	post-void residual volume	0 1	< 150 ml ≥ 150 ml	β_8 = 4.71	β_8 = 1.19

The general formula for the calculation of the appropriateness of surgery or 5-α-reductase inhibitor after failure of α-blocker treatment (versus watchful waiting):
Appropriateness = $100\% * 1/(1+ e^{-(\beta_0 + \Sigma(\beta_1 * X_1 \ldots \beta_8 * X_8))})$.
A cut-off value of 25% was taken in order to determine actual appropriateness (<25%: not appropriate; ≥ 25% appropriate).



CHAPTER 4

Please be seated?

Position-related differences in voiding in men with lower urinary tract symptoms

- a 20-fold replicated randomized single case experiment

ISRCTN number: ISRCTN62736854

Norg RJC, Porteglis PJM, van Schayck CP, van de Beek C, Knottnerus JA. Please be seated? position-related differences in voiding in men with prostatic hyperplasia - a 20-fold replicated single case study. *Submitted*.

Abstract

background:

Lifestyle advices are frequently given to men with lower urinary tract symptoms suggestive for benign prostatic hyperplasia (LUTS). Despite the lack of evidence, they can be found in guidelines and many freely accessible medical websites in the United States. One is to urinate sitting instead of standing.

objective:

We assessed the differences in urinary flow parameters between these two positions.

methods and materials:

20 randomised single case experiments followed by a meta-analysis. Measurements were done in the participants' own toilet facilities ('natural micturition environment') among a sample of a population-based cohort of elderly men with moderate LUTS (International Prostate Symptom Score 8-19). After a four-day training period, 40 consecutive urinations were randomly allocated to the sitting (N=20) and standing (N=20) position. Subjects were blind with respect to the voiding position up till the moment they entered the toilet. Main outcome measures were individual and overall differences in maximum urinary flow (primary outcome measure), average urinary flow, voided volume, and voiding time (secondary outcome measures). The randomisation test was used for the analysis.

results:

Standing, maximum urinary flow was higher in 14 and lower in 6 subjects. Overall, maximum urinary flow was 0.55 ml/s higher (p=0.02). Average urinary flow was also higher (0.55 ml/s, p=0.006). Voided volume tended to be higher (11 ml, p=0.06), while voiding time was about equal (0.38s, p=0.75).

conclusion:

A general recommendation for elderly men with LUTS to urinate sitting is not correct.

The replicated single-case design is feasible and combines validity with high statistical power and efficiency.

Introduction

Lower urinary tract symptoms suggestive of benign prostatic hyperplasia (LUTS) are a common and increasing health problem in the ageing male.[1] More than 40% of older men experience these symptoms, such as decreased strength of the urinary stream, feeling of incomplete emptying of the bladder, frequent voiding, and dribbling after urination. The natural course is variable.[2] For patients with moderate symptoms treatment often starts with periodic reassessments (watchful waiting)[3], often supplemented with a large variety of lifestyle advices, few of which are based on evidence.[4,5] Although the short-term and long-term effects often are unclear[5], such lifestyle advices can also be found in guidelines[3,6-8] and on hospital[9-11] and accredited[12] medical websites[9,13,14] freely accessible to the United States public. One is the advice to urinate sitting instead of standing.[8-11,13,14] The idea is that sitting would help relax, make micturition easier and facilitate more complete bladder emptying. This idea is based on 'common knowledge' and uncontrolled 'clinical experience'. Although some evidence exists for healthy young men[15], the only study in male subjects with LUTS points at the contrary.[16] Both maximum urinary flow (Q_{max}) and voided volume (VV) tended to be lower while sitting (0.7 ml/s and 21.5 ml respectively). Post-void residual volume was slightly higher.[16] This therefore contradicts the aforementioned suggestion that sitting would be better.

Because LUTS develop gradually over many years, even small everyday differences may, on the long term, be relevant. A difference of 0.7 ml/s in maximum urinary flow is comparable with the decline in about five years.[17-19] Patients who realize this may be willing to change their voiding habits, even though they would not be able to sense the difference. Easily applicable, non-pharmacological interventions are reasonable, even if the effects seem small.

The usual way to investigate such an effect of changing voiding position would be a cross-over design. A group of patients is tested in a controlled clinical setting in a standardized way, once or a small number of times. The difference in mean maximum urinary flow between the two positions is measured. However, if one wants to generalize the conclusions to daily life, one has to realize that the standardized research setting differs from the 'natural' setting at home, for instance because men may be more stressed.[20] Furthermore, at home the circumstances in which men urinate vary a lot with respect to e.g. bladder filling and time of day.[21,22] All these circumstances should be included to provide a good representation of the real-life situation. Therefore we placed uroflowmeters at home, and chose a specific cross-over design, the replicated randomised single case experiment. Since there is no carry-over effect from one measurement to the next we could easily include a large series of measurements under all circumstances, thereby reducing bias. Each subject

served as his own control, giving a precise result. The experiment was repeated (replicated) among a group subjects from the indicated population. The individual results were pooled in a meta-analysis to give an overall result, suitable for generalization.[23]

Methods and materials

study population

Subjects were recruited by mail from a well-documented population-based cohort of elderly men.[24] Eligible were all subjects with stable moderate LUTS, i.e. an International Prostate Symptom Score (IPSS) ≥ 8 and ≤ 19 on all five measurements during the previous 3 years. Excluded were those with a history that might interfere with study compliance (e.g. dementia), or anatomy and physiology of the lower urinary tract (e.g. surgery of prostate or lower abdomen, or medication such as α-blockers, 5-α-reductase inhibitors, diuretics and anticholinergics). We invited subjects until 20 respondents were included.

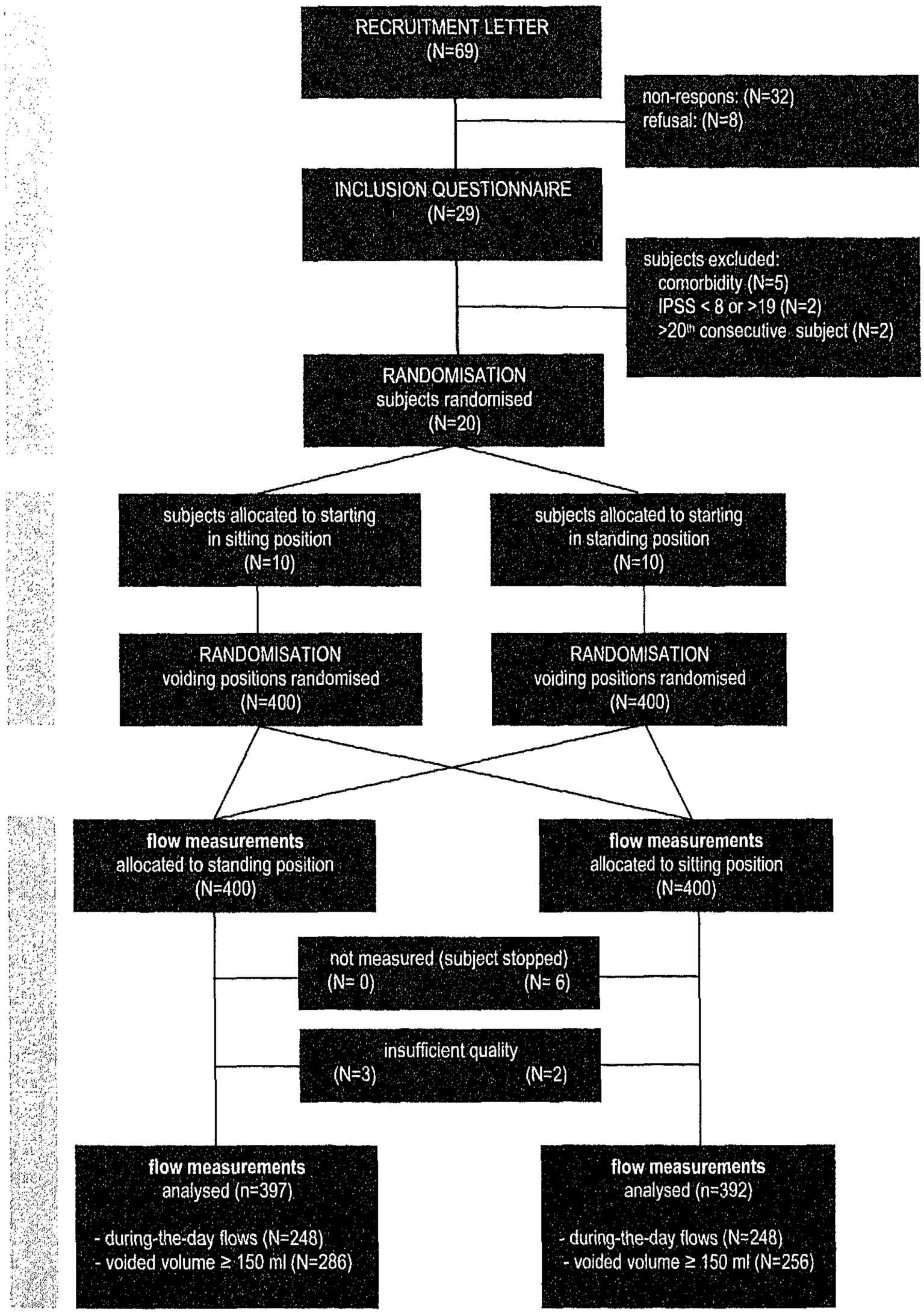
design

20-Fold replicated randomised single case experiment (figure 4.1, see page 59). Estimating a minimally clinically relevant difference was tricky. Simple, non-pharmacological lifestyle changes may be worthwhile, even if small differences can be expected. We chose a minimally clinically relevant difference in maximum urinary flow of 3 ml/s on the individual level. This is about what most men can perceive as a difference in urinary flow. For the group level, we chose 0.5 ml/s, being the average decline in 4 to 5 years (quite hypothetically to be translated in the time that the start of medication could be postponed). Using the estimates of Sonke et al.[22] 40 measurements per subject would be needed. Therefore, the subjects collected a consecutive series of 40 uroflow measurements at home, randomly assigned to the standing (N=20) and sitting (N=20) position. This experimental period was preceded by a four-day run-in period (two days for each position), so that the subjects could get accustomed to the procedures, and measurements during the experimental period would not be influenced by a learning effect.[25,26]

outcome measures

Primary outcome measure: difference in maximum urinary flow (Q_{max}diff). Other outcome measures: difference in average urinary flow (Q_{ave}diff), voided volume (VVdiff) and voiding time (Vtdiff). Differences were calculated as the value in the standing minus value in the sitting voiding position.

figure 4.1. Summary of the study protocol and recruitment of the study population



analysis

primary analysis

Results per subject were calculated using the non-directional (two-sided) randomisation test[23] (SCRT 1.1 software[27,28]; 20.000 randomisations). For the overall result a directional meta-analysis is appropriate[29], testing the hypotheses 'sitting is favourable' ($Q_{max}diff<0$) and 'standing is favourable' ($Q_{max}diff>0$). Using the then required adjusted $\alpha=0.025$, an overall $Q_{max}diff=0.5$ ml/s could be detected with 20 subjects.[29]

secondary analysis

Secondary (exploratory) analyses investigated the association of the difference in maximum urinary flow with possible determinants: age, IPSS score, bother (Danish Prostate Symptom Score, Dan-PSS[30]), mean Q_{max}, and pre-study preference for a certain voiding position (Spearman's correlation coefficient (r_s), Mann-Whitney U test; SPSS 12.0.2 for Windows).

subset analysis

Subset analyses were planned to check for the consistency of the data, not because we expected clinically meaningful differences. We first analysed only the 'spontaneous during-the-day measurements' to eliminate bias from (possibly incompletely randomised) incomplete bladder-filling or sleep. Flows within a timeframe of 30 minutes before and after sleeping or defecation were excluded. Secondly, we included only measurements with a voided volume ≥150 ml. This is common practice in research concerning LUTS, and regarded necessary to acquire reproducible results.[3]

measurements

questionnaires

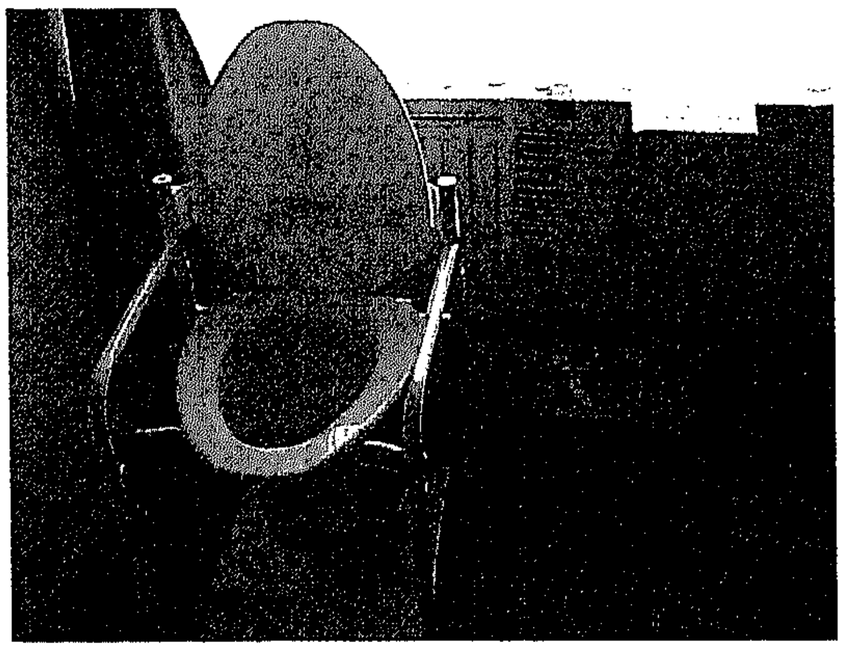
Before inclusion the subjects completed a questionnaire consisting of the IPSS, Dan-PSS, and questions about medical history, medication use, and pre-study preference for a certain voiding position. After the study a second questionnaire evaluated their experiences.

urinary flow measurements

Participants urinated in a uroflowmeter (Urodyn 1000, Danica/Medtronic®) placed under a foldable micturition chair (Dantec/Medtronic®) adjusted to the subjects' height (figure 4.2, see page 61). They received 40 numbered sealed opaque envelopes with voiding instructions (standing or sitting). These were to be kept in their toilet, and opened only just before voiding. Thus, the moment of

urination would not be influenced by knowledge of the voiding position. The random sequence of the voiding instructions was previously generated with an SPSS randomization procedure. Participants were blinded to the order and exact number of instructions for each position. The subjects directly put a print of the flow recording in the numbered envelope on which they registered time of urination and voided volume.

figure 4.2. The micturition chair and the uroflowmeter



frequency-volume chart ('diary')

The subjects registered the hours and volumes of the urinations, and the hours of defecation, bedtime and waking-up. If someone had to defecate and urinate at the same time, he was asked to first urinate in the uroflowmeter, and secondly to defecate on the conventional toilet. The participants registered all times of voiding, also when not at home. The participants were asked to register on 'normal' days, i.e. days with usual activities, drinking and voiding habits, etcetera.

data quality control

The combination of the frequency-volume chart and the envelopes with the printouts of the flow recordings functioned as a check to prevent mix-up of the order and time of the flow recordings.

We used the criteria described by Witjes et al.[31] to eliminate artefacts. Two investigators (RJCN and PJMP) independently assessed all recordings blinded for voiding position. If the difference in maximum urinary flow between the readers was ≥ 1 ml/s they independently reassessed the flows to check for possible misinterpretation. Afterwards, the mean of both readings was taken for the analyses.

Results

study population

The characteristics of the study population are described in table 4.1 (see page 63). 69 men were invited; 37 men responded; 8 of them explicitly refused and 7 did not meet the inclusion and exclusion criteria; 22 subjects were eligible, of whom the first 20 respondents by date of received informed consent were included. When at home, two subjects always voided sitting; eight always voided standing. Outdoors, almost everybody (N=19) always voided standing. Eight

subjects had no specific preference for a certain voiding position; five preferred the sitting and seven the standing position. Those who did not respond or refused were significantly older than the participants (3.7 years, p=0.04), but did not differ with respect to mean last recorded flow or symptom score.

data quality control

We found two discrepancies between the official randomisation order and the actual order as registered by the participants in their diaries. According to the intention to treat principle we used the official randomisation order in the analyses. One subject found the procedure too time-consuming to complete the study, and stopped after having collected 34 measurements.

Artefacts required major corrections (>3 ml/s) of the maximum urinary flow in about 10% of the flow recordings, and made five measurements non-interpretable. The visual readings of the flow recordings by the two investigators were almost perfectly correlated (R=0.996), although there was a systematic difference of 0.43 ml/s (SD 0.46, range -4.0 to +0.9 ml/s). The correlation between the visual readings and the automated registration was R=0.83 for both investigators.

results of the 20 experiments and the meta-analysis

primary analysis

The results are shown in table 4.2 (page 64). The overall difference in maximum urinary flow (Q_{max}diff) was 0.55 ml/s (p=0.02), ranging from -1.64 to +2.80 ml/s. In 14 subjects mean Q_{max} was higher when standing, in 6 subjects when sitting. In five subjects these differences reached statistical significance (four in favour of standing, one in favour of sitting). The overall difference in average urinary flow (Q_{ave}diff) was also 0.55 ml/s (range -0.64 to +3.79; p=0.006). The overall difference in voided volume (Vvdiff) and voiding time (Vtdiff) were 11.0 ml (-40.5 to +87.8 ml; p=0.06) and 0.38 s (-13.4 to +14.8 s; p=0.75), respectively.

secondary analysis

The Dan-PSS correlated with the difference in maximum urinary flow (Q_{max}diff): for those with more bothersome symptoms standing was more favourable (r_s=0.45, p=0.049). No relation was found with age (r_s=-0.21, p=0.38), IPSS score (r_s=0.35, p=0.13), mean Q_{max} (r_s=0.32, p=0.17), or pre-study preference (p=0.81).

table 4.1. Characteristics of the study population

subject	characteristics				
	age (years)	IPSS	Dan-PSS	mean Q_{max} (ml/s)	preferred voiding position
a	67	19	25	11.8	sitting
b	60	14	7	8.3	no
c	66	10	12	13.1	no
d	69	13	3	8.0	no
e	61	15	25	18.6	no
f	61	13	3	14.3	sitting
g	75	16	31	10.4	standing
h	67	13	1	13.9	no
i	79	19	30	11.8	standing
j	76	16	20	5.4	standing
k	77	18	24	11.8	sitting
l	65	15	16	21.4	no
m	60	11	9	16.2	sitting
n	68	10	4	15.4	standing
o	63	12	14	6.7	standing
p	60	19	35	17.3	standing
q	62	19	31	13.7	sitting
r	77	15	6	3.7	no
s	65	12	7	9.3	standing
t	75	16	20	7.5	no
overall	**68.1**	**14.8**	**16.2**	**11.9**	

IPSS = International Prostate Symptom Score; Dan-PSS = Danish Prostate Symptom Score; Q_{max} = maximum urinary flow

subset analysis

Both subset analyses showed similar results to those of the primary analysis. Including only 'spontaneous during-the-day measurements' the overall Q_{max}diff was 0.79 ml/s (p=0.14; 496 flows). Using only the 542 flows with a voided volume ≥150 ml, the overall Q_{max}diff was 0.32 ml/s (p=0.20). The number of available flows in the standing position was significantly higher than in the sitting position (286 versus 256, p=0.03).

table 4.2. Results of the 20 single case experiments and the meta-analysis

subject	outcome measures #							
	Q_{max}diff (ml/s)	*p-value* *	Q_{ave}diff (ml/s)	*p-value* *	VVdiff (ml)	*p-value* *	Vtdiff (s)	*p-value* *
a	2.80	*.00*	1.37	*.00*	-40.5	*.14*	-13.4	*.02*
b	1.78	*.01*	0.92	*.01*	40.2	*.20*	-1.9	*.78*
c	-0.07	*.93*	0.27	*.52*	28.9	*.49*	1.2	*.86*
d	-1.05	*.20*	0.25	*.75*	18.6	*.39*	2.7	*.68*
e	1.99	*.18*	1.44	*.10*	6.7	*.87*	-1.2	*.93*
f	-1.36	*.15*	-0.64	*.31*	16.1	*.56*	1.3	*.82*
g	1.85	*.00*	1.44	*.00*	-1.2	*1.00*	-6.5	*.02*
h	0.33	*.77*	0.88	*.04*	-14.0	*.35*	-5.9	*.35*
I	1.04	*.12*	0.67	*.51*	17.5	*.33*	2.6	*.45*
j	-0.76	*.08*	-0.21	*.39*	8.9	*.57*	9.8	*.27*
k	0.14	*.86*	-0.30	*.40*	6.9	*.94*	3.0	*.61*
l	1.11	*.38*	0.75	*.29*	-17.0	*.80*	-3.6	*.58*
m	1.11	*.30*	0.65	*.19*	-26.6	*.59*	-5.3	*.41*
n	0.70	*.25*	0.48	*.38*	6.1	*.74*	-.5	*.95*
o	0.05	*.93*	0.24	*.53*	10.8	*.52*	-1.7	*.87*
p	1.40	*.28*	3.79	*.00*	24.3	*.50*	11.7	*.01*
q	0.97	*.43*	-0.46	*.42*	87.8	*.09*	14.8	*.03*
r	1.24	*.00*	0.46	*.00*	60.8	*.00*	6.7	*.54*
s	-0.55	*.33*	-0.44	*.31*	1.2	*.95*	2.2	*.67*
t	-1.64	*.00*	-0.57	*.06*	-15.2	*.43*	-8.5	*.47*
overall	**0.55**	***p=0.02*** **	**0.55**	***p=0.006*** **	**11.0**	***p=0.06*** **	**0.38**	***p=0.75*** **

Q_{max}diff = difference in maximum urinary flow; Q_{ave}diff: difference in average urinary flow; VVdiff: = difference in voided volume; Vtdiff = difference in voiding time.
Differences are calculated as: value of the standing minus value of the sitting voiding position.
* Non-directional p-value (individual subjects); related to the null hypothesis of equality between the two voiding positions (e.g. Q_{max}diff=0).
** Directional p-value (meta-analysis); presented p-values are related to the one-sided hypothesis (e.g. overall Q_{max}diff>0), and should therefore be interpreted using an adjusted significance level (α=0.025).

Discussion

main results

Both maximum and average urinary flow were higher when subjects were urinating standing. The voided volume tended also to be higher. A higher bother score (Dan-PSS) correlated with a higher maximum urinary flow when standing. A correlation with age or pre-study preference in voiding position was not found.

strenghts and limitations

We chose the single case design as a specific form of a crossover design to improve precision. Intra-individual differences due to e.g. the circadian rhythm, differences in bladder filling and general variability, were overcome by the large number of measurements per subject. In the home-based setting the privacy of the participants during the measurements was fully respected, and a minimal influence of the experiment on the normal micturition pattern and habits guaranteed. Due to the randomisation the statistical power was higher than in other studies evaluating voiding positions. We replicated the single case experiments in a sample of the indicated population: aged men with stable moderate lower urinary tract symptoms.

The correlation between Q_{max}diff and Dan-PSS score was not hypothesized beforehand. Our sample size calculation was not aimed at finding such correlations in our exploratory secondary analyses. There may well be important correlations with age, symptoms or mean Q_{max} that were not found in our study.

Although a voided volume ≥150 ml is regarded necessary to acquire reproducible results[3], for our aim we chose otherwise: we were not interested in possible maximum urinary flow at a standardized pre-voiding bladder volume, but in actual real-life flow. In a home-based setting one can expect a considerable number of low-volume flows.[21] Such urinations - in our study 30% - are part of the real life situation in patients with LUTS. These flows were not equally divided between the two voiding positions, or between the subjects. Subjects with a higher number of low-volume flows are generally more obstructed.[21] Selection of high-volume flows would therefore in our study lead to unnecessary loss of statistical power, and non-generalisable and biased results, with consequently an underestimation of the real effect. A sound analysis should include all available measurements, like we did in our primary analysis, provided possible differences in voided volume cannot be attributed to external influences like bedtime and defecation. The first subset analysis shows that our design met these requirements.

We could - of course - not measure the post-void residual volume. This would be impossible without disturbing the normal voiding pattern in the home-based setting. One may argue that -given the large number and the randomisation of the measurements - the pre-voiding bladder volume in both positions should be

equal. The position-related difference in voided volumes may than serve as proxy of position-related differences in residual volume.

Urinary flow measurements may show artefacts, especially with respect to the maximum urinary flow. The number of artefacts may be up to 18%, even in well-controlled settings.[21,22] The need for visual correction has been described before.[21,31] Many studies however, fail to report how artefacts are handled. We used a transparent way of visual correction (two independent investigators blinded for voiding position using rules described by others).

what is already known?

Our literature search revealed four studies[15,16,32,33] in which the standing and the sitting voiding position were compared in men. In healthy young volunteers, two studies[32,33] found no differences, while one[15] found a difference in favour of sitting. One study included subjects with LUTS suggestive for benign prostatic hyperplasia.[16] The (non-significant) results suggest that standing would be the favourable voiding position. Compared to our study, the subjects were younger and had higher symptom scores. They had relatively large prostates. Relatively few flow measurements were used ('at least two'); measurements with a voided volume <150 ml were excluded. This may have reduced the ability to detect existing differences.

what this study adds

Overall, the standing voiding position is best. Both maximum and average urinary flow were higher in the standing position. Apparently, unlike young healthy men, the urinary outflow of patients with moderate LUTS is generally more obstructed while urinating sitting. This raises etiological and physiological questions. We find it hard to imagine that position-related differences in bladder contraction would be responsible for the differences in urinary flow. Abdominal straining may have had some influence, but cannot be responsible for the total effect.[34]

The overall effect size seems small. One might therefore conclude that it is hardly relevant. However, the effect is about 35-50% of the net effect of α-blockers[35,36] We think such a general mean improvement in flow parameters is clinically relevant, especially since it involves an easily applicable lifestyle intervention with a day-to-day effect.

Like Unsal[16], we found a tendency for an increased voided volume in the standing position in subjects with LUTS. Two of the 20 subjects even had a significant difference in voided volume >+60 ml. A post-void residual volume >100 ml results in an increased risk of acute urinary retention.[37] If we may attribute a (precisely measured) position-related difference in voided volume to a difference in post-void residual volume, the clinical consequence would be that sitting increases the risk of acute urinary retention. This conclusion warrants

confirmation by other studies, and a long-term follow-up.

Although the position-related differences in flow parameters were quite large in some individuals, they were generally too small to be noticed by the subject himself. This explains the lack of correlation between pre-study preference and the difference in maximum urinary flow found. Patient's experience therefore cannot guide a more individualized lifestyle advice. The positive correlation suggests that the bother score may help. Otherwise, a short period of comparative home-based urinary flow measurements would be possible. Such an intervention may be an effective, non-invasive and cheap possibility to improve the outcome of patients with moderate LUTS, and should be studied in a clinical trial.

Our study shows that the choice of the design has an important impact on the conclusions drawn by researchers. Our replicated single case study was able to detect statistically significant differences that were smaller than the non-significant differences found by Yamanishi[32] in an almost equal number of subjects, and Unsal[16,33] in about twice as many. Our design was efficient and relatively cheap. Single case experiments can be performed if the intervention has no carry-over effect and the condition under study is relatively stable during the study period. In our view, such experiments should be considered more often as a possible design to investigate new research questions for which funding of (otherwise needed) large studies is difficult to obtain, which often is the case in lifestyle interventions.

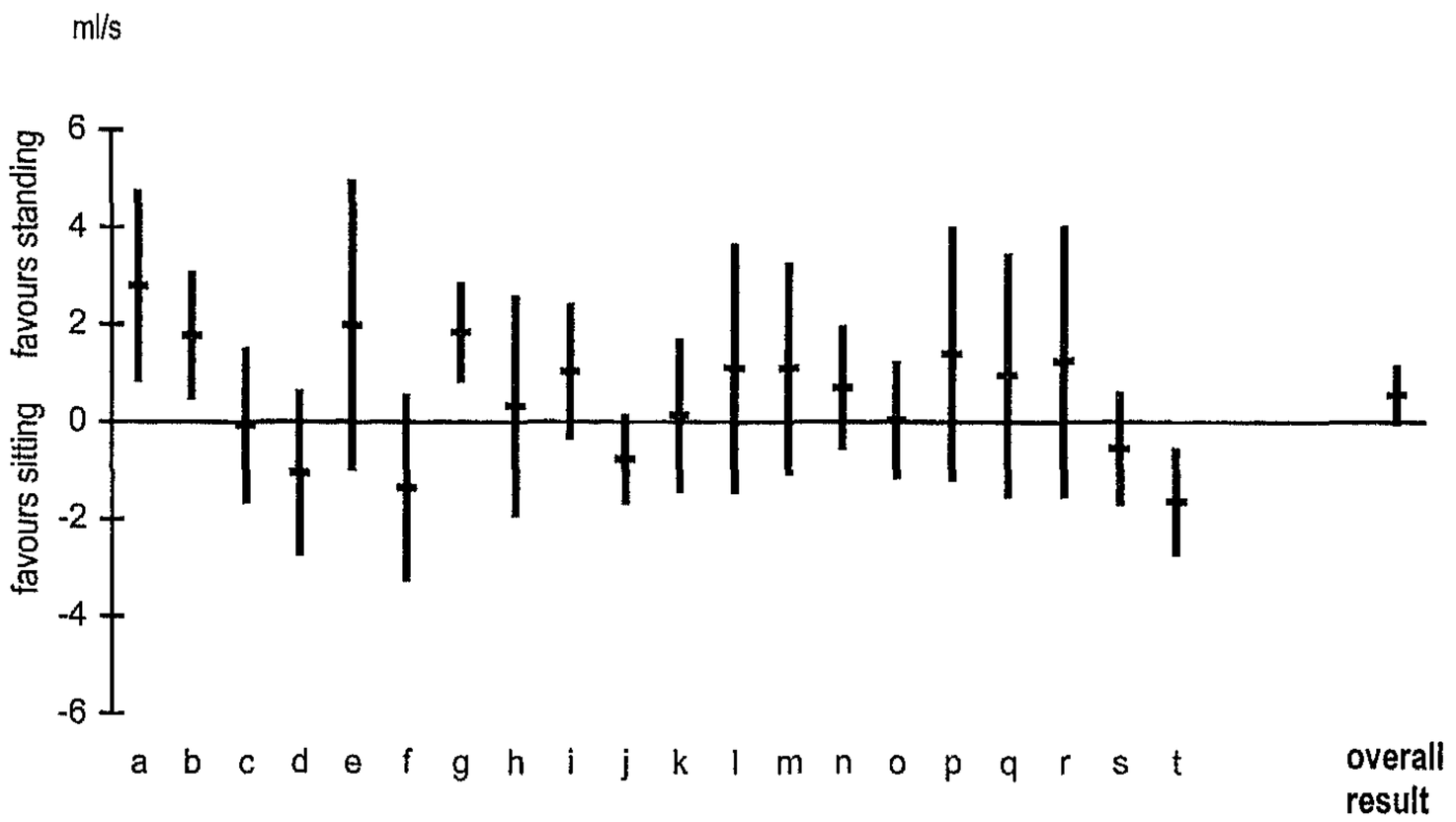
Our study focused on physiologic differences between voiding positions in elderly men with moderate LUTS. We must be aware that there may be other reasons to choose a certain voiding position, such as comfort, hygiene or religion. We did not investigate to what extent such arguments determine a patient's voiding position, or whether an advice of his doctor could alter this. However, we conclude that there is no basis for a general advice to urinate sitting in patients with lower urinary tract symptoms.

References

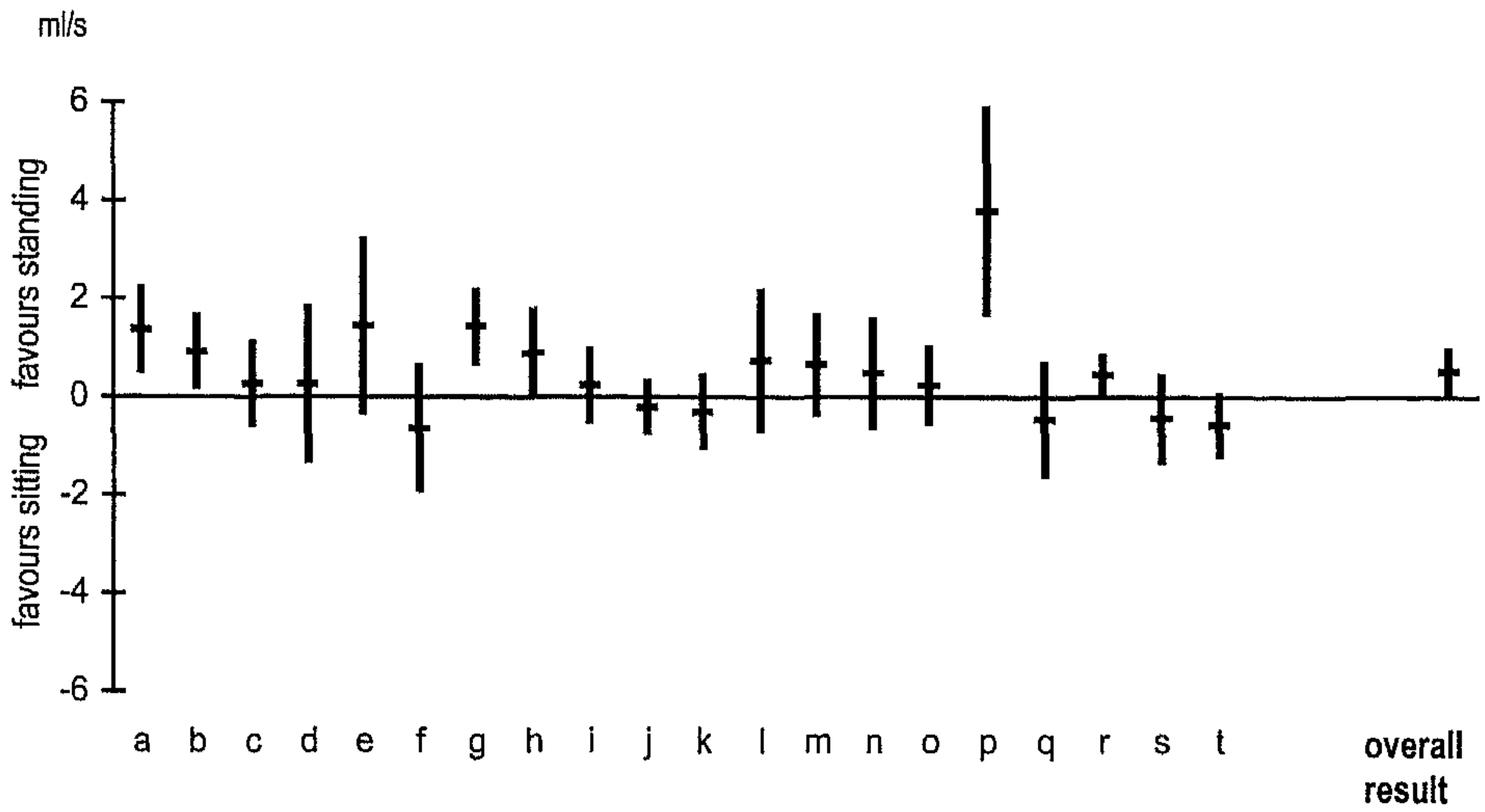
1. Wolfs GG, Knottnerus JA, Janknegt RA. Prevalence and detection of micturition problems among 2,734 elderly men. J Urol 1994;152(5 Pt 1):1467-70.
2. American Urological Association. Guideline on the management of Benign Prostatic Hyperplasia (BPH), 2003 (Accessible via www.auanet.org).
3. De la Rosette JJ, Madersbacher S, Alizivatos G, Rioja Sanz C, Emberton M, Nordling J. Guideline on benign prostatic hyperplasia: European Association of Urology, Dec 2004.
4. Platz EA, Kawachi I, Rimm EB, Colditz GA, Stampfer MJ, Willett WC, et al. Physical activity and benign prostatic hyperplasia. Arch Intern Med 1998;158(21):2349-56.
5. Brown CT, Van der Meulen J, Mundy AR, Emberton M. Lifestyle and behavioural interventions for men on watchful waiting with uncomplicated lower urinary tract symptoms: a national multidisciplinary survey. BJU Int 2003;92(1):53-7.
6. Hirst G, Butler J, Lajoie V. Clinical Practice Guidelines. The management of uncomplicated lower urinary tract symptoms in men: National Health an Medical Research Council of Australia, 2000.
7. Anonymous. Referral advice. A guide to appropriate referral from general to specialist services. London, UK: National Institute for Clinical Excellence, 2001.
8. Klomp MF, Gercama AJ, de Jonge-Wubben JG, Mulders AH, Romeijnders AC, Rosmalen CF, et al. NHG-Standaard Bemoeilijkte mictie bij oudere mannen. Huisarts en Wetenschap 1997;37:357-65.
9. Mayo Clinic. http://www.mayoclinic.com/health/prostate-gland-enlargement/DS00027/UPDATEAPP=0&DSECTION=9 (Accessed Nov. 16 2007).
10. Smith TL. Bathroom tips for benign prostatic hyperplasia, 2004. http://www.sjmercyhealth.org/body.cfm?xyzpdqabc=0&id=431&action=detail&AEProductID=HW_Catholic&AEArticleID=hw55013&AEArticleType=Frame#hw55013-sec (Accessed Nov. 16 2007).
11. Brooks City-Base San Antonio, TX: US Airforce, the 59th Medical Squadron. http://www.brooks.af.mil/web/consult_service/waiver%20guide/Genito-Urinary/Benign%20Prostatic%20Hypertrophy.htm (Accessed Nov. 16 2007).
12. Health on the Net foundation. www.hon.ch (Accessed Nov. 16 2007).
13. Helping Yourself: DrugDigest. http://www.drugdigest.org/DD/HC/HelpingYourself/0,4052,550246,00.html (Accessed Nov. 16 2007).
14. http://www.webmd.com/hw-popup/bathroom-tips-for-benign-prostatic-hyperplasia (Accessed Nov. 16 2007).
15. Eryildirim B, Tarhan F, Kuyumcuoglu U, Erbay E, Pembegul N. Position-related changes in uroflowmetric parameters in healthy young men. Neurourol Urodyn 2006;25(3):249-51..
16. Unsal A, Cimentepe E. Effect of voiding position on uroflowmetric parameters and post-void residual urine volume in patients with benign prostatic hyperplasia. Scand J Urol Nephrol 2004;38(3):240-2.
17. Drach GW, Layton TN, Binard WJ. Male peak urinary flow rate: relationships to volume voided and age. J Urol 1979;122(2):210-4.
18. Haylen BT, Ashby D, Sutherst JR, Frazer MI, West CR. Maximum and average urine flow rates in normal male and female populations - the Liverpool nomograms. Br J Urol 1989;64:30-8.
19. Madersbacher S, Klingler HC, Schatzl G, Stulnig T, Schmidbauer CP, Marberger M. Age related urodynamic changes in patients with benign prostatic hyperplasia. J Urol 1996;156(5):1662-7.
20. Golomb J, Lindner A, Siegel Y, Korczak D. Variability and circadian changes in home uroflowmetry in patients with benign prostatic hyperplasia compared to normal controls. J Urol 1992;147(4):1044-7.
21. Witjes WP, Wijkstra H, Debruyne FM, de la Rosette JJ. Quantitative assessment of uroflow: is there a circadian rhythm? Urology 1997;50(2):221-8.
22. Sonke GS, Klemeney LA, Verbeek AL, Kortmann BB, Debruyne FM, de la Rosette JJ. Low reproducibility of maximum urinary flow rate determined by portable flowmetry. Neurourol Urodyn 1999;18(3):183-91.
23. Onghena P, Edgington ES. Customization of pain treatments: single-case design and analysis. Clin J Pain 2005;21(1):56-68; discussion 69-72.
24. Norg RJ, Portegijs PJ, van de Beek K, van Schayck O, Knottnerus JA. A decision aid for GPs for the treatment of elderly male patients with lower urinary tract symptoms (LUTS). Fam Pract 2006;23(6):666-73.
25. Reynard JM, Peters TJ, Lim C, Abrams P. The value of multiple free-flow studies in men with lower urinary tract symptoms. Br J Urol 1996;77(6):813-8.
26. Jepsen JV, Leverson G, Bruskewitz RC. Variability in urinary flow rate and prostate volume: an investigation using the placebo arm of a drug trial. J Urol 1998;160(5):1689-94.

27. Onghena P, van Damme G. SCRT 1.1: Single Case Randomization Tests. Behavior Research Methods, Instruments, & Computers 1994;26(4):369.
28. Onghena P, Edgington ES. Randomization tests for restricted alternating treatments designs. Behav Res Ther 1994;32(7):783-6.
29. Van den Noortgate W, Onghena P. Hierarchical linear models for the quantitative integration of effect sizes in single-case research. Behav Res Methods Instrum Comput 2003;35(1):1-10.
30. Hansen BJ, Flyger H, Brasso K, Schou J, Nordling J, Thorup Andersen J, et al. Validation of the self-administered Danish Prostatic Symptom Score (DAN-PSS-1) system for use in benign prostatic hyperplasia. Br J Urol 1995;76(4):451-8.
31. Witjes WP, de la Rosette JJ, Zerbib M, Vignoli GC, Geffriaud C, Debruyne FM, et al. Computerized artifact detection and correction of uroflow curves: towards a more consistent quantitative assessment of maximum flow. Eur Urol 1998;33(1):54-63.
32. Yamanishi T, Yasuda K, Sakakibara R, Hattori T, Minamide M, Yuki T, et al. Variation in urinary flow according to voiding position in normal males. Neurourol Urodyn 1999;18(6):553-7.
33. Unsal A, Cimentepe E. Voiding position does not affect uroflowmetric parameters and post-void residual urine volume in healthy volunteers. Scand J Urol Nephrol 2004;38:469-71.
34. Reynard JM, Peters TJ, Lamond E, Abrams P. The significance of abdominal straining in men with lower urinary tract symptoms. Br J Urol 1995;75(2):148-53.
35. Wilt TJ, Mac Donald R, Rutks I. Tamsulosin for benign prostatic hyperplasia. Cochrane Database Syst Rev 2003(1):CD002081.
36. Wilt TJ, Howe RW, Rutks IR, MacDonald R. Terazosin for benign prostatic hyperplasia. Cochrane Database Syst Rev 2002(4):CD003851.
37. McNeill SA, Hargreave TB, Geffriaud-Ricouard C, Santoni J, Roehrborn CG. Postvoid residual urine in patients with lower urinary tract symptoms suggestive of benign prostatic hyperplasia: pooled analysis of eleven controlled studies with alfuzosin. Urology 2001;57(3):459-65.

appendix 4.1. Reconstructed differences and 95% confidence intervals in maximum urinary flow (Q_{max}diff) between the standing and sitting voiding position

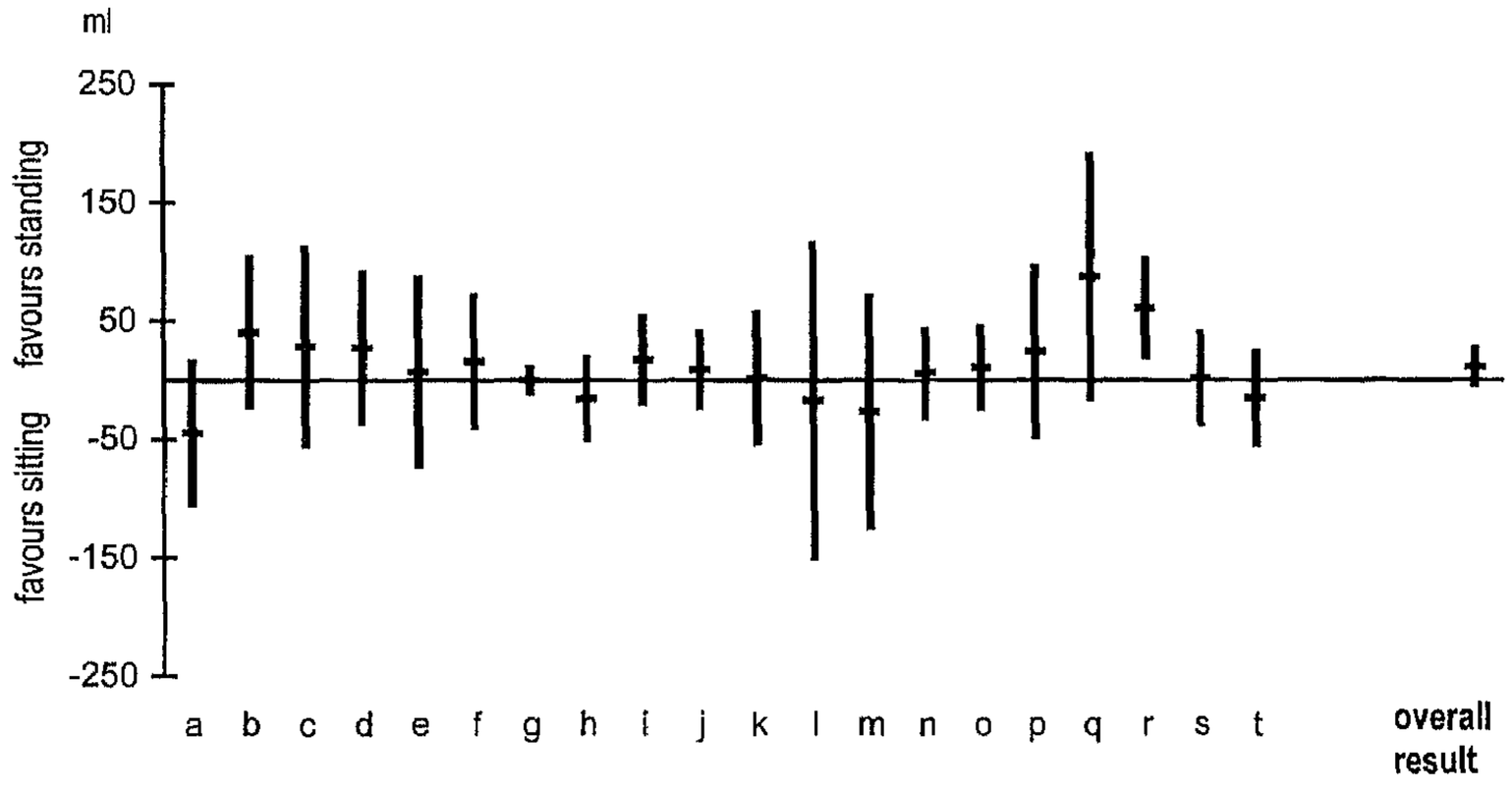


appendix 4.2. Reconstructed differences and 95% confidence intervals in average urinary flow (Q_{ave}diff) between the standing and sitting voiding position

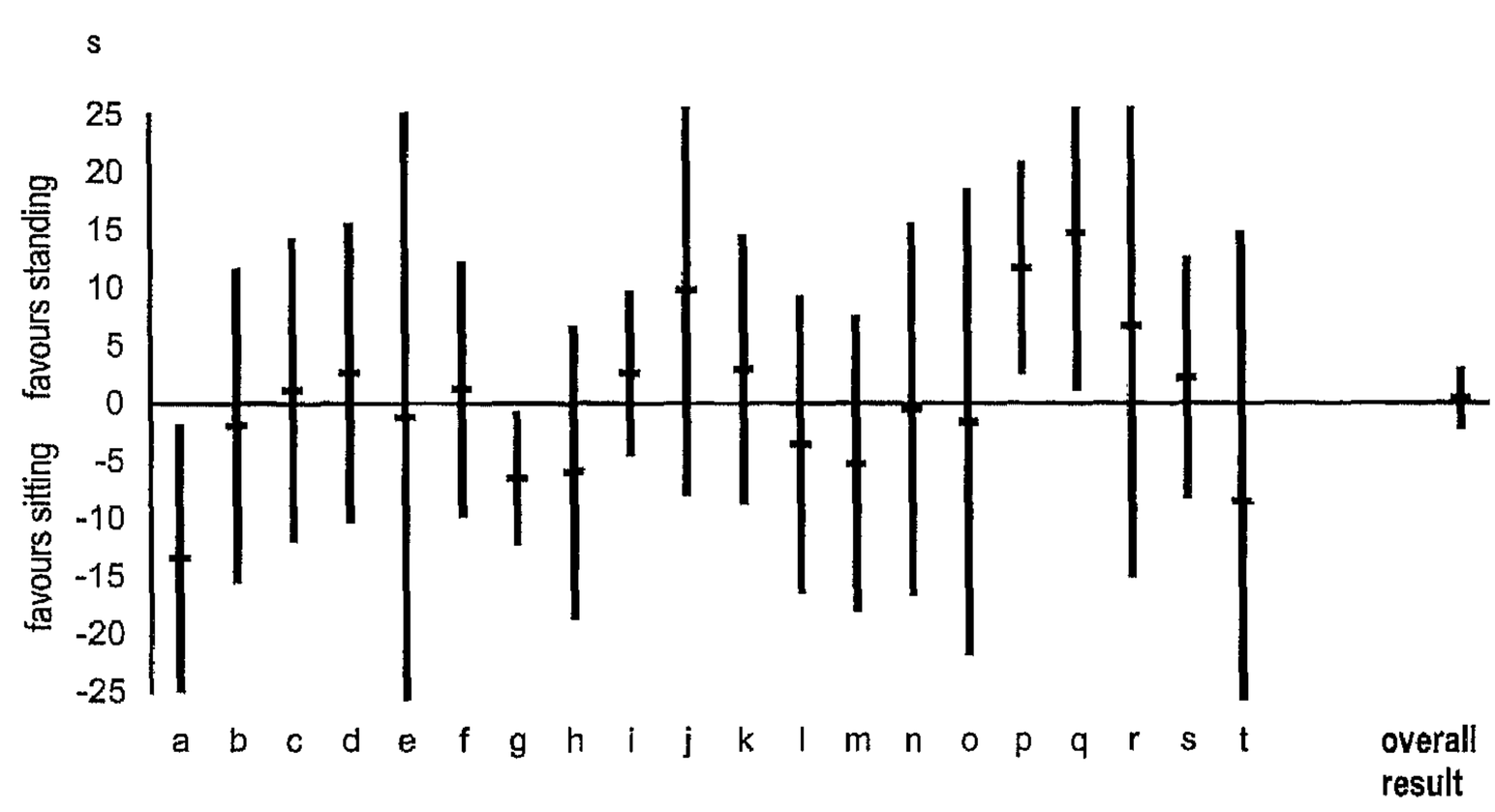


Note! The figures shown in this appendix are meant to visualise the results. However, they are based on tables that are used in a different statistical test procedure (t-test) than has been used in the analysis of this study (randomisation test). They should therefore be interpreted with caution. The lines represent the mean and (constructed) 95% confidence intervals of all possible differences between the measured values of Q_{max} and Q_{ave} resp. in the standing and sitting position.

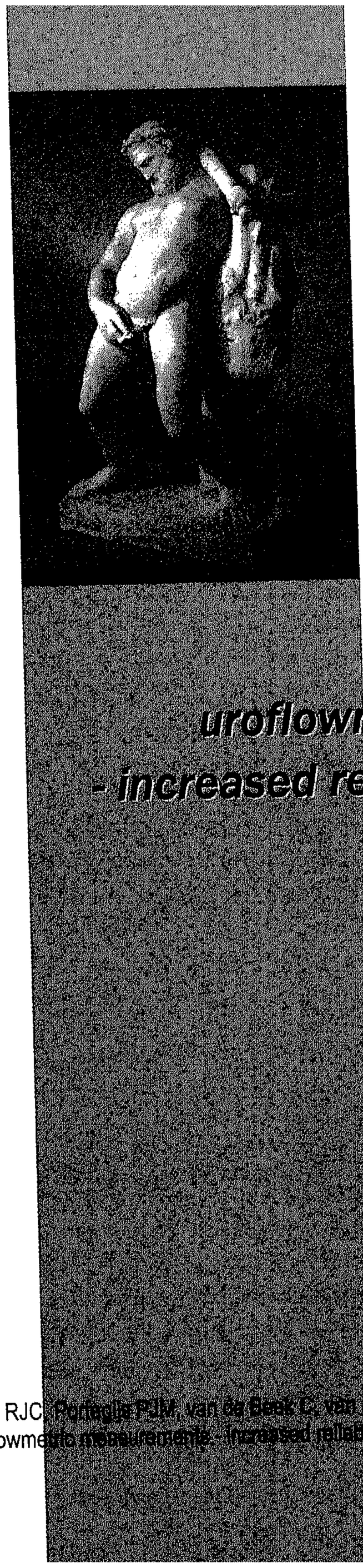
appendix 4.3. Reconstructed differences and 95% confidence intervals in voided volume (VVdiff) between the standing and sitting voiding position



appendix 4.4. Reconstructed differences and 95% confidence intervals in voiding time (Vtdiff) between the standing and sitting voiding position



Note! The figures shown in this appendix are meant to visualise the results. However, they are based on tables that are used in a different statistical test procedure (t-test) than has been used in the analysis of this study (randomisation test). They should therefore be interpreted with caution. The lines represent the mean and (constructed) 95% confidence intervals of all possible differences between the measured values of VV and Vt resp. in the standing and sitting position.



CHAPTER 5

Objective reading of uroflowmetric measurements - increased reliability and precision

Norg RJC, Portegijs PJM, van de Beek C, van Schayck CP, Knottnerus JA. Objective reading of uroflowmetric measurements - increased reliability and precision. *Submitted*

Abstract

background:

Uroflowmetry in patients with Lower Urinary Tract Symptoms is susceptible to artefacts. Visual reading and correction of these artefacts is therefore needed to guarantee the validity of test results, but is to some extent subjective. Automated reading would make this process more objective, quicker and therefore easier, both in the clinical and in the research setting.

objective:

To develop an automated tool to correct for artefacts.

methods and materials:

design: modelling study comparing visual analysis and several computer-assisted smoothing functions of uroflow measurements.

study material: A total of 754 uroflow measurements of 20 male subjects with moderate LUTS (IPSS 8-19). The maximum urinary flow (Q_{max}) was visually read by two independent readers and the mean was taken as the reference standard.

development: A correction algorithm was developed to correct for artefacts. The algorithm used a first smoothing function of one second and a second smoothing function of four seconds in case of important artificial peaks. These peaks were identified using the first and second derivative of the flow curve. Opening pressure artefacts were corrected by ignoring the first 1.2 seconds of the flow curve.

outcome measures: 1. The correlation between the visual readings and the corrected Q_{max}; 2. difference between the visual and corrected Q_{max}.

results:

The correlation between the original data and reference reading was R=0.82. Correction of artefacts led to a correlation of 0.99. The median difference between the visual reading and original Q_{max} was 0.54 ml/s ($P_{2.5}$,$P_{97.5}$ = -1.03 , +6.91 ml/s); after correction it was 0.14 ml/s ($P_{2.5}$, $P_{97.5}$ = -1.59, +0.91 ml/s).

conclusion:

The algorithm we developed closely approximated the visual readings of two independent readers. The algorithm-reference difference was almost the same as the difference between the two readers. Since such an algorithm is more objective and efficient (cheaper and less time consuming), such an algorithm should be made available in all uroflow equipment.

Introduction

Uroflowmetry is widely used in the diagnostic evaluation of patients with lower urinary tract symptoms (LUTS). Although uroflowmetry is not able to discern between all possible causes of the symptoms, it is a recommended tool in the management and follow-up of patients with LUTS.[1,2] A normal urinary flow obviates the need for invasive testing. Urinary flow recordings can be used to objectify the possible decrease of urinary flow over time in a follow-up regimen, and to evaluate the effect of treatment.

Maximum urinary flow (Q_{max}) is the most important parameter measured with uroflowmetry. A major disadvantage is its intra-individual variability. Some of the variability has a 'biological' origin, such as different bladder filling, or a learning effect in repeated measurements. But apart from this, artefacts can influence the test results. Such artefacts can be caused by increased opening pressure, abdominal pressure, bumps against the flow meter or manipulation of the penis during voiding leading to narrowing of the prepuce or swaying of the urine flow. Even in well-controlled situations a considerable number of such artefacts remain.[3-5] Depending on the type of flow meter, up to 68 or 94% of the flows may show such artefacts.[5] Consequently, all test results must be checked visually before they can be registered, and it is recommended that at least two flow curves are obtained at a time. [1,3,4] In a clinical setting this is, however, not very practical. Patients need some time to drink and produce a full bladder for a second measurement. In a research setting artefacts lead to a decrease of statistical power and therefore necessitate a higher number of measurements or a larger study population, both leading to increased costs. Furthermore, visual reading implies a (more or less) subjective interpretation of the study data. It is time-consuming and (thereby) costly.[6]

There is no extensive literature describing a correct and validated way of reading uroflowmetric curves. It is common to describe the way in which flow recordings are read correctly as 'the way in which an experienced urologist would do this'.[6] It is unknown how thoroughly visual artefact correction is done in daily practice. Although objective data gathering is a key element in clinical research, most papers of randomised controlled trials do not describe this. The way in which variability and artefacts are handled is hardly ever made explicit, even if peak urinary flow is an outcome measure.

Having performed a uroflowmetric study that required precise reading of a large number of uroflow measurements, we decided to use these data to find out whether a computerized analysis would be able to read/interpret equally well.[7] Two researchers had already read all flow readings. We aimed to develop a simple algorithm that closely approximated these visual recordings.

Methods and materials

patient data

We used the data of a series (N=20) of randomised single case studies comparing the urinary flow measures in standing and sitting voiding position.[7] The subjects were elderly male volunteers (≥ 59 years of age) with stable moderate LUTS (International Prostate Symptom Score (IPSS) ≥ 8 and ≤ 19 during the previous 4 years). Each subject voided, in random sequence, 20 times sitting and 20 times standing in a uroflowmeter (Urodyn 1000 mictiograph). Measurements were done by the subject himself, at home, in about a week. All measurements were included; no restrictions were made with respect to voiding time or voided volume.

flow curves

The flow curves were available as paper prints of the mictiograph. We scanned all flows with a UMAX flatbed scanner at 300 dpi, and subsequently read (digitally 'remastered') the flow curves using a custom made curve recognition software program, written in MatLab 7.2 (Mathworks®, www.mathworks.com). The data were then stored into a database for analysis (SPSS 12.0 for Windows). In order to check the scanning process, we compared the original data as given by the mictiograph with the digital data in the database. Outliers - where the digitally stored Q_{max} differed > 2 ml/s from the original Q_{max} - were analysed separately (additional analysis). All other flows were used for the primary analysis.

reference standard

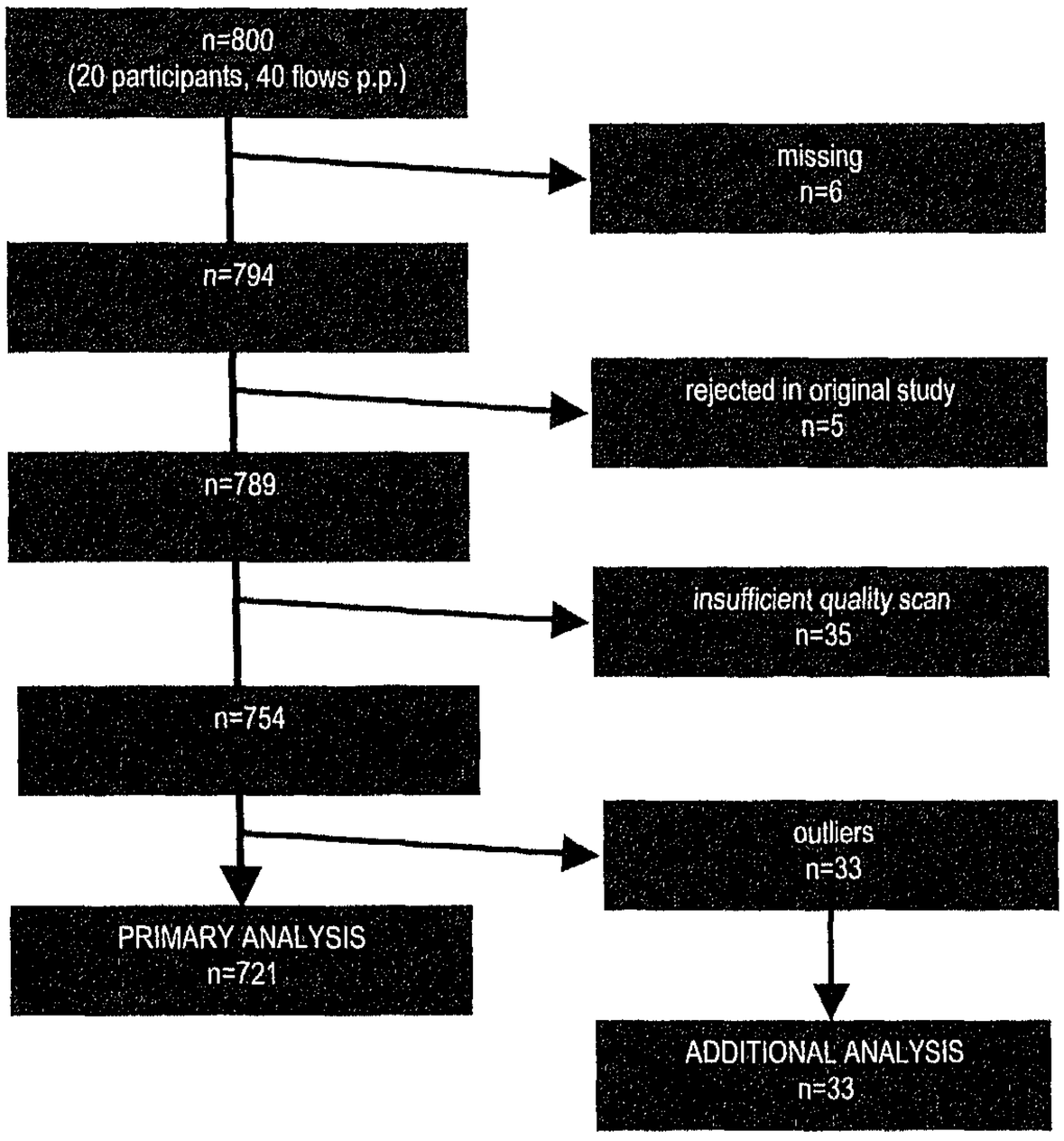
Few publications are available that describe how visual reading can or should be done. We used the method described by Witjes et al.[5] All paper flow curves were visually read for Q_{max} by two independent readers (RJCN and PJMP). If the difference between the readers was ≥ 1 ml/sec both investigators reassessed the flows to check for possible misinterpretation. Afterwards, the mean of both readings was taken for the analyses and served as the reference standard ('gold standard').

correction of artefacts

Artefacts can arise from several factors. There may be a high opening pressure. This produces a high peak early in the curve. Patients may also squeeze the penis or prepuce during voiding causing large spikes later on in the curve often followed by compensating troughs. They may move their penis from the centre of the flow meter to the side and let the stream touch the rim before reaching the measurement area. Finally, there often is some 'noise' signal, i.e. small high-frequency disturbances of the signal throughout the whole flow curve.

We aimed at developing an algorithm that was able to recognise such artefacts, correct them, and show the thus corrected curve.

figure 5.1. Flow chart of the datamanagement



correction algorithm

High frequency artefacts due to 'noise' can be removed by 'smoothing' the curve with a moving average. A moving average replaces a data point with the mean value (average) of all data points within a certain time frame before and after that particular data point. This way, high frequency disturbances of the signal can be levelled to a more stable signal (figure 5.2b, see page 78). In order to reduce high frequency disturbances of the signal, a small time frame (small moving average) is sufficient; moving averages using larger time frames tend to underestimate 'true' peak values. We used a moving average of 1 second.

Artefacts due to *increased opening pressure* are dealt with very effectively by deleting all data within 1.2 seconds after the start of the flow curve. We programmed a linear alternative curve (figure 5.2c, see page 78).

figure 5.2. An original, remastered and corrected flow curve (participant nr 14, flow nr 13).

figure 5.2a. Urodyn® mictiograph: Q_{max} = *23.3 ml/s*

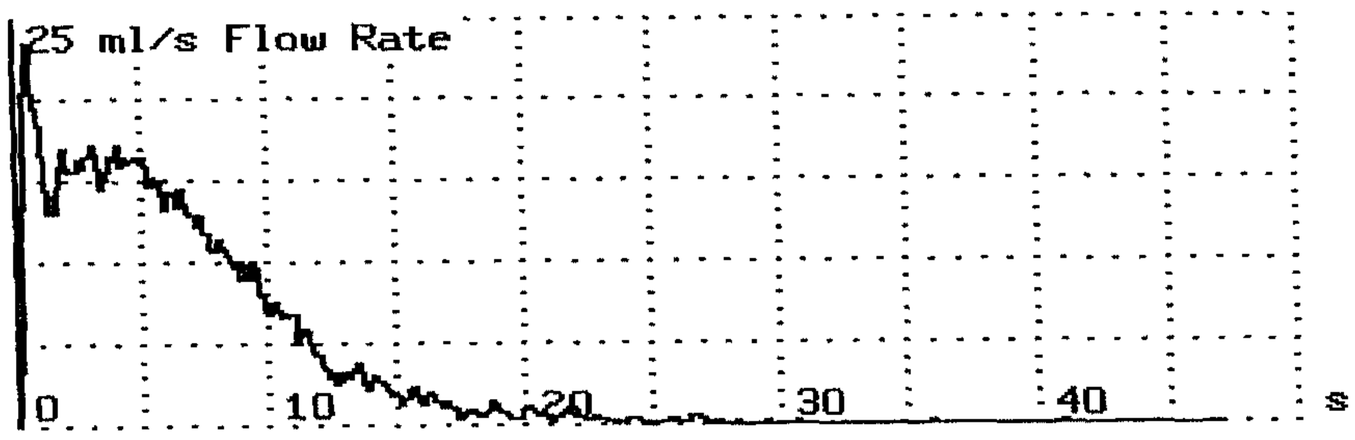
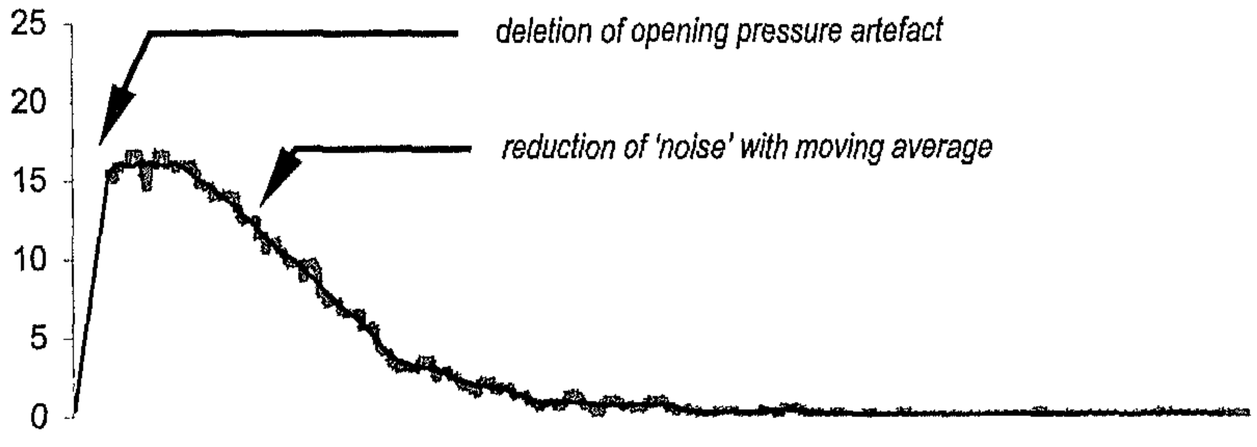


figure 5.2b. Reconstructed flow recording used in the analyses

figure 5.2c. Flow recording after corrections: Q_{max} = *16.6 ml/s*



Larger artefacts later on in the curve are most difficult to deal with. In order to recognise such artefacts we chose to take the first and second derivative of the curve. The *slope* of the original curve (at a particular point in time) comes back as the *value* of the first derivative at that point. And the *curvature* of the original curve (which is the *slope* of the first derivative) is the *value* of the second derivative. A steep slope can be an indication of an artefact, because it represents a strong increase (or decrease) of the urinary flow. High values (positive or negative) of the second derivative indicate sudden changes in the slope of the flow curve as occur at troughs and peaks. In other words, troughs (and peaks) in the curve of the second derivative may indicate peaks (and troughs) in the original flow curve. A deeper trough in the second derivative indicates sharper peaks in the flow curves (figure 5.3b, see page 80). Peak values of the first derivative (steep flow curve) followed by troughs in the second derivative (sharp peaks in the flow curve) indicate artefacts. Such artificial peaks with a maximum length of two seconds should be corrected, as should complementary troughs of at maximum two seconds that may follow such a peak. Where such peaks and troughs were identified the data of the flow curve were therefore replaced by the values of the moving average of a large time frame of four seconds (figure 5.3c, see page 80). Theoretically, the use of only the second derivative should be enough to identify peaks. However, this way too many peaks were identified that had little bearing on the value of Q_{max}. In other words: too many peaks were identified. Therefore, we used the first derivative to select among all identified peaks the *relevant* peaks which subsequently were corrected.

primary analysis

In order to analyse the effects of each step of our corrections we calculated the correlation between the corrected data and the visual readings, and calculated the median difference between the visual and corrected values of Q_{max}, as well as the $P_{2.5}$ and $P_{97.5}$. These lower and upper limits include 95% of the observations. This outcome was used as an indicator of increased precision.
For the effects on individual subjects, we repeated these calculations per subject. However, we used the P_{10} and P_{90}, comprising 80% of the flow measurements, because the number of flows per subject was limited. Outliers would unduly influence $P_{2.5}$ and $P_{97.5}$.

additional analysis

For this analysis a calculation of the difference between the visual reading and the corrected data would not be fair. Therefore we calculated only the increase in the correlation coefficient as a measure of increased precision for this subset of the data.

figure 5.3. An original and corrected flow curve (participant number 2, flow number 21).

figure 5.3a. Urodyn® mictiograph: Q_{max} = *10.3 ml/s*

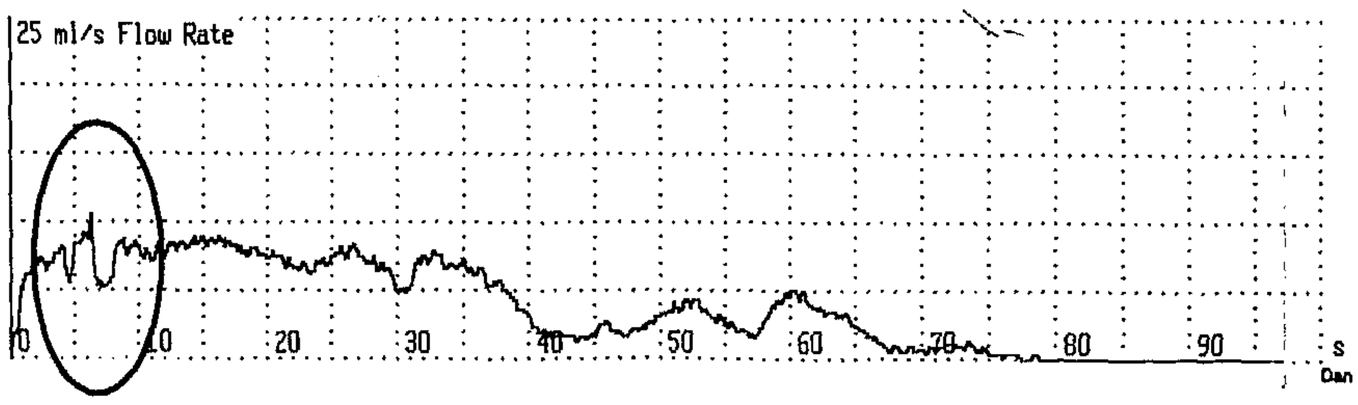


figure 5.3b. Flow curve and second derivative, indicating peaks an troughs in the flow curve

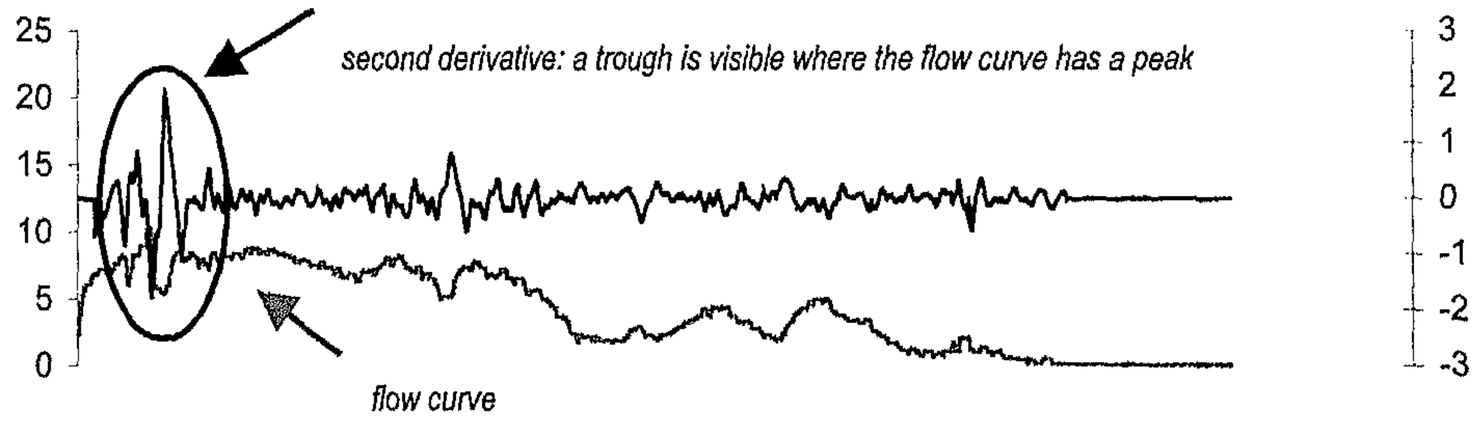


figure 5.3c. Flow recording after corrections: Q_{max} = *8.7 ml/s*

Results

data management

study material

A total of 800 flows should have been recorded in the original study (figure 5.1, page 77). However, one participant had not completed all measurements in the study, so 794 uroflows were available. Five flows had been rejected due to (visually) non-interpretable results. Technical problems with scanning the mictiograph's thermal paper prints led to 754 flows in our study sample.

quality of the data

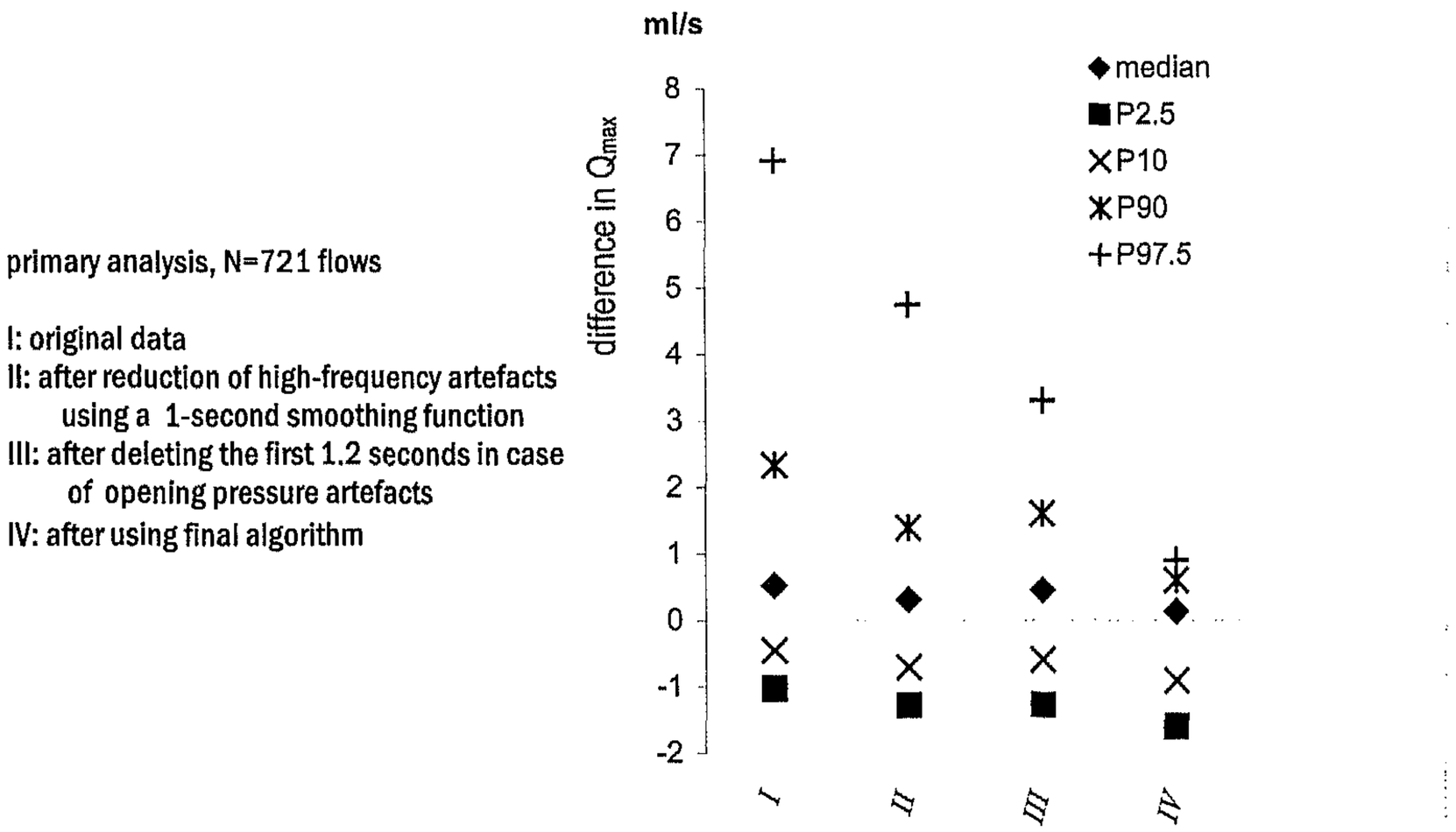
The correlation between the original (computed by the mictiograph) and the remastered Q_{max} was 0.98, indicating adequate scanning and curve recognition. The correlations of the original Q_{max} (0.83) and the remastered Q_{max} (0.82) with the visual readings (the reference standard) were almost equal. This shows that no selection bias had occurred due to the technical problems: both uroflows with and without artefacts were equally included. So we concluded that we could use the remastered data as a basis for further analysis.

analysis

primary analysis

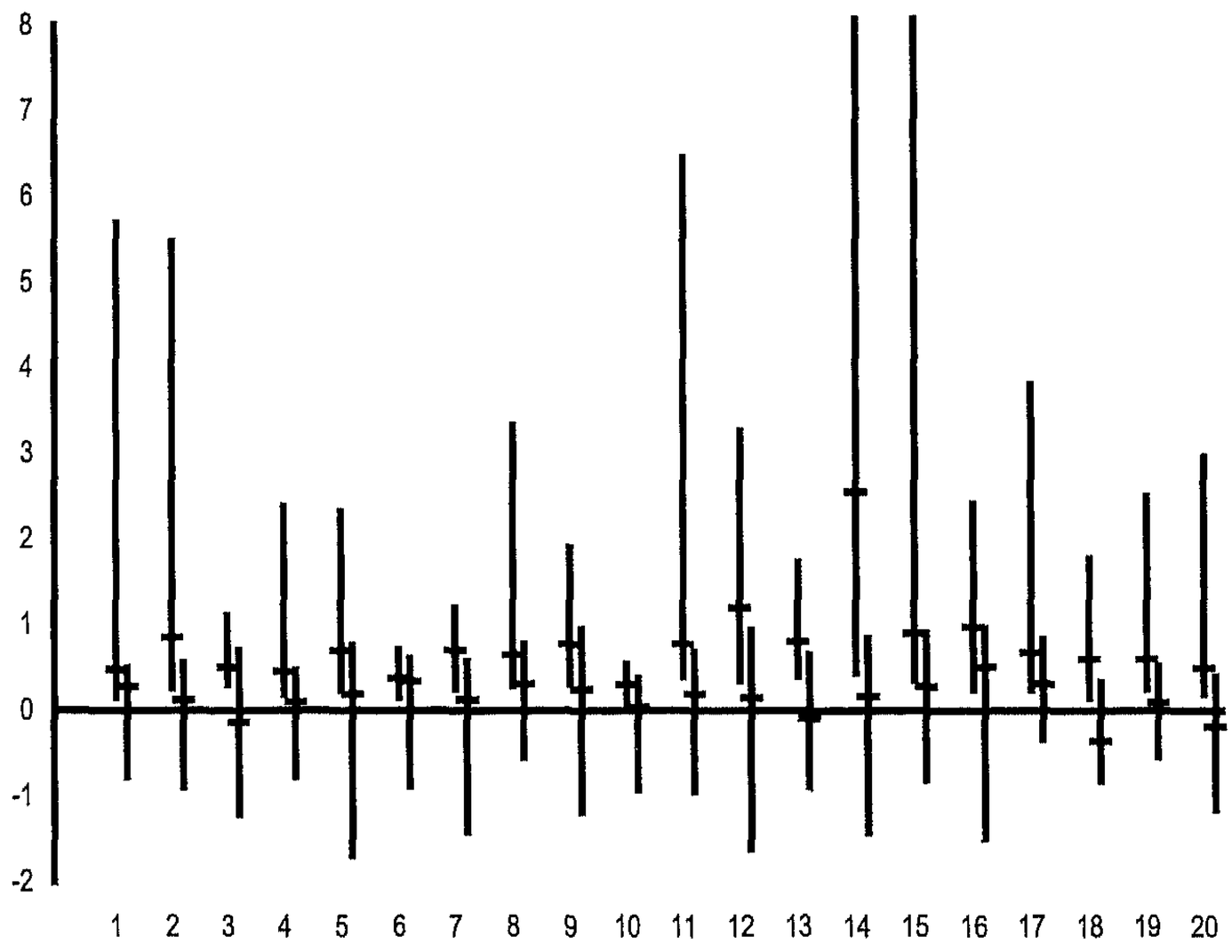
721 Flows could be included in the primary analysis. The correlation between the remastered data and reference reading was R=0.82. The median difference between the reference reading and the remastered Q_{max} was 0.54 ml/sec; 95% of the observations of the difference between Qmax according to the visual reading and the smoothed curve fell between $P_{2.5}$ = -1.03 ml/s and $P_{97.5}$ = +6.91 ml/s (figure 5.4). Reduction of the noise signal using a smoothing function led to an increased correlation of 0.951, and a increase in precision ($P_{2.5}$, $P_{97.5}$ = -1.27, +4.76 ml/s). Subsequent deletion of the opening pressure artefacts increased the correlation to 0.963 and precision to $P_{2.5}$, $P_{97.5}$ = -1.27, +3.31 ml/s.

figure 5.4. Differences (median, $P_{2.5}$, P_{10}, P_{90}, $P_{97.5}$) of the maximum urinary flow according to the original data and the corrected flow curves, versus the visual reading (reference standard).



Finally, *correction of artefacts led to a correlation of 0.99 and a precision of* $P_{2.5}$, $P_{97.5}$ = -1.59, +0.91 ml/s. The median difference between the corrected reading and the remastered Q_{max} was 0.14 ml/s.
The effects per subject are shown in figure 5.5. The difference between the reference standard's Q_{max} and the corrected Q_{max} is smaller than that between the reference standard's Q_{max} and the Q_{max} as given by the mictiograph in all subjects. The point estimate of the mean Q_{max} therefore is more accurate. In most, but not all, subjects the range of all Q_{max}'s also narrows.

figure 5.5. Differences (median, P_{10}, P_{90}) of the maximum urinary flow according to the original data and the corrected flow curves, versus the visual reading (reference standard), per subject.



For each subject, the left line refers to the original values of Q_{max} (median, P_{10}, P_{90}); the right line to the corrected Q_{max} (median, P_{10}, P_{90}).

additional analysis

33 Flows could be included in the secondary analysis. The correlation between the visual reading and the corrected flows increased from 0.861 to 0.961.

Discussion

main findings

A substantial number of flow curves need to be corrected to obtain a correct test result. A smoothing function decreasing high-frequency noise signal followed by an analysis of larger artefacts based on the first and second derivative and correction with a four-second moving average already leads to a sufficient result. The algorithm we developed closely approximated the visual readings of two independent readers. In fact, the algorithm-reference difference (median 0.14 ml/s, $P_{2.5}$, $P_{97.5}$ = -1.59, +0.91 ml/s) was almost the same as the difference of the two readers after individual reading and a consecutive consensus meeting (median 0,4 ml/s $P_{2.5}$, $P_{97.5}$ = -0.2 +1.4 ml/s).

methodological considerations

We had to scan the paper prints and 'remaster' the original data. As a result we could not include all flows. However, we did not find an indication that this has led to (selection) bias. It is very likely that a slightly different algorithm would perform better in other study samples. Of course, the results of this study should be validated in similar studies, including populations with different characteristics, especially patients using medication and more symptomatic patients.

comparison with reference standard

The reference standard used was based on the mean of the visual readings by two independent researchers, using the same previously defined set of criteria. The results of these two readings were generally congruous. Nevertheless there were some differences. The median difference between both readers was 0.4 ml/s ($P_{2.5}$, $P_{97.5}$ = -0.2, +1.4 ml/s). The difference between the results of the correction algorithm and the reference standard used is nearly the difference between the results of two independent readers.

In most subjects the range between the lowest and highest Q_{max} was reduced by correction. However, in some subjects this increased. This may be due to 'overcorrection' (which tends to a lower Q_{max}). But it may also be that those subjects created peaks in a very consistent way, e.g. by increasing abdominal pressure. Removal of those peaks may then reveal the true value of Q_{max}. The home-based collection of the data makes it difficult to explore this thoroughly.

relevance for clinical practice

In daily clinical practice, the gain is most distinct in the logistics. Current guidelines require uroflow measurements to be performed 'at least twice' in order to get a valid test result. Intra-individual variability is one of the reasons for this. Artefacts are another. However, multiple flow measurements are time-

consuming and make logistics inefficient. If most artefacts can be corrected automatically, immediate visual checks would not be necessary. Flow measurement could be delegated to lower classed staff more easily. Furthermore, test results will be more accurate for longitudinal follow-up.

relevance for clinical research

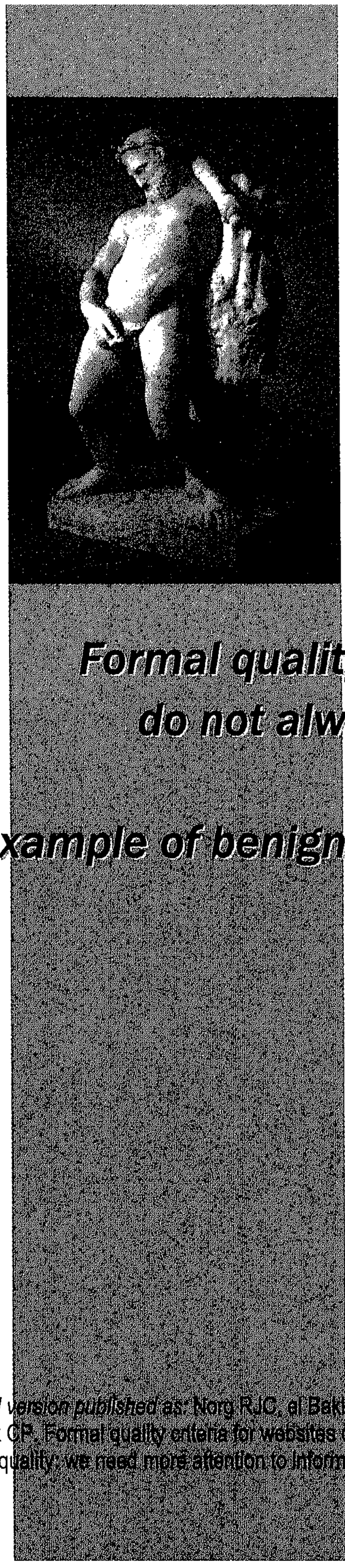
Artefacts may cause an increase in variability in measurements both within subjects and between subjects. If the number or amplitude of artefacts correlate with other determinants, and are directional - selectively increasing or decreasing the value of the test result - they also cause bias. Variability and bias reduce the ability to find true relations between two variables, or true differences in outcome in separate groups under study in clinical trials. Variability reduces statistical power. According to the CONSORT statement, any research protocol of clinical trials should state any methods to enhance the quality of measurements.[8] It is also a matter of ethics to optimise precision in order to reduce the waste of study material and prevent unnecessary demands on participating volunteers. Our algorithm shows that this can be done simply and objectively. Correction of artefacts reduces the standard deviation and increases the number of measurements to be included in the analysis. Any reduction of the standard deviation has a squared effect on the number of measurements (or participants) needed: if the standard deviation of the corrected data is half the standard deviation of the uncorrected data, the number of measurements needed is only a quarter compared to a situation in which uncorrected data were used. A standard automated way of correcting the data is preferable. It reduces subjectivity; blinding of the 'corrector' is not needed. It has therefore increased validity, is less time-consuming and cheap.

future research and development

More refined techniques are available for the recognition of graphical data and curves than the techniques used in the present study. Some additional gain can therefore probably be won. Our effort shows that it is worthwhile to include such algorithms in all uroflow meters. We would prefer a situation in which both the original data and the corrected flow curve are printed or displayed together with original and corrected characteristics (of which the Q_{max} is the most relevant, of course), so that clinicians and researchers have access to all the relevant information. Unless there are obvious mistakes made by the algorithm, the test result after standardized correction should then be used.

References

1. De la Rosette JJ, Madersbacher S, Alizivatos G, Rioja Sanz C, Emberton M, Nordling J. *Guideline on benign prostatic hyperplasia*: European Association of Urology, Dec 2004.
2. American Urological Association. Guideline on the management of Benign Prostatic Hyperplasia (BPH), 2003 (Accessible via www.auanet.org).
3. Jorgensen JB, Mortensen T, Hummelmose T, Sjorslev J. Mechanical versus visual evaluation of urinary flow curves and patterns. *Urol Int* 1993;51(1):15-8.
4. Grino PB, Bruskewitz R, Blaivas JG, Siroky MB, Andersen JT, Cook T, et al. Maximum urinary flow rate by uroflowmetry: automatic or visual interpretation. *J Urol* 1993;149(2):339-41.
5. Witjes WP, de la Rosette JJ, Zerbib M, Vignoli GC, Geffriaud C, Debruyne FM, et al. Computerized artifact detection and correction of uroflow curves: towards a more consistent quantitative assessment of maximum flow. *Eur Urol* 1998;33(1):54-63.
6. Abrams P, Griffiths D, Höffner K, Liao L, Schäfer W, Tubaro A, et al. *The urodynamic assessment of lower urinary tract symptoms*. 5th ed, 2000.
7. Norg RJ, Portegijs PJM, Van de Beek C, Onghena P, Van Schayck CP, Knottnerus JA. Please be seated? Position-related differences in micturition quality in elderly men with moderate lower urinary tract symptoms (LUTS): a replicated randomised single case study and meta-analysis. *Submitted.*
8. http://www.consort-statement.org/ (Accessed Nov 16 2007).



CHAPTER 6

Formal quality criteria for websites do not always reflect the quality of their content - the example of benign prostatic hyperplasia

modified version published as: Norg RJC, el Bakkali N, Portegijs PJM, Knottnerus JA, van Schayck CP. Formal quality criteria for websites do not guarantee complete health information of good quality: we need more attention to information on prognosis. *EJGP*, 2007;13(3):164-6.

Abstract

background:

Concerns about the quality of health information on the Internet have lead to the introduction of quality criteria for websites. Most of these deal with the presentation and with formal control of the trustworthiness of the information (formal quality criteria). The quality of the actual content is not measured (content quality).

objective:

This study assessed whether these formal quality criteria reflect content quality.

methods and materials:

In June 2003 we conducted 54 separate searches using 27 search terms regarding benign prostatic hyperplasia with two search engines. We followed the first 40 weblinks of each search. The resulting sample of 251 websites was scored for completeness and accuracy using the NHS prodigy guidance as the reference standard. We scored websites on information concerning aetiology, diagnosis, therapy, and prognosis. The scores were compared between websites that did and did not meet formal criteria, websites found using lay versus professional search terms, and HON-accredited and non-accredited sites (HON = Health on the Net foundation).

results:

Most websites complied with at least 7 out of 10 quality criteria and offered fairly complete and almost always accurate information. Accredited sites scored better than non-accredited sites. In general, there was a positive correlation between meeting the formal criteria and the content quality. However, there were some noticeable exceptions. Furthermore, information regarding the prognosis was relatively sparse.

conclusion:

Presently used formal criteria are indeed useful. However, consumers must be cautious for selective information.

Introduction

The Internet is considered both a medium of unknown freedom and a potential threat to people's health.[1-4] It offers both professionals and lay consumers free access to an increasing number of information brochures, guidelines, research articles, and informal patient-to-patient contacts. The estimated more than 100.000 health websites are used by up to 30-40% of Europeans and Americans.[5,6] Most people look for a description of diseases and use this as a tool to better communicate with their doctors.[7,8]

However, concerns about incorrect or selective information have also led to a call for regulation. Several parties have taken initiatives to measure the quality of health websites and to promote high quality health websites.[9] These initiatives have resulted in website judgements by experts (i.e. professionals)[10], self-evaluation scoring lists for patients (i.e. lay consumers)[11] and self-regulation systems by website providers.[12] Recently, the European Union has established common quality criteria for health related websites.[13] These criteria "focus on the reliability of a website as a potential vehicle for health related messages, rather than on the substance and content of the messages themselves".[13] The assumption is that there is a strong positive correlation between thus defined reliability (further referred to as 'formal quality') and the quality of the content. It is unclear how strong the correlation really is, especially with regard to common ailments in general practice.

We investigated this using websites regarding benign prostatic hyperplasia (BPH). We chose this subject because it is common in the open population with a relatively low doctor's consultation rate. The complaints start in an age group (45-65 years) where many are familiar with the Internet. Most men are reluctant to visit their doctor for this "ailment" and might want to look for the information in private.[14]

Methods and materials

measurement instruments

Formal quality: The European Union has formulated Quality Criteria for Health Related Websites.[13] These are divided into six groups of quality criteria: transparency and honesty, authority, privacy and data protection, updating of information, accountability and accessibility. The criteria are described in general terms and need to be made operational in order to be able to use them to score websites. We selected a 9-item checklist of formal criteria with which we scored all websites. The resulting nine separate scores for each site were also combined to an overall score (formal score) ranging from 0 to 10 (see table 6.1, page 90).

Content quality was assessed on four major topics separately: (a) the aetiology and pathophysiology of BPH, (b) the diagnostic procedures, (c) the therapeutic options, and (d) the prognosis of BPH after therapy including natural course of the disease, no therapy, or watchful waiting. Each of the items (a) to (d) were scored on 5-point Likert scales for completeness (ranging from 0 (no information) to 4 (complete information)) and accuracy (ranging from -4 (4 or more errors) to 0 (no errors)). All scores were combined to an overall score (content score) ranging from 0 to 10 (see table 6.1). The checklist was pretested in a separate study, and showed sufficient inter-rater reliability. The NHS Guidance on benign prostatic hyperplasia [15] was used as the reference standard.

table 6.1. checklist for scoring health websites

formal items			min. score	max. score
1	registration necessary to acquire information?		0	4
2	source of information clear?		0	4
3	commercial or otherwise 'interested party' (independent source)?		0	4
4	website up to date?		0	4
5	clear for whom website is made (target group)?		0	4
6	clear aim of website?		0	4
7	good readability?		0	4
8	clear explanation and use of jargon?		0	4
9	clear presentation on website?		0	4
***content items** *			min. score	max. score
a	aetiology	1. completeness	0	4
		2. accuracy	-4	0
b	diagnosis	1. completeness	0	4
		2. accuracy	-4	0
c	therapy	1. completeness	0	4
		2. accuracy	-4	0
d	prognosis	1. completeness	0	4
		2. accuracy	-4	0

formal score: the sum of score 1 to 9 divided by 3.6 (resulting in a scale from 0 to 10)
content score: the sum of score a1 to d2 divided by 1.6 (resulting in a scale from 0 to 10). Negative content scores were set to zero (This emphasizes the need of accurate information over completeness)
* content domain defined by NHS Guidance on benign prostatic hyperplasia[15]

website selection

In order to compose a broad sample of websites we used 27 search terms regarding BPH with the two most popular search engines in the UK (www.google.com and www.msn.com).[16] We asked two male volunteers (English native speakers, ca. 50 years old, no medical background) to suggest search

terms. Apart from their 10 search terms (further referred to as lay search terms) we used 17 other search terms (referred to as professional search terms). Most of these are symptoms of BPH. This procedure generated over a million links to webpages (weblinks).

table 6.2. Number of weblinks and relevant websites, and content scores per search term

	search term	number of weblinks*	relevant (certified) sites*		content score median (P_{25},P_{75})	
'lay' search terms						
1	bladder problems	561878	4	(1)	7.5	(7.5, 9.4)
2	peeing problems	39589	2	(1)	3.4	(-,-)
3	difficulty peeing	8730	2	(0)	6.9	(-,-)
4	difficulty urinating	51054	8	(5)	10.0	(8.0, 10.0)
5	problems urinating	79469	17	(4)	7.5	(4.4, 10.0)
6	prostate enlargement	150118	29	(6)	5.0	(2.8, 9.7)
7	prostate problems	570073	27	(3)	5.0	(2.5, 7.5)
8	urinary problems	509654	12	(7)	8.4	(7.5, 10.0)
9	urination problems	161913	16	(2)	3.8	(2.0, 5.6)
10	urine dribbling	16760	17	(2)	6.9	(2.8, 7.5)
'professional' search terms						
11	acute urinary retention	50023	12	(3)	7.5	(7.5, 10.0)
12	benign prostatic hyperplasia	85475	31	(8)	8.8	(6.9, 10.0)
13	benign prostatic hypertrophy	35281	34	(4)	6.9	(5.6, 10.0)
14	benign prostatic obstruction	16005	30	(7)	7.5	(5.0, 10.0)
15	bladder outflow obstruction	5677	8	(1)	8.8	(5.6, 10.0)
16	bph	194451	24	(3)	8.1	(5.0, 10.0)
17	dysuria	42012	7	(3)	10.0	(10.0,10.0)
18	hesitant voiding	774	11	(4)	8.8	(6.9, 10.0)
19	incontinence	558482	14	(4)	7.5	(5.0, 10.0)
20	micturition problems	7021	1	(1)	10.0	(-,-)
21	nocturia	19037	16	(6)	10.0	(6.9, 10.0)
22	overflow incontinence	10735	26	(6)	7.8	(5.6, 10.0)
23	post-voiding dribbling	76	10	(1)	5.0	(2.2, 7.7)
24	prolonged voiding	7012	12	(3)	8.8	(7.0, 10.0)
25	urgency	1309221	5	(1)	10.0	(7.5,10.0)
26	urinary frequency	208600	14	(1)	10.0	(6.9, 10.0)
27	weak urinary stream	13067	20	(5)	6.3	(3.9, 10.0)

*** combined number found with google.com and msn.com**
(-,-) no meaningful quartiles due to small numbers

Per search term we included only the first 40 weblinks for the analysis leading to an initial sample of 2,160 weblinks. Subsequently, we screened for duplicate links and excluded dead links, and links to websites about prostate cancer only, or links about micturition problems in animals. For safety reasons we excluded

websites that required other programs to be installed on the computer. Eventually, 251 websites were assessed. All searches and assessments were done by the second author in June 2003.
We registered whether sites were accessed following weblinks that belonged to the first 20 weblinks or to the second 20 weblinks (or perhaps both, as could be the case with duplicate links). We also registered whether websites were accredited by the Health on the Net Foundation (HON). This is the oldest and probably best-known self-regulation system.[12] Its set of quality criteria is currently used by more than 3,000 Internet sites worldwide.

statistical analysis

For each formal criterion we calculated the median scores of the sites that did and did not meet the criterion. We also compared the content scores of websites accessed following the first 20 versus the second 20 weblinks, websites resulting from lay versus professional search terms, and HON-accredited versus non-accredited sites. In all comparisons non-parametric tests were used (Kruskall-Wallis and Mann-Whitney U where applicable).
In order to assess whether the four major content topics (aetiology, diagnosis, therapy, prognosis) were equally present we compared their mean scores for completeness. We also compared these for HON-accredited and non-accredited sites separately (Wilcoxon sign rank test).

Results

Except for one, all used search terms led to at least one accredited website. (table 6.2, page 91) Accredited sites scored better than non-accredited sites. Eighteen accredited sites scored the maximum 10 points, and thirteen scored 7.5. Accreditation provided no guarantee, however: five websites (12%) scored between 2.5 and 5.0, despite high formal scores.
Information of the natural course of the disease and the prognosis after treatment was relatively more absent or sparse. The mean scores for completeness were (on a range from 0 to 4) 3.1, 2.8, and 3.1 for aetiologic, diagnostic, and therapeutic information versus 1.4 for prognostic information ($p<0.001$). For HON-accredited sites all scores were about 0.6 points higher, but the differences were almost the same: accredited sites scored 3.7, 3.5, and 3.7 vs. 2.1 ($p<0.001$); non-accredited sites scored 3.0, 2.7, and 3.0 vs. 1.3 respectively ($p<0.001$).
Websites accessed following weblinks that were among the first 20 weblinks only had marginally better content scores than websites accessed following links listed by the search engine at rank 21 to 40 only. Websites accessed following links that were only found using professional search terms had better content scores than websites found following weblinks provided by lay search terms.

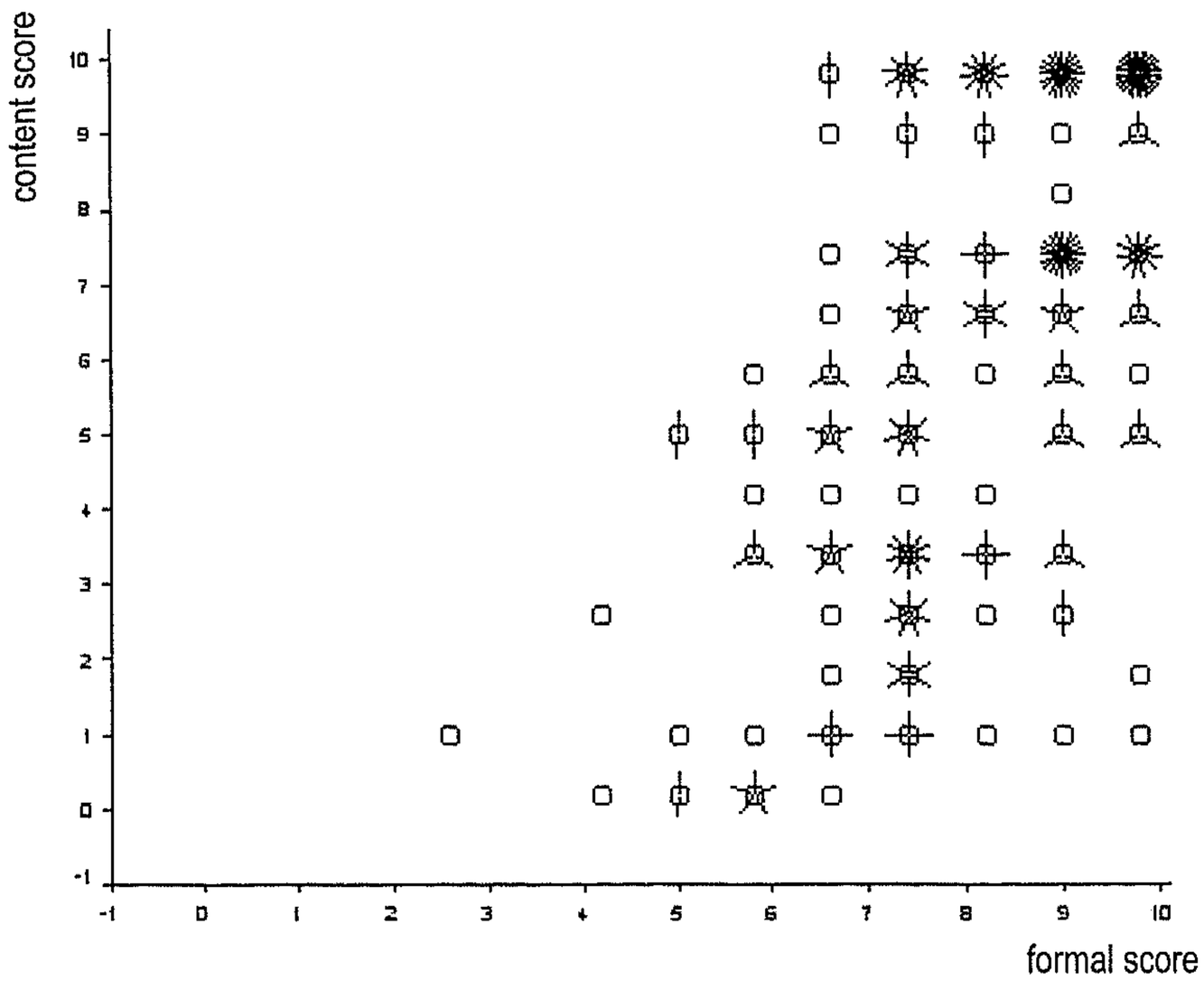
table 6.3. Formal quality criteria and content quality

item		*yes**			*no**			
		number	content score median (P_{25},P_{75})		number	content score median (P_{25},P_{75})		
1	no registration necessary	247	6.9	(3.8, 10.0)	4	10.0	(5.3,10.0)	NS
2	clear source	193	7.5	(5.0, 10.0)	58	4.1	(2.5, 5.6)	
3	independent source #	152	7.5	(5.6, 10.0)	41	6.3	(3.8, 9.1	
4	up to date	141	7.5	(6.3, 10.0)	110	5.0	(2.5, 7.5)	
5	clear target group	215	7.5	(5.0, 10.0)	36	3.8	(0.6, 5.6)	
6	clear aim	247	7.5	(3.8, 10.0)	4	3.8	(1.6, 5.5)	NS
7	good readability	178	7.5	(4.4, 10.0)	73	6.3	(2.8, 10.0)	NS
8	clear explanation of jargon	149	7.5	(5.0, 10.0)	102	5.0	(2.5,7.5)	
9	clear presentation	239	7.5	(3.8, 10.0)	12	2.8	(0.8, 5.5)	
I	first 20 only $	92	7.5	(3.8, 10.0)	78	5.6	(3.0, 9.4)	
II	lay terms only %	70	4.4	(2.5, 7.5)	148	7.5	(5.0, 10.0)	
III	certified by HON	42	7.5	(7.5, 10.0)	209	6.3	(3.8,10.0)	

* score 3 or 4 (yes) versus score 0 or 1 or 2 (no)
# based on 193 websites (58 sites had no source mentioned)
$ first 20 weblinks only versus second 20 weblinks only
% lay search terms only versus professional search terms only
NS: $p \geq 0.05$; all other differences: $p<0.05$

The median formal score was 8.3 (P25 7.0, P75 9.3). This indicates that 3 out of 4 websites met at least 7 of the 10 quality criteria. The median content score was 6.9 (P25 3.8, P75 10.0). Figure 6.1 shows the distribution of both scores.
In general, meeting a formal criterion was associated with higher content scores in contrast to not meeting the criterion, except for the criterion 'no registration necessary'. Only 4 websites (2%) required registration, however. Aim and presentation of the website were unclear in a small minority of cases as well, explaining the lack of (significant) difference in content scores. In most cases the readability score was high. High readability scores tended to be associated with higher content scores. In 41% of the cases there was no clear explanation of jargon. 56% of the websites had recently been updated. In 44% of the cases this was not the case, or not mentioned. On almost a quarter of the websites we found no indication of the source of the information provided. In these cases the quality of the information was significantly lower. The same holds for websites with a source of information we qualified as not independent.

figure 6.1. Correlation of formal and content quality scores



The number of observations is reflected by the number of leafs of the 'sunflowers'.

Discussion

The findings of our study show the expected positive relation between established quality criteria and the quality of the actual content of websites. But we also found some noticeable exceptions. Attention is therefore necessary. The correlation is fair, but not strong enough to guarantee complete and correct information. Even of the accredited sites 12% were rated with half the maximum score or less. Fortunately, this was a result of incomplete rather than incorrect information. However, selective incomplete information can also be misleading. For instance, the relative sparse information about prognosis and natural course of the disease could lead to an understanding that surgical or pharmaceutical intervention would be necessary in almost all cases. For many men however, a watchful waiting strategy for BPH is often a first choice procedure.

There are some limitations to our study. Our sample was large in comparison to most other studies concerning the quality of information on the Internet. Consequently, for budgetary and practical reasons only one investigator scored all websites. The websites were scored on-screen. We did not print the text so that the scorer would be blinded with regard to accreditation status. Therefore we can not exclude bias in favour of accredited sites caused by this.

This study used websites concerning BPH. It is possible that the conclusions may differ with other medical subjects. Furthermore, although it is almost impossible to draw up a complete set of possible search terms our selection of search terms could have 'biased' our findings. Also, although most people use search engines to look for medical information, we can not extend the findings to other Internet sources of information for patients like portals or guideways, or to online patient networks.[17]

Most consumers only look at the first 10 weblinks. If they do not find the information they look for there they tend to use other search strategies or search terms. In our study, the mean content scores as well as the number of relevant sites show much variation for different search terms. Websites found with professional search terms were rated with higher content scores than those found using lay search terms. We therefore conclude that switching search terms is a sensible thing to do. Especially when the second search term is a professional term acquired when visiting a former website. We also conclude that this should be an incentive for providers of health information to use more lay terms in their texts. After all, this will enhance maximum retrieval and impact of their sites.

The Internet unites contradictory functions. One is to provide free access to a large database of useful information. This function is the one consumers use when looking for trustworthy and accurate information on medical subjects. Accreditation supports this function. The other is to enable people to present freely their own opinions, insights and beliefs to the world. As a result the Internet offers a large variety of verifiable and non-verifiable 'facts'. This function of the Internet is the one people use when looking for alternatives for the generally accepted 'standard' medical information.

This implies that retrieval of 100% trustworthy information is a utopian dream. Supporters of the second function might even call it a utopian nightmare: a monopoly of information controlled by 'authorities'. A fierce debate arose when the World Health Organization (WHO) presented a proposal for a new top level domain *.health*.[18] Some also are afraid of a conservative effect, where dissemination of new techniques and insights would be delayed.[19] However, we think that the public are very aware of the choice they have between regular and complementary medicine. Accordingly, in most cases they know whether they look for official information or for alternatives. On the Internet, the difference

between both types of information (official and alternatives) sometimes is hard to see, even for professionals. In order to be able to make their personal informed decisions it is important that consumers can recognize 'trustworthy' information (i.e. information generally accepted by health professionals) from other (complementary) information. Any sign of trustworthiness (like accreditations) should be directly identifiable, based on widely recognized authority, and simple. A new top-level domain *.health* would meet all these requirements.

what does this study say about information regarding BPH on the Internet?

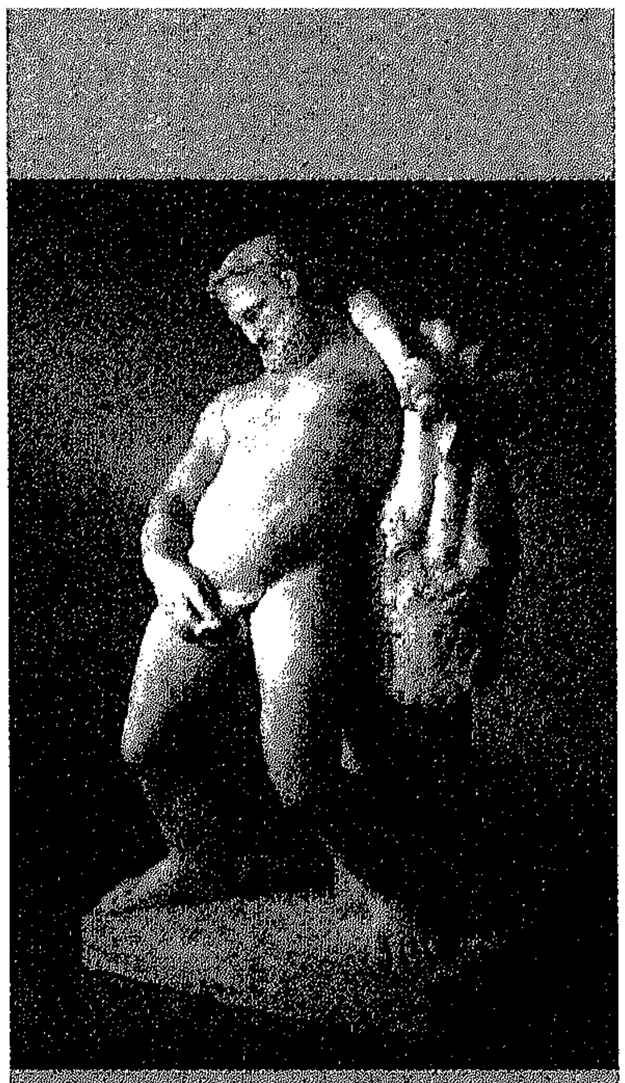
In general the websites that were found with the first 40 weblinks show satisfactory scores. Other studies on other medical subjects show ratings of accuracy and completeness ranging from 15% to 85%.[17] These results are difficult to compare because no specification was made between aetiologic, diagnostic, therapeutic and prognostic information. In this study we found relatively little attention for prognostic information. For male micturition problems this information is however crucial. The natural course of BPH is often favourable. Symptoms improve in many patients by the combination of time and the assurance that no cancer is present.[20] In contrast, surgery is powerful in relieving patients from severe symptoms, but may have serious adverse effects. It is therefore important that men with lower urinary tract symptoms have balanced and complete information on the prognosis. Otherwise accurate but selective information can indeed leave consumers wrong-footed.

conclusion

Our data show that quality accreditations like the HON code can indeed be of help. Although the content itself is not directly part of the accreditation criteria, in comparison to non-accredited sites the information on accredited sites is more in accordance with current guidelines. However, in our sample of websites regarding benign prostatic hyperplasia prognostic information was relatively sparse. Consumers therefore may be left with wrong impressions due to selective information.

References

1. Anderson JG, Rainey MR, Eysenbach G. The impact of CyberHealthcare on the physician-patient relationship. J Med Syst 2003;27(1):67-84.
2. Eysenbach G, Diepgen TL. Towards quality management of medical information on the Internet: evaluation, labelling, and filtering of information. BMJ 1998;317(7171):1496-500.
3. Heathfield H, Pitty D, Hanka R. Evaluating information technology in health care: barriers and challenges. BMJ 1998;316(7149):1959-61.
4. Jadad AR, Gagliardi A. Rating health information on the Internet: navigating to knowledge or to Babel? JAMA 1998;279(8):611-4.
5. Eysenbach G, Powell J, Kuss O, Sa ER. Empirical studies assessing the quality of health information for consumers on the world wide web: a systematic review. JAMA 2002;287(20):2691-700.
6. Eaton L. A third of Europeans and almost half of Americans use Internet for health information. BMJ 2002;325(7371):989b-.
7. Provost M, Perri M, Baujard V, Boyer C. Opinions and e-Health Behaviours of Patients and Health Professionals in the U.S.A. and Europe (online tables). 2003. Available from: http://www.hon.ch/Survey/Spring2002/res.html. (Accessed Jan 19 2004).
8. Boyer C, Provost M, Baujard V. Highlights of the 8th HON Survey of Health and Medical Internet Users. 2002. Available from: http://www.hon.ch/Survey/8th_HON_results.html (Accessed Jan 19 2004).
9. Gagliardi A, Jadad AR. Examination of instruments used to rate quality of health information on the Internet: chronicle of a voyage with an unclear destination. BMJ 2002;324(7337):569-73.
10. Available from: http://www.medcircle.org/ (Accessed Jan 19 2004).
11. Available from: http://www.discern.org.uk/ (Accessed Jan 19 2004).
12. Available from: http://www.hon.ch/ (Accessed Jan 19 2004).
13. Commission of the European Communities. eEurope 2002: Quality Criteria for Health Related Websites. J Med Internet Res 2002;4(3):E15.
14. Wolfs GG, Knottnerus JA, Janknegt RA. Prevalence and detection of micturition problems among 2,734 elderly men. J Urol 1994;152(5 Pt 1):1467-70.
15. Prodigy Guidance of Benign Prostatic Hyperplasia. 2001 2003 January [cited 2003 May 19]; Available from: http://www.prodigy.nhs.uk/guidance.asp?gt=Prostate - benign_hyperplasia
16. searchenginewatch. Available from: http://www.searchenginewatch.com (Accessed May 19 2003)
17. Ferguson T. From patients to end users. BMJ 2002;324(7337):555-6.
18. WHO. WHO proposal would raise quality of Internet health information - dot health could soon be as well known as dot com. 2000 November 13 2000 [Available from: http://www.who.int/inf-pr-2000/en/pr2000-72.html (Accessed 2004 June 15).
19. Proposal for ".health" Internet domain - Report by the Secretariat: Executive Board of the World Health Organization; 2003.
20. Brown CT, Van der Meulen J, Mundy AR, Emberton M. Lifestyle and behavioural interventions for men on watchful waiting with uncomplicated lower urinary tract symptoms: a national multidisciplinary survey. BJU Int 2003;92(1):53-7.



CHAPTER 7

General practitioners, their patients and the Internet

- a qualitative study of opinions and attitudes, and their interrelation with the doctor-patient relationship

Norg RJC, Bleuzé A, Portegijs PJM, Knottnerus JA, van Schayck CP. General practitioners, their patients and the Internet - a qualitative study of opinions and attitudes, and their interrelation with the doctor-patient relationship. *Submitted*

Abstract

background:

The potentially large influence of the Internet on the doctor-patient relationship has received little attention in research.

objective:

To assess the opinions and experiences of both patients and general practitioners with regard to the use of the Internet for health information purposes, and the interrelation with the doctor-patient relationship.

methods and materials:

design: qualitative study. Setting: General practice. Study populations and data-collection: (1) 24 semi-structured interviews with patients who had used the Internet for health information purposes during the previous year. (2) Semi-structured group discussion by 9 GPs with different personal characteristics and Internet experience.

results:

(1) Internet information is not discussed in case of minor complaints, if no news is expected, or if the patient has searched after the consultation. Considerations to discuss the information are doubts about the information, the desire to play a more active role during the consultation, and to take responsibility for one's own health. Patients often mention it informally. (2) The GPs discussed their emotional reactions, consequences for daily routines, and the consultation techniques. Most GPs have a positive attitude towards patients using the Internet. Some wish to stimulate this; others are concerned, however, about being overruled, and consequently becoming an instrument instead of an independent professional.

conclusion:

Most patients are positive about the Internet; GPs are less unanimous. However, they do not seem to be aware of each other's opinions. As a consequence, both patients and GPs seem to miss opportunities by not openly discussing Internet information.

Introduction

The use of the Internet by patients increases. Some expect this will lead to well-informed patients with consequently healthier lifestyles and better health status.[1] Others fear patients may become frightened by misleading, selective, or incorrect information.[2-4] This has resulted in a call for clear, easily accessible high-quality health information, recognisable by quality certificates or a specific domain.[4-7] Whatever one's opinion may be, Internet health information is an inevitable reality with a possibly profound impact on daily practice.

Patients using the Internet expect to become more confident to talk to their doctor, to better understand their condition, and to be encouraged to follow their doctor's advice.[8] Over 80% of these patients think they can easily obtain health information of good quality. Over 80% of the doctors rate the ability of their patients to appraise the quality of Internet health information as only fair or poor.[9] Diverting opinions about - and perceptions of - the value, trustworthiness, and quality of Internet health information between doctors and patients may burden the doctor-patient relationship.[10,11] We investigated and described these opinions and perceptions in a qualitative study among both patients and their general practitioners.

Methods and materials

study populations and data collection

We aimed at combining the views of both patient and general practitioners. We started off by generating concepts from interviews with patients, and subsequently used these as the basis of a discussion by GPs.

patients: postal questionnaire and telephone interview

The purpose of the study was to include a broad variety of people with regard to possible determinants of Internet use and of the quality of the doctor-patient relationship. In June 2004 we asked general practitioners (GPs) from 11 practices to participate in the study. These included both active Internet users and non-users, men and women, full-timers and part-timers, with different patient profiles (socio-economic status, urban and rural, age, native and non-native Dutch). Eventually, six general practices (13 GPs) participated. A random sample of 60 patients per practice aged between 18 and 70 received a postal questionnaire about Internet use for health information purposes; the perceived attitude of their general practitioner towards Internet use; and possible contacts with their doctor. Non-responders received a reminder six weeks later.

Subjects who had searched the Internet for information concerning their own health during the previous 12 months were asked informed consent for a telephone survey. This semi-structured interview addressed two general themes

not presented here (search behaviour; and the actions taken based on the information found), and three themes specifically aimed at the purpose of this study: (1) whether they mentioned their Internet activities to their GP and why they did (not); (2) the perceived attitude of the general practitioner towards Internet use by his/her patients; (3) the relationship between the subject and his/her GP. Participants were encouraged to elaborate on every topic in depth. The interviews were performed by AB, tape-recorded, typed verbatim and coded independently by RN and AB. The resulting coding systems were very similar and were integrated into one overall analysis in a consensus meeting.

general practitioners: semi-structured group discussion

Based on the analysis of the patient data we constructed a questionnaire for general practitioners. For each of the 11 practices we had asked to cooperate, we asked one GP to fill in this questionnaire, and to participate in a group discussion in which the questionnaire was used as a starting point. Ten questionnaires were returned, and 8 GPs participated in the discussion. (4 from participating, and 4 from non-participating practices). The group discussion was led by RN. Transcriptions were used for the analysis. The aim during this group discussion was to elicit attitudes and opinions, and experiences of the GPs by stimulating them to react on each other.

Results

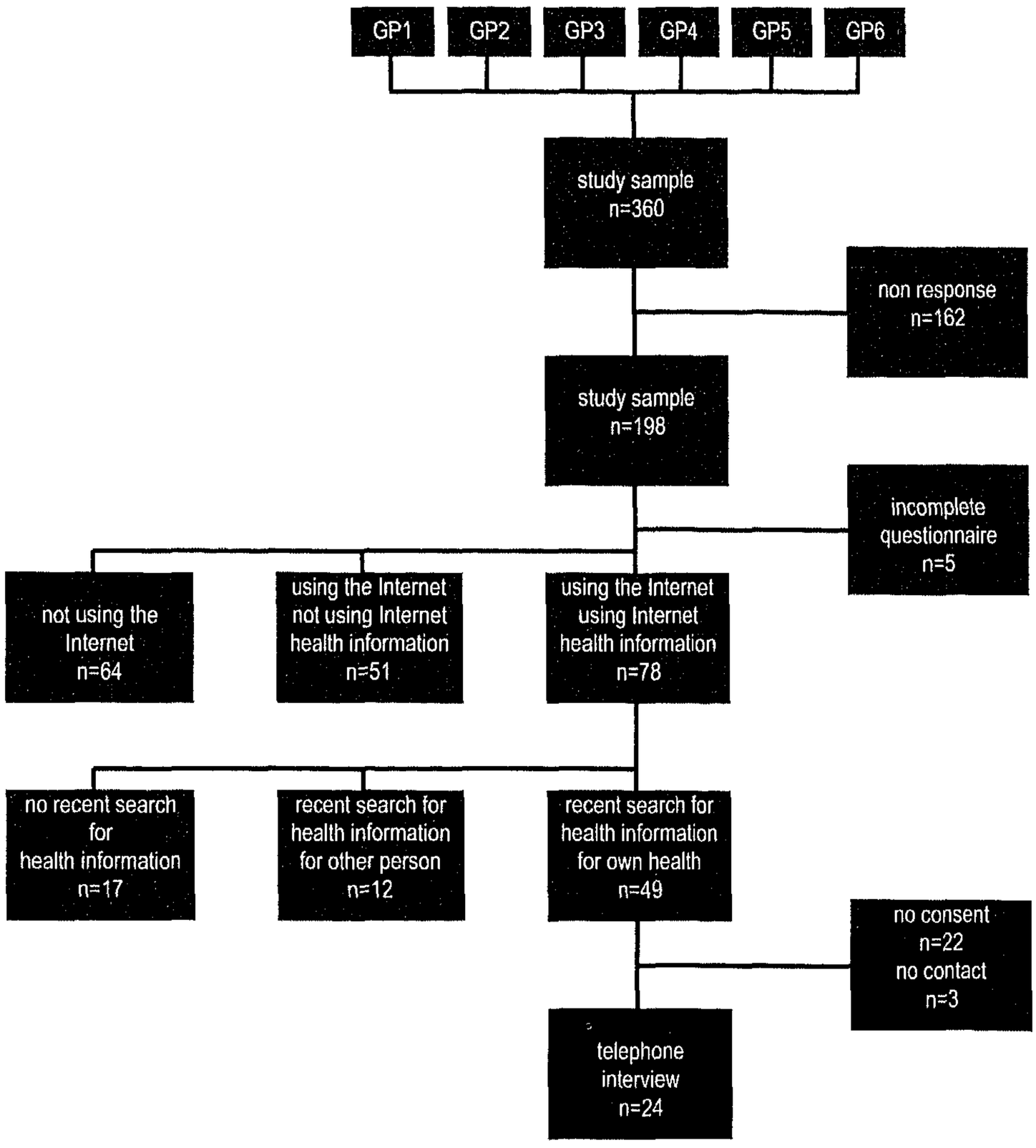
patient data

The recruitment procedure of the study population is summarized in figure 7.1 (page 103). Table 7.1 shows characteristics of the study population. All categories of age, sex, educational level and practice profile were present.

table 7.1. Characteristics of the study population (telephone interview), N = 24.

characteristics	categories	number of subjects
general practice	1	2
	2	2
	3	5
	4	7
	5	4
	6	4
age	18-30	8
	31-50	12
	51-70	4
sex	male	14
	female	10
educational level	none/primary school	1
	low level secondary school	6
	high level secondary school	9
	academic	8

figure 7.1. Recruitment of the study population



GP = general practitioner

table 7.2. Summary of the results of the qualitative analysis

patient data (telephone interviews)

theme 1. Reasons (not) to mention an Internet search to the general practitioner

1a. Why people mention their internet search to the general practitioner
- 1a.1. incongruous information on different internet sources
- 1a.2. more active role
- 1a.3. discussion on higher level / more in-depth information

1b. Why people do not mention their Internet search to the general practitioner
- 1b.1. no visit to GP
  - no news from their GP expected
  - type of complaint is not within the GPs domain
  - complaint is not serious enough
- 1b.2. search after visit to GP
  - confirmation about what has been told
  - more information than has been given
- 1b.3. (supposed) attitude of GP
- 1b.4. disturbed relationship with GP
- 1b.5. forgotten to mention it / no specific reason

1c. How people mention their internet search to the general practitioner
- 1c.1. explicitly
- 1c.2. informally

theme 2.The perceived attitude of GP towards the Internet

- 2.1. positive
- 2.2. negative
- 2.3. unclear

theme 3.The interrelation with the doctor-patient relationship

- 3.1. content
  - more in-dept information
- 3.2. procedure
  - more equality between doctor and patient
  - more efficient

GP data (group discussion)

theme 3.The interrelation with the doctor-patient relationship

3a. emotional reactions
- 3a.1. challenged
- 3a.2. affirmed

3b. consequences for daily work / routine
- 3b.1. instrumental use of GP by patients
- 3b.2. Internet functions as the reference standard
- 3b.3. the consultation reaches a 'higher level'.
- 3b.4. the Internet helps GPs to educate patients

3c. use of communication techniques

questionnaire

78 respondents (39%) had searched for health information, of whom 49 (25%) had done so in relation to their own health problems in the previous year. 51 (26%) subjects had not used the Internet for health purposes. Half of these had had no health questions; a quarter did not trust the Internet. Lack of experience or availability of the Internet was hardly ever the reason.

70% of all respondents had no idea what the attitude of their GP was towards the Internet. Only 4 subjects thought it to be negative. Respondents who had discussed Internet information with their GP reported markedly more often a positive attitude of their GP.

interviews

The results of the interviews that are presented here address the reasons (not) to mention an Internet search to the GP (theme 1), the perceived attitude of the GP (theme 2), and the doctor-patient relationship (theme 3). The basic structure of the results is shown in table 7.2 (see page 104).

theme 1. Reasons (not) to mention an Internet search to the general practitioner

1a. People mention their Internet search to the general practitioner in case of incongruous information from different sources. GPs may be asked to decide on the truth. Discussing health information may also be an expression of the responsibility for one's own health and health decisions, or a possibility to get a more active role in the conversation with their GP (quote [p266]).

> quote [p266]:
>
> 'A very open conversation, just like a normal conversation. But, there is something to think about, not only brought forward by him. Because that is what I don't like about such a situation. That the advice comes from the same man that has the knowledge, ánd has weighted all pros and cons for you as well. In the end it's all about your own health. That is also your own responsibility.'

1b. People do *not* mention their Internet search to the general practitioner, simply because they decide not to go to the general practitioner. Sometimes a problem is regarded yet a medical problem but nevertheless outside the domain of the GP, either because the general practitioner is regarded to address only physical, or only psychosocial complaints. Also, complementary medicine may be seen as outside the general practitioner's domain.

Patients also may qualify their problem as not serious enough, or expect no news from their GP since the information from several different websites is congruous.

Many subjects look for Internet information *after* their visit to the general practitioner. They reported the need for more elaborate information, or a confirmation of what had been told.
An (assumed) negative attitude of the GP or a disturbed doctor-patient relationship were also reasons not to ask the doctor's opinion or mention the health information. Finally, sometimes the contact did not require a reference to the information at hand, or the patient simply forgot to mention it when it would have been appropriate.

1c. Some patients had explicitly mentioned the Internet information. In case of contradictory information this may be the only way to get clarity. If someone wants to accomplish a change of practice by the doctor, an explicit reference to Internet may be considered useful, like in the case of this mother who tried to convince her doctors that she could breastfeed her diabetic child (quote [p235]).

> quote [p235]:
>
> 'Then I said: I've checked it out. It is somewhat harder, but I want to try. It 'll succeed. It went wrong in the hospital afterwards, but later, at the consultation, I said to the doctors "Boys, go looking at the Internet, at the research there." (...) It has been discussed in the breastfeeding team of that hospital, like "We should pay more attention to this." '

However, in most cases an informal way is chosen without mentioning the source of the information. This way of presenting the information is deliberately chosen if the patient expects a negative reaction or does not know what reaction to expect. Patients are afraid to challenge their doctor if they speak of it openly (quote [p014]).

> quote [p014]:
>
> 'Because I think my GP would not be open for this (...) I would be able to ask him more direct questions, because I am prepared. But I would not tell him I read this or that on the Internet."

theme 2. The perceived attitude of the general practitioner towards the use of the Internet by patients.
Most general practitioners do not seem to give a clue about their opinion towards the Internet. Quote [p492] shows a typical remark coming from several patients with different general practitioners.

> quote [p492]:
>
> 'I don't think he has a problem with it'. (...) 'He doesn't reveal that much'.

When invited to speculate most of those who said they had no idea about the opinion of the doctor supposed their GP probably would be negative, since patients might be more annoying, time-consuming, or frightened without reason. Only one subject had an outspoken negative experience (quote [p423]).

> quote [p423]:
>
> 'Negative, how did you notice that?'
>
> 'Because, when I start talking about it, he immediately reacts with "Aah, the Internet..." '

theme 3. The interrelation with the doctor-patient relationship.
Informed patients feel taken more seriously and given more space to contribute to the consultation. The result of this is that patients get more in-dept information from the GP. Along with this patients experience more equality between doctor and patient. Both results are shown in quote [p235].

> quote [p235]:
>
> 'Since you know more they take you more seriously. Normally the explanation stays somewhat superficial, but if they notice that you have gone into the subject, they show more.'

The consultation may become more efficient (quote [p106]).

> quote [p106]:
>
> 'Anyway, he didn't have to show me all the pictures and drawings.... I already knew. (...) It saves time.'

GP data

theme 1. Reasons (not) to mention an Internet search to the general practitioner
For most GPs the relatively large proportion of patients that look on the Internet after the consultation was an eye-opener. One GP recognised the situation in which a patient tries to get the GP's opinion in an informal way (quote [gp2]).

> quote [gp2]:
>
> 'If you are attentive you just remark "And what did the Internet tell you?"'

theme 2. The perceived attitude of the general practitioner towards the use of the Internet by patients.

Most GPs showed a somewhat to very positive attitude towards the use of the Internet for health information by their patients, but they hardly ever advised their patients to do so, except for one specific website about venereal diseases. Although they recognised that their hard-copy information leaflets often were outdated, and many up-to-date versions are published on the Internet, none of the GPs referred to websites.

theme 3. The interrelation with the doctor-patient relationship

The general practitioners' discussion focussed on three items:

- 3a. emotional reactions caused by patients bringing Internet information;
- 3b. expected or possible consequences for daily routine; and
- 3c. meta-communicative aspects and use of communication skills.

3a. The patients appeared to provoke both positive and negative emotional reactions (quote [gp1]/[gp2]).

> quote [gp1]/[gp2]:
>
> [gp1]: 'When someone comes saying "Doctor, I've got this disease, what are we going to do about it?" I feel walked over - emotionally'.
>
> [gp2]: 'I actually have the feeling "They come to check their knowledge against mine". That is an affirmation of me as a doctor.'

3b. The prospect of patients determining what the GP should or should not do, based on unknown Internet authorities, caused some discomfort among the GPs. They feared to loose the ability to use their own intellect, knowledge, and skills. In effect: an instrumental use of the general practitioner (quote [gp1]).

> quote [gp1]:
>
> 'Sometimes I see the future: a whole morning referring patients who have diagnosed themselves from the Internet'.

In line with this is the situation in which the Internet functions as the reference standard to which the GP (or any doctor) is measured (quote [gp8]).

> quote [gp8]:
>
> 'I had a patient with cancer who wanted to go to a specialised cancer hospital somewhere else. Until he came back to me with a print from the Internet. He said "Look, the specialist here [in the local hospital] does exactly the same as on the Internet. I'd rather stay here then."'.

The information the patient brings along is more detailed, more specific, and more related to the actual complaints or diseases of the patient. This differs from the classic ('old') situation in which a newspaper article or TV programme discusses general aspects of a certain disease that may, but also may not, apply to the specific situation of the patient. The Internet offers patients the possibility to select only the information that is applicable to their individual situation from a large database of knowledge and experiences.
The result may be that the consultation reaches a 'higher level'. It gets more interesting because the GP is able to use more of his specific expertise and knowledge, instead of routinely conducting daily work. This way, the Internet helps general practitioners to educate patients (quote [gp7]).

quote [gp7]:

[gp7]: 'Most people who come with Internet information are intelligent and interested people. You have to stimulate this. Aren't we busy to try to learn our patients to be the doctor themselves?'.

3c. The theory regarding consultation and communication techniques seems clear: a GP should be aware of the background-knowledge; reasons for encounter; the questions of the patients in each consultation; and whether or not patients have used the Internet. In practice, this often may not be so simple.

quote:

[gp1]:'When someone has made his own diagnosis based on the Internet, the consultation proceeds very differently. Then I think something like ...uh ... that I am busy talking him out of it. That is very different from a situation in which someone asks me what it is he suffers from. I start differently.'

[gp2]: 'But if you ask: "what is it you want from me?" and he says: "I want to know if I'm correct", you can say "Okay, I'll see to it", and start to do your normal work again.'

[gp1]: 'Yes, okay, you're quite right. But what if they do not want to know whether they're correct, but just assume they are. Then, you can not start your normal ways again.'

Discussion

summary of main findings

We investigated the interrelation between Internet use by patients and the doctor-patient relationship and identified several factors that mutually influence each other. These are complaint-related (e.g. minor versus 'really serious' complaints), Internet-related (e.g. congruous versus incongruous information), patient-related (e.g. expectations) and doctor-related (e.g. basic attitude, use of communication techniques). Assumed and expected possible negative effects go hand in hand with mainly positive real experiences. Many doctors and patients seem to be reluctant to discuss Internet information openly.

strengths and limitations of this study

There are many possible determinants of both Internet use and the quality of the doctor-patient relationship. It is difficult to compose a specific target population for purposeful sampling. We therefore chose for a random sample of patients from a targeted population of GPs. By including a sufficient number of patients per GP in the sample, selecting the thus identified relevant population, and subsequently interviewing every eligible subject we could contact, we were able to include a mix of subjects with different socio-economic status, age and sex. On this level we reached data saturation. However, at the level of GPs we cannot guarantee data saturation, despite their different individual characteristics.

We had to meticulously discern between the *motives* of patients and the *results* of their actions. For instance, one may want to play an active role in the consultation, and not succeed; and one may play an active role without having had the explicit wish to do so. General practitioners and their patients often have longstanding relationships, also relating to other health questions. With our data it is impossible to prove a one-way linear causal effect. The reality is far more complex. Therefore we use the term 'interrelation' instead of 'influence' to describe the relation.

what is already known

In recent years the use of the Internet for health information purposes has become an increasing field of research. Most research has focused on the number of patients using the Internet, and characteristics of Internet users, the kind of information people search for, the search-strategies used, and the appraisal of the trustworthiness of websites. Less research has been done into the effect of the Internet on the doctor-patient relationship. Quantitative studies have shown that doctors may feel challenged with a negative impact on the doctor-patient relationship.[9] Patients who qualify their doctors' level of care as fair or poor, or feel their doctors provide them with too little information are more likely to search the Internet for health information.[8,12] Health professionals

think that their patients are too optimistic about their search skills. They also are rather pessimistic about their own IT literacy. Both connotations seem to contribute to their negative feelings towards the Internet. Overall the effect of the Internet on the doctor-patient relationship is described as 'subtle' rather than 'dramatic'.[11]

what this study adds

Our results support these conclusions. However, our sample of GPs seems to be somewhat more positive than the health professionals studied by Hart et al. This may be due to the fact that our study was performed approximately 3 years later. Both GPs and patients may have adapted more to the new reality of the Internet. Nevertheless, it is still seldom discussed between doctors and patients, despite the fact that both doctor and patient may benefit if the use of the Internet as a source of additional information for patients would be an accepted topic of conversation. Patients may benefit since they can become a more active contributor to the consultation. Doctors may benefit since they can start on a higher level of knowledge of the patient. It may enrich the working day of the GP and stimulate his own knowledge. More understanding by the patient generally leads to more satisfaction and appreciation.[8] However, there must be a clear explicit or informal mutual understanding about what Internet sources are acceptable. Certification of websites and clarity with respect to the website nomenclature (for instance with a domain name *dot health*) may contribute to this.

For most patients it is hard to integrate all information and to translate it properly to one's own situation. This is pre-eminently the expertise of doctors. They may want to discuss Internet information and are willing to use their broader knowledge and experience as the reference standard against which this information can be judged. Our study shows that (patients of) doctors who are aware of this, and use their communication skills to overcome negative feelings, may benefit most of the possibilities of the Internet.

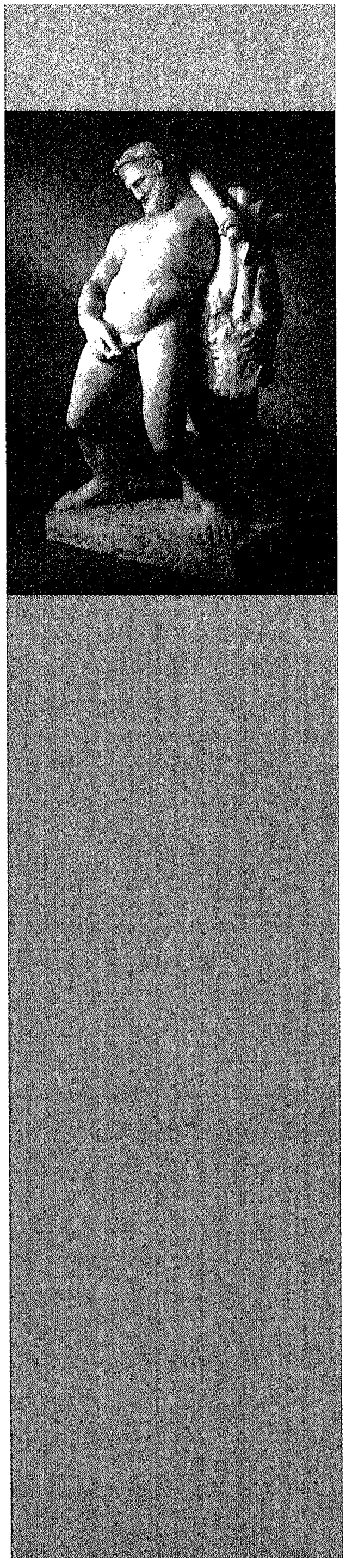
The response of doctors with regard to Internet information and other sources of health information like direct advertising or newspaper articles show similarities, but also marked differences. [13,14] The Internet is used when a patient wants to search information on a specific topic at a specific moment in time; advertisements and newspaper articles are read when published. In our study GPs rated the Internet information as more adequate and relevant to the individual situation of the patient than 'classic' forms of information, and appreciated this positively.

implications for clinical practice

Most patients are positive about the Internet; GPs are less unanimous. However, they do not seem to be aware of each other's opinions. Both doctors and patients seem to miss opportunities. If they would discuss Internet information more openly, they could make optimal use of the expertise of GPs: the ability to integrate all kinds of information and apply this to an individual patient in his context, based on a longstanding continuous relationship.

References

1. Anderson JG, Rainey MR, Eysenbach G. The impact of CyberHealthcare on the physician-patient relationship. J Med Syst 2003;27(1):67-84.
2. Eysenbach G, Diepgen TL. Towards quality management of medical information on the internet: evaluation, labelling, and filtering of information. Bmj 1998;317(7171):1496-500.
3. Heathfield H, Pitty D, Hanka R. Evaluating information technology in health care: barriers and challenges. Bmj 1998;316(7149):1959-61.
4. Jadad AR, Gagliardi A. Rating health information on the Internet: navigating to knowledge or to Babel? Jama 1998;279(8):611-4.
5. www.hon.ch (Accessed Nov 16 2007).
6. Commission of the European Communities. eEurope 2002: Quality Criteria for Health Related Websites. J Med Internet Res 2002;4(3):E15.
7. World Health Organization. WHO proposal would raise quality of internet health information - dot health could soon be as well known as dot com, 2000.
8. Murray E, Lo B, Pollack L, Donelan K, Catania J, White M, et al. The impact of health information on the internet on the physician-patient relationship: patient perceptions. Arch Intern Med 2003;163(14):1727-34.
9. Murray E, Lo B, Pollack L, Donelan K, Catania J, Lee K, et al. The impact of health information on the Internet on health care and the physician-patient relationship: national U.S. survey among 1.050 U.S. physicians. J Med Internet Res 2003;5(3):e17.
10. Quintana Y, Feightner JW, Wathen CN, Sangster LM, Marshall JN. Preventive health information on the Internet. Qualitative study of consumers' perspectives. Can Fam Physician 2001;47:1759-65.
11. Hart A, Henwood F, Wyatt S. The role of the Internet in patient-practitioner relationships: findings from a qualitative research study. J Med Internet Res 2004;6(3):e36.
12. Sciamanna CN, Clark MA, Diaz JA, Newton S. Filling the gaps in physician communication. The role of the Internet among primary care patients. Int J Med Inf 2003;72(1-3):1-8.
13. Murray E, Lo B, Pollack L, Donelan K, Lee K. Direct-to-consumer advertising: physicians' views of its effects on quality of care and the doctor-patient relationship. J Am Board Fam Pract 2003;16(6):513-24.
14. Murray E, Lo B, Pollack L, Donelan K, Lee K. Direct-to-consumer advertising: public perceptions of its effects on health behaviors, health care, and the doctor-patient relationship. J Am Board Fam Pract 2004;17(1):6-18.



CHAPTER 8

General discussion

General discussion

putting things in perspective

Research often requires changing one's perspective. One has to change one's view on things in order to move further in the field of research, develop new hypotheses and add new findings and insights to the existing body of knowledge. In this thesis we chose some less conventional approaches to study the problem of lower urinary tract symptoms in ageing men. This is a result of my view on general practice, the role of general practitioners in the (Dutch) medical system and - by consequence - my view on the problem of LUTS within this context. These backgrounds were already outlined in the introduction to this thesis. In this general discussion I will place the main findings of the studies in this broader context.

First I will describe the general practitioner with respect to lower urinary tract symptoms as a case-manager. As a case-manager the general practitioner acts somewhat differently than in the classical role with its sequence of diagnosis and consequential therapeutic action. The understanding of the case-manager role is important for a correct understanding of chapter 2.

Then I will take a side-road addressing the role of the general practitioner in the diagnosis of prostate cancer. My view on this role has important consequences for the studies described in this thesis, especially the development of the decision aid (chapter 2) and the randomised controlled trial (chapter 3). I will argue that prostatic cancer casts its shadow on decisions made in clinical practice. This should be recognised in clinical research.

The randomised controlled trial of the effectiveness of the treatment protocol was based on an interpretation of the guidelines as they were at the time of the design and conduct of the trial. In essence, those guidelines haven't changed much. They leave much room for interpretation, especially for the relatively mild cases that present themselves with micturition problems in primary care. The consequences of this will be discussed. One is that the patient plays a key role in the diagnostic process and in the choice of treatment. A well-informed patient is crucial here. Therefore this discussion contains some remarks about how patients are informed. Although much effort is made to provide patients with good quality information - as shown in chapter 6 and 7 - several shortcomings are still present.

It is remarkable that many non-pharmacological treatments are being given in daily practice by all kinds of health care providers. Some of these treatments can even be found in guidelines. Few have been investigated. Our single case study investigated one of these lifestyle advices. The results will be discussed in the broader context of guidelines and non-pharmacological treatment. I will argue that much work should be done here.

Several reviewers of leading medical journals have criticized the methodology used in our studies. Although in the relevant chapters the choices of the design and methodology used have been accounted for - with their strengths and limitations - I will address some of these criticisms too. Some general comments may teach us lessons for the future.

I will close this chapter formulating my view on the subject as a general practitioner working in both research and practice. It appears that the conclusions one wants to draw depend very much on one's perspective.

main findings of this thesis

Chapter 2 describes the development of checklists for the treatment in primary care and the referral to secondary care of patients with lower urinary tract symptoms. We described the preliminary analysis of the usefulness of these checklists and presented an indication of the number of subjects treated and referred if such checklists would be used. These numbers were very acceptable: using the checklists the general practitioner can follow a rational treatment strategy without gross overprescription of medication or undertreatment of patients.

A key element in this chapter is the explicit choice either the doctor or the patient (or both) has to make with respect to the preferable sequence of diagnostic evaluation and therapeutic intervention. One may choose a complete diagnostic evaluation before starting first-line therapy. Alternatively, one may skip several diagnostic investigations and start therapy first, presuming that if this works those diagnostic evaluations are not necessary to reach the goal: symptom relief. If, however, this therapy should fail further diagnostic evaluation may still be an option.

When deciding whom to refer for surgery, general practitioners want to reach two goals. They do not want to withhold probably beneficial therapeutic options for their patients. On the other hand they do not want to refer many patients who eventually will not be operated. In other words: general practitioners must be selective, but not too selective.

Many guidelines provide lists of items that should be considered when deciding to refer or not to refer. Less often, these items can be weighted. However, few doctors will defend that they consider all items equally important. Few patients show 'the full picture'; several items may not be present. Therefore, those items that are present should be (and often less explicitly are) weighed when considering what to do. Our checklists show that age is considered much more important when deciding about surgery (relative to e.g. symptom score) than when medication is considered.

It is important to make such implicit considerations more explicit. Credits should be given to Stoevelaar and McConnell for having done this. This was necessary

to be able to improve not only the treatment policy of urologists, but also that of general practitioners. The latter now can anticipate on the expected treatment policy of their colleagues; they can now make an optimal selection for referral and prescription of medication. I assume that this may also improve the satisfaction of both professionals since this way they will be able to make optimal use of their own and their colleague's expertise.

Chapter 3 presents the results of a randomised controlled trial of a treatment protocol for lower urinary tract symptoms in general practice. The main conclusion is that many more intervention patients receive active treatment (mostly medication) than usual care patients, but that nevertheless patient outcome (symptom score, maximum urinary flow) after two years is not better. Watchful waiting should remain an important option for men with lower urinary tract symptoms.

100 Subjects participated in the intervention arm of the trial. This started with a proposal for a three-month period of tamsulosin use. Although all men knowingly volunteered in participating in a study on treatment of lower urinary tract symptoms, 13 men did not want medication. Apparently, even in this self-selected group of trial participants a considerable number of subjects regarded medication 'too much for too little'. 11 Men stopped their medication due to side effects. This also is a considerable number. It was higher than we expected and is an important finding for the everyday treatment of LUTS. Apparently, it is important and worthwhile to re-evaluate the effectiveness of the treatment in an early stage and take note of possible side effects, (e.g. after one or two weeks - although also later on subjects stopped their medication because of side effects).

After the first three-month period of tamsulosin use about half of the men showed sufficient improvement of the symptom score. These men continued their medication for another six months and were then asked to stop their medication for another three-month period. Afterwards the symptom score was re-measured. A quarter did not show worsening of their symptoms. This also is an important finding suggesting that either the medication had no effect at all (and the measured improvement was not real), or the medication had cured those subjects once and for all, or that the symptom score is not that good a measure after all, e.g. because the subjects could not correctly fill in the symptom score. Several other explanations may also be hypothesized. Whatever the explanation may be, it is important not only to measure the effect of prescribing α-blockers, but also to evaluate whether this effect is a true effect that requires longstanding re-prescriptions.

Chapter 4 shows that in general sitting while urinating does not improve urinary flow. It contradicts suggestions that sitting would facilitate bladder emptying and improve the outflow of urine during micturition. 20 Men participated in this study. In fourteen subjects the mean maximum urinary flow was higher when standing. In six it was lower. The overall difference in maximum urinary flow between standing and sitting was 0.55 ml/s in favour of standing (p=0.02). The average urinary flow was also 0.55 ml/s higher (p=0.006), while the voided volume showed a similar tendency (11 ml higher when standing, p=0.06). Micturition time was similar. While sitting, two subjects had a mean voided volume > 60 ml lower than while standing, suggesting that they leave more residue after micturition when sitting. Such an enlarged residue is clinically relevant because it correlates with higher risk of recurrent urinary tract infections and acute urinary retention.

We acknowledge that there may be other valid reasons for men to sit on the toilet. These reasons are not within the domain of the general practitioner who acts as a professional who advises patients based on clinical evidence and learnedness. (Yet, it is of course relevant for general practitioners who seek advices that are in keeping with their patients' preferences.) At the moment there is no hard evidence that supports the theory that sitting while urinating facilitates bladder emptying. Consequently, we should not give the *medical* advice to urinate this way.

Chapter 5 shows that the correction of artefacts in flow measurements can be automated and standardized with simple means. Visual correction is simple, but is inherently subjective. It is difficult to standardise and maintain standardised during a longer period. An algorithm incorporated in the uroflowmeter would overcome this. It will enhance the objectivity of data collection in clinical research and increase statistical power. It may also facilitate more efficient logistics in clinical practice.

The value of this study for general practice is limited since uroflow measurements are not regularly conducted in primary care. But maybe it is an example for other researchers, in primary care as well, to scrutinize their procedures.

In chapter 6 we conclude that the Internet contains many web sites with health information. With regular search terms and the most popular search engines one can retrieve millions of web sites regarding micturition problems in the elderly. Many are well documented and provide consumers/patients with good quality information. However, many websites are also focusing on a part of the problem, e.g. medication, herbal or surgical solutions. Such selective information may be as misleading as incorrect information. Accreditation of web sites may

be a good way to improve the quality of health information web sites. It insures minimum quality standards. But accreditation focuses on 'formal', procedural aspects of quality. Content quality aspects are not directly secured this way. It is no guarantee for completeness of the information provided. One can argue that the Internet should contain a 'corner' where selected, high-quality information is available to the public. A top-level domain such as *.health* could be such a virtual place. However, the Internet is also an (almost) free sphere, where freedom of speech, and the right to express and get acquainted with opinions and beliefs that dissent with the 'official' truths, are key values. In our open Western societies, the availability of *only* 'high-quality' websites may therefore be considered both a Utopian dream and a Utopian nightmare.

Chapter 7 describes some practical views from both patients and general practitioners on health information in the media, especially the Internet. It shows the increase in information brought in by the patient, but - according to the general practitioners - also a change in the type of information. The information the patient brings in is more detailed, more specific, and more related to the actual complaints or diseases of the patient. This differs from the classic ('old') situation in which a newspaper article or television programme discusses general aspects of a certain disease that may, but also may not, apply to the specific situation of the patient. The Internet offers patients the opportunity to select the information that is applicable to their individual situation from a large database of knowledge and experiences.
The result may be that the consultation reaches a 'higher level'. The consultation becomes more interesting because the general practitioner is more able to use his specific expertise and knowledge, instead of routinely conducting daily work. And patients can get more and more adequate information. This way, the Internet helps GPs to educate patients. However, both patients and doctors also seem to miss opportunities in the consultation room. They hardly discuss the 'background' information during the consultation. They should be more open discussing health information from the Internet.
Furthermore, patients use the Internet as a means to get a second opinion, or as a source used for 'recapitulation' of everything that has been said by the general practitioner during the consultation. As a result, the use of the Internet is often *after* the patient has visited the doctor, not before. General practitioners seem to be hardly aware of this phenomenon.

specific methodological considerations
Each chapter of this thesis describes a different study. A general remark on the appropriateness of the methodology would be therefore almost chapter specific, since both quantitative and qualitative methodologies were used, and both a

'classic' randomised controlled trial and a replicated randomised single case study. Quite extensive descriptions of the strenghts and limitations of these studies are given in the respective chapters. More general remarks can be found in this discussion chapter when e.g. the role of prostate cancer detection and PSA-measurement in the definition of the inclusion criteria are discussed, or the current relevance of the treatment protocol. The most important chapter specific methodological considerations will be (re)considered here.

chapter 2

The development of the checklists for the decision on treatment of lower urinary tract symptoms was based largely on the anticipated treatment policy of urologists. This anticipated policy was derived from the consensus among an international panel of specialists. This is of course subject to changes due to new insights and developments in medicine (some are discussed later on in this chapter). Local urologists may have dissenting opinions. If one follows this line of thinking consequently, one may conclude that local general practitioners should create their own checklists, based on the expected treatment policy of their local urologist(s). This would make application in daily practice difficult. Probably, however, this is more a theoretical problem than a real one.

The development of the checklists took place in the same population as the testing of their value. They should however, be validated in another primary care population. Furthermore, we chose for a watchful waiting strategy when the beneficial effect of a treatment was considered 'uncertain', arguing that re-evaluation could be done at a later stage. This means that general practitioners should indeed follow their patients with lower urinary tract symptoms and not consider the case 'closed' after having given their advice.

chapter 3

The population included in the randomised controlled trial on the effectiveness of our treatment protocol for lower urinary tract symptoms in elderly men was heterogeneous. It reflects the population in general practice. However, these patients were invited to participate. Patients with LUTS who consult their general practitioner decide themselves whether and when to do so and thereby are a self-selected subgroup. The characteristics of this subgroup may vary in time and place. The enormous increase in the number of prescriptions for LUTS in both primary and secondary care illustrates this. However, this trend moves towards - not from - the situation in our study. The self-selection phenomenon becomes 'less selective'.

Although we based our treatment protocol on an internationally validated expert opinion, other treatment strategies may also be appropriate. This is further discussed in the paragraph that deals with recent (developments in) guidelines.

chapter 4

The single-case methodology combines precision with internal validity. By using a (for single case studies) large random sample of the indicated population we can also assume that the results can be generalized to the whole of the primary care population. However, we could not measure the effect on symptom score or post-void residual volume directly, because symptoms are not sufficiently sensitive to change and repeated measurement of post-void residual volume was not possible at home. Some possibly relevant outcome measures could therefore not be studied.

chapter 5

The data used for the study described in chapter 5 were not collected for the purpose of this study. To acquire the dataset for this analysis we had to scan the paper prints and 'remaster' the original data. Consequently, we could not include all flows. However, we did not find an indication that this has led to selection bias. But it would have been better if we would have been able to capture the data from the uroflow meter directly. This would also be the situation in which such an algorithm would be used.

Our study data consisted of many measurements among a small group of participants. Similar studies in other datasets from larger study populations are necessary to verify and refine the results of this study (a sensible commercial party who would build in this feature in his uroflowmeter will do this).

chapter 6

One of the main limitations of this study is that we developed our own measurement instrument. This makes it hard to compare the results with other studies. Furthermore, we had only one investigator, who was blinded for the status of accreditation. We could have printed the web page and have it scored separately from this information.

Our study showed that most websites are good in providing health information. Completeness of information is a very important part of health education. Selective information may have similar negative effects as has incorrect information. It is not only about telling the truth, but also about telling the whole truth. Unfortunately, we did not investigate whether websites did or did not tell 'nothing but the truth'.

chapter 7

We chose for a random sample of subjects for inclusion in our study. This is unusual in a qualitative study. In qualitative research one often chooses for a 'convenience sample', being the population that is easy to reach for the researcher, or, alternatively, for a purposeful sampling technique. In this latter

way one includes subjects according to their specific characteristics. The researcher tries to include a sufficient number of subjects with different characteristics to be able to guarantee that all kinds of opinions are represented in the study population. Random sampling is often considered inappropriate. However, it is - in my opinion - also not wrong to use a random sampling technique, if possible. Sometimes, it may even be better to do so.
Random sampling has the disadvantage that it may lead to inclusion of a large number of 'similar' subjects in the study population that do not contribute to the wealth of data desired by the researcher. It then is an inefficient way to do qualitative research. It may also distract the researcher from dissenting opinions, if the mainstream of the study population shows a similar opinion.
However, in certain circumstances random sampling may be a good - even preferable - way of forming a study population for qualitative research. We concluded that there are many possible determinants of both Internet use and the quality of the doctor-patient relationship. Some of these determinants are patient-specific - such as age, gender, experience with Internet use in general and for health information in particular, educational level, health status (which is a variety in itself), marital status, having/not having children. Others are doctor-specific, such as age, gender, experience in general practice, having practiced in the current practice for a long or short period, working full-time or part-time, alone or in a group practice. Some relate to the combination of patient and doctor, such as congruous or different age, sex, health and disease experiences.
When a lot of such (possibly) relevant determinants are known, it is difficult to compose a specific target population for purposeful sampling. When one suspects yet *unknown* determinants to be present, it is impossible. Random sampling is then - in my view - the best alternative. Therefore we chose for a random sample of patients registered with a targeted population of GPs. By including a sufficient number of patients per GP in the sample, selecting the thus identified relevant population, and subsequently interviewing every eligible subject we could contact, we were able to include a mix of subjects with different socio-economic status, age and sex. On this level we reached data saturation. However, at the level of general practitioners we cannot guarantee data saturation, despite their different individual characteristics, because the sample of general practitioners was too small.

the role of the general practitioner: a case-manager?

diagnosis or case?

The role of the general practitioner in the medical system varies with respect to the condition or type of condition at hand. The roles the general practitioner plays can be distinguished in the classic distinction: prevention, diagnosis,

therapy and prognosis. The past years there has been growing awareness that all kinds of mixes of these roles can be seen in daily practice. An important cause for such mixed roles is uncertainty. This uncertainty can itself be caused by lack of medical knowledge (e.g. due to lack of education or training of the general practitioner), lack of scientific knowledge (e.g. because it has not been studied or study results are equivocal), or by the unknowable aspects of human nature. Such unknowable aspects may have different names, such as biologic variation, intra-individual variation, inter-individual variation, random error or unexplained variance, depending on the way one looks at the matter. These unknowable aspects are refractory in all analyses of patients and their medical conditions, but the pursuit of an explaining theory that reduces such refractory uncertainty goes on.

With respect to lower urinary tract symptoms in elderly men we have seen several concepts passing by during the past decades, each with its own nomenclature. Some of these concepts were addressed in the introduction of this thesis. Essential for general practitioners nowadays is that the concept is mainly descriptive, and individual preferences of patients play an important role in the choice of treatment.

As already mentioned in the introduction, most guidelines share - sometimes more implicitly, sometimes explicitly - the following three elements when dealing with lower urinary tract symptoms:

1. A urinary tract infection or prostate cancer should be excluded as possible cause for lower urinary tract symptoms;
2. Treatment should focus on the reduction of (bother due to) symptoms;
3. Treatment options should be explained; the choice for a specific treatment should be made by the patient.

In such a context, the general practitioner does not come to a defined medical diagnosis, but to a defined medical problem. In a 'normal' diagnostic process, the general practitioner is an interpreter who reassures the patient by recognising the complaints as being of a medical nature and translates these into an established diagnosis, naming a certain disease entity that leads to one or more *treatment options*. But with lower urinary tract symptoms, a definite diagnosis is not possible. In fact, the concept supposes that other diagnoses have to some extent been ruled out, leaving symptom and bother (scores) as the leading trails. When a diagnosis is not available within the possibilities of the general practitioners' working context, the role shifts from an interpreter who translates the complaints into an established diagnosis to an interpreter who translates the complaints into an established medical problem, naming a descriptive concept that leads to one or more *case-management strategies*. Such case-management strategies require not only knowledge on what is the matter (in a classical sense: the diagnosis, but here we should call it the medical

problem at hand), and what can be done (in a classical sense: the treatment options), but also what *will* be done if certain steps are taken. Case-management strategies require knowledge of the actions other actors in the medical field will take when they will see this particular patient. The actions of these other actors may include further diagnostic investigations. Anticipation on the most probable outcome of such investigations and even refraining from such investigations given the known or estimated probability of the outcome of such investigations, is an important characteristic of the difference between the diagnosis-therapy sequence and the case-management strategy. The latter may include much more uncertainty while increasing efficiency and reducing delay of initial treatment, patient burden by diagnostic procedures and costs.
This is in fact the rationale behind chapter 2, the development of a decision aid for the case-management of male lower urinary tract symptoms in general practice. General practitioners should make a shift from a diagnosis to a patient management strategy that includes the expected actions other actors, i.e. the urologists, will take when they will see this particular patient based on a symptom-score based working hypothesis rather than an established diagnosis.

(how) should case-management strategies be investigated?
In diagnostic research it has long been established that the value of a diagnostic variable for the eventual diagnosis depends on the presence or absence of other diagnostic variables. Diagnostic research therefore often uses a more comprehensive approach than testing a single diagnostic variable; it includes other diagnostic variables as well. Stepwise multivariate regression is a statistical technique that can be used to calculated the added value of a given diagnostic variable in the presence or absence of potential other variables.
In therapeutic research such approaches are less frequent, but can also be found in multi-arm trials where two (or more) therapies and a combination treatment are tested against each other and/or to a placebo treatment. Such a design acknowledges that the effectiveness of a second treatment may be additive, less than additive or synergistic.
Investigation of case-management strategies requires a comprehensive approach that includes the possibility for sequential therapeutic interventions and/or the weighing of possible benefits and risks of different therapeutic approaches. It should be acknowledged that during the 'work-up phase' of the doctor there are several moments in which a decision can be made that determines the outcome of the case-management process. Alternative diagnostic routes and alternative treatment options should be kept in mind when a research project is designed. One should also acknowledge the setting in which the case-management process is conducted.
Studies into such case-management strategies may support multidisciplinary

guidelines and interdisciplinary consensus statements between primary care physicians and specialists.

prostatic cancer: its role in decision making in general practice

Chapter 3 describes our randomised controlled trial into the effectiveness of a treatment protocol for LUTS in the elderly male in the setting of the general practitioner. Included in this trial were men with objectified LUTS, i.e. men with scores on the International Prostate Symptom Score higher than a generally accepted cut-off point (IPSS>7). *Not included* in the trial were men with 'known or increased risk of prostate cancer' (see page 40). These seven words may not directly catch one's eye, but they are crucial in the context of general practice. In chapter 2 I describe the development of a decision aid for the treatment and referral of patients with lower urinary tract symptoms. Included were only men with a PSA value < 10 ng/ml.

It was our aim to be able to generalize the results of these two studies to the situation in daily practice. That means that we apparently supposed that the PSA-value would be available to the general practitioner in all patients. Is that assumption justified? And if so, why?

current theory and practice in prostate cancer detection

Although screening for prostate cancer is not (yet) recommended, case finding is more or less current practice. The recent revision of the guideline on LUTS in elderly men by the Dutch College of General Practitioners states that routine PSA testing cannot be recommended.[2] However, it is current practice. Most general practitioners are testing prostate specific antigen (PSA) if requested by the patient and many do it in every man who presents with LUTS for the first time. This means that this particular part of the guideline is often not followed in their daily practice by general practitioners. Why is that?

to test or not to test, is that still the question?

The question whether to grant a request for prostate cancer testing using the PSA-test can be approached from different angles. One may look at it from a macro-economic or health care system perspective. Doing that, the question is easily translated into the question: to screen or not to screen for prostate cancer (using the PSA-test). Then, the work of Wilson and Jungner[1] becomes relevant. In their own words:

'The object of screening for disease is to discover those among the apparently well who are in fact suffering from disease. (...) In developed countries (...) it would seem that the practice of screening for disease should be widespread. >

> That it is not so to the extent that might be expected is due to a number of factors, among them the cost of screening and the tendency in the medical profession to wait for patients rather than actively to look for disease in the population. Another factor undoubtedly is inadequate knowledge of the principles and practice of screening for disease.'

Wilson and Jungner formulated 10 criteria for screening programmes (see box). Following these criteria one may conclude that mass screening cannot be recommended since it has not been shown that it would do more good than harm - like the Health Evidence Network of the World Health Organization Regional Office for Europe did in their recent publication. The Health Council of The Netherlands calls the recent rise in the number of PSA tests among patients who present themselves with lower urinary tract symptoms to their doctors as 'morbid growth'[3], which - from a linguistic point of view - is an interesting term for a cancer test. Should this money-and-space-occupying process be stopped?

criteria of Wilson and Jungner[1]

1. The condition sought should be an important health problem
2. There should be an accepted treatment for patients with recognized disease
3. Facilities for diagnosis and treatment should be available
4. There should be a recognizable latent or early symptomatic stage
5. There should be a suitable test or examination
6. The test should be acceptable to the population
7. The natural history of the condition, including development from latent to declared disease, should be adequately understood
8. There should be an agreed policy on whom to treat as patients
9. The cost of case-finding (including diagnosis an treatment of patients diagnosed) should be economically balanced in relation to expenditure on medical care as a whole.
10. Case-finding should be a continuing process and not a "once and for all" project.

reasons for mass individual screening for prostate cancer

In order to answer this question one should also look at the individual perspective. From this point-of-view several criteria which are relevant when deciding on mass screening programmes are less important or unimportant when deciding on an individual (screening) test.

Programmes for screening prostate cancer do not meet criteria 2, 3, 4, 8, 9 and 10. However, in individual cases criterion 3, 9 and 10 are not relevant, especially if the patient persists in his request after having been informed properly and has sufficient means or is adequately insured. Criteria 2, 4 and 8 are to some extent

subjective: whether to treat or not to treat given a certain outcome of the PSA-test (and possible additional tests) may be discussed with and decided by the patient.

Testing that can not be recommended on a general population level can therefore nevertheless be very sensible and rational in individual cases, even if such individual cases turn out to be a large part of the population. The 2004 guideline of the Dutch College of General Practitioners recognises this and recommends not to screen for prostate cancer, but nevertheless perform the test (along with digital rectal examination, the other screening test) when the patient after having been informed persists in his request. This recommendation fits in with current practice.

other reasons for mass testing

'Doctors in the field' may also have other valid reasons for performing the tests. The statistical truth about the correlation between LUTS and prostate cancer among scientists differs from the psychological truth about this relation among the lay public. Prostate cancer is an important cause of death in elderly men, and thereby much feared. Impaired voiding quality as manifested by LUTS is often felt as a first possible warning sign for prostate cancer.[2] The fear of prostate cancer often is the question-behind-the-question for a patient presenting himself with LUTS to the GP.[2] It has been shown that the burden of LUTS can be reduced if doctors exclude prostate cancer and consequently reassure their patients that the symptoms are not caused by a serious disease. Wise doctors will always explicitly pay attention to prostate cancer during a first consultation about LUTS. It is unwise to treat LUTS and prostate cancer as totally different topics during such a first contact.

Futhermore, guidelines for lower urinary tract symptoms require a - limited - effort to exclude prostate cancer. In daily practice this means that general practitioners can choose between digital rectal examination and PSA testing. Digital rectal examination and PSA have comparable sensitivity and specificity in a research setting (PSA often is even slightly better).[4] The daily care effectiveness of PSA in prostate cancer detection is therefore at least comparable and probably better than digital rectal examination. Doctors may feel safer performing such an objective test, especially if they do not have much experience with rectal examination for exclusion of prostate cancer. Therefore, PSA may be useful as a first, limited tool.

PSA values are easier to follow-up in a multi-doctor setting, and are therefore a good starting point for a follow-up strategy.

Notably, the guideline of the European Association of Urology states:

> 'Digital rectal examination has been used in all major screening trials but its actual impact in the early diagnosis of prostate cancer has been questioned'[5]

and

> 'The measurement of PSA is recommended when a diagnosis of prostatic carcinoma will change the decision made about which therapeutic option to use'.[5]

Reading these two recommendations in the EAU guidelines, one may easily conclude that PSA-testing for the purpose of exclusion of early stages of cancer in a general practice setting, is superior to digital rectal examination.
Finally, the American Cancer Society promotes screening:

> 'Both the prostate-specific antigen (PSA) blood test and digital rectal examination (DRE) should be offered annually, beginning at age 50, to men who have at least a 10-year life expectancy'. Men at high risk should even start with testing earlier (40-45 years).[6]

There are therefore also leading medical associations who support mass testing.

consequences for research in general practice

Researchers in general practice cannot ignore this phenomenon when carrying out a pragmatic trial such as the one described in chapter 3, or developing a case-management strategy as described in chapter 2. If we would have included a more unselected case-mix in our analyses for the decision aid, we would have had subjects in our analyses that in daily practice would have been referred for further analysis of an elevated PSA-level and the risk of prostatic cancer. The analyses would then have been contaminated and their results biased by patient characteristics - such as an enlarged prostate - that co-occur with such an elevated PSA-level. The result would then not have been representative for those patients in which in real life referral or treatment in primary care is still a matter of choice. The same holds for the target population of the randomised trial in chapter 3. Those patients who certainly would have been referred for further follow-up diagnostics for prostatic cancer are most probably also initially treated for their lower urinary tract symptoms by the urologist. Therefore, the treatment protocol will not be applied to them and this protocol should therefore not be tested with these subjects included in the study population.

our treatment protocol versus current and near future treatment guidelines

Guidelines are joint expert opinion. They require interpretation of the available evidence by experts, and are the result of consensus procedures. Guidelines should be used in the health care setting for which they are made. Consequently,

guidelines may differ in time or place even if the facts are the same. Also, they may differ for varying groups of health care professionals. Guidelines for general practitioners often differ from those for specialists because of the different case-mix of patients these doctors care for. Furthermore, the health care system in different countries may lead to different weighing of the facts in guidelines. A more 'defensive' health care system, i.e. a system in which the liability of doctors for missed diagnoses is greater, may lead to a more extensive diagnostic process, whereas a system that has more financial restraints will include more (or different) cost-effectiveness trade-offs leading to a less extensive diagnostic work-up routine. A similar reasoning holds for treatment. This leads to the question: how useful was the pragmatic trial described in chapter 3?

alternatives for our treatment protocol?

Key question is whether the treatment protocol is situational or has general validity. Can the results be translated to the situation in other countries, and will they last? Or will new developments change the policy of doctors thereby making the conclusions of the study obsolete? Are other guidelines much different? Can important changes be expected soon?

It is clear that α-blockers have got a more prominent place in the treatment of lower urinary tract symptoms. 5-α-reductase inhibitors are also making a 'comeback' since the MTOPS study has shown that combination therapy of an α-blocker and a 5-α-reductase inhibitor has statistically significant added value above a single medication treatment after five years of follow-up.[7] The risk of acute urinary retention and clinical progression that necessitates surgery can be reduced. However, this effect is small in terms of absolute risk reduction, since the risk of acute urinary retention is only about 0.5 per 100 patient years.

Currently, α-blockers are considered a suitable option for the treatment of men with moderate lower urinary tract symptoms. A 5-α-reductase inhibitor may also be suitable if the prostate volume exceeds 30-40 ml.[8] However, the patient should be thoroughly informed about the possibly small effect and the specific side effects of the medications.[8] Although these medications can be useful, apparently they do not lead to overwhelming enthusiasm when costs and benefits are weighted.

When we designed the study, the results of the long-term follow-up of the MTOPS study were not available. Combination therapy was considered not to be beneficial above single medication treatment. Therefore, it had no place in our treatment protocol. We might have reconsidered this if our study would have been conducted in the present circumstances. However, this would then only regard men with larger prostates (> 30-40 ml). One would presume that valid data about prostate size would (or could) be known to the general practitioner at the time of inclusion. In the current circumstances this would mean that either

an open-access ultrasound unit is available to measure prostate size by means of echography, or data are collected with digital rectal examination. Alternatively, some cut-off point of the PSA value could be taken as a workable proxy for prostate size estimation. All these three options have their limitations that make it hard to incorporate combination therapy in treatment protocols in general practice.

Another radically different approach would have been to randomise patients not on symptom score but on bother score. All guidelines are divided by either symptom severity (giving the appropriate treatment modalities for each grade of symptom severity) or by treatment option (defining the proper groups for whom these treatment modalities are indicated). Most guidelines state that bother is a key element in the decision which treatment to choose. Especially for a study among patients in general practice one may also choose for bother as the only (or most important) criterion on which patients are randomised. After all, the International Prostate Symptom Score is more a symptom *frequency* score than a symptom *severity* score.

the evidence for non-pharmacological treatment

Many lifestyle advices are currently given in daily practice. Many more can be found in other sources of health information such as magazines and the Internet. The effects of such non-pharmacological interventions are unknown as are the side effects. Most of them are not based on scientific evidence. Few have been investigated at all. This can partially be explained by a lack of financial 'drive' to investigate such lifestyle advices. Non-pharmacological interventions are (regarded) not profitable while medication is. The importance of marketing in the increase of attention for LUTS is great.[5]

Some of them nevertheless have gained a more or less official status, because they have been adopted by The European Association of Urology and the American Urological Association. An example is the reduction of fluid intake in the evening to prevent nocturia. How effective is this advice? What are the consequences for the remainder of the day, where a minimum fluid intake of 1500 ml is still regarded necessary? Can this advice and the finding that increased urine output due to increased water intake may improve bladder functioning[9], both be true? Does it interact with other medical therapies such as diuretics? And are there circumstances in which such an advice should not be given? All such questions need to be answered if one would like medical practice to be as evidence-based as possible. One may assume that 'common knowledge' and 'reason' may also be good foundations for medical practice, but the replicated single case experiment showed that this is in fact not always justified.

In chapter 4 we stipulate that one can better urinate standing than sitting. The overall difference in urinary flow between these voiding positions is clinically relevant since it resembles the net (added) effect of medication (above placebo) in up to five years. This is of course only true if
1. urinary flow is an important aspect of the decision to start medication;
2. patients can one way or the other experience the effect of a good voiding position;
3. the effect is stable and lasts for years; and
4. the effect on urinary flow is not countered by a contrary effect, e.g. on bother.
If patients find it hard or inconvenient to urinate standing-up, or spill more urine outside the toilet because of a weak stream, the lifestyle advice will be useless. These aspects were not studied, although we wanted to.

reflections on the advantages of the methodology used for research in general practice

pragmatic studies
The studies described in chapter 2 (decision aid), 3 (effectiveness of a treatment protocol) can be regarded as pragmatic studies. Chapter 4 (voiding position) we would classify as an explanatory trial that, however, was designed much like a pragmatic trial.
In chapter 2 we used 'GP data' to stratify the patients into the different groups, but the expected treatment strategy of the specialist was based on the information that would be available in a secondary care setting. Prostate size as estimated by digital rectal examination in the primary care setting was used ('GP data'). For the determination of the expected treatment strategy of the specialist we used the data of the ultrasound prostate size measurement.
If general practitioners would choose for more certainty about prostate size and therefore refer for ultrasound prostate size measurement, the case-management strategy would become totally different. A very different indicated population would remain under the care of the general practitioner. If, on the other hand, an open-access ultrasound facility would become available for such measurements, the indicated population under the care of the general practitioner might increase, and partially consist of people with different characteristics.
One of the main characteristics of pragmatic trials is that the study population included resembles the real life population as much as possible by allowing that an in many ways heterogeneous group of subjects is included. Key element is that the thus created study population is easily identifiable with as limited means as possible (or reasonable given the situation of the doctor for whom the results of the study are meant). Pragmatic studies therefore have pragmatic in-

and exclusion criteria. In the randomised controlled trial of chapter 3 the study population was included on symptom scores alone. Increased risk of prostatic cancer was an exclusion criterion as were potentially contraindications of the medication used in the treatment protocol under study. But the diagnosis was not supported by either uroflowmetry or pressure-flow studies, nor by a minimum prostate size as are most medication trials for lower urinary tract symptoms in men. This approach lies closely to the pragmatic way in which most general practitioners deal with such patients.

The purpose of pragmatic studies is to generate results that are relevant and useful under usual care circumstances.[10] Although they are more useful, they can be more difficult to finance and conduct, because they need to be larger than 'explanatory trials', as the variability inherent in their unselected approach reduces power.[10] Explanatory trials may be useful for answering questions about whether an intervention can work under ideal circumstances.

The main advantage of the results of explanatory trials is that the ideal circumstances are easier to re-create. Experiments can therefore easier be repeated. The results of a study can therefore easily be supported or falsified by a following experiment. However, use of the results in a daily practice situation ('usual care') supposes that usual care does not differ much from the ideal situation as used in the experiments. This supposition is often hardly supported by evidence. In contrary, daily practice appears to vary much and often does not resemble the 'optimal' situations as they should be according to guidelines or 'best practice' situations (which often also do not resemble the ideal circumstances of explanatory trials). This gap between the ideal situation and the daily practice situation often is not discussed at all. Explanatory trials generate an 'abstracted truth'.

Results of pragmatic studies resemble the 'real life truth' much more. But it is self-evident that the results can also not be generalised to every doctors' daily practice when usual care shows much variation. Since circumstances differ in time and place pragmatic trials are more difficult to re-conduct. Supportive or falsifying evidence is much more difficult to acquire.

It is also much more difficult to perform a meta-analysis on such trials because the heterogeneity of the studies makes it hardly possible to correctly summarise the results into a overall result. Since meta-analyses are regarded the highest form of scientific truth, providing the strongest evidence, one may even assume that pragmatic studies are not very scientific. Science presupposes an experiment that can be redone. If pragmatic trials are hard to redo, they 'just' use scientific methodology, but do not add to the scientific body of knowledge, one could say.

Nevertheless, such studies are necessary because it is not likely that in the near

future a major breakthrough will occur with respect to the treatment with medication or surgery of lower urinary tract symptoms that will alter the treatment policy of general practitioners. Therefore, in order to make usual care as evidence based as possible one has not only to provide doctors with guidelines based on scientific evidence, but ideally also with tested guidelines with proven effectiveness ('evidence-founded' guidelines).
Such studies are also useful to determine the number of patients that are 'covered' by the (evidence-based results in) guidelines. Using the results of chapter 2 and 3 the number of patients in certain strategy-options can be calculated; for individuals the chance that a particular treatment is indeed carried out and will be effective can be determined (and be weighed and play a role in the individual's decision making process).
Finally, these studies are scientific if one tries to explain why real life results differ from ideal situation results, and uses such an explanation as a hypothesis in further research.

single case studies
Single case studies can be defined as designed experiments in which one entity is observed repeatedly during a certain period under different levels ('treatments', or 'interventions' - including the control condition) of at least one independent variable. The primary aim of a single case study is to prove or falsify a relation between two variables within a single subject. These variables may be related to etiologic or diagnostic questions, or to the effectiveness of a specific treatment. It may be that one observes a certain disease status and exposure to a possible causing agent and would like to know whether these are related. A single case study may generate new hypotheses that should afterwards be studied in studies with groups of subjects. One may also know from large-scale group studies that a certain remedy is not effective in all patients, but may be of use in hard-to-define subgroups. In such a situation it may be prudent to test the effectiveness of this treatment in a particular patient before prescribing it for an undefined but long period of time. Such a prescription may be unwarranted because of the costs, long-term side effects or interactions with other treatments, or because the treatment is unpleasant.
Single case and group experimental research are therefore complementary, and - in some ways - even interdependent.[11]
There are, however, some limitations to single case studies that should be kept in mind. The intervention should be suitable for repeated administration; the outcome measure should be sensitive to change in relation to the observation period, but robust to the effects of possible confounders. Furthermore, measurements should be timed so that the intervention under study has no carry-over effect. A carry-over effect means that the effect of the intervention (or

the control condition) lasts till after the change between intervention en subsequent control condition (v.v.) and is still present at the time of the following measurement. In other words: the effect of the treatment is still present at the time that the subject is measured in a presumably non-treated status. In such a situation the measurement of health status during the control condition is 'contaminated' by the effect of the intervention under study. The treatment effect might then be underestimated.

Another prerequisite for single case studies is that the condition is more or less stable over time. Single case studies can be performed with conditions that are fluctuating over time, but this requires much more measurements. As a result, the methodology is most suitable for chronic diseases.

advantages for general practitioners to perform single case studies

Hypothesis generating single case studies require recognition of a seemingly coincidental co-occurrence or sequential occurrence of two phenomena that forms the basis for the study question. Specialists take care of a more or less homogeneous patient group. This may have the advantage that certain patterns are easier to recognise. On the other hand, general practitioners may experience the advantage of having a broader knowledge of aspects of their patients' lives. This means that seemingly coincidental co-occurring phenomena may be easier to recognise. This may lead to unexpected findings.[12] Single case studies require intensive guidance of the participant during the time of data collection. This can relatively easily be done in a primary care setting where the (physical) distance between the participant's home and the general practitioner's surgery often is small.

Chronic, stable diseases are one of the core businesses of the general practitioner. Therefore, the general practitioner is in a good position to perform single case studies. Stimulation of single case studies by general practitioners has been advocated before. [13,14] This regarded mainly the targeting of medication to their patients. It is certainly possible, however, to perform explanatory (multiple) single case studies aiming to generate, confirm or falsify explanatory hypotheses and research questions.

meta-analysis of single case studies: it can be done, but should it be done?

Given the nature of single case studies it would appear inconsequent to pool the results of the studied individuals together into a meta-analysis. Some argue that a wide variation of efficacy of an intervention under study across patients in replicated single case studies makes such studies heterogeneous. Therefore - similar to systematic reviews - it would be 'pointless' to pool such heterogeneous results to a 'mean' effect. Only if the studies show a 'negligible variation' in efficacy of the results, pooling may seem sensible. However, in such cases single

case studies are considered 'not necessary'.[13]
I question that this point of view is correct. The similarity with systematic reviews is appealing. One should bear in mind that heterogeneity does not depend on the differences in outcome, but results from differences in design, patient recruitment, and methods used in the different trials as well as the way these are conducted. If single case studies are properly replicated, both methodology and conduction of the trials are homogeneous. If, furthermore, one is able to recruit an unselected, representative sample of the indicated population such a meta-analysis is very defendable. The only remaining question is whether the group of subjects included is large enough to make sure that (most of the) relevant characteristics of the target population are included in the study population. In other words: the study population of the meta-analysed single case studies as a whole should meet the criteria of external validity, and be large enough to produce a sufficiently precise result (adequate power of the meta-analysis).
Our single case study meets these requirements. Nevertheless, the results should not be regarded as the definite answer. As (nearly) always, additional studies are necessary to confirm or falsify the results.

ethics and clinical research
During the conduct of the studies of this thesis there has been an important formalisation of the ethical codes that researchers (should) use. Legislation has become more stringent both in Europe as a whole and in The Netherlands. The Central Committee on Research Involving Human Subjects (CCMO) has produced regulation. Procedures of our local medical ethics committee have changed several times. Registration of randomised clinical trials is obligatory for pharmacological studies used as evidence in registration procedures and has become a prerequisite for publication in leading medical journals. Publication of trials should follow a standardised scheme laid down in the so called CONSORT statement.[15]
A personal observation is that several items of e.g. the CONSORT statement are apparently considered as more important than others. Much effort is required to prove that selection bias has not occurred, or to show where possible selection (bias) has occurred. Flow charts of the recruitment of participants and the flow of these subjects through the study protocol can be found in all current study reports. Practical questions about how data are handled and how problems with e.g. artefacts, or missing data are solved are hardly to be found. This, despite the fact that CONSORT statement explicitly requires that:

> 'authors should give full details of how the primary and secondary outcomes were measured and whether any particular steps were taken to increase the reliability of the measurements'.[16] They 'should also should also discuss any imprecision of the results'.[17]

The CONSORT statement focuses primarily on the reduction of bias in study reports. It is regarded as a quality criterion. In my view, it also has important ethical connotation. As Knottnerus and Bouter rightly state:

> 'To collect sufficient evidence to detect a clinically relevant difference between both interventions, if present, a minimum number of subjects per group should be included. At the same time, from an ethical point of view and to warrant an efficient use of resources, it is also required that the sample size be not larger than needed'.[18]

Efficient use of resources, be these money, material or (especially) voluntarily participating people, should play a more important role, both in the reporting of studies and -earlier on- in the design. It should get more attention in the design (and conduct) of clinical studies, and - consequently - in the reviewing process of both medical journals and medical ethical committees. The Central Committee on Research Involving Human Subjects states as one of the (scientific, not ethical (!)) criteria for the evaluation of medical studies by medical ethical committees that as few participants as possible are included.[19] Implementation of simple measures to enhance the precision - be these regarded as quality improvement or ethical requirement - such as the algorithm described in chapter 5 should therefore be promoted.

the concept of the study domain

the outline of the study domain

In the introduction to this thesis I presented the outlines of the study domain, along with the context of the present studies. We discussed several ways to look at the topic. I showed that the concept of impaired micturition in the elderly male has undergone a shift from a diagnosis-based orientation (benign prostatic hyperplasia) to a working hypothesis-based orientation. This latter orientation focuses on presenting symptoms (lower urinary tract symptoms) and acknowledges uncertainty about the specific cause of the symptoms.

conclusions to be drawn from this thesis

General practitioners tend to support a patient with a chronic ailment rather than to treat a well-defined disease. This creates room for a more individualised approach rather than following a treatment protocol. In chapter 2 this is expressed as two management strategies: a referral first strategy and a medication first strategy. We underline that in principle these strategies are equally defendable. The choice between both strategies should be based on the preference of both the doctor and the patient. The patient has a strong influence here. There is no 'good' and 'bad' strategy. There is a choice between two reasonable strategies. However, this does not mean that all possible strategies are reasonable. It is - within a context of a gatekeeper system as the Dutch health care system - reasonable to anticipate on the expected treatment choices of one's colleagues working in the secondary care setting. It is not reasonable to either refer just anyone or to always prescribe medication if one knows that easier or better options (such as not prescribing medication at all) are available. But most of all, it is not reasonable *not* to discuss the options with their advantages and disadvantages with the patient and to try to reach a common understanding about the strategy of choice.

Can the presented checklists be used in daily practice? No and yes. No, because most general practitioners already have several 'checklists' and score forms in their drawers such as the International Prostate Symptom Score, with score forms for depression, anxiety, low back pain and several other diseases as well. In daily practice, a general practitioner will only use a limited number of such forms. So one may conclude that the checklists presented in chapter 2 are of scientific interest, but of limited value in daily practice. However, when incorporated in a guideline the usefulness may become clear. It may show that several possible criteria for patient management decisions are not of equal value, but that some (e.g. severe symptoms) are much more important than others (e.g. age).

In chapter 3 we described the failure of our randomised controlled trial to prove that a 'more active' treatment protocol leads to a better outcome in terms of both symptoms and urinary flow parameters. In the studied treatment protocol more patients were considered eligible for medical treatment. One may say that the failure of this treatment protocol to improve health outcome at the group level shows that the individual approach currently conducted by the general practitioners in the study is correct. Or at least one may conclude that it is justified since there is no better alternative. Such an individual approach is preferable, because it resembles daily practice as much as possible.

The results of chapter 4 too led us to the conclusion that an individual approach could be useful when an advice concerning voiding position would be given. The hypothesis that sitting while urinating would facilitate bladder emptying and

consequently would lead to better urinary flow parameters could not be proven. On the contrary, in general the opposite conclusion should be drawn. However, there were also subjects for whom sitting was - or tended to be - better. An individualised approach may therefore be preferable. The conduct of the replicated randomised single case experiment was relatively easy. The participants were able to collect complete and good quality data in very little time. The fact that measurements could be done in their own homes certainly contributed to this. We therefore suppose that this can be done easily in daily practice as well.

Chapter 5 should be an incentive for all researchers in the field of medicine, and in particular in the field of general practice. It should also be an appeal to members of ethics committees and review boards of funding agencies. A close look at the methods to increase cost-effectiveness and *human resource-effectiveness* should become/remain an integrated part of their work.

Chapter 6 and 7 may teach a lesson to all general practitioners. But also to those teaching general practitioners in training. Nowadays, good communication skills imply that general practitioners have an open mind for the (at least in the view of the patient himself) well-informed patient.

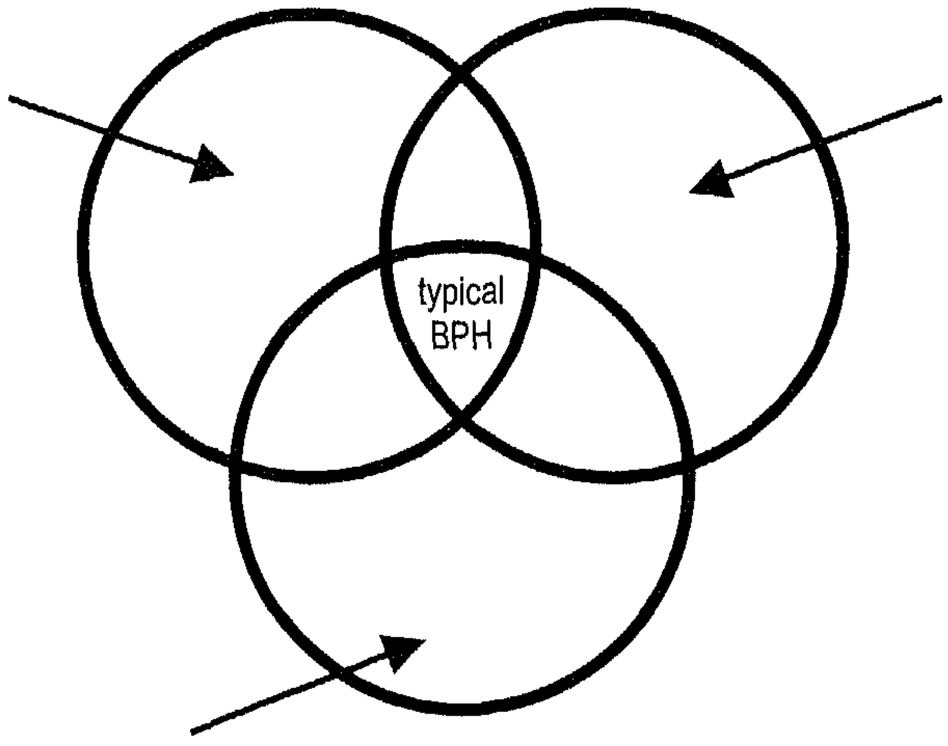
recommendations for future research

What should we be investigating in near future research projects? Certainly, we should keep looking at the matter from different angles. For daily general practice a new chemical entity that would prove to be a very effective medicine would of course have a large impact. But as long as this does not happen, we should place the individual patient more in the middle of our focus. Many non-pharmacological treatments are given. It has already been mentioned that few have been investigated. That is a pity. Expected results, but most of all unexpected ones from such studies may lead to new hypotheses that open new paths for further research. Why is it that we *think* that sitting is beneficial, but *find* that standing often is better? What does it tell us about the way we evaluate patients, assume things, register selectively, or look at the (patho)physiology of the lower urinary tract? This also implies that a new round of qualitative research may be needed. Especially, I would be interested to know why 'severe symptoms' and 'severe bother' can be such different entities? The 'discongruent subjects' (see figure 8.1, page 140) are the most interesting to study: severely bothered by mild or moderate symptoms (as assessed by the IPSS), or, alternatively, hardly bothered though having severe symptoms (as assessed by the IPSS).

I already pointed out that part of this can be explained by a misinterpretation of the IPSS, which is *in strictu sensu* a symptom *frequency* score and not a symptom *severity* score. This also implies that future randomised controlled

trials into the efficacy, but especially into the effectiveness of treatments for lower urinary tract symptoms should randomise patients stratified on bother score rather than on symptom frequency score.

figure 8.1. Figure 1.1 revisited. Discongruent cases may inspire to new insights



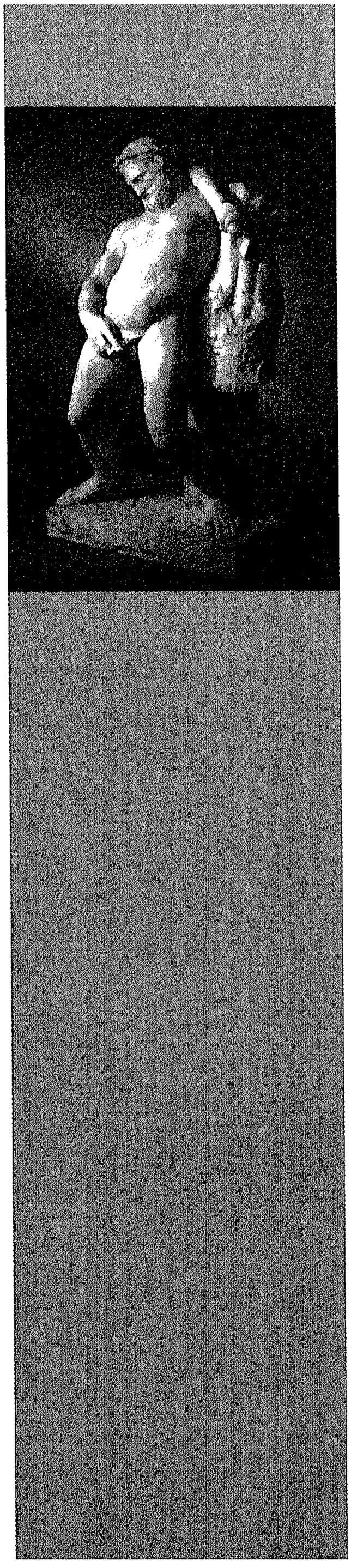
Finally, we know that the world is changing and patients become more informed. But how do patients anticipate on and respond to doctors who are less informed than they are (supposed) to be? What do patients do if the Internet tells them differently than their doctor did before? And how do doctors respond to the teaching rather than the asking patient? How should they respond? And how do we teach them to respond properly? Can this be done with the 'usual skills' or does this require specific training or special attitudes? And - more philosophical - where does the specific expertise of the general practitioner start, and where does it end?

final remarks

Several consequences of this point of view should be borne in mind. In daily practice an important focus of therapy lies with indicating prognosis. In a follow-up regimen (watchful waiting) the doctor should indeed be watchful, not only waiting. And during this period of time lifestyle advices become more important in dealing with the bother of the uncured ailment. Such lifestyle advices fit in with the patient's (ability to change his) lifestyle as much as possible. Good and reliable information is necessary. All interventions, both with medicines as with lifestyle advices, need an evidence-base. And they should be published. The Internet is a much used and suitable instrument. Most patients seem to know how to use it. Unfortunately, it is underused in the consultation room where GPs and patients can come to a shared decision based on reasonable management strategy alternatives. After all, shared decision making between doctors and patients is always a process of bringing together different perspectives.

References

1. Wilson J, Jungner Y. Principles and practices of screening for disease. Geneva: World Health Organization, 1968.
2. Wolters RJ, Spigt MG, van Reedt Dorlant FH, Gercama AJ, Klomp MLF, Romeijnders ACM, et al. NHG-Standaard Bemoeilijkte mictie bij oudere mannen. *Huisarts en Wetenschap* 2004;47(12):571-86.
3. Gezondheidsraad. Jaarbericht bevolkingsonderzoek 2006. Den Haag: Geondheidsraad, 2006.
4. Klomp MF, Gercama AJ, de Jonge-Wubben JG, Mulders AH, Romeijnders AC, Rosmalen CF, et al. NHG-Standaard Bemoeilijkte mictie bij oudere mannen. *Huisarts en Wetenschap* 1997;37:357-65.
5. De la Rosette JJ, Madersbacher S, Alizivatos G, Rioja Sanz C, Emberton M, Nordling J. *Guideline on benign prostatic hyperplasia*: European Association of Urology, Dec 2004.
6. Smith RA, Cokkinides V, Eyre HJ. American Cancer Society guidelines for the early detection of cancer, 2006. *CA Cancer J Clin* 2006;56(1):11-25.
7. McConnell JD, Roehrborn CG, Bautista OM, Andriole GL, Jr., Dixon CM, Kusek JW, et al. The long-term effect of doxazosin, finasteride, and combination therapy on the clinical progression of benign prostatic hyperplasia. *N Engl J Med* 2003;349(25):2387-98.
8. Nederlandse vereniging voor Urologie. Richtlijn diagnostiek en behandeling van LUTS/BPH. 2005.
9. Spigt M, van Schayck O, Knipschild P, Westerterp K, van de Beek C, van Kerrebroeck P, et al. Is it possible to improve elderly male bladder function by having them drink more water? A randomized trial of effects of increased fluid intake/urine output on male lower urinary tract function. *Urology* 2006;68(5):1031-6.
10. Zwarenstein M, Oxman A. Why are so few randomized trials useful, and what can we do about it? *J Clin Epidemiol* 2006;59(11):1125-6.
11. Onghena P, Edgington ES. Customization of pain treatments: single-case design and analysis. *Clin J Pain* 2005;21(1):56-68; discussion 69-72.
12. Stricker BH, Heijermans HS, Braat H, Norg J. Fever induced by labetalol. *Jama* 1986;256(5):619-20.
13. Wegman ACM. N-of-1 trials in general practice. 'tailoring treatments to individual patients'. Thesis. Vrije Universiteit Amsterdam, 2005.
14. Vandenbroucke JP. [The N-of-1 trial: the ideal study design that is underused]. *Ned Tijdschr Geneeskd* 2006;150(51):2794-5.
15. http://www.consort-statement.org/ (Accessed Jun 15 2007).
16. http://www.consort-statement.org/examples6b.htm (Accessed Jun 15 2007).
17. http://www.consort-statement.org/Statement/examples20.htm (Accessed Jun 15 2007).
18. Knottnerus JA, Bouter LM. The ethics of sample size: two-sided testing and one-sided thinking. *J Clin Epidemiol* 2001;54(2):109-10.
19. http://www.ccmo-online.nl/hipe/uploads/downloads_catp/Toetsingshandleiding_2002.pdf (Accessed Jun 15 2007).



Summary

Summary

the context of the studies in this thesis

For most people who have grown out of their early childhood diapers micturition and urinary continence have become self-evident phenomena. However, during life the quality of urine storage and voiding often changes. These changes take place very, very slowly. The evolution from 'health' to 'disease' - if we may call it a disease at all - takes place very gradually over decades. Like other urogenital ailments, micturition problems often are a taboo. Most people hesitate to talk about it with others, be they friends, relatives or health care professionals. Micturition is a private matter, and a solitary one. People lack a proper reference standard for self-assessment of the normality of one's micturition behaviour. With regard to micturition the optimal state of health is clear. However, the state of 'disease' is not defined.

Medical models require physical and/or mental targets for intervention by doctors as explanatory mechanism. In former days, doctors spoke of the 'disease' as benign prostatic hyperplasia (BPH) with corresponding symptoms of prostatism. Later, when there appeared to be only a poor to moderate correlation between symptoms and 'objective' parameters measured with urinary flow measurement, urodynamics and ultrasound, the medical profession started to use a first abstraction of the anatomic concept: *clinical* BPH. At present, in the predominant medical model the 'disease' is called 'lower urinary tract symptoms', which is a second step in the abstraction from the physical target of medical intervention to the conceptualisation of the 'disease'. This term immediately shows that this concept consists of a combination of symptoms rather than a well-defined disease. Consequently the target of the diagnostic work-up has returned from the general debate in medical science to the consulting room where the physician meets the individual patient.

General practitioners are specialists in uncertainty medicine. With limited technical means they try to come from the reason for encounter of the patient to a reasonable working strategy. This strategy is based on the most likely hypothesis in the absence of serious other possible conditions rather than on a well-established diagnosis.

In the past decades three main trends have influenced the management of lower urinary tract symptoms in general practice. As a result the average man who consults his doctor for lower urinary tract symptoms is younger, and less symptomatic. In conclusion: GPs treat new groups of patients with new stages of the 'disease' with new medication in a new setting.

Most guidelines for the diagnosis and treatment of lower urinary tract symptoms (in general practice) share three elements:
1. A urinary tract infection or prostate cancer should be excluded;
2. Treatment should focus on the reduction of (bother due to) symptoms;
3. Treatment options should be explained; the choice for a specific treatment should be made by the patient.
Over-the-counter preparations and non-pharmacological interventions are often used. The Dutch College of General Practitioners (NHG) recommends amongst others to 'void sitting'. This recommendation is not evidence-based and has hardly been investigated.

a decision aid for general practitioners for the treatment of elderly male patients with lower urinary tract symptoms
Chapter 2 describes the development of checklists for the treatment in primary care and the referral to secondary care of patients with lower urinary tract symptoms. We described the preliminary analysis of the usefulness of these checklists and presented an indication of the number of subjects treated and referred if such checklists would be used. These numbers were very acceptable: using the checklists the general practitioner can follow a rational treatment strategy without gross overprescription of medication or undertreatment of patients.
A key element in this chapter is the explicit choice either the doctor or the patient (or both) has to make with respect to the preferable sequence of diagnostic evaluation and therapeutic intervention. One may choose a complete diagnostic evaluation before starting first-line therapy. Alternatively, one may skip several diagnostic investigations and start therapy first, presuming that if this works these diagnostic evaluations are not necessary to reach the goal: symptom relief. If, however, this therapy should fail further diagnostic evaluation may still be an option.
When deciding whom to refer for surgery, general practitioners want to reach two goals. They do not want to withhold probably beneficial therapeutic options for their patients. On the other hand they do not want to refer many patients who eventually will not be operated. In other words: general practitioners must be selective, but not too selective.
Many guidelines provide lists of items that should be considered when deciding to refer or not refer. Less often, these items can be weighed. However, few doctors will defend that they find all items equally important. Few patients show 'the full picture'; several items may not be present. Therefore, those items that are present should be (and often less explicitly are) weighed when considering what to do. Our checklists show that age is considered much more important an issue when deciding about surgery (relative to e.g. symptom score) than when

medication is considered.
It is important to make such implicit considerations more explicit. Credits should be given to Stoevelaar and McConnell for having done this. This was necessary to be able to improve not only the treatment policy of urologists, but also that of general practitioners. The latter now can anticipate on the expected treatment policy of their colleagues; they now can make an optimal selection for referral and prescription of medication. I assume that this may also improve the satisfaction of both professionals since this way they will be able to make optimal use of their own and their colleague's expertise.

the effectiveness of a treatment protocol for male urinary tract symptoms in general practice
Chapter 3 presents the results of a randomised controlled trial of a treatment protocol for lower urinary tract symptoms in general practice. The main conclusion is that many more intervention patients receive active treatment (mostly medication) than usual care patients, but that nevertheless patient outcome (symptom score, maximum urinary flow) after two years is not better. Watchful waiting should remain an important option for men with lower urinary tract symptoms.
100 Subjects participated in the intervention arm of the trial. This started with a proposal for a three-month period of tamsulosin use. Although all men knowingly volunteered in participating in a study after treatment of lower urinary tract symptoms, 13 men did not want medication. Apparently, even in this self-selected group of trial participants a considerable number of subjects regarded medication 'too much for too little'. 11 Men stopped their medication due to side effects. This also is a considerable number. It was higher than we expected and is an important finding for the everyday treatment of LUTS. Apparently, it is important and worthwhile to re-evaluate the effectiveness of the treatment in an early stage and take note of possible side effects, e.g. after one or two weeks. Later on, however, subjects stopped their medication because of side effects as well.
After the first three-month period of tamsulosin use about half of the men showed sufficient improvement of the symptom score. These men continued their medication for another six months and were then asked to stop their medication for a three-month period. Afterwards the symptom score was re-measured. A quarter did not show worsening of their symptoms. This also is an important finding suggesting that either the medication had no effect at all (and the measured improvement was not real), or the medication had cured those subjects once and for all, or that the symptom score is not that good a measure after all, e.g. because the subjects could not correctly fill-in the symptom score. Several other explanations may also be hypothesized. Whatever the explanation

may be, it is important not only to measure the effect of prescribing α-blockers, but also to evaluate whether this effect is a true effect that requires longstanding re-prescriptions.

position-related differences in voiding in men with lower urinary tract symptoms
Chapter 4 shows that in general sitting while urinating does not improve urinary flow. It contradicts suggestions that sitting would facilitate bladder emptying and improve the urine flow during micturition. 20 Men participated in this study. In fourteen subjects the mean maximum urinary flow was higher when standing. In six it was lower. The overall difference in maximum urinary flow between standing and sitting was 0.55 ml/s in favour of standing ($p=0.02$). The average urinary flow was also 0.55 ml/s higher ($p=0.006$), while the voided volume showed a similar tendency (11 ml higher when standing, $p=0.06$). Micturition time was similar. While sitting, two subjects had a mean voided volume > 60 ml lower than while standing, suggesting that they leave more residue after micturition when sitting. Such an enlarged residue is clinically relevant because it correlates with higher risk of recurrent urinary tract infections and acute urinary retention.
We acknowledge that there may be other valid reasons for men to sit on the toilet. These reasons are not within the domain of the general practitioner who acts as a professional who advises patients based on clinical evidence and learnedness. At the moment there is no hard evidence that supports the theory that sitting while urinating facilitates bladder emptying. Consequently, we should not give the *medical* advice to urinate this way.

objective reading of uroflowmetric measurements in order to increase reliability and precision
Chapter 5 shows that the correction of artefacts in flow measurements can be automated and standardized with simple means. Visual correction is simple, but inherently subjective. It is difficult to standardise and maintain standardised during a longer period. An algorithm incorporated in the uroflowmeter would overcome this. It will enhance the objectivity of data collection in clinical research and increase statistical power. It may also facilitate more efficient logistics in clinical practice. This study may be an example for other researchers, also those in primary care, to scrutinize their procedures.

formal quality criteria for health websites and quality of the content
In chapter 6 we conclude that the Internet contains many web sites with health information. With regular search terms and the most popular search engines one can retrieve millions of web sites regarding micturition problems in the elderly. Many are well documented and provide consumers/patients with good quality information. However, many websites are also focusing on a part of the

problem, e.g. medication, herbal or surgical solutions. Such selective information may be as misleading as incorrect information. Accreditation of web sites may be a good way to improve the quality of health information web sites. It insures minimum quality standards. But accreditation focuses on 'formal', procedural aspects of quality. Content quality aspects are not directly secured this way. It is no guarantee for completeness of the information provided. One can argue that the Internet should contain a 'corner' where selected, high-quality information is available to the public. A top-level domain such as *.health* could be such a virtual place. However, the Internet is also an (almost) free sphere, where freedom of speech, and the right to express and get acquainted with opinions and beliefs that dissent with the 'official' truths, are key values. In our open Western societies, the availability of *only* 'high-quality' websites may therefore be considered both a Utopian dream and a Utopian nightmare.

general practitioners, their patients and the Internet

Chapter 7 describes some practical views from both patients and general practitioners on health information in the media, especially the Internet. It shows the increase in information brought in by the patient, but - according to the general practitioners - also a change in the type of information. The information the patient brings in is more detailed, more specific, and more related to the actual complaints or diseases of the patient. This differs from the classic ('old') situation in which a newspaper article or television programme discusses general aspects of a certain disease that may, but also may not, apply to the specific situation of the patient. The Internet offers patients the possibility to select the information that is applicable to their individual situation from a large database of knowledge and experiences.

The result may be that the consultation reaches a 'higher level'. The consultation becomes more interesting because the general practitioner is more able to use his specific expertise and knowledge, instead of routinely conducting daily work. And patients can get more and more adequate information. This way, the Internet helps GPs to educate patients. However, both patients and doctors also seem to miss opportunities in the consultation room. They hardly discuss the 'background' information during the consultation. They should be more open discussing health information from the Internet.

Furthermore, patients use the Internet as a means to get a second opinion, or as a source used for 'recapitulation' of everything that has been said by the general practitioner during the consultation. As a result, the use of the Internet is often *after* the patient has visited the doctor, not before. General practitioners seem to be hardly aware of this phenomenon.

general view

the role of the general practitioner

The role of the general practitioner in the medical system varies with respect to the condition or type of condition at hand. In the context of lower urinary tract symptoms the general practitioner does not come to a defined medical diagnosis, but to a defined medical problem. The concept supposes that other diagnoses have to some extent been ruled out, leaving symptom and bother (scores) as the leading trails. The role of the general practitioner shifts from an interpreter who translates the complaints into an established diagnosis to an interpreter who translates the complaints into an established medical problem, naming a descriptive concept that leads to one or more *case-management strategies*. Such case-management strategies require not only knowledge on what is the matter (in a classical sense: the diagnosis, but here we should call it the medical problem at hand), and what can be done (in a classical sense: the treatment options), but also what *will* be done if certain steps are taken.

Investigation of case-management strategies - as was done in the studies described in chapter 2 and 3 - requires a comprehensive approach that includes the possibility for sequential therapeutic interventions and/or the weighing of possible benefits and risks of different therapeutic approaches.

Studies into such case-management strategies may support multidisciplinary guidelines and interdisciplinary consensus statements between primary care physicians and specialists.

the role of prostatic cancer

Prostatic cancer plays an important role in decisions concerning management of lower urinary tract symptoms in general practice. There are several valid reasons for mass individual screening for prostatic cancer in patients with LUTS. Some guidelines recommend periodic testing. The statistical truth about the correlation between LUTS and prostate cancer among scientists differs from the psychological truth about this relation among the lay public. Prostate cancer is an important cause of death among elderly men, and thereby much feared. Consequently, researchers in general practice cannot ignore this phenomenon when carrying out a pragmatic trial such as the one described in chapter 3, or developing a case-management strategy as described in chapter 2.

the evidence for non-pharmacological treatment

Many lifestyle advices are currently given in daily practice. Many more can be found in other sources of health information such as magazines and the Internet. The effects of such non-pharmacological interventions are unknown as

are the side effects. Most of them are not based on scientific evidence. Few have been investigated at all. This can partially be explained by a lack of financial 'drive' to investigate such lifestyle advices. Non-pharmacological interventions are (regarded) not profitable while medication is. One may assume that 'common knowledge' and 'reason' may also be good foundations for medical practice, but the replicated single case experiment showed that this is in fact not always justified.

pragmatic studies

The purpose of pragmatic studies is to generate results that are relevant and useful under usual care circumstances.

Such studies are may also be useful to determine the number of patients that are 'covered' by the (evidence-based results in) guidelines. Using the results of chapter 2 and 3 the number of patients in certain strategy-options can be calculated; for individuals the chance that a particular treatment is indeed carried out and will be effective can be determined (and be weighted and play a role in the individual's decision making process).

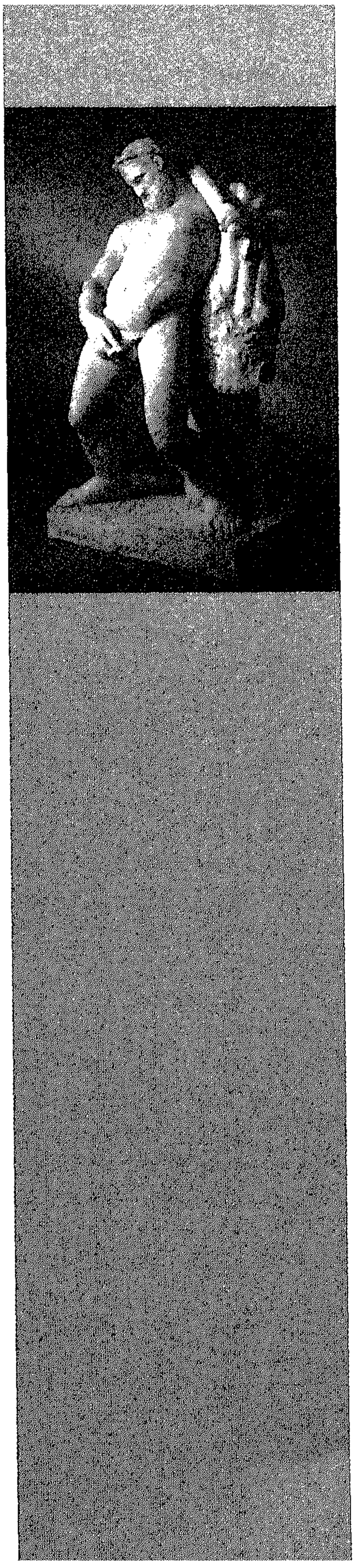
single case studies

Single case studies can be defined as experiments in which an entity is observed repeatedly during a certain period under different levels ('treatments', or 'interventions' - including the control condition) of at least one independent variable. The primary aim of a single case study is to prove or falsify a relation between two variables in a single subject.

General practitioners have the advantage of having a broader knowledge of aspects of their patients' lives. This means that seemingly coincidental co-occurring phenomena may be easier to recognise. Chronic, stable diseases are a core business of the general practitioner. Therefore, the general practitioner is in a good position to perform single case studies. It is certainly possible to perform explanatory (multiple) single case studies aiming to generate, confirm or falsify hypotheses and research questions.

ethics and research

Efficient use of resources, be these money, material or (especially) voluntarily participating people, should play a more important role, both in the reporting of studies and -earlier on- in the design. It should get more attention during the design (and conduct) of clinical studies, and - consequently - in the reviewing process of both medical journals and medical ethical committees. Implementation of simple measures to enhance the precision such as the algorithm described in chapter 5 should therefore be promoted.



Samenvatting

Samenvatting

de context van de onderzoeken in dit proefschrift

Voor de meeste mensen die hun luiers ontgroeid zijn, zijn mictie en continentie vanzelfsprekendheden geworden. Gedurende het leven verandert dit zeer geleidelijk. De kwaliteit van het urineren neemt af. Maar wanneer is dit niet meer normaal? Zowel arts als patiënt staat hier voor een dilemma. Bijkomend probleem is dat het urineren een privé-aangelegenheid is. Dit maakt het extra moeilijk om een referentiekader aan te geven en te bepalen wat 'normaal' is en wat niet. De definitie van 'ziekte' is derhalve niet goed omschreven.

medische modellen

Medische modellen hebben fysieke en/of psychische behandeldoelen nodig als een verklarend mechanisme. Voorheen werd door artsen gesproken over de 'ziekte' benigne prostaat hyperplasie (BPH) met bijbehorende klachten van prostatisme. Toen bleek dat er slechts een geringe correlatie bestaat tussen klachten en 'objectieve' parameters zoals uroflowmetrie, urodynamica en echografie, ontstond een abstractie van het anatomisch model: *klinische* BPH. Tegenwoordig wordt in het meest gebruikte medische model over de 'ziekte' gesproken als 'lower urinary tract symptoms' (LUTS). Daarmee bestaat het concept nog slechts uit een combinatie van klachten. Het diagnostisch traject is hiermee niet meer gericht op het stellen van een objectieve diagnose, maar is verplaatst naar de spreekkamer waar de arts de individuele patiënt voor zich heeft.

de rol van huisartsen

Huisartsen zijn specialisten in onzekerheids-geneeskunde. Met beperkte technische mogelijkheden proberen ze van de hulpvraag van de patiënt te komen tot een rationele strategie. Deze werkwijze is meer gebaseerd op de meest waarschijnlijk hypothese met uitsluiten van mogelijke ernstige condities, dan op een goed onderbouwde diagnose.

trends

In de afgelopen decennia zijn drie belangrijke trends van invloed geweest op de benadering van LUTS in de huisartsenpraktijk. Als gevolg hiervan is de gemiddelde man die zijn huisarts consulteert met deze klachten jonger, en minder symptomatisch. Huisartsen behandelen een nieuwe patiëntengroep met nieuwe stadia van de 'ziekte' met nieuwe medicatie in een nieuwe setting.

richtlijnen

De meeste richtlijnen voor de diagnostiek en behandeling van LUTS (in de huisartsenpraktijk) hebben drie elementen gemeen:

1. Een urineweginfectie en prostaatcarcinooom dienen uitgesloten te worden
2. Behandeling moet gericht zijn op het verminderen van (klachten veroorzaakt door) de symptomen;
3. Behandelopties moeten worden uitgelegd; de keuze voor een specifieke behandeling moet door de patiënt gemaakt worden.

Niet-medicamenteuze adviezen worden vaak gegeven. Zo beveelt de NHG-standaard onder meer aan om 'zittend te plassen'. Deze aanbeveling is niet evidence-based en er is nauwelijks onderzoek naar gedaan.

a decision aid for general practitioners for the treatment of elderly male patients with lower urinary tract symptoms

Hoofdstuk 2 beschrijft de ontwikkeling van checklists voor de behandeling van patiënten met LUTS in de eerste lijn en de verwijzing naar de tweede lijn. We bespreken de waarde van deze checklists en geven een indicatie van het aantal personen dat behandeld en verwezen zou worden bij het gebruik van zulke checklists. Deze aantallen waren zeer acceptabel: gebruik makend van de checklists kan de huisarts een rationele behandelstrategie volgen zonder excessief medicatie voor te schrijven en zonder onderbehandeling van patiënten. Centraal in dit hoofdstuk staat de expliciete keuze die arts en patiënt moeten maken ten aanzien van de volgorde van de diagnostische en therapeutische mogelijkheden. Zo kan men ervoor kiezen eerst het diagnostische traject volledig te bewandelen alvorens te starten met de therapie. Men kan er ook voor kiezen alvast met medicatie te starten in de veronderstelling dat als dat werkt verdere diagnostiek overbodig is.

Als huisartsen verwijzen willen ze hun patiënt deze mogelijk effectieve behandeling niet onthouden, maar tegelijk willen ze ook niet teveel verwijzen. Ze moeten selectief zijn, maar ook weer niet tè.

Veel richtlijnen bevatten lijstjes met aandachtspunten die bij deze beslissing kunnen worden betrokken, maar die zijn doorgaans niet allemaal even belangrijk. Weinig patiënten tonen alle kenmerken van de aandoening. Dat betekent dat bepaalde items uit dergelijke lijstjes tegen elkaar afgewogen moeten worden. Bij onze checklist zien we bijvoorbeeld dat de leeftijd een belangrijker item is bij de beslissing om wel of niet te opereren dan bij de keuze om al dan niet medicatie voor te schrijven. Het is wel van belang dergelijke vaak impliciet blijvende overwegingen te kunnen expliciteren, zodat zowel huisarts als uroloog daarvan kunnen profiteren.

the effectiveness of a treatment protocol for male lower urinary tract symptoms in general practice

Hoofdstuk 3 laat de resultaten zien van een randomised controlled trial naar een behandelprotocol voor LUTS. De belangrijkste conclusie is dat watchful waiting als aanpak bij de behandeling van plasklachten belangrijk blijft. In de interventie-arm van ons onderzoek kregen veel meer mensen medicatie dan in de usual-care arm. Toch bleek uiteindelijk nauwelijks een verschil zichtbaar op de uitkomstmaten.

100 deelnemers deden mee aan de interventie-arm van de studie. Deze begon met een recept tamsulosine voor drie maanden. Ook al deden allen bewust en vrijwillig mee aan de studie, toch wilden 13 personen geen medicatie. Blijkbaar vond zelfs in deze groep een substantieel aantal mannen medicatie 'teveel van het goede'. 11 personen stopten met de medicatie vanwege bijwerkingen. Dit is een aanzienlijk aantal, hoger dan verwacht. Blijkbaar is het de moeite waard om de effectiviteit van de behandeling in een vroeg stadium te herevalueren en te letten op mogelijke bijwerkingen (bijv. na één of twee weken).

Na de eerste 3 maanden was bij ongeveer de helft van de mannen voldoende verbetering in de symptoomscore opgetreden. Zij continueerden de medicatie voor nog zes maanden en stopten daarna voor 3 maanden met de medicatie. Hierna werd de symptoomscore weer gemeten. Een kwart van de personen liet geen verslechtering van de symptoomscore zien. Dit zou kunnen suggereren dat ofwel de medicatie helemaal geen effect heeft gehad (en de gemeten verbetering dus niet echt was), ofwel de medicatie de ziekte voor eens en altijd genezen heeft, ofwel dat de symptoomscore op zichzelf niet deugt, bijv. omdat de deelnemers de symptoomscore niet goed in konden vullen. Er zijn ook nog andere suggesties mogelijk. Wat de verklaring ook moge zijn, het is belangrijk om niet alleen het effect van α-blokkers te meten, maar ook te kijken of het een echt effect is dat langdurige voorschriften rechtvaardigt.

position-related differences in voiding in men with lower urinary tract symptoms

Hoofdstuk 4 laat zien dat zittend plassen over het algemeen de flow van de urinestraal niet verbetert. Zitten zou de blaaslediging vergemakkelijken, maar dit kon niet worden bevestigd. 20 mannen namen deel aan deze studie. Bij veertien personen was de gemiddelde maximale urineflow hoger bij staand plassen. Bij zes personen was deze lager. Het overall verschil in maximale flow tussen staan en zitten was 0.55 ml/s ten gunste van staan (p=0.02). De gemiddelde flow was eveneens 0.55 ml/s hoger (P=0.006), terwijl het mictievolume een zelfde tendens liet zien (11 ml hoger bij staande mictie, p=0.06). De mictietijd was hetzelfde.

Zittend hadden 2 personen een gemiddeld mictievolume >60 ml lager dan staand, suggererend dat er een groter residu-volume is na de mictie. Zulk een groot residu-volume is klinisch relevant omdat het een groter risico geeft op urineweginfectie en acute urineretentie.
We zijn ons ervan bewust dat er andere redenen kunnen zijn voor mannen om op het toilet te gaan zitten. Deze redenen liggen niet op het terrein van de huisarts die werkt als een professional die adviseert op basis van clinical evidence en empirie. Op dit moment is er geen hard bewijs dat de theorie ondersteunt dat zittend urineren de blaasontlediging bevordert. Aldus zouden we niet het *medisch* advies moeten geven op deze manier te plassen.

objective reading of uroflowmetric measurements in order to increase reliability and precision
Hoofdstuk 5 laat zien dat de correctie van artefacten in uroflowmetrie-metingen geautomatiseerd en gestandaardiseerd kan worden met eenvoudige hulpmiddelen. Visuele correctie is eenvoudig, maar subjectief. Een algoritme dat in de uroflowmeter wordt ingebouwd zou dit kunnen verhelpen. Het zal de objectiviteit van dataverzameling in klinisch onderzoek vergroten en de statische power versterken. Eveneens wordt een efficiëntere logistiek in de klinische praktijk mogelijk. Deze studie laat zien dat het zeer zinvol kan zijn om procedures nog eens onder de loep te nemen.

formal quality criteria for health websites and quality of the content
In hoofdstuk 6 concluderen we dat het internet veel websites met gezondheidsinformatie bevat. Met reguliere zoektermen en de meest populaire zoekmachines kan men miljoenen websites vinden over plasproblemen bij ouderen. Veel ervan zijn goed gedocumenteerd en geven consumenten of patiënten informatie van goede kwaliteit. Veel websites focussen echter op slechts een gedeelte van het probleem (of van de oplossing). Deze selectieve informatie kan misleidend zijn.
Accreditatie van websites zou een goede manier kunnen zijn om de kwaliteit van de informatie van websites te verbeteren. Het stelt minimale kwaliteitseisen. Maar de huidige accreditatie van websites is gericht op 'formele', procedurele aspecten van kwaliteit. De kwaliteit van de inhoud is niet altijd verzekerd op deze manier. Het zou goed zijn als het internet een 'forum' biedt dat geselecteerde, kwalitatief goede informatie voor het brede publiek toegankelijk maakt, bijvoorbeeld onder de domeinnaam *.health*. Het feit dat het internet een (vrijwel) vrij toegankelijk medium is, waar vrijheid van meningsuiting en het recht om meningen en overtuigingen te uiten en vernemen die afwijken van de 'officiële waarheid' basale waarden zijn, maakt dit onmogelijk.

general practitioners, their patients and the Internet

Hoofdstuk 7 beschrijft enkele praktische gezichtspunten van patiënten en huisartsen over gezondheidsinformatie in de media, met name het internet. Het laat zien dat de patiënt in toenemende mate informatie van het internet gebruikt. Er is een verandering in type informatie waarneembaar. De door een patiënt naar zijn arts meegebrachte informatie is meer gedetailleerd en meer toegespitst op de actuele toestand van de patiënt dan voorheen. Dit verschilt van de oude situatie waarbij in een krantenartikel of televisieprogramma algemene aspecten van een bepaalde ziekte aan de orde komen, die niet altijd volledig van toepassing zijn op de specifieke situatie van de patiënt. Het internet biedt de mogelijkheid om die informatie te selecteren die van toepassing is op de individuele situatie.

Het consult kan hierdoor nuttiger en leerzamer worden omdat de huisarts meer in staat is zijn kennis en expertise in te zetten in plaats van routinematig te handelen. De patiënt verkrijgt op deze manier meer en beter toegespitste informatie. Op deze manier helpt het internet de huisarts om de patiënt voor te lichten. Het blijkt echter dat zowel patiënten als huisartsen deze kansen wat weinig benutten. De achtergrondinformatie die via het internet verkregen is wordt nauwelijks besproken.

Daarnaast gebruiken patiënten het internet als *second opinion*, of om nog eens na te zoeken wat er nu allemaal gezegd is. Vaak wordt het internet geraadpleegd *na* het bezoek aan de dokter, en niet ervoor. De onderzochte huisartsen waren zich hiervan nauwelijks bewust.

slotbeschouwing

De rol van de huisarts varieert per aandoening. Voor wat betreft de LUTS gaat de klassieke rol van vertaler van een klacht naar een ziekteconcept niet goed op, omdat het ziekteconcept zelf bestaat uit een verzameling symptomen. Dat betekent dat de huisarts een andere rol krijgt, namelijk die van case-manager. Hij moet het beloop van de klacht bij de patiënt meer centraal stellen, daarbij in het oog houdend wat zijn (beperkte) mogelijkheden zijn en welke mogelijkheden urologen hebben. Op het beleid van deze laatsten moet hij vervolgens in zijn eigen beleid anticiperen.

Onderzoek naar dit soort case-management strategieën vergt een brede aanpak met aandacht voor opeenvolgende therapeutische interventies en verschillende therapeutische benaderingen.

de rol van prostaatkanker

Prostaatkanker speelt een belangrijke rol bij de besluitvorming rond LUTS. Er zijn valide redenen om te screenen en/of periodiek te testen op prostaatcarcinoom. Prostaatkanker is een belangrijke doodsoorzaak bij oudere mannen en wordt

gevreesd door de patiënt met mictieklachten. Hier moet aandacht aan worden besteed, ook al is er nauwelijks een statistische correlatie tussen LUTS en prostaatcarcinoom.

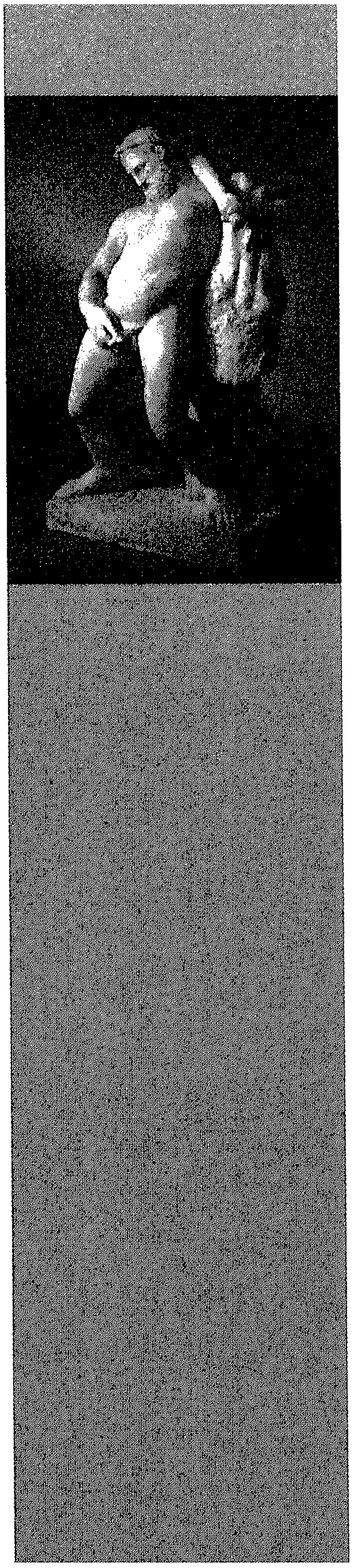
het bewijs voor niet-medicamenteuze interventies
In de dagelijkse praktijk worden veel lifestyle adviezen gegeven. Daarbij komen nog de adviezen die uit de media en van het internet gehaald worden. Er is slechts weinig onderzoek gedaan naar de werking en bijwerking van niet-medicaementeuze interventies. Deels komt dit waarschijnlijk door het ontbreken van een financiële stimulans voor dergelijk onderzoek. De single case studie laat zien dat adviezen die gebaseerd zijn op 'algemene kennis' en 'logica' niet altijd juist blijken te zijn.

pragmatische studies
Pragmatische studies streven naar resultaten die relevant zijn voor en toepasbaar zijn in de dagelijkse praktijk. Eveneens laten deze studies zien op hoeveel patiënten de (evidence based uitkomsten van) richtlijnen van toepassing zijn. Met de resultaten van hoofdstuk 2 en 3 kunnen patiëntenaantallen berekend worden voor bepaalde strategie-opties. Voor individuen kan de kans bepaald worden dat een therapie inderdaad wordt uitgevoerd en effectief zal zijn. Dit kan gebruikt worden bij de individuele besluitvorming.

single case studies
Single case studies kunnen worden gezien als experimenten waarbij herhaaldelijk en onder verschillende omstandigheden (behandelingen, interventies, maar ook controle) gedurende een bepaalde tijd naar het onderzoeksobject gekeken wordt. Het primaire doel is het aantonen of afwijzen van een relatie tussen twee variabelen bij één subject.
Huisartsen hebben het voordeel dat ze dikwijls een bredere kennis hebben over (het leven van) hun patiënt. Hierdoor kunnen ze gemakkelijker verbanden of co-factoren herkennen. Chronische, stabiele ziekten (het terrein van de huisarts) lenen zich bij uitstek voor single case studies. Het is mogelijk om met (multiple) single case studies hypotheses en research vragen te generaliseren, te bevestigen of af te wijzen.

ethiek en onderzoek
Een efficiënt gebruik van middelen, of het nu geld, materiaal of (vooral) vrijwillig deelnemende mensen betreft, zou veel meer aandacht moeten krijgen. Bij het ontwerpen en uitvoeren van klinische studies en zeker ook bij het beoordelingsproces van tijdschriften en ethische commissies zou dit meer

centraal moeten staan. De implementatie van simpele maatregelen om de precisie te vergroten zoals het algoritme beschreven in hoofdstuk 5 zou meer gestimuleerd moeten worden.



Woord van dank

Woord van dank

Laat ik beginnen met te zeggen dat mijn dank groot is. Ik ben blij dat ik zo geschapen en gevormd ben, en dat ik de kansen gekregen heb, dat dit mogelijk is geworden. Over het geschapen zijn is geen wetenschappelijk zinnig woord te zeggen, maar voor de vorming en het bieden van de kansen zijn vele mensen verantwoordelijk. Van al diegenen kan ik hier slechts een beperkt aantal noemen. Als ik ze voor mijzelf op een rijtje zet, dan lijkt zo'n proefschrift haast een inkopper.

Kijkend naar het universitaire traject, dan begint het in 1997. Ik begon aan een (verhoopt) AIOTHO-traject in een warm bad. Want als zodanig voelde het binnentreden in de capaciteitsgroep huisartsgeneeskunde aan. Geïnteresseerde collegae, deuren altijd open, zoniet letterlijk dan toch zeker figuurlijk. Stimulerende begeleiders, *Jelle Stoffers* en *Arno Hoes* (UMCU), in toen nog zowel Maastricht als Rotterdam, met wie ik heftige en uiterst leerzame discussies had over hoe het onderwerp hartfalen te onderzoeken; een kamergenoot, *Piet Portegijs*, die bereid was zijn eigen zienswijze op het werken aan een onderzoek te pas en te onpas kenbaar te maken. En heel eventjes deelde ik de computer en het bureau met een afzwaaiende promovendus, *George Wolfs*.

Een proefschrift ontstaat uit een plan. En dit proefschrift ziet er grotendeels anders uit dan eerder voorzien. In dat dynamische proces 'steken en bloeien de rozen'. Veel dank gaat uit naar degenen die met mij die weg zijn gegaan.

André Knottnerus, mijn eerste promotor. Je bent geestelijk vader van het project, die mee de pragmatische trial in de steigers zette, mijn sterke twijfels bij de haalbaarheid van mijn eigen idee voor de decision aid wegnam, en zo de intellectuele uitdagingen aanreikte en met me verkende. Blijf, zoals te weinigen, bij tijd en wijle de kamers binnengaan van de leden van de capaciteitsgroep, jouw medewerkers, en vragen hoe het met de mens achter het onderzoek gaat.

Onno van Schayck, mijn tweede promotor. Jij hebt de eerste, moeilijke startperiode begeleid. Van een promotieplan van twee promovendi op een drietal nauw verweven onderzoeken naar twee zelfstandige promoties op losse projecten. Dat viel niet mee, maar de stimulansen bleven, zowel inhoudelijk als in de aansturing. Je toonde mij het focus, maar stimuleerde me ook om met wetenschapsstudenten zijpaden te gaan. Met kritische vragen, waarvan ik veel geleerd heb, en die ik altijd gewaardeerd heb.

Piet Portegijs, co-promotor, luis in de pels, horzel, slak en uil ineen. Je hobby's leidden mij vaak over ongedachte wegen met parallellen die alleen jij voor volstrekt logisch houdt. Zoals die tussen een windmolen en de blaas; of die tussen een derdewereldwatervoorzieningsproject en een uroflowmeter. Jij liet me de polytome logistische regressieanalyses tot op het bot uitbenen, om me na enkele weken tot de conclusie te laten komen dat die niet de meest geschikte analysemethode voor hoofdstuk 2 waren (waarvan je vervolgens overigens nog moeilijk te overtuigen was). Je gaf ook de gouden tip die de last-minute herziening van het onderzoeksvoorstel voor de mictiehoudingstudie in gang zette. Wil je daar een volgende keer niet weer pas mee aankomen als het onderzoeksvoorstel een dag voor verzending naar ZonMw getekend en al klaar ligt? Dan heb ik iets meer tijd om me te verdiepen in onbekende methodologie; en hoeven we niet weer een powerberekening uit te voeren die niet deugt maar wel klopt... (achteraf gezien). Maar het meest dankbaar ben ik (nu) dat je ervoor gezorgd hebt dat ik doorgegaan ben, toen ik er eigenlijk mee had willen kappen.

Cees van de Beek, uroloog, trainer van de echografiste, mede-ontwerper van de trial en kritisch beoordelaar-vanuit-de-urologische-praktijk van de manuscripten en ideeën die vanaf de andere kant van de straat kwamen. Jouw inhoudelijke suggesties waren zeer welkom, alsmede de gulle lach waarmee die werden omgeven. En het was leuk om te zien hoe je naarmate het proefschrift meer vorm begon te krijgen enthousiaster werd over de inhoud ervan.

Voor de diverse studenten die als wetenschapsstagiaire een stukje van mijn pad meewandelden heb ik een bijzonder plekje in mijn herinneringen ingeruimd. *Maud Kertzman, Marcelle Wedemeijer, Tim van Dorst , Lucy de Bruin* als eersten. Later, toen ik het begeleiden ook beter in de vingers gekregen had, *Nadia el Bakkali*, die een belangrijke bijdrage leverde aan hoofdstuk 6, gepresenteerd op het congres van de European General Practitioners Research Network in Verona. *Angelique Bleuzé,* ik vergeet niet dat je op het congres van de EGPRN in Göttingen alom lof ontving voor jouw presentatie, die volgde uit het onderzoek van hoofdstuk 7. Ik vind het leuk om te merken dat jullie overwegen huisartsgeneeskunde én huisartsgeneeskundig onderzoek in jullie toekomstplannen een plaatsje te geven. Angelique, *Martijn Brakman*, onvermoeibaar meedenker over de praktische uitvoering van de mictiehoudingstudie. Door jou weet ik wat een scheidingstrafo is (zelfs Piet was daar niet opgekomen!).

Mijn collega-aiotho's van het eerste en latere uur: kamer- en lotgenoten. *Annemiek Nijholt, Katinka Prince, Loes van Bokhoven, Marjolein de Vries, Niels Chavannes, Patrick van Limpt, Paul Houben, Rogier Hopstaken* en *Wemke Veldhuijzen*. We hebben samen het pad van het promoveren en de huisartsopleiding moeten verkennen, gezamenlijk en ieder op de eigen manier. Met ons maandelijkse lotgenotencontact, waarin ideeën werden gesloopt en bekroond, concepten van artikelen publicabel gemaakt werden en vooral veel steun werd gegeven. We moeten onze jaarlijkse etentjes blijven voortzetten (net als de etentjes tussendoor om om het even welke andere 'goede' redenen).

Zonder assistentie kan geen promovendus werken. En ik ben gezegend geweest met een keur aan onderzoeksassistentes, in min of meer chronologische volgorde: *Karin Aretz, Marga van der Aa, Susanne Hanssen, Vinca Linssen* en *Paulien Versteeg.* Jullie werkten met veel inzet aan dataverzameling (echografie), databestanden en veel, heel veel vragenlijsten. Niet altijd was mijn waardering daarvoor zichtbaar, maar die was en is er wel.

Mark Spigt, met jou ben ik begonnen aan het project. We hebben goed gewerkt samen, tot het moment dat je koos voor een eigen weg, helaas.

Veldwerk in de huisartsgeneeskunde zonder huisartsen en praktijkassistentes die de deuren openen, letterlijk en figuurlijk, is onmogelijk. Veertien praktijken werkten mee. Op sommige momenten was ik kind aan huis; bij geen ging ik met tegenzin op bezoek. Dank daarom aan *J. Brull in Berg en Terblijt, P. Hulshof en R. Costongs in Voerendaal, V. Kaiser en J. Veldhuizen in Eijgelshoven, J. Naus en J. Mannens in Herkenbosch, P. Passage in Kerkrade, H. Reinink in Maastricht, P. Ruers, P. Winters, L. Ruers en S. Hobma in Sittard, O. Sikkes-Mekel in Geleen, A. Bokelman in Eindhoven, J. van Dongen te America, J. Hobbelen en W. van der Plas te Heerlen, M. Meindersma-Vluggen te Vaals, A. Schuivens te Blerick en A. Visschedijk en A. Massuger te Heel* en al hun praktijkassistentes.

Een AIOTHO-traject houdt vooraleer ook een opleiding tot huisarts in. Vaak moet die zich mee plooien aan de eisen van het onderzoek. Daarvoor is flexibiliteit en steun vanuit de huisartsopleiding onontbeerlijk.

Yvonne van Leeuwen, jouw warme enthousiasmerende en paden-effenende inzet als hoofd van de huisartsopleiding kan niet overschat worden. *Paul Ram*, in het grootste deel van de opleidingsperiode adjunct-hoofd, verantwoordelijk voor de uitvoering van het beleid, plooier van de soms knellende regeltjes: af en toe geen gemakkelijke taak! *Geert-Jan Dinant* als AIOTHO-mentor. Dat valt niet mee, met al die eigenwijze types!

Math Muris en *Maria Mulder*, mijn eerste groepsbegeleiders, groep 64. Dank voor het geschenk dat jullie mij tijdens mijn eerste TE-dagen gaven. Ik zet hem (in gedachten) nog geregeld op! *Stijn de Vries* en *Gerard van Mierlo*, begeleiders van mijn tweede opleidingsgroep in het staartje van mijn opleiding, als exoot in een nieuwe groep. Stijn, ik koos je niet voor niets als mijn huisarts.

Geertje van Zanten, ik was jouw eerste huisarts-in-opleiding. Jouw broddellapje. Bij mij moest je het opleiden nog leren, zei je zelf. Dus het was maar goed dat ik een AIOTHO was, die konden wat hebben. Volgens mijn latere opleiders lag er toch een stevig fundament na dat jaar. Ik schuif graag nog een keer bij jou, Alies en Charlotte aan tafel aan. Nu kan dat gelukkig zonder semafoon op zak.

Martin Link, opleider in de tweede periode. Jij wilde eigenlijk geen haio, maar bood me toch onderdak en leerde me mijn consultvoering aan te scherpen.

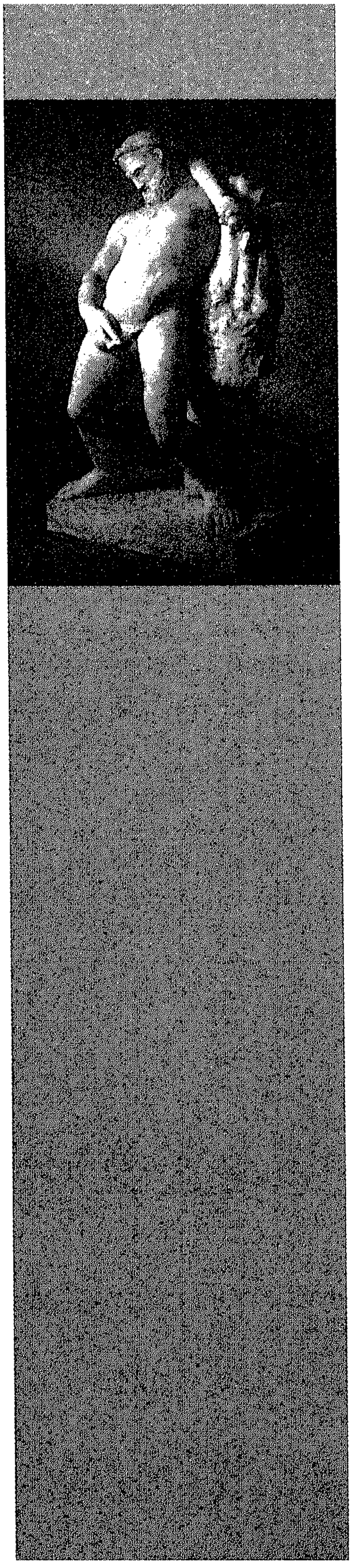
Anton Smulders en *Yvonne Smeets*, duo-opleiders in jaar 3. Geen diensten meer van huis uit, maar op een huisartsenpost. Anton, nu als jouw collega-bestuurslid van de STAK Huisartsenzorg Midden-Limburg zie ik wat je deed in de tijd dat ik op je wachtte voor de nabespreking. Ik bewonder je vermogen om zowel veel voor de huisartsgeneeskunde in zijn algemeenheid te doen, als dicht bij je patiënten te staan. Yvonne, ik ben erg blij dat we elkaar nu als collegae maandelijks blijven zien in onze toetsgroep.

Leden van Toetsie, ook jullie bijdrage aan mijn proefschrift kan niet zonder dank blijven. Ik waardeer het zeer dat mijn jeugdige, wetenschappelijke inslag zich mag voegen bij de mengeling van huisartsgeneeskundige kwaliteiten die in jullie zijn verenigd.

Oma Bimbam, hoe vaak hebt u niet de gaten in de agenda's van Caroline en mij gestopt?

Waar zou ik zijn zonder mijn vader en moeder, broers, mijn C'line, Hans, Lucas en Matthijs?

Die vraag beantwoord ik elders...



Curriculum vitae

The author

Roelf Johannes Cornelis Norg was born in Zutphen on December 10, 1969. He completed secondary school at the Baudartius College in Zutphen. In 1987 he started his study medicine at the Catholic University Nijmegen (at present: UMC St. Radboud). In 1990 he started to study law at the same university. Also in 1990 he won a prize of the Royal Dutch Medical Association (KNMG) for an article about ethics and transplantation medicine. In 1994 he graduated from Law School with as specialties administrative law, forensic psychiatry and health law; and in 1996 from Medical School.

Between July 1996 and October 1997 he worked at the Elkerliek Hospital in Helmond in the Internal Medicine, Cardiology and Pulmonology departments.

In November 1997 he started his work as reṡearcher in the Department of General Practice at the Universiteit Maastricht. In 1999 he started a combined vocational training and research programme. In the context of this programme the studies presented in this thesis were conducted.

In June 2003 he was registered as a General Practitioner. From January 2006 onwards he is working in a practice in Haelen, presently together with his wife Caroline Schulpen. They have three children: Hans (1999), Lucas (2001) and Matthijs (2004).

De auteur

Roelf Johannes Cornelis Norg werd geboren in Zutphen op 10 december 1969. Hij volgde het VWO-β aan het Baudartius College te Zutphen. In 1987 begon hij met de opleiding geneeskunde aan de Katholieke Universiteit Nijmegen (nu: UMC St. Radboud). In 1990 begon hij aan dezelfde universiteit met de opleiding Rechtsgeleerdheid. Ook in 1990 kreeg hij een prijs van de KNMG voor een artikel over ethiek en transplantatiegeneeskunde. In augustus 1993 behaalde hij het doctoraal examen geneeskunde; in mei 1994 het doctoraal examen rechtsgeleerdheid met als afstudeervakken bestuursrecht, forensische psychiatrie en gezondheidsrecht. In juli 1996 legde hij met goed gevolg het artsexamen af.

Daarna werkte hij tussen juli 1996 en oktober 1997 als arts interne geneeskunde, cardiologie en longziekten in het Elkerliek Ziekenhuis te Helmond. Aansluitend begon hij als wetenschappelijk onderzoeker aan de capaciteitsgroep huisartsgeneeskunde van de Universiteit Maastricht.

In januari 1999 startte hij daar met het zogeheten AIOTHO-traject, de opleiding tot huisarts en onderzoeker. Hierin werden de opleiding tot specialist-huisartsgeneeskunde en de wetenschappelijke opleiding tot doctor verenigd.

In juni 2003 werd de huisartsopleiding afgerond en volgde de registratie als huisarts.

In januari 2006 vestigde hij zich als huisarts in Haelen, alwaar hij nu in maatschapsverband werkt met zijn echtgenote, Caroline Schulpen. Samen hebben zij drie kinderen: Hans (1999). Lucas (2001) en Matthijs (2004).