Diversity among patients in medical practice: Challenges and implications for clinical research

Nicolien Wieringa, Anita Hardon,
Karien Stronks, Amâde M'charek (eds.)

UNIVERSITEIT VAN AMSTERDAM
ZonMw ondersteunt
ZonMw
am

This report was prepared by the Universiteit van Amsterdam, in a collaborative project of the Amsterdam School of Social science Research and the Academic Medical Centre. The project was commissioned by the Netherlands Organisation for Health Research and Development (ZonMw) in The Hague, The Netherlands.

Amsterdam
February 2005

ISBN 90-901965-1

Layout: Noor Oosterhof
Cover photograph: “Living together in diversity. Learning to accept (and respect) our differences”, Kyoto, Japan. Nicolien Wieringa © 2004

Preferred citation:
N.F. Wieringa, A.P Hardon, K. Stronks, A.A. M’charek (eds). Diversity among patients in medical practice: challenges and implications for clinical research. Amsterdam: Universiteit van Amsterdam, 2005.

Diversity among patients in medical practice:

Challenges and implications for clinical research

Contents

Diversity among patients in medical practice: main findings and recommendations for clinical research

Anita Hardon, Nicolien Wieringa, Karien Stronks, Amâde M'charek

Main findings

This report is the result of an exploratory project, commissioned by the Netherlands Organisation for Health Research and Development (ZonMw), to examine factors that facilitate and constrain a focus on "diversity" in clinical research. To what extent do individual traits and circumstances influence health outcomes, and are these sufficiently investigated in clinical research? The project specifically aimed at assessing:

1. Why clinical research needs to take diversity as point of departure;
2. What conceptual, practical, ethical and methodological constraints hamper an appropriate consideration of diversity in clinical research; and
3. Which novel strategies can be used to facilitate more systematic attention for diversity in clinical research?

Our approach was multidisciplinary, and involved clinicians, epidemiologists, ethicists, sociologists and anthropologists. Jointly we wrote six reviews, exploring these questions from different angles. One of our first tasks was to agree on a definition of clinical research. We defined clinical research broadly as exploratory research on the aetiology of diseases and on health perceptions, and observational (quantitative and qualitative) or experimental research on the diagnosis, treatment or prevention of diseases. The assumption underlying the project was that diversity by age, sex and ethnicity in health and health outcomes is not sufficiently acknowledged in clinical research. However, in the course of the project, we found that many other dimensions of diversity exist and might need to be considered for clinical research to be relevant to health and health outcomes in different populations and individuals.

Review 1[1], reflecting an epidemiological perspective, argues that clinical research needs to take diversity as point of departure. It shows how in many diseases the aetiology, prognosis, disease perception and/or effects of interventions are modified by age, sex/gender and/or ethnicity.

Review 2 uses an anthropological perspective to examine how diversity is dealt with in clinical practice. This enquiry reveals that diversity issues reach beyond the classifications of sex/gender, age and ethnicity. Many different kinds of diversity may be relevant to patients in every day medical practice, and their relevance may change over time. The review argues that the 'local knowledge' of patients and clinicians needs to be acknowledged and utilised, to learn more about diversity issues that matter.

Review 3 explores factors that constrain and facilitate attention to diversity in clinical research, and it focused mainly on the history of randomised controlled trials (RCTs). The review argues that RCTs, as a side-effect, have had a homogenizing influence on medical practice. The widespread adoption of RCTs as gold standard for clinical research has led to a paradigm shift from an individual difference approach to a biological reductionist point of departure that all humans are equal biologically unless and until differences can be demonstrated. An analysis of recent trials in a number of diseases, for which diversity issues are known to be relevant, revealed that sub-group analysis by age, sex of ethnicity is rarely done. The paper points to barriers that constrain a focus on diversity in clinical research, including the dominance of pharmaceutical industry sponsorship of RCTs. It makes proposals for methodological reform.

Review 4 takes an ethical perspective, and shows how ethical debates on clinical research have focused on patient protection, i.e. doing no harm and informed consent. It describes how ethicists nowadays acknowledge that unequal access to clinical trials by different population groups is also an important ethical issue. The review suggests that clinical researchers need to consider health benefits of participation in trials and confront exclusionary mechanisms.

Review 5, from a clinical epidemiology perspective discusses the methodological implications of focusing on diversity in RCTs. It specifically describes how RCTs can test differences in the outcomes of interventions between populations (effect modification in epidemiological terms), caused by biological and/or socio-cultural differences. Since present research practices generally aim at including homogeneous populations, many barriers are identified that may need to be addressed.

The report concludes *(Review 6)* by exploring novel strategies to deal with diversity in clinical research, including (1) new methods to capture

[1] More detailed summaries are given at the beginning of each of the reviews presented in this report.

dimensions of diversity that matter to patients and practitioners; (2) mechanisms for incorporating diversity issues into the research agenda; (3) ways of involving patients and health practitioners in all stages of clinical research; and (4) suggestions for broadening the scope of methods used in clinical research.

Final drafts of these six reviews were critically reviewed by peers and discussed during an international expert meeting, which was held in November 2004. Our aim was to bring together people from a variety of relevant backgrounds and expertise, to reflect on the findings of the reviews and generate a set of concrete recommendations (see appendix for the list of participants, and the agenda of the meeting).

The expert meeting generated insights on the inherent strengths and limitations of clinical research, for example concerning the tension between evidence generated from artificially composed groups in RCTs and evidence from observational research on individual patient needs, experiences and views. It was acknowledged that the diversity issues that matter to patients and clinicians are varied, and change over time and from one context to another, depending on characteristics of the patient, their lifestyles, and the state of his or her disease.

An important issue raised during the discussions was that diversity is insufficiently addressed in academic medical training and in medical standards, despite an increasing body of knowledge on the different ways in which factors such as sex, age, and ethnicity affect aetiology, diagnosis and treatment of disease.

The results of our reviews and the discussions at the expert meeting suggest that clinical researchers should take as point of departure that differences in diagnosis and/or the effects of treatment by sex, age and ethnicity can be caused by complex interplays of biological, socio-cultural, economic, behavioural and environmental factors. Thus, when differences between subgroups are found in RCTs, underlying causes of these differences need to be further explored (if these are not yet known).

The reviews and subsequent discussions identified many more challenges to the incorporation of diversity into clinical research. Firstly, it became clear that not only patients differ, but medical practices are diverse too. They change over time and from one health care context to another. *Review 1* for example describes a shift from diastolic to systolic blood pressure as diagnostic procedure for cardiovascular risk in the past 15 years, based on longitudinal studies which revealed that after the age of 60 diastolic blood pressure tends to decrease, while systolic blood pressure continues to rise. *Review 2* shows how diabetes II is dealt with differently in first-line general practitioners clinics and in out-patient departments of hospitals. General practitioners tend to base their diagnosis of diabetes on their knowledge of diversity in *populations at risk*. For example, a cluster of vague symptoms in a young woman of Moroccan

origin (considered at risk for diabetes) are a reason for GPs to conduct a test for diabetes II. A young white female with the same symptoms will not be tested for diabetes II. By contrast, in out-patient departments all patients have already been diagnosed with diabetes II, and here practitioners focus not on risk *populations* but on *individual* differences in lifestyle and living. They need to do this to achieve optimal treatment outcomes.

The reviews further suggest that a diversity of methods and ethical principles need to be considered. *Reviews 2* and *3* argue that a mix of qualitative and quantitative methods are needed in addition to the gold standard of RCTs in intervention research, to adequately address diversity issues that matter to patients and health practitioners. Then *Review 4* calls on ethical committees to consider new insights; not only the principles of doing no harm and informed consent, but also the rights of different populations to participate in trials, in order for them to benefit from the results of the trials. *Review 5* challenges clinicians conducting RCTs to consider the various ways in which they can contribute to a better understanding of effect modifications. And *Review 6* put forward a large variety of institutional reforms.

Our multidisciplinary team of researchers found this a challenging and rewarding project. From the start it was recognised that social scientists view diversity issues differently than clinical researchers. Social scientists acknowledge a wide range of socio-cultural, behavioural, economic, and contextual factors which determine health outcomes in everyday life. They rely on a variety of qualitative and quantitative methods to explore such diversity issues, using mainly observational designs. Clinical researchers tend to decontextualise the effects of treatments, and seek biological explanations for differences. They consider one experimental method, the RCT, to be the gold standard, and they aim at defining a limited number of endpoints to be measured, and a limited set of subgroups for analysis.

It was something of an achievement that the different disciplinary perspectives of the experts involved at the international meeting resulted in a single list of recommendations. There was unanimous acceptance of the proposition that a focus on diversity in clinical research should lead to better health outcomes. This consensus holds a promise for the future development of this work.

The biggest challenge is however yet to come: putting our recommendations into practice. Critically, mechanisms need to be developed to prioritise from the broad range of diversities that we have explored, diversities that matter in everyday lives of patients and are relevant to the medical practice of health professionals. Putting our recommendations into practice requires a collaborative effort between different actors in health care. As researchers we are committed to this process.

Recommendations

There is a broad consensus that health care practice should be evidence based. To implement evidence based medicine for all groups in society, and therefore to promote quality of health care, clinical research should take human diversity into account. The outcomes of this project underline that diversity is not sufficiently addressed in clinical research.

The following recommendations were formulated and agreed upon at the project's expert meeting in November 2004. We have formulated eight recommendations for changes in the way clinical research is conducted (Section I) and ten recommendations for institutional change (Section II).

Section I. Incorporating diversity into clinical research

I.1 Considerations of diversity in clinical research should take as point of departure that diversity in health and health outcomes can be caused by complex interplays of biological, socio-cultural, environmental, behavioural and health care factors. Age, sex and ethnicity alone are not sufficient categories or entries for research on diversity.

I.2 When differences by age, sex, ethnic origin or other dimensions of diversity are found through subgroup analysis in clinical research, researchers should be encouraged to further investigate possible underlying causes of the differences if these are not known, including an analysis of the role of biological, socio-cultural, behavioural, environmental and health care factors.

I.3 To address diversity in health and health outcomes hypothesis generating and hypothesis testing research is needed. Mechanisms and methods need to be developed to ensure that relevant diversity

hypotheses are generated, and then analysed in focused biomedical and socio-cultural studies (to explore underlying causes) and tested in RCTs (to determine effect modifications).

I.4 A mix of qualitative and quantitative methods, including for example observational cohort studies, Delphi methods, and ethnographic research, are needed to capture hypotheses on diversity issues that matter to patients in their daily lives and to health professionals in clinical practice.

I.5 Randomized controlled trials (RCTs) are important tools in studying benefits of medical interventions and effect modification. In the presence of hypothesized diversity, RCTs might be designed and used to specifically address the treatment effect in relevant subgroups, and/or to enlarge the original trial with members from other subgroups, such that subgroup analysis can be performed.

I.6 Innovations in RCT methodology are needed. In conducting trials, researchers should go beyond measurement of pre-defined endpoints, and explore unexpected phenomena and variance in effects, including potential outliers. Such observations within RCTs can lead to diversity hypotheses for further analysis and testing.

I.7 Diversity-sensitive parameters and end-points need to be included in RCTs, wherever possible and appropriate.

I.8 Methods for patient selection in clinical research need to be addressed to create opportunities for participation of diverse patient groups.

Section II. Institutional mechanisms and arrangements

II.1 Diversity issues need to be mainstreamed into all phases of funding programs for health research from commissioning to implementation.

II.2 Funding agencies should encourage multidisciplinary studies, using a combination of qualitative and quantitative methods, which aim at generating and analysing diversity relevant hypotheses.

II.3 Agencies funding research should ensure that relevant mechanisms and methods for substantial patient/consumer and health practitioner involvement exist in all stages of clinical research to ensure that experiential knowledge and clinical observations inform:

- the selection of priorities for diversity relevant research,
- the setting of objectives and diversity-sensitive endpoints for clinical studies,
- the preparation, and conduct of clinical studies
- the analysis of data,

- the formulation of conclusions, and
- the dissemination and implementation of the results.

II.4 An inventory of collaborative studies involving patients and health practitioners in the formulation and conduct of clinical research should be set up, as a basis for learning from the processes and results (see invo.org.uk database for results in the UK).

II.5 Funding agencies should review their project criteria and procedures to ensure that applicants for clinical research funds take account of the significance of diversity issues. Applicants should specifically be encouraged:
- to support their proposals with a concise, systematic review of the relevant evidence on diversity. Literature research methods have to be developed for such systematic reviews;
- to present research protocols and patient selection mechanisms that create opportunities for diverse population groups to participate in clinical research where relevant;
- to make explicit the ways in which patient/consumers and health practitioners will be involved in all stages of the clinical studies.

II.6 ZonMw should develop strategic alliances with the EU and other research agencies in The Netherlands and abroad to enhance implementation of this diversity agenda.

II.7 Agencies that appraise RCT results (Health Care Insurance Board, Drug Regulatory Agency, and organizations developing guidelines for medical practice) should require and systematically review subgroup analysis in RCTs.

II.8 Existing knowledge on diversity issues needs to be translated into educational curricula.

II.9 Further development of expertise on how to implement diversity issues into clinical research and into health care practice is essential.

II.10 ZonMw and other funding agencies should regularly monitor, evaluate and feedback outcomes of their efforts to promote this diversity research agenda.

Review 1

Diversity from an epidemiological perspective: looking for underlying causes and changing merits

Nicolien Wieringa, Menno Reijneveld, Karien Stronks

Contents

1

Diversity from an epidemiological perspective: looking for underlying causes and changing merits

Summary

This review addresses diversity from an epidemiological perspective. Six diseases were selected to study epidemiological differences between women and men, younger and older patients, and people of varying ethnic origin, and their underlying causes. The diseases included the attention-deficit hyperactive disorder (ADHD), asthma, gastric cancer, hypertension, HIV/AIDS and osteoporosis. Differences in aetiology, prognosis, disease presentation, perception and effect modification were studied as a function of sex/gender, age and ethnicity. Also, research implications concerning diversity issues were collected from the literature. Furthermore, we studied how diversity issues have shaped disease concepts over time. As a result of changing opinions about disease concepts, new epidemiological differences and effect modification become relevant to research and medical practice.

Differences in disease manifestation in relation to sex/gender, age and ethnicity were found for all diseases, caused by a large variation in underlying biological and socio-cultural mechanisms. Interactions result in complex differences in disease patterns. Effect modification of diagnostic, therapeutic or preventive interventions were documented in a number of cases, but more often gaps in knowledge were identified as relevant areas for the development of sex/gender, age or ethnic specific interventions.

The analysis of changes in disease concepts of ADHD and hypertension illustrate how opinions about diversity issues determine which populations receive which kinds of health care. The case studies also address the social context in which diagnoses function, for example in drug development and research. The relevance of studying diversity can therefore be expressed in medical, social and political terms.

How can these results contribute to change approaches to diversity in clinical research? First, it can be expected that diversity issues are relevant in many diseases. At all levels of research programming, assessment and conduct, expertise it is necessary to address these issues. Second, it can be expected that in areas where clinical insights are changing, diversity issues may emerge. Evaluation of such processes and underlying causes can provide relevant input to research. Third, the question can be raised whether the use of sex, age and ethnicity is the most 'natural' and effective categorisation to distinguish between groups of patients. A more nuanced understanding of differences and similarities between patients should address the underlying biomedical and socio-cultural mechanisms.

1.1 Introduction

"Women, especially those older than 65 years, delay longer than do men before seeking medical treatment for symptoms of an acute myocardial infarction. (...) Effective treatment is time dependent as mortality and morbidity rise with each hour of delay. (...) Three categories emerged to explain why women delay in seeking treatment: 1) clinical, 2) socio-demographic, and 3) psychosocial factors. These factors were found to be multifaceted and complex" (Lefler 2004). This example illustrates diversity in disease manifestation and its effects on outcomes of health care – and its underlying complexities. From an epidemiological perspective, diversity is a clinical relevant issue if differences between patient groups warrant differentiation in health care.

In this review, we use an epidemiological framework to study diversity issues. Therefore, we analyse sex, age and ethnicity-related differences in disease manifestation and effect modification in six common diseases. The aim of this analysis is twofold. First, we want to illustrate that many clinically relevant diversity issues are present. Second, since the goal of this project is to identify novel approaches to diversity in clinical research, we use the findings to discuss approaches to select relevant diversity issues for research programming.

The epidemiological framework applied in this review uses three perspectives of clinical relevance of diversity. The first refers to differences in aetiology and prognosis of diseases as function of sex, age and/or ethnic origin: if diseases differ in their causes and prognosis among various populations, it may be relevant to develop different approaches or interventions in health care. For example, the relatively high cardiovascular mortality of the Surinamese population in The Netherlands is thought to be related to a higher genetic susceptibility for and exposure to unhealthy lifestyles, as compared to the indigenous Dutch population (Bindraban 2003). Such differences are particularly relevant to define target populations for early detection of diseases, screening programmes and targeted preventive interventions.

The second way to define the clinical relevance of diversity is when outcomes of diagnostic procedures and of preventive and treatment interventions are modified by characteristics of people like age, sex/gender or ethnic origin. In epidemiological terms, this is referred to as effect modification. In general, this may have different causes, i.e. biomedical or socio-cultural. An example of effect modification is the higher susceptibility of women to torsades des pointes, which may cause a potentially life threatening side effect of QT-prolonging drugs such as quinidine or tricyclic antidepressant medications (Elming 2003).

In the third place, differences in health outcomes may be caused by differences in health perception and practices between men and women, elderly and younger patients and patients from various ethnic or cultural backgrounds. The above example of sex/gender differences in seeking medical care with myocardial infarction is well known. Important reasons for these sex differences are atypical presentation of symptoms in women, attribution or labelling of symptoms to the heart and perceived seriousness of the symptoms (Lefler 2004). Moreover, qualifying the symptoms of myocardial infarction in women as 'atypical' is in itself a reason for gender differentiation.

Increasingly, the relevance of diversity issues is being recognised in clinical practice and research. An important development in this respect is the endorsement of evidence-based medicine (EBM). This has fuelled the debate on the limited generalisability of findings in clinical research. Also, the patients' and consumers' movements have shown health problems in specific patient groups, in particular women, minorities/disadvantaged groups, elderly and children, to be neglected in clinical research and have documented sub-optimal care. The notion of clinical research as social process, with clinical methods and knowledge constantly being shaped and re-shaped, is increasingly acknowledged.

In this review, we selected six diseases to analyse epidemiological differences between men and women, younger and older persons and people of varying ethnic origin and their underlying causes. The analysis addresses the following questions: 1) How do aetiology, prognosis, disease presentation and perception differ as a function of sex/gender, age and ethnicity? 2) Which data on effect modification in diagnosis, prevention or treatment are present? 3) Which research implications concerning diversity issues of sex/gender, age and ethnicity are mentioned in the literature? In the second part of this review, we illustrate clinical research as social process by analysing how diversity issues have shaped disease concepts over time. As a result of changing opinions about disease concepts, new questions about epidemiological differences and effect modification emerge and become relevant to research and medical practice.

Using biomedical literature, we studied differences in disease manifestations of six common diseases in relation to the above patient characteristics and their consequences for diagnosis, treatment and

prevention (section 1.3.1). These outcomes are also used to identify gaps in clinical knowledge. In addition, two of the six diseases were selected to study in detail how knowledge on diversity issues has shaped the disease concepts over time (section 1.3.2).

1.2 Methodology

Literature search was applied to investigate the above questions. For this purpose Medline was used, therefore focussing this review on biomedical knowledge regarding diversity issues. Six diseases were selected for the analysis. These included various medical fields, i.e. immunology (asthma, HIV/AIDS), the cardiovascular system (hypertension), cancer (gastric cancer), behavioural problems (ADHD), and diseases specifically in the elderly (osteoporosis).

For the first analysis, English or Dutch literature was searched on the relationships between sex, age and ethnicity on the one hand, and differences in aetiology, prognosis, symptoms, incidence, prevalence, effect modification or differences in disease perceptions on the other hand. The search strategy is included in the appendix. A limitation was applied to include review articles only, therefore allowing the analysis to focus on aggregated data regarding diversity issues. Also, implications for clinical research on diversity issues were collected from the literature. On a number of issues no (review) data were found. If available, relevant original papers were collected. This method was applied in particular on disease perceptions, HIV/AIDS and gastric cancer. The website of the National Institute of Public Health and Environmental Protection (RIVM) was consulted for data on the incidence and prevalence of diseases in The Netherlands.

To study changes of diseases concepts in relation to sex/gender, age and ethnicity, two of the six diseases were selected of which diagnostic criteria are currently under debate. Additionally, the cases were chosen to include different medical domains, i.e. one which uses measurable biomedical parameters for diagnosis, as opposed to the use of opinion-based diagnostic tools. Based on these criteria the attention-deficit hyperactivity disorder (ADHD) and hypertension were selected.

For the case studies, Medline was used to identify English and Dutch literature on sex, age and ethnicity in relation to disease definitions and classifications. The search strategy is included in the appendix. Also, Dutch literature and opinion based appraisals (also ‘grey’ literature) of relevant issues were collected and analysed. Disease classification systems and practice guidelines are important sources of information to study disease definitions. For ADHD the US-based DSM classification system for psychiatric diseases was chosen, because this is used in many countries including the Netherlands. For hypertension, the Dutch guidelines for general practitioners and medical specialists were chosen. Since much of

the research in hypertension has been performed in the US, the US-based guidelines from the Joint National Committee on prevention, detection, evaluation and treatment of high blood pressure (JNC) were also included in the analysis.

The terminology which is used in The Netherlands to indicate a person's ethnic background often differs from the terms used in literature originating from other countries. In this review we have chosen to follow the terminology used in the original source of literature. Also, we use the term sex/gender. Sex refers to biological differences, whereas the term gender denotes the social construction of sexual identities and positions. Since many of the diversity aspects under consideration in this review combine biological and social meanings, we prefer to use the term sex/gender together, unless specific reference is made to biological sex differences.

1.3 Results

1.3.1 Variability in diseases and effect modification

Data on sex, age and ethnicity as determinants of disease manifestation and effect modification of ADHD, asthma, gastric cancer, HIV/AIDS, hypertension and osteoporosis are presented, and research implications to focus on diversity issues as reported in the literature.

ADHD

(Gingerich 1998; Kooij 2003; Linden 2004; Milich 2001; Spencer 1998; Spencer 2002)

Disease manifestation
ADHD is thought to be caused by a complex combination of environmental, genetic and biological factors. More boys than girls are diagnosed with ADHD: the prevalence in The Netherlands is 2.0/1000 boys and 0.3/1000 girls. The boy-girl ratios vary from 2:1 in community populations, to 9:1 in clinically referred populations, indicating a gender-based referral bias. This bias is thought to originate in sex/gender differences in the symptoms of ADHD: boys present more externalising problem behaviour (aggression), while girls are socialised to internalise problems (anxiety, low self esteem). Because of the central role of hyperactivity symptoms in the diagnosis of ADHD, there may be under-diagnosis in girls.

Boys are more prevalent in all subtypes of ADHD, but even more in the combined subtype, whereas girls are more likely to have the inattentive subtype. As a result of sex/gender differences in symptoms, boys are more likely to be referred to special education and may be more prone to criminal behaviour than girls, but the latter experience more learning problems. These findings indicate sex/gender differences in presentation and prognosis of ADHD with implications for diagnosis and treatment.

Cultural differences influence normative issues of child behaviour, terminology, diagnostic criteria and assessment methodology and were found to interact with ethnic differences. There is a lack of research regarding cultural or ethnic differences in presentation and diagnosis, but the scarce data do not indicate that the subtypes of ADHD have different relative prevalence's in various cultural or ethnic groups. Because opinions about normal/abnormal child behaviour and societal tolerance are culture specific, diagnostic criteria for ADHD are not likely to be globally uniform.

ADHD is now recognised to persist into adulthood. Symptoms of hyperactivity and impulsivity may decay, but inattention tends to persist. Deficits in cognition, executive function, and difficulties with organisation and time management may be present. As a result of ADHD, adults may encounter educational and social problems, limiting their abilities for work. If the diagnosis is made at adult age, those involved are likely to have experienced serious functional problems.

Effect modification
Effect modification of drug treatment for ADHD has not been documented in the literature, but studies generally included white boys only.

Research implications
These findings indicate that sex/gender, age at diagnosis and cultural or ethnic differences are important variables for the diagnosis and prognosis of ADHD. Development of diagnostic criteria is necessary that take into account sex/gender, age and cultural variations, either in the nature of the criteria or the thresholds in their application. Research is also needed to develop health care for children and adults with ADHD from a lifespan perspective, addressing various age, sex and culture specific aspects of the disorder. Concerning prevention, important aspects refer to identification of ADHD at young age and adequate care at school and at home. There is a need to study possible effect modification of drug treatment on subtypes of ADHD in various populations. Also, long term safety data, such as growth inhibition and sex specific side effects, have not been studied sufficiently.

Asthma

(Balzano 2001; Becklake 1999; Caracta 2003; Lemanske 2002; Lieu 1997; NHG standaard Astma 2004; Osman 2003; Poos 2003; Renwick 1999; Shek 2000; Spahn 2004; Von Hertzen 2004; Von Mutius 2001; Vrieze 2003)

Disease manifestation
Asthma is a chronic inflammatory disease of the airways, caused by a complex interaction of cells, mediators and cytokines. Overall, asthma is higher prevalent in women than in men, but the pattern changes over age. Childhood asthma is more frequent in boys, because they have a relatively slow airway development with growth of lung volume, as compared to girls. In young children, two wheezing phenotypes can be found. Wheezing without atopy or family history of asthma suggests an aetiology of

functional developments (relatively often seen in young boys), as opposed to children with family history of asthma, eczema and immunological reactions. The prognosis of the first group is better, the latter predisposes to recurrent asthma in adulthood.

The prevalence of asthma changes in teenagers. At the age of 20, there is female preponderance for atopy, increased bronchial responsiveness. Asthma starting during adulthood is more severe than childhood-onset asthma. When the disease starts at menopause or in old age, it generally has a severe nature. With hospital admissions as indicator for severity, the female-male ratio is 3:1 in patients aged between 20-50 year, and 2.5:1 in those over 50 years. Overall, different factors contribute to the aetiology of asthma in men and women, and the prognosis of asthma in women is worse than in men.

Levels of female sex hormones correlate directly with clinical and functional features of asthma and affect several immunological mechanisms; compared to men women present more pronounced immune responses and higher reactions to auto-antigens. Thirty to 40% of women with asthma report perimenstrual worsening of symptoms. Also, women are more susceptible to the effects of tobacco smoke than men are. Breathlessness may be perceived more sensitive but less specific in women, as compared to men, because airway function is subject to cyclical hormonal variations. Dyspnoe, a key element in quality of life scales, may be perceived by women as a more global indicator of health than by men. Gender differences in reporting of sputum production and swallowing phlegm relate to cultural factors.

Asthma and atopic conditions, such as hay fever, increase largely after migration. This is the case for migration from the tropics to a moderate climate, but also the other way around. A plausible explanation is that the genetic composition is appropriate in the original environment, but a new one may offer very different immunological exposures. Additionally, there is evidence that the number and relative importance of asthma-susceptibility genes varies between ethnic groups, resulting in different genetic risks for asthma and atopy. Whether this variation is reflected in differences in the prognosis, in addition to age and sex differences, is unknown.

Effect modification

The FEV1 (forced expiratory volume in one second) is generally used in the diagnosis of asthma. In young children and elderly, effect modification of this diagnostic tool may occur. FEV1-values of children with asthma are likely to fall in the normal range and the diagnosis is either missed or the severity underestimated. In young children, other diagnostic procedures may express airway lability better, for example FEV1 following pre- and post-bronchodilator challenge tests. In elderly, asthma tends to present non-specific and may be confused with chronic obstructive pulmonary disease. Demonstrating reversibility of the obstruction is required, but

physicians were found to be reluctant to perform these tests in elderly, especially in the case of co-morbidity. The physical signs associated with acute asthma are diminished with increasing age, which may contribute to underestimation of disease severity.

At all ages, inhaled corticosteroids are the most important treatment, but it is unclear at what age to start therapy. In elderly, the function of ß-receptors diminishes, therefore reducing the sensitivity to ß-2 agonists. Because of this effect modification, other inhaled bronchodilators may be first choice in elderly. Elderly people also tend to have more problems in correct use of inhalators in asthma, of which few have been formally assessed in elderly.

Research implications

Stratification by sex and age is necessary in descriptive and aetiological studies of the occurrence, risk factors and natural history of asthma. In studies of adolescents, gender differences will depend on the relative proportions of children who have reached puberty.

For diagnostic tools, an important issue is how to standardise for sex and age in comparisons across various groups because of age and sex-related differences in lung and airway size, and perception of symptoms. There is a need for research of perimenstrual asthma, including development of criteria, validated symptom scores, treatment and possible effect modification.

Many therapeutic questions in children remain, for example concerning long-term side effects of inhaled corticosteroids (inhibition of lung growth and bone development). It is also unclear whether long-acting ß-2 agonists are safe and effective in children under the age of four. Development of clinical tools of symptom perception in children is important for effective asthma management.

Clinical trials of drugs in asthma should be analysed in a gender specific manner. For women, information on their reproductive history and current status (pre- or postmenopausal and whether or not on oral contraception or hormonal replacement therapy) should be taken into account in the evaluation of treatment outcomes.

Gastric cancer

(Gill 2003; Roder 2002; Wijnhoven 2002; Yao 2002; Yao 2003)

Disease manifestation

Gastric cancer is the second leading cause of cancer deaths worldwide. The location of gastric cancer varies by sex and ethnic origin. Women develop gastric cancer mostly in the lower and smaller part of stomach, whereas the highest proportion of men develops it in the upper (proximal) part of the stomach. Proximal gastric tumours are thought to have a distinct aetiology from those in lower parts of the stomach. Symptoms of

gastric cancer are related to its location, dysphagia occurs predominantly in patients with proximal cancer, while abdominal pain, nausea/vomiting and early satiety are associated with non-proximal cancer. The prognosis of gastric cancer was found to be slightly better in women, probably associated to sex-differences in location.

The prevalence of gastric cancer varies between countries, the highest rates are found in Asia. The aetiology of gastric cancer is associated with dietary habits; in particular high salt intake is a risk factor. Asian populations have low proportions of proximal gastric tumours. Prognosis based on survival outcomes differs considerably and was found lower in Western than in Asian countries, likely to be associated with differences in tumour location. Age-related differences at diagnosis were found to interact with ethnic differences: Asian patients were younger at diagnosis.

Effect modification

Concerning possible effect modification in the treatment of gastric cancer, various factors appear to interact and no clear picture exists. In particular, ethnicity-related differences in survival after gastrectomy were largely attributable to differences in tumour location. Therapeutic approaches remain aggressive for potentially curable gastric tumours, irrespective of the ethnic background of the patient.

Research implications

Treatment studies using survival as outcome measure need to stratify by sex and ethnicity as a result of the variation in prognosis in relation to localisation of the tumour and underlying differences in aetiology.

HIV/AIDS

(De Wolf 2003; Gibb 2003; Gilad 2003; Haks 2002; London 2000; Loutfy 2004; Luzuriaga 2004; Seal 2004; Smith 2003)

Disease manifestation

The incidence of HIV/AIDS is higher in men than in women, 1.01 vs. 0.17/100,000 in The Netherlands. In many countries however, infection rates among women are rapidly growing. Women have a greater susceptibility to HIV infection than men do. Transmission routes and symptoms of AIDS are partly sex-specific. Transmission in women is predominantly heterosexually and a small proportion through intravenous drug use, as compared to more homosexual transmission in men and a higher proportion through drug use. Sex-specific HIV symptoms concern the viral load, that is app. 50% lower in women during a time in which CD4+ cell count is relatively preserved. AIDS symptoms specific for women are gynaecological neoplasms, high prevalence of other STD's, and pelvic inflammatory disease. Women are at higher risk for complications: sexually transmitted diseases (STD's), toxoplasmosis, herpes simplex, dental/oral lesions; and at lower risk for Karposi sarcoma and Epstein-Barr infection.

The prognosis of HIV/AIDS is similar between men and women, but age and co-morbidity related with implications for diagnosis and prevention. HIV infection at higher age progresses sooner into AIDS. It develops in men generally at a higher age, 30-45 years, as compared to women. Young women in their 20 to 30s are especially at risk for infection and develop AIDS between 25 and 40 years of age. Pregnancy does not seem to influence the rate of progression to AIDS and prognosis. Children can be infected with HIV during pregnancy, but chances are higher during birth and breast feeding. Infection rates depend on the mother's viral load and are lower with treatment. Progression of HIV into AIDS is higher in children than in young adults, and the disease is more aggressive.

In The Netherlands, a higher proportion of HIV infected women than men is migrant (78% vs. 47%), originating from HIV-endemic countries in the Sub-Sahara region. Cultural differences account largely for differences in transmission routes. Ethnic differences appear to be greatest during the early stages of infection: lower viral load and lower CD4+ cell count, but comparable disease progression were found in predominantly male population of Blacks, as compared to Caucasians. Possible explanations for these findings may concern the test essays used and both viral and host factors (infection with different subtypes of HIV; genetic variation in chemokine production and other immune responses). Present data show no differences in progression rates between ethnic groups. Knowledge about HIV/AIDS, transmission routes and risk perception vary among populations and are associated with educational levels, literacy, cultural values, reproductive norms and behaviour.

Effect modification

Age differences in disease progression have resulted in paediatric treatment guidelines. For adults, treatment recommendations are generally based on plasma viral load and standardized with data primarily from white males. According to current guidelines, sex and ethnicity-related differences in AIDS risks may lead to less eligibility of women and non-whites for treatment after seroconversion.

Effect modification was found in the toxicity of antiretroviral drugs. In the Netherlands, women switched antiretroviral drug regimens more often than men because of toxicity. Drug use during pregnancy poses specific risks, but limited data are available of toxicity. Hepatic and pancreatic toxicity, exacerbation of pregnancy-associated hyperglycaemia and developing drug resistance were found in pregnant women. In children, mitochondrial dysfunction has been observed in relatively large numbers. Prevention strategies need to be specified for various populations and take into account different perceptions, educational levels and cultural differences in reproductive norms and behaviour.

Research implications

The development of sex and ethnic specific treatment guidelines is recommended, based on CD4+ cell counts in stead of viral load. Also, treatment effects (and possible effect modification) need to be studied

according to age, sex, and ethnic specific standards. Women may achieve a faster and more durable response to treatment compared with men, but this finding needs further research.

Many research questions remain for drug treatment in children, in particular regarding pharmacokinetics, appropriate formulations and long-term studies of side effects, developmental problems and effectiveness. Potential short- and long-term complications of treatment for mother and child during pregnancy are important topics for continued research. It is recommended that all children exposed to antiretroviral agents *in utero* are followed in a registry to help answer relevant questions.

Sex, age and ethnic-specific outcome measures of disease progression are relevant for application in research and practice. An increasing number of non-Dutch persons is treated with highly active antiretroviral therapy (HAART), but with higher failure rates. Therapeutic monitoring and research into adherence in these specific patient groups is relevant.

Hypertension

(Asmar 2003; Bindraban 2003; Brewster 2004; Brondolo 2003; Franklin 1999; Franklin 2002; Jamerson 2004; Kannel 2003; Leest 2003; Linden 2004; NIH 2004; Safar 2004)

Disease manifestation

High blood pressure is one of the risk factors for cardiovascular disease. The prevalence of hypertension in all ages groups is 43.6/1000 men and 70.4/1000 women in The Netherlands. Prevalence increases at older age, and interacts with sex: at age 20-60 hypertension is present in 24% of men and 19% of women. Above 65, more females than males are hypertensive, 42% versus 38%.

Sex-related causes of hypertension include eclampsia during pregnancy, and the use of oral contraceptives. With respect to the aetiology of essential hypertension, it is speculated that before menopause the levels of female sex hormones result in the lower prevalence of hypertension as compared to age-matched men. Due to differences in body size and faster heart rates, women have a lower aortic blood pressure at all ages as compared to men. After menopause, arterial stiffening differs from that in men. Women experience higher levels of systolic blood pressure with ageing and white-coat hypertension is higher prevalent in women than in men.

The nature of hypertension changes with increasing age. Up to 50 years, the systolic and the diastolic blood pressure (SBP, DBP) track together; after 60 SBP continues to rise and DBP decreases. The pulse pressure, defined as the difference between SBP and DBP, rises in particular in women, leading to a high prevalence of isolated systolic hypertension (ISH) in elderly women. Large-artery stiffening and changes in wave reflection account for these changes. The prognosis of hypertension is worse in those with the highest pulse pressure. These aetiological differences result

in sex-related differences in prognosis: women develop cardiovascular events generally at ten year older age as compared to men.

People from African descent develop hypertension at younger age. It is often more severe, more resistant to treatment and more likely to be fatal at younger age, than in other ethnic groups. Ethnic differences in the aetiology of hypertension are found in salt sensitivity and plasma renin activity, calcium regulation of sodium transport and vascular reactivity to stress. Social stress, in particular racism, has been hypothesised to account for some of the higher prevalence of hypertension in Blacks in the US. Asians have smaller body stature than Europeans causing differences in blood pressure and hypertension. The higher cardiovascular mortality among people originating from South-Asia and the Caribbean, including Surinamese, is linked to the higher genetic susceptibility for cardiovascular risk factors (hypertension, diabetes, obesity, hypertriglyceridemia) and unfavourable lifestyles. Thus, ethnicity and age interact in cross-cultural variation in hypertension.

In sum, the aetiology and prognosis of hypertension are a function sex, age and ethnicity, with black patients particularly at risk for severe hypertension.

Effect modification

The aim of antihypertensive treatment is prevention of cardiovascular disease. Benefits of treatment in women are seen primarily in the prevention of strokes, whereas in men it prevents coronary events and stroke. Lifestyle changes, such as low salt intake, have different effects among various ethnic groups. Due to age-related changes in hypertension, side effect profiles in elderly are likely to differ from those in younger patients.

Black persons may respond different to specific antihypertensive drugs, in comparison to other ethnic groups. Effect modification, in terms of diminished efficacy in black patients, was found for ß-blockers and ACE-inhibitors. Calcium-channel blockers were the only drug type that effectively lowered blood pressure across subgroups of black patients with varying severity of hypertension. However, when pressure control is achieved, there is no evidence that morbidity and mortality outcomes depend on the drug chosen for initial therapy.

Research implications

At present, a shift is taking place in the treatment of hypertension and clinical testing of antihypertensive drugs from DBP as primary outcome measure toward SBP. Also, the pulse pressure has potentially important implications for the approach to treatment and prevention of cardiovascular diseases, and subsequently research. Most of the uncontrolled hypertension is in elderly with ISH. Research for optimal drug choices in various patient groups is necessary, and strategies for implementation of the new standards for treatment. To study possible effect modification of

antihypertensive drugs, stratification of patient groups by age, sex and ethnic origin is needed.

Osteoporosis

(Ebeling 1998; Jackson 2001; Linden 2004; Meadows 2004; Melton 2001; NIH 2001; Seeman 2002)

Disease manifestation

Osteoporosis is characterised by a reduction of bone density which is associated with skeletal fragility and an increased fracture risk following minor trauma. In the Netherlands, the overall prevalence is 7.3/1000 for females and 1.0/1000 for men. In all age groups, the incidence and prevalence are higher in women than in men, ranging from a factor 2 in the age group 25-44, to a factor 5 in persons older than 75 years.

Sex differences in the aetiology of osteoporosis are present. Women start with a smaller skeleton at peak in the third decade of life. Trabecular bone loss leads to loss of connectivity in bone microarchitecture in women, whereas in men the structure becomes thinner. Also, women have less cross sectional areas of periosteal bone than men. Bone loss accelerates in old age because the reduced mineralised mass of bone is subject to the same or larger volume being removed. Consequently, structural damage and fragility increase out of proportion to the reduction of bone mass. As a result, a higher proportion of (elderly) women than men have bone size and volumetric bone mineral density reduced to below a critical level.

Suboptimal bone growth in childhood and adolescence is as important as later bone loss in the development of osteoporosis. The long-term effects (prognosis) on bone health of risks at young age are unknown. These risks include: premature and low birth weight, use of corticosteroids, conditions associated with malabsorption and malnutrition (anorexia nervosa), and hypogonadal states.

Primary and secondary causes of osteoporosis differ between men and women. Among men, 30-60% of the cases are associated with secondary causes (hypogonadism, use of corticosteroids, alcoholism); in peri-menstrual women this figure is 50% (too low oestrogen levels, use of corticosteroids, thyroid hormone excess, anticonvulsant therapy). In post-menopausal women oestrogen deficiency is the leading cause of osteoporosis.

At puberty, the increase of trabecular bone mineral density is similar in men and women of the same ethnic origin, but is greater in African American than in white populations. Black people have thicker trabeculae, but comparable tissue density. Cultural variation in dietary habits and calcium consumption may limit attainment of optimal bone mass. Additionally, lactose maldigestion is highly prevalent in Asians, African Americans and Native Americans and is associated with low calcium intake. The prevalence of osteoporosis is increased among symptomatic

lactose maldigesters, which may have implications for preventive strategies. These various biological and environmental differences lead to the probability that a 50-year-old will have a hip fracture during life is 14% for a white women and 5-6% for a white man. The risk for African Americans is lower, 6% for women and 3% for men aged 50 years. In the US, white postmenopausal women experience three quarter of all hip fractures and have the highest age-adjusted incidence of fracture.
Hip and vertebral fractures are a problem for women in their late 70s and 80s; wrist fractures in their late 50s to early 70s, all other fractures (pelvis and rib) are a problem throughout menopause. Fear, anxiety and depression are often reported in women with osteoporosis. Little data exist on relationships between fractures and psychological and social well-being, either in men or women.

Effect modification
The WHO defines osteoporosis as bone density 2,5 SD below the mean for young white adult women. Diagnostic criteria have not been developed for various age groups, men and persons of various ethnic origins, despite age, sex and ethnicity-related differences in bone development and aetiology of osteoporosis.

Treatment and prevention strategies for osteoporosis are partly age, sex and ethnic specific, and should address specific hormonal and digestive issues. Drug treatment in men and young adults with secondary causes of osteoporosis has been poorly studied and data on possible effect modification are lacking.

Research implications
Although osteoporosis affects mainly women because of differences in aetiology, it is recognised that more research is necessary to address issues of prevention, diagnosis and treatment from a male perspective, in addition to addressing specific age and ethnic factors. Age, sex and ethnic specific diagnostic criteria for osteoporosis need to be developed. In disorders that impede bone development, interventions to maximize peak bone mass in girls and boys and across ethnic groups need to be defined. The quality of life is significantly impaired in osteoporosis. There is a need to characterise and validate quality of life tools in patients across sex, age and ethnicity to identify connected health problems.

Evidence linking lactose maldigestion and decreased calcium intakes to the aetiology of osteoporosis includes data from small studies that were conducted in Caucasian populations. Research on ethnically diverse populations is necessary to better understand how this condition influences the risk for osteoporosis and to develop adequate preventive and treatment strategies that take into account cultural dietary patterns.

1.3.2 Diversity issues shaping disease concepts

In order to address the health problems of a population, it is necessary to relate epidemiological patterns to medical supply. The classification of diseases and diagnostic tools form important policy instruments to identify a population's health needs and differences among various subpopulations. Therefore, it is interesting to analyse how disease concepts have developed in relation to biomedical knowledge on diversity and its consequences for health care.

ADHD and hypertension contrast on an important feature. By definition, ADHD is an opinion-based diagnosis, because there is no objective biomedical test or measurement. Blood pressure on the other hand, is a biomedical parameter which can be specifically measured and professional definitions are subsequently used to define hypertension. In both diagnoses, sex/gender, age and ethnicity are important categories of diversity. In the case studies on ADHD and hypertension we will analyse how these categories function to shape diagnostic criteria and professional opinions. Which biomedical knowledge about diversity is relevant and why?

Changes in disease concepts of ADHD

At present, a central question in the diagnosis of ADHD refers to whether hyperactivity, impulsivity and inattentiveness belong to the same or distinct disorders (Milich 2001). How this question is relevant to diversity issues will be analysed using the US-based DSM criteria for psychiatric diseases. Also, we analyse the current changes in opinion concerning ADHD in adults and in various ethnic groups. The analysis thus illustrates how disease concepts function to define who receive which type of health care.

Development of the disease concept
Minimal brain damage and the hyperkinetic syndrome were introduced in DSM-II in 1968, both on the same symptoms. In 1980, the term attention deficit disorder (ADD) was introduced. Also, the disorder was subdivided to distinguish individuals with hyperactivity, from those without hyperactivity. Diagnostic criteria were included to focus on impulsivity and inattention. These inclusions can be seen as a broadening of the diagnosis. In 1987, the term ADHD was introduced. It was conceptualised as a unidimensional category without subtypes. DSM-IV, published in 1994, introduced the present subdivision known as ADHD/predominantly hyperactive-impulsive type (ADHD/HI), ADHD/predominantly inattentive type (ADHD/I) and ADHD/C, which is a combination of the former two. The combined subtype is the most prevalent, and is often referred to when using the term ADHD (Pieters 2002).

Consequences for sex/gender differences in prevalence
How can we understand the above changes in relation to the small, but growing body of evidence about differences in symptom presentation in girls and boys?

In Western industrialised countries, boys and girls socialise differently. Hyperactive traits in girls are socially less accepted than in boys, causing girls to exert greater internalising behaviours, depression and anxiety disorders (Gingerich 1998; Jackson 2004). From this gender difference in what is considered normal/abnormal behaviour for boys and girls, it follows that the central role of hyperactivity in ADHD is likely to function as gender-bias for the detection of the disorder in girls. Relative to boys, girls show lower levels of hyperactivity, lower rates of other externalising behaviour, but greater intellectual impairment (Gaub 1997). Overall, they are less likely to meet the criteria for ADHD. The broadening of the concept of ADHD in 1980 to include inattentiveness without hyperactivity may be regarded as a way to address these gender-based shortcomings; the new criteria were more likely to address symptoms in girls.

In support of this hypothesis are the referral bias and the differential patterns of the combined and inattentive subtypes of ADHD, depending on whether population or clinic samples are examined. In community samples, boy/girl ratios of ADHD are 2:1 or 3:1. ADHD/I is approximately twice as prevalent as ADHD/C. In clinical samples, the boy/girl ratios may be as high as 9:1. ADHD/C is approximately 1,5 times more prevalent than ADHD/I, thus revealing that many more boys with hyperactivity symptoms are referred for specialist care. An implication of the focus on hyperactivity is that specific aspects of behavioural and learning disorders in girls have not been fully addressed in health care and research (Gaub 1997; Milich 2001). Only recently it was suggested to adjust the symptom cut-off scores of the DSM-IV criteria for sex. Because the majority of children in the DSM-IV field trial were male, the symptom threshold chosen is most appropriate to males (Barkley 2003).

Despite the relative differences in presentation of ADHD in boys and girls, the purpose of adaptations in the DSM criteria was to include both genders. This is now changing. Milich and co-workers raised the question whether differences between hyperactivity and impulsivity subtypes of ADHD as compared to the inattentive subtype should be valued, instead of the commonalities between them. In particular they noted that the subtypes may experience very different types of attention problems, almost half of the patients with ADHD/I fail to meet the criterion for early age of onset, and a higher proportion of the ADHD/I subgroup is non-responsive to drug treatment. In studies of children with ADHD, those with the inattentive subtype are often excluded and many research questions remain. Because of the high proportion of girls with ADHD/I, such a distinction would have a particular impact on the diagnosis of behavioural problems in girls. The authors recognise the positive effect of

distinguishing ADHD/I as a distinct disorder in terms of increased research interests (Milich 2001).

Development of an age-related disease concept

Whereas ADHD has always been recognised as a disorder affecting boys and girls, this is not the case for adults. ADHD was originally conceptualised as a childhood disorder, and is only recently recognised to persist in adulthood. ADHD appears to represent a 'dysmaturity' of specific parts of the brain. It is likely that genetic factors contribute significantly to most cases of ADHD, in addition to environmental factors (McArdle 2004). In approximately one-third of the children the disorder persists fully in adulthood, while in 50-60% one or more aggravating symptoms are present (Kooij 2003). Longitudinal studies show that symptoms of hyperactivity and impulsivity may decay, but inattention tends to persist.

Systematic research into adult ADHD is as recent as the 1990's. It is believed that ADHD in adults was only recently recognised, because child clinicians do not usually follow up patients into adulthood and structured diagnostic interviews in adult psychiatric settings often do not include ADHD (Weiss 2003; Weisfelt 2001; Biederman 1994; Kooij 1996). A complicating factor for diagnosing ADHD in adults is that the DSM-IV criteria were developed for and validated among children aged 4-16 and this study population was predominantly male (Kooij 2003; Barkley 2003). Clearly, health care is presently not fully addressing the problems of adults with ADHD. It can be expected that a next version of DSM will include age specific criteria, thus facilitating recognition of ADHD in adults.

ADHD and ethnicity

Ethnicity appears to be a problematic factor in relation to ADHD. When studying differences in prevalence and symptoms of ADHD between various ethnic groups, the interaction with underlying cultural differences is apparent. It is generally acknowledged that terminology, behavioural characteristics, diagnostic criteria and assessment methodology may vary among ethnic (or cultural) groups. Thus, cultural variation can be recognised in social opinions about problematic behaviour as well as in professional opinions.

To demonstrate the hypothesis that professionals from several countries differ in diagnostic methodology, researchers asked experienced clinicians from four countries to rate children in terms of ADHD symptomology (reference in Gingerich 1998). Although the clinicians rated children based on identical videotapes and used identical rating criteria, differences in character and severity of diagnoses were significant. For example, clinicians from China and Indonesia gave higher ratings of hyperactivity than did clinicians from the US and Japan. When studying the relationships between ethnicity and ADHD, it is imperative to draw careful conclusions. Studies in the UK have shown that Black children may score relatively low on ADHD, but much higher than other ethnic groups on conduct disorders (Evans 2004). The author questions whether we are more likely to

attribute behavioural disturbances to ADHD in Whites, and to conduct disorders in Blacks.

Contextualised consequences

As is the case with sex/gender differences, social factors may influence the degree of hyperactivity that is seen as a problem in different cultures. Therefore, an intriguing cultural phenomenon is the use of US-based DSM criteria for ADHD in many countries including the Netherlands, despite well-known variation in behavioural criteria, norms of deviance, assessment methodology and treatment. The DSM criteria have not developed along the lines of cultural variation, for example regarding threshold values, although differences in behavioural dimensions have been identified across various ethnic and cultural groups (Barkley 2003). Another relevant cultural aspect of ADHD is the context in which the diagnosis functions. In the US, children diagnosed with ADHD are considered disabled and therefore eligible for special education services. Together with the growth of school-based health clinics, the reform of special education may have increased the flow of children diagnosed with and treated for ADHD (Olfson 2003; Bussing 2003). These developments are likely to have influenced public debates about over- or under-diagnosis and medicalisation of children with ADHD in the US. Recognition of this specific context appears to be relevant when comparing the incidence and prevalence data of ADHD in the US with those in other countries.

In conclusion, ADHD is a disorder characterised by opinion-based diagnostic criteria. This case study shows that developments in these criteria determine who receives which type of care – they have large social impact. In particular, recognition of sex/gender differences in the presentation of behavioural, social and educational problems in boys and girls have questioned the diagnostic criteria of ADHD. Age-related differences remained unrecognised primarily because of the organisation of health care and opinions about the nature of the disorder. Regarding ethnic differences, the interaction with underlying cultural differences in values and assessment of child behaviour and its implications is apparent. Interestingly, cultural diversity questions the uniformity of diagnostic criteria for ADHD much more than sex/gender differences do. The social context, in which the diagnosis functions, appears to be highly relevant to understand the public debate about over- and under diagnosis of ADHD.

Changes of disease concepts of hypertension

The classification of hypertension has changed fundamentally during the last two decades. Age-related differences in the aetiology of hypertension have contributed to these conceptual changes, which can be understood in the context of drug development. Presently, other diversity issues in the manifestation of hypertension have not resulted in changes of the disease concept, although this may be the case in future.

Development of the disease concept

Although hypertension has been recognised as a risk factor for cardiovascular disease (CVD) for more than a century, there has been an ongoing debate regarding the relative importance of the diastolic and the systolic blood pressure (DBP, SBP) in predicting the future risk of CVD. The 1988 US guideline of the Joint National Committee on prevention, detection, evaluation and treatment of high blood pressure (JNC) classified hypertension in adults on the elevation of DBP, whereas 15 years later, the key message of JNC-VII is that in persons older than 50 years, elevated SBP is a much more important risk factor for CVD than DBP (JNC-VII 2003). The 1991 Dutch NHG-guideline differentiated between three severity levels of hypertension, all based on elevated DBP (Binsbergen 1991). Isolated systolic hypertension (ISH) was also known as a risk factor for CVD, but the guideline referred to insufficient scientific evidence to support pharmacotherapy and for this reason ISH was not included. The current Dutch guidelines for treatment of hypertension also changed to determine cardiovascular risk assessment on SBP (CBO 2000; Walma 2003).

Consequences for age differences in prevalence

How is this major shift in the concept of hypertension related to diversity issues? The prevalence of hypertension increases with age. In addition, the nature of hypertension changes. From age 30 to 50 years, SBP and DBP track together in a nearly parallel manner. After the age of 60, DBP decreases, while SBP continues to rise (Franklin 1999). Thus, the differences between SBP and DBP, known as the pulse pressure (PP), continue to rise at older age. The rise in SBP and PP in middle-aged and elderly subjects is due primarily to an increase in large-artery stiffness and an associated increase in wave reflection. As early as 1971, a trend was found of declining relative importance of DBP with a corresponding increase in the importance of SBP with advancing age as predictors of CHD. More recently, it was found that in patients younger than 50 years DBP was the strongest predictor of CHD risk, between 50 and 59 years was a transition period in which DBP, SBP and PP indexes were comparable predictors, and from 60 years on, DBP was negatively related to CHD risk, so that PP became superior to SBP (Franklin 2001). In subjects with identical SBP levels, those with ISH are at greater risk for CHD than those with combined systolic-diastolic hypertension (Franklin 1999).

Contextualised clinical research

To understand the reasons why it has taken such a long period of time to acknowledge the role of SBP, it is necessary to look at the broader context in which the diagnosis and treatment of hypertension have developed.

Since the 1950s, modern anti-hypertensives have been marketed and used. Regulatory approval of anti-hypertensives is based on the demonstration of efficacy, i.e. blood pressure reduction. DBP was chosen as the primary blood pressure component for the selection of trial

participants and as main target for evaluation of drug efficacy (Wieringa 1999). Two reasons account for this choice. First, DBP measurement is more reproducible than SBP, and smaller trial samples can be used (Deedwania 2002). Secondly, given the historical overrepresentation on relatively young, male patients in clinical trials, the choice for DBP was imperative and remained so (Wieringa 1999). The mean age of patients in efficacy trials of antihypertensive drugs lies around 55-57 years. In these relatively young populations, an elevated DBP can still be found, but at older ages the prevalence of ISH increases largely. Drug registration files of anti-hypertensives, marketed in the Netherlands between 1984 and 1995, showed a remarkable lack of efficacy data in elderly patients (Wieringa 1999). From a pharmacological point of view this can be understood, because studying anti-hypertensives in elderly implies that the SBP should be the guiding parameter, not DBP. Consequently, comparison of efficacy data between younger and older hypertensive patients is a much more complex matter, than just comparing levels of DBP reduction; it requires other research standards, thus, introducing new barriers for drug regulation.

Drug regulatory agencies have acknowledged the significance of isolated or predominant systolic hypertension, and concluded that this demands explicit evaluation of the effect of a drug on SBP (EMEA 2000). However, since its drafting in 2000, this guideline has not been accepted by the pharmaceutical industry, indicating their reluctance to shift antihypertensive efficacy testing from DBP to SBP.

The effects of the focus on DBP in efficacy trials for medical practice cannot be underestimated. Medical practitioners still focus on DBP (Basile 2003). As recent as 1988, the US guidelines for hypertension stated that in most elderly patients with ISH pharmacotherapy was not warranted (JNC-IV 1988). Nowadays ISH is acknowledged as the most common form of hypertension, and also the most difficult to treat because the aim is to reduce SBP and minimise the reduction of DBP. However, the available drugs were designed to reduce DBP and no antihypertensive drugs have specifically been developed for ISH. Many patients have uncontrolled ISH and physicians' knowledge of, and attitudes towards, vigorous SBP control are a problem (Kannel 2003). For elderly with ISH, the first study to demonstrate treatment effectiveness was published in 1991 (SHEP 1991). Deedwania concludes that the reductions in cardiovascular events in this study were far more impressive than in any other previous trial in hypertension, emphasizing that the lack of attention to SBP might be the reason for the relatively small decrease in heart disease observed in those studies (Deedwania 2002).

Development of sex and ethnic specific disease concepts?
With respect to hypertension in women and various ethnic groups, aetiological differences in comparison to white men have been documented. These originate in biological and environmental differences. Epidemiologically, women develop CVD approximately 10 years later than

men do. The role of female sex hormones in hypertension has been studied extensively, but no clear effects on the prevention of CVD have been found (NIH 2004). Non-hormonal haemodynamic sex differences can be found in the smaller body size and faster heart rates of women, affecting different blood pressure levels compared to men at all ages. Ethnic variation in blood pressure may result from differences in body size, but also from developmental differences. People originating from Africa develop hypertension which is characterised by low renin levels, high salt sensitivity, early onset and high severity, as compared to whites. Also, the prevalence of other cardiovascular risk factors, such as obesity, diabetes and lipid levels may vary among different ethnic groups. Thus, cardiovascular risk profiles vary according to sex and ethnicity. Treatment guidelines for hypertension are increasingly based on methods for cardiovascular risk assessment. At present, only sex differences have been incorporated into such methods, although it is acknowledged that ethnicity should also play a role (CBO 2000; Bindraban 2003).

Special populations

In the light of these sex, age and ethnic differences, the terminology used to describe patient categories in practice guidelines is of interest. As is the case in the 1991 Dutch guideline for general practitioners, the JNC guidelines since long recognise that hypertension is most prevalent in elderly patients, and that ISH is most common in this population. However, subsequent guidelines refer to elderly as a 'special population' for the treatment of hypertension. This definition also includes hypertension in racial and ethnic minorities, children and adolescents, (pregnant) women and in patients with cardiovascular and other coexisting diseases (JNC-IV 1988; JNC-V 1993; JNC-VI 1997; JNC-VII 2003). However, based on epidemiological data and the understanding of the larger association with CVD risk of the components of hypertension in elderly, it seems more accurate to define younger, hypertensive, white males as the 'special population' in hypertension.

In conclusion, hypertension is a disorder characterised by measurable parameters for diagnosis. This case study shows the major shift in opinion about the disease concept of hypertension from DBP to SBP as primary component for diagnosis that took place during the last 15 years. The shift's origin can be found in the results of longitudinal population studies. In contrast to antihypertensive drug regulatory studies, which focus on assessment of efficacy in younger populations, the longitudinal studies followed patients over time. These data, in combination with basic research into underlying mechanisms of hypertension in ageing, have led to a better understanding of the age-dependent development of risk factors for CVD. Thus, it can be concluded that differences in the nature of hypertension in elderly and younger populations have shaped the changing disease concepts. The prevalence of hypertension shows an interaction between age and sex; the shift toward SBP as primary blood pressure component therefore effects diagnosis and treatment of hypertension in women. In this case study we did not find evidence to support that sex- or

ethnicity-related aetiological differences have influenced diagnostic criteria for hypertension. From a medical point of view it seems relevant to differentiate between hypertension with high and low renin activity (Alderman 2004). Such a categorisation would allow better assessment of preventive and treatment interventions, affecting in particular people from African origin.

1.4 Discussion and conclusions

In this review, we focused on diversity issues from an epidemiological perspective. In medical practice, differences in aetiology, prognosis, disease presentation, perceptions and effect modification are relevant to consider, because they may affect the quality of health care. The findings highlight various issues for differentiation of treatment guidelines. The research implications also illustrate many directions for new hypothesis generation to study diversity. In the second part of this review, we analysed the clinical and social impact of diversity issues on disease concepts. The findings illustrate that opinions about disease concepts change over time, allowing new insights to be implemented into medical practice and clinical research.

Different types of diseases were selected for this study. The six diseases differ for example in their causes, genetic (asthma) vs. acquired (HIV/AIDS), and physical (hypertension) vs. behavioural (ADHD). Also, the socio-cultural contexts of the diseases are different. Life styles are important in hypertension, osteoporosis, gastric cancer and HIV/AIDS, but in each disease different aspects of life styles are relevant. Therefore, this selection of diseases may function to illustrate a variety of diversity issues, their complexities and interrelatedness.

For all six diseases, the aetiology, prognosis, disease presentation and perception were found to differ as function of sex/gender, age and/or ethnicity. Biological sex differences in development were found to influence asthma in children, the aetiology of hypertension, and bone growth as starting point for osteoporosis risk. Also in later stages of life, the role of sex hormones remained apparent in the aetiology and prognosis of these diseases. In HIV/AIDS, sex differences account for differences in susceptibility for infection, as well as specific disease symptoms and complications. Gender differences in socialisation influence the manifestation of ADHD. In gastric cancer, sex differences in localisation are present, although the underlying causes are unknown. Prognosis was found to be dependent on localisation of the tumour. In summary, sex differences in disease manifestation were found for all diseases.

Age is likely to influence the development of many diseases, because of underlying biological processes. However, specific age-related developmental aspects were found in asthma, ADHD, hypertension, osteoporosis and HIV/AIDS. The effects of these differences may refer to

the diagnosis, age-dependent risk differentiation, and differences in prognosis. The data revealed for example that asthma is more difficult to detect in elderly people, diagnostic criteria have only been developed for ADHD in children, and different parameters are needed for hypertension in elderly than in younger patients. In osteoporosis, suboptimal bone development at young age increases the risk for disease development at older age. The prognosis of HIV/AIDS is age-dependent, with the highest disease progression in children and older adults. In summary, age-related factors of disease manifestation were present for all diseases except gastric cancer.

Diversity with respect to ethnic differences was found to interact with genetic and/or socio-cultural differences in all diseases, but in various ways. Gastric cancer is associated with dietary habits, which differ between countries. Ethnic differences in bone development and lactose maldigestion were found, but inverse relationships exist concerning the risk for developing osteoporosis. In asthma, ethnic differences appear to become particularly relevant after migration, because of underlying differences in development of the immune system. In ADHD, ethnic differences strongly interacted with socio-cultural variation in opinions about child behaviour and diagnostic criteria. The aetiology of hypertension differs among various ethnic groups, leading to differences in severity and prognosis. Finally, in HIV/AIDS ethnic differences were found in disease parameters.

In conclusion, for all diseases aetiological and prognostic differences by age, sex/gender or ethnicity were found, leading to various patterns in disease manifestations. A large variation exists in underlying biological and socio-cultural causes. In addition, these data show that underlying causes of differences may interact. For example in osteoporosis, sex-related differences interact with ethnic differences in bone development and deterioration, leading to highly complex differences in disease patterns. Another example is asthma, where sex hormonal and sex-related developmental differences interact in various ways during life time. Therefore, addressing the relationships between patient characteristics and the underlying mechanisms may provide meaningful ways to include the complexities of diversity issues in research.

Effect modification of diagnostic, preventive or therapeutic interventions was documented in a number of cases. Lower levels of efficacy with antihypertensive drugs in patients from African origin, and ß-2 agonists in elderly patients with asthma are examples of effect modification in therapeutic drug treatment. These findings are relevant for treatment guidelines in specific populations. Effect modification may also occur on the side effect profiles of drugs. Because drugs were found to be tested in specific populations, for example drugs for ADHD predominantly in white boys and for osteoporosis predominantly in women, limited data on side effect profiles in other populations are available. In conclusion, effect modification was found in a number of therapeutic interventions, but more

often gaps in knowledge were identified in relevant areas for sex, age and/or ethnic specific interventions.

Research implications were identified on many diversity issues. The recommendation to develop age, sex and ethnic specific diagnostic criteria for ADHD, asthma, HIV/AIDS and osteoporosis was argued for in the literature because of underlying differences in disease manifestations. Also, the development of preventive strategies for osteoporosis and HIV/AIDS that take into account sex, age and ethnic differences was mentioned. In osteoporosis, the high prevalence of lactose maldigestion in many ethnic groups underscores the need to differentiate nutritional policies for optimal bone development.

The analysis of changes in disease concepts and diagnostic criteria of ADHD and hypertension illustrates how opinions about diversity issues determine which populations receive which kinds of health care. In both cases, disease concepts functioned to include (and exclude) specific categories of patients. In ADHD, disease criteria function as gender bias to identifying girls with the disorder, as well as adults. In the case of hypertension, disease criteria excluded the diagnosis in elderly, because of specific disease characteristics in these patients. Professional opinions changed significantly over time with respect to the disorders in these categories of patients, thus illustrating the social constructs of the disease concepts.

The case studies also illustrate the social context of diagnoses. The ADHD case shows that, in the US, identification of ADHD functions in the educational system to generate specific facilities and funding. Thus, public debates about under- and over-diagnosis of the disorder have a specific social and political context. In the extrapolation of epidemiological figures of ADHD to other countries, it is therefore relevant to consider this context. The hypertension case illustrates that diversity issues of aging have influenced developments in diagnostic criteria. It can be questioned why aetiological and prognostic differences in hypertension with respect to ethnicity have not shaped the diagnostic criteria in a comparable way as aging has. The social context of race and ethnicity in the US is likely to play a role in these processes.

Our results have implications for all involved in clinical research and the implementation of findings in medical practice. The clinical relevance of studying diversity can be expressed in medical, social and political terms. In medical practice, evidence of the effects of sex/gender, age and ethnicity on health and health outcomes of care add to a better targeting of medical care on an individual level. This review illustrates that diversity issues not only apply to variation in a certain disease measure, but may also refer to the parameter itself. The recent shift in diagnostic criteria of hypertension forms a good example of professional adjustment, although at present physicians have not fully implemented the new guidelines.

The social relevance of studying diversity is based on the democratic principles of our society. In a country with a multicultural population that considers all inhabitants to have equal rights and access to health care, it can be expected that relevant differences are considered in health care policy, including the conduct of clinical research. These social reasons have political implications. Addressing diversity issues on a political level is necessary to guide funding of clinical research.

Given these results, how can we contribute to change approaches to diversity in clinical research? First, it can be expected that diversity issues are clinically relevant in many diseases. Awareness and expertise to identify these issues for research is therefore important at all levels of programming, assessment and conduct of clinical research. A number of research areas were found, in particular differentiation of diagnostic criteria and interventions, the identification of under-researched areas, and increasing our understanding of biomedical and socio-cultural differences in diseases and health care.

Second, we showed that which diversity is considered relevant may change over time. Therefore, it can be expected that in areas where clinical changes are taking place, such as important changes in practice guidelines, diversity issues may emerge. Evaluation of these changes and their underlying causes may provide relevant input to define diversity issues for research. An example is the change in parameters to define hypertension, but also the neglect to define hypertension in terms of low and high renin activity in practice guidelines. This differentiation is likely to be meaningful to further define optimal treatment schedules for the patient groups which generally present with low renin activity, i.e. patients from African descent and elderly. Thus, placing clinical developments in their historical and social context may show which avenues of diversity have been pursued and which may have future potential.

Third and last, it is relevant to question whether the use of sex, age and ethnicity is the most 'natural' and effective categorisation to distinguish between groups of patients. A more nuanced understanding of differences and similarities between people is likely to address the underlying biological and/or socio-cultural mechanisms. In other words, sex, age and ethnicity function as surrogate classifiers of relevant biomedical and socio-cultural differences. For example, in osteoporosis an important sex difference of bone loss refers to the mechanism of bone deterioration. Thus, it can be expected that studying the efficacy of medication on these different mechanisms is likely to provide highly relevant knowledge for the treatment of various patient groups. Using sex as a categorisation for biomedical differences therefore functions as a surrogate to compare mechanisms of action. Identification of the relationship between the patient characteristic and the underlying biomedical and/or socio-cultural causes may help to overcome social barriers in identifying research questions. At present, developments in genomics and pharmacogenomics illustrate new approaches for identification of research populations on

biological variables, such as fast and slow metabolisers of drugs. Many more genetic differences in metabolising are likely to interact with sex and ethnicity. Therefore, new avenues are being explored to define research populations. These approaches are relevant to consider when studying diversity in clinical research. Moreover, it should be kept in mind that categorisations in clinical research have functioned to exclude specific populations, and it will remain necessary to address these socio-cultural and political effects of studying diversity.

References

Alderman MH, Cohen HW, Sealey JE, Laragh JH. Plasma renin activity levels in hypertensive persons: their wide range and lack of suppression in diabetic and in most elderly patients. Am J Hypertens 2004; 17: 1-7.

Asmar R, Safar ME, Quenaeu P. Pulse pressure. An important tool in cardiovascular pharmacology and therapeutics. Drugs 2003; 63(10): 927-932.

Balzano G, Fuschillo S, Melillo G, Bonini S. Asthma and sex hormones. Allergy 2001; 56: 13-20.

Barkley RA. Issues in the diagnosis of attention-deficit/hyperactivity disorder in children. Brain Develop 2003; 25: 77-83.

Basile JN. Systolic blood pressure elevation. It's where the action is. Drugs Aging 2003; 20(4): 287-288.

Becklake MR, Kauffmann F. Gender differences in airway behaviours over the human life span. Thorax 1999; 54: 1119-1138.

Biederman J, Faraone SV, Spencer T, Wilens T, Mick E, Lapey KA. Gender differences in sample of adults with attention deficit hyperactivity disorder. Psych Res 1994; 53: 13-29.

Bindabran NR, Stronks K, Klazinga NS. Cardiovasculaire risicofactoren bij Surinamers in Nederland: een literatuuroverzicht. Ned Tijdschr Geneeskd 2003; 147(33): 1591-1594.

Binsbergen JJ van, Grundmeyer HGLM, van den Hoogen JPH, van Kruysdijk M, Prins A, van Ree JW, Thomas S. NHG-standaard Hypertensie. Huisarts Wet 1991; 34(8): 389-395.

Brewster LM, van Montfrans GA, Kleijnen J. Systematic review: antihypertensive drug therapy in black patients. Ann Intern Med 2004; 141(8): 614-627.

Brondolo E, Rieppi R, Kelly KP, Gerin W, Perceived racism and blood pressure: a review of the literature and conceptual and methodological critique. Ann Behav Med 2003; 25(1):55-65.

Bussing R, Gary FA, Mills TL, Garvan CW. Parental explanatory models of ADHD. Gender and cultural variations. Soc Psychiatr Epidemiol 2003; 38: 563-575.

Caracta CF. Gender differences in pulmonary disease. Mt Sinai J Med 2003; 70(4): 215-234.

CBO Kwaliteitsinstituut voor de Gezondheidszorg. Herziening richtlijn hoge bloeddruk. Van Zuiden Communications BV, Alphen aan de Rijn 2000.

Deedwania PC. The changing face of hypertension. Is systolic blood pressure the final answer? Arch Intern Med 2002; 162: 506-507.

Ebeling PR. Osteoporosis in men. New insights into aetiology, pathogenesis, prevention and management. Drugs & Aging 1998; 13(6): 421-434.

Elming H. Sonne J, Lublin HKF. The importance of the QT interval: a review of the literature. Acta Psychiatr Scand 2003; 107: 96-101.

EMEA. ICH principles document for clinical evaluation of new hypertensive drugs (CMPM/ICH/541/00). London, 29 June 2000.

Evans R. Ethnic differences in ADHD and the mad/bad debate (letter). Am J Psychiatry 2004; 161 (5): 932.

Franklin SS, Kahn SA, Wong ND, Larson MG, Levy D. Is pulse pressure useful in predicting risk for coronary heart disease? The Framingham heart study. Circulation 1999; 100:354-360.

Franklin SS, Larson MG, Khan SA, Wong ND, Leip EP, Kannel WB, Levy D. Does the relation of blood pressure to coronary heart disease risk change with aging? Circulation 2001; 103: 1245-1249.

Franklin SS. Definition and epidemiology of hypertensive cardiovascular disease in women: the size of the problem. J Hypertens 2002; 20(suppl 2): S3-S5.

Gaub M, Carlson CL. Gender differences in ADHD: A meta-analysis and critical review. J Am Acad Child & Adolesc Psych 1997; 36(8): 1036-1997.

Gibb DM, Duong T, Tookey PA, Sharland M, Tudor-Williams G, Novelli V et al. Decline in mortality, AIDS, and hospital admissions in perinatally HIV-1 infected children in the United Kingdom and Ireland. BMJ 2003; 327, 1019-1024.

Gilad J, Walfisch A, Borer A, Schlaeffer F. Gender differences and sex-specific manifestations associated with human immunodeficiency virus infection in women. Eur J Obs Gyn Reprod Biol 2003; 109: 199-205.

Gill S, Shah A, Le N, Cook F, Yoshida EM. Asian ethnicity-related differences in gastric cancer presentation and outcome among patients treated at a Canadian cancer center. J Clin Oncol 2003; 21: 2070-2076.

Gingerich KJ, Turnock P, Liftin JK, Rosén LA. Diversity and attention deficit hyperactivity disorder. J Clin Psychol 1998; 54(4): 415-426.

Haks K, Op de Coul ELM, van de Laar MJW. Hoe vaak komt AIDS en HIV-infectie voor en hoeveel mensen sterven eraan? In: Volksgezondheid Toekomst Verkenning, Nationaal Kompas Volksgezondheid. Bilthoven: RIVM, http://www.nationaalkompas.nl Gezondheidstoestand\Ziekten en aandoeningen\Infectieziekten en parasitaire ziekten\AIDS en HIV-infectie, 27 maart 2002.

Jackson KA, Savaiano DA. Lactose maldigestion, calcium intake and osteoporosis in African-, Asian-, and Hispanic-Americans. J Coll Nutr 2001; 20(2): 198S-207S.

Jackson DA, King AR. Gender differences in the effects of oppositional behavior on teacher ratings of ADHD symptoms. J Abnormal Child Psychol 2004; 32(2): 215-224.

Jamerson KA. The disproportionate impact of hypertensive cardiovascular disease in African Americans: getting to the heart of the issue. J Clin Hypertens 2004; 6(4 suppl 1): 4–10.

JNC-IV. The 1988 report of the Joint National Committee on detection, evaluation, and treatment of high blood pressure (JNC 4). Arch Intern Med 1988; 148: 1023-1038.

JNC-V. The fifth report of the Joint National Committee on detection, evaluation, and treatment of high blood pressure (JNC 5). Arch Intern Med 1993; 153: 154-183.

JNC-VI. The sixth report of the Joint National Committee on prevention, detection, evaluation, and treatment of high blood pressure (JNC 6). Arch Intern Med 1997; 157: 2413-2439.

JNC-VII. The seventh report of the Joint National Committee on the prevention, detection, evaluation, and treatment of high blood pressure (JNC 7). Resetting the hypertension sails. Hypertension 2003; 41: 1178-1179.

Kannel WB. Prevalence and implications of uncontrolled systolic hypertension. Drugs Aging 2003 ; 20(4): 277-286.

Kooij JJS, Goedkoop JG, Gunning WB. Aandachtstekortstoornis met hyperactiviteit op volwassen leeftijd; implicaties voor diagnostiek en behandeling. Ned Tijdschr Geneesk 1996; 140(37): 1848-1851.

Kooij JJS. ADHD bij volwassenen. Inleiding in diagnostiek en behandeling. Swets & Zeitlinger BV, Lisse 2003.

Leest LATM van, Verschuren WMM. Bloeddruk samengevat. In: Volksgezondheid Toekomst Verkenning, Nationaal Kompas Volksgezondheid. Bilthoven: RIVM, http://www.nationaalkompas.nl Determinanten\Persoonskenmerken\Bloeddruk, 5 december 2003

Lefler LL, Bondy KN. Women's delay in seeking treatment with myocardial infarction. A meta-synthesis. J Cardiovasc Nurs 2004; 19(4): 251-268.

Lemanske RF. Issues of understanding pediatric asthma: epidemiology and genetics. J Allergy Clin Immunol 2002; 109: S521-524.

Linden MW van der, Westert GP, de Bakker DH, Schellevis FG. Tweede nationale studie naar ziekten en verrichtingen in de huisartsenpraktijk. Klachten en aandoeningen in de bevolking en in de huisartsenpraktijk. Nivel, Utrecht 2004.

Lieu TA, Quesenberry CP, Capra AM, Sorel Me, Martin KE, Mendoza GR. Outpatient management practices associated with reduced risk of pediatric asthma hospitalization and emergency department visits. Pediatr 1997; 100(3): 334-341.

London AS, Robles A. The co-occurrence of correct and incorrect HIV transmission knowledge and perceived risk for HIV among women of childbearing age. Soc Sci Med 2000; 51: 1267-1278.

Loutfy MR, Walmsley SL. Treatment of HIV infection in pregnanct women. Antiretroviral management options. Drugs 2004; 64(5): 471-488.

Luzuriaga K, McManus M, Mofenson L, Britto P, Graham B, Sullivan JL. A trial of three anti-retroviral regimens in HIV-1-infected children. N Engl J Med 2004; 350: 2471-2480

McArdle P. Attention-deficit hyperactivity disorder and life-span development. Br J Psych 2004;184:468-469.

Meadows LM, Mrkonjic LA, Petersen KMA, Lagendyk LE. After the fall: Women's views of fractures in relation to bone health and midlife. Women & Health 2004; 39(2):47-62.

Melton LJ, Orwoll ES, Wasnich RD. Does bone density predict fractures comparably in men and women? Osteoporosis Int 2001; 12: 707-709.

Milich R, Balentine AC, Lynam DR. ADHD Combined Type and ADHD Predominantly Inattentive Type are distinct and unrelated disorders. Clin Psychol Sci Prac 2001; 8:463-488.

NHG standaard Astma. Website: http://nhg.artsennet.nl/upload/104/standaarden/M26/start.htm. Consulted on 17-10-2004.

NIH Consensus Development Panel on Osteoporosis Prevention, Diagnosis, and Therapy. Osteoporosis prevention, diagnosis, and therapy. JAMA 2001; 285(6):785-795.

NIH website: www.nlm.nih.gov/databases/alerts/estrogen_alone.html Clinical alert: NIH asks participants in women's health initiative estrogen-alone study to stop study pills, begin follow-up phase. Consulted on May 25 2004.

Olfson M, Gameroff MJ, Marcus SC, Jensen PS. National trends in the treatment of attention deficit hyperactivity disorder. Am J Psychiatry 2003; 160(6): 1071-1077.

Osman M. Therapeutic implications of sex differences in asthma and atopy. Arch Dis Child 2003; 88: 587-590.

Pieters T, te Hennepe M, de Lange M. Pillen & Psyche: culturele eb- en vloedbewegingen. Medicamenteus ingrijpen in de psyche. Rathenau Instituut, Den Haag 2002.

Poos MJJC, Gijsen R. Prevalentie, incidentie en sterfte naar leeftijd en geslacht. In: Volksgezondheid Toekomst Verkenning, Nationaal Kompas Volksgezondheid. Bilthoven: RIVM, http://www.nationaalkompas.nl Gezondheidstoestand\Ziekten en aandoeningen\ziekten van de ademhalingswegen\astma en chronische obstructieve longziekten (COPD), 14 mei 2003.

Renwick DS, Connolly MJ. Improving outcomes in elderly patients with asthma. Drugs Aging 1999; 14(1): 1-9.

Roder DM. The epidemiology of gastric cancer. Gastric Cancer 2002; 5 (suppl 1): 5-11.

Safar ME, Smulyan H. Hypertension in women. Am J Hypertens 2004; 17: 82-87.

Seal DW, Ehrhardt AA. HIV-prevention-related sexual health promotion for heterosexual men in the United States: pitfalls and recommendations. Arch Sex Behav 2004; 33(3): 211-222.

Seeman E. Pathogenesis of bone fragility in women and men. Lancet 2002; 359: 1941-1850.

Shek LP, Tay AH, Goh DL, Lee BW. Ethnic differences in genetic susceptibility to atopy and asthma in Singapore. Ann Acad Med Singapore 2000; 29(3): 346-350.

SHEP Cooperative Research Group. Prevention of stroke by antihypertensive drug treatment in older persons with isolated systolic hypertension. Final results of the systolic hypertension in the elderly program (SHEP). JAMA 1991; 265(24): 3255-3264.

Smith PR, Sarner L, Murphy M, James B, Thomas JM, Skinner CJ, Aitken C. Ethnicity and discordance in plasma HIV-1 RNA viral load and CD4+ lymphocyte count in a cohort of HIV-1-infected individuals. J Clin Virol 2003; 26: 101-107.

Spahn JD, Cherniack R, Paull K, Gelfand E. Is forced expiratory volume in one second the best measure of severity in childhood asthma? Am J Respir Crit Care Med 2004; 169: 784-786.

Spencer T, Biederman J, Wilens T. Growth deficits in children with Attention Deficit Hyperactivity Disorder. Pediatrics 1998; 102: 501-506.

Spencer TJ, Biederman J, Wilens TE, Faraone SV. Overview and neurobiology of Attention-Deficit/Hyperactivity Disorder. J Clin Psychiatry 2002; 63 (suppl 12): 3-9.

Von Hertzen LC, Haahtela T. Immunization and atopy: possible implications of ethnicity. J Allergy Clin Immunol 2004; 113: 401-406.

Von Mutius E. Paediatric origins of adult lung disease. Thorax 2001; 56: 153-157.

Vrieze A, Postma DS, Kerstjens HAM. Perimenstrual asthma: a syndrome without known cause or cure. J Allergy Clin Immunol 2003; 112: 271-282.

Walma EP, Thomas S, Prins A, Grundmeyer HGLM, van der Laan JR, Wiersma Tj. NHG-standaard Hypertensie (derde herziening). Huisarts Wet 2003; 26(8): 435-449.

Weisfelt M, Schrier AC, de Leeuw MC. Druk gedrag bij volwassenen; mogelijk aandachtstekort-hyperactiviteitstoornis (ADHD). Ned Tijdschr Geneesk 2001; 145(31): 1481-1484.

Weiss M, Murray C. Assessment and management of attention-deficit hyperactivity disorder in adults. CMAJ 2003; 168(6): 715-722.

Wieringa NF. Cardiovascular drugs: discrepancies in demographics between pre- and post-registration use. Eur J Clin Pharmacol 1999; 55: 537-544.

Wijnhoven BPL, Poos MJJC. Hoe vaak komt maagkanker voor en hoeveel mensen sterven eraan? In: Volksgezondheid Toekomst Verkenning, Nationaal Kompas Volksgezondheid. Blithoven: RIVM, http://www.nationaalkompas.nl Gezondheidstoestand\Ziekten en aandoeningen\Nieuwvormingen\Maagkanker, 5 november 2002.

Wolf F de. Summary of monitoring of human immunodeficiency virus (HIV) in the Netherlands. Report 2003. Amsterdam: Stichting HIV Monitoring.

Yao JC, Schnirer II, Reddy S, Chang S, Najam A, Yu C, Giacco G, Hess K, Rashid A, Xie K, Lynch P, Ajani JA. Effects of sex and racial/ethnic group on the pattern of gastric cancer localization. Gastric Cancer 2002; 5: 208-212.

Yao JC, Mansfield PF, Pisters PWT, Feig BW, Janjan NA, Crane C, Ajani JA. Combined-modality therapy for gastric cancer. Seminars in Surgical Oncology 2003; 21: 223-227.

Appendix

Search strategy for Medline on age, sex, ethnicity as determinants:

"disease name"
AND sex factor
AND all following terms separately: aetiology/prognosis/incidence/prevalence /symptom/perception/diagnosis/treatment/prevention/effect modification
AND age factor
AND all above terms
AND ethnic factor
AND all above terms

The last selection term to be applied was "review". In case no reviews were found, where possible a selection of relevant primary studies was made. Literature included was published in 1995 or later.

Search strategy for Medline on changes in disease concepts:

"disease name"
AND disease concept
AND all following terms separately: sex factor/gender/age factor/ethnic factor
AND classification
AND all above terms

Primary studies and reviews were included and no limitation was made to year of publication. Also, reference lists of retrieved papers were screened for relevant papers.

Review 2

Diversity in clinical practice: which differences matter?

Amâde M'charek, Mirjam Kohinor, Ronald Stolk[2]

[2] The authors thank Jeanette Pols and Annemarie Mol for comments on a previous version of this paper and suggestions for improving it. This paper has also benefited from discussions during the "Invitational expert meeting on diversity in clinical research" (Bergen, November 2004). We thank all participants for their feedback.

Contents

2

Diversity in clinical practice: which differences matter?

Summary

Diabetes Mellitus (DM) has taken epidemic forms in The Netherlands and the western world. The exponential growth is mainly caused by the rise in patients with Type 2 diabetes. Until recently this variant of DM was referred to as adult onset diabetes since it was classically described as a disease of the elderly. But times are changing, not only because more people become diabetic, but also at a younger age. The goal of this review is to find out which diversity matters in clinical practise and what can be done to recognise, acknowledge and then use this diversity to provide better care.

In this paper we first review the international literature about DM2 with the following questions in mind:

1. What is diversity in clinical practice?
2. Which differences are enacted in that practice?
3. And how does this enable or inhibit other differences?

In the literature it is clear that hyperglycaemia is a central problem in DM, but it is unclear whether this is caused by genetic predisposition, eating habits, lifestyle, body composition or something else. With its travel trough medical literature the shape of DM changes, and different differences matter in every discipline. Diversity is in itself highly diverse and dynamic. In our ethnographic study of diabetes care in a general hospital in Amsterdam we found that many different care givers are involved general practitioners, internists, dieticians, podiatrists, diabetes-nurses. And again with every caregiver the disease changes. Based on 'practical knowledge' practitioners add relevant versions of diversity that do not receive much

attention in clinical research. Moreover, this clinical knowledge is not clearly articulated.
Diversity is more than sex, age and ethnicity; it is not stable and cannot be reduced to a number of fixed variables. The relevance of versions of diversity depend on its contexts. Ethnographic approaches to clinical research may provide insight and help to establish links between research and clinical practice.

2.1 Introduction

This review deals with Diabetes Mellitus type 2 and diversity in The Netherlands. In The Netherlands as well as in other countries diabetes has taken epidemic shapes. Precise estimations for the prevalence are not available, but it is estimated that the number of patients diagnosed with diabetes in The Netherlands is around 450.000 (of which 385.000 have type 2 diabetes mellitus). And each year 60.000 new patients are diagnosed. Next to the magnitude of the problem, diabetes is also a complex disease. This holds also for the care that patients receive. In terms of prevalence and diagnosis diabetes has moved way beyond the nice distinction between juvenile (type 1) and adult-onset diabetes (type2). For example ever more children are diagnosed with type 2 diabetes.
In the treatment of diabetes many different care givers are involved: among others, general practitioners, internists, dieticians, podiatrists, diabetes-nurses, but also the patient, who has an important role in the management of the disease. In addition, within the medical disciplines involved, various differences are made for the purpose of diagnoses and treatment, such as genes and environment, males and females, young and old, but also in terms of weight, ethnicity, eating habits, exercise, feet-care, good or bad shoes, blood-sugar-level, medication, etc, etc. All these differences, which can be found in the literature and elsewhere, cannot be accounted for at the same time in any given practices. Hence our question: which differences are made to matter, where? With reference to diversity in diabetes research and care, our goal is to shed light on the following three questions:

1. What is diversity in clinical practice?
2. Which differences are enacted in that practice?
3. And how does this enable or inhibit other differences?

Instead of reducing difference to one, or even many, biological characteristics, we take difference as the outcome of a heterogeneous medical practice. Our point of departure is therefore: differences (or diversity) in medicine do not simply refer to entities (genes, hormones, etc.) in human bodies. In practice many more humans and things are involved in making differences or similarities, such as patients/ professionals and their knowledge, techniques, protocols, or usual ways of doing work in the clinic. We therefore take difference as an effect of such complex interactions. The differences that matter (in both senses of the

word) may be stable or fluid. The normative question is whether this enables good care for every body, or whether and how it should be changed.

Method

Our analyses are based on two studies, a literature review and a short ethnographic study. In this paper we first review the (international) literature on prevention of and intervention in diabetes. Our aim here is to trace similarities and differences with reference to patients or persons at risk, practices of good or failing care, knowledge and information about diabetes, and with reference to health and healthy living.

We have searched broadly for literature on the Internet (mainly MedLine, but at a later stage also Google) and have focused on ethnicity, adolescent and genetics. The following entries have initially been used for that purpose:

- diabe* + netherlands
- diabetes + type 2
- diabetes + type 2 + children
- diabetes + genetics
- diabetes + type 2 + genetics
- type 1 + diabetes + genetics
- diabetes + type 2 + children + adolescent
- diabetes + ethnicity
- diabetes + UK

In addition, reports, such as those of the Health Council of The Netherlands, the Dutch Diabetes Federation and the Dutch College of General Practitioners (NHG) and the American Diabetes Association (ADA), as well as references that we have found in the already retrieved literature have been gathered.

Secondly we report on a focused ethnographic study of diabetes cares in a general hospital in Amsterdam, The Netherlands. Because of the size of diabetes care in The Netherlands, the care for these patients is usually provided by so called Diabetes Care Teams (diabetes behandel teams). In these teams the diabetes-nurse has been introduced as a mediator between different medical practices and a monitor of the various interventions. The chain of care in these teams differs for type 1 and type 2 diabetes. Whereas the professional care for patients with type 1 diabetes is located in the hospital (with the internist), that of patients with type 2 is usually covered by the GP. In both cases there is a strong collaboration with the dietician, the podiatrists, and the diabetes-nurse.

In our study we have interviewed an internist, two diabetic nurses, and a GP. All are part of such a Diabetes Care Team. We have gathered reports of former studies conducted in the hospital where the internist and diabetic nurses are based and educational material for the patients (such as videotapes and leaflets). Next to the interviews we have observed consultations at the same hospital where diabetes patients were seen by the internist. These sites as well as the interviewees are made anonymous

in our presentation. Also, the material presented and analyzed is not intended to be representative, but should rather be taken as illustrative of the complexity of clinical practice.

Goal of the review
The review offers an analysis of the interference between different practices of care (as laid down in the literature or in different contexts in the chain of care). By tracing the various ways in which differences are enacted it evaluates the dominance of some practices over others. This may, so we hope, contribute to an understanding of diversity that goes beyond biological characteristics and which is more attune with practices of care and health. Moreover, evaluating practices, this review aims at tracing possibilities for introducing novel ways of making differences in order to facilitate good care for more patients.
And finally, we contemplate on the impact of our findings for clinical research practices, and on the limits and possibilities to account for more diversity in particular research projects. These results will be elaborated on in review 6.

Organization of the Paper
In this paper we review the literature by answering the following questions:

- What is diabetes?
- Where can it be found?
- How is it handled?

Based on these questions we identify different sites in the literature where the answers may change in tone. In the second part, the empirical findings are analyzed in accord with the following questions:

- What is good care and for whom?
- Does diversity matter, and how?
- Which differences are imported to the medical practice and what can we learn from these about diversity in clinical research?

The paper is concluded with a discussion based on the last question mentioned.

2.2 Which differences matter? DM2 in the literature

In the second half of the year 2004 the Dutch Ministry of Health has rung the alarm bells by calling diabetes type 2 (DM2) an insidious epidemic disease (‘een sluipende volksziekte’). More effort of insurance companies, health professionals and researchers were needed as to diagnose persons at risk at an early stage and to provide better care (NRC 2004). The Health Council of The Netherlands has recently completed a study on the (costs and) benefits of screening for DM2 (Gezondheidsraad 2004) and most

recently the Netherlands Organization for Health Research and Development (ZonMw) has devoted a special issue of its periodical Mediator on DM2 and the coordination of research and care. In other words diabetes type 2 is receiving much attention in Dutch medical health care and policy. In discussions diabetes is related to various 'populations', separated by age, by ethnicity, life style and socio-economic status, or weight. Differences and similarities are thus made as to distinguish between groups of persons at risk and others that are not. In this paper we take such differences into account, but show how they shift and change depending on the context addressed, but also how other differences may enter research, diagnosis and intervention. The relevance of difference, as we will show, is not given and which difference matter is dependent on where diabetes is found, i.e. where in the trajectory of care. As indicated above our discussion of the literature is organized around the three questions: what is diabetes, where can it be found and how is it handled?

2.2.1 What is diabetes?

The answer to the question 'what is diabetes' seems straight forward; one simply has to open a medical handbook to find out. One would for example read that it is "a condition characterized by a raised concentration of glucose in the blood because of a deficiency in the production and/or action of insulin, a pancreatic hormone made in special cells called the islet cells of Langerhans" (Black's Medical Dictionary, 1999: 141). This then tells us that diabetes is related to the (mal) functioning of the pancreas resulting in a disturbed blood glucose level and that the hormone insulin is crucial in that. There are however many different kinds of medical literature and many different emphases placed while dealing with diabetes. Here we will consider insulin and blood glucose level, genes and environment, prevalence and population, and patients and managers.

Insulin and blood glucose level

Standard medical literature on diabetes usually starts by distinguishing type1 from type2 diabetes. DM1 is characterized by the absence of insulin production due to auto-immune destruction of the beta-cells in the pancreas. DM2, previously known as 'non-insulin-dependent diabetes mellitus', is a disease where the pancreas does produce insulin, but the cells of the body are less sensitive to the action of insulin and fail to take up the glucose from the blood. The result is an increased concentration of glucose: diabetes.

Although diabetes is defined by an increased glucose level in the blood, the main concerns with the disease are the long term complications which result in increased mortality. Complications are divided into micro vascular (notably eye, kidney and toes) and macro vascular (notably myocardial infarction and stroke).

Genes and environment

There are many debates about the genetic basis of diabetes. Some literature seems to provide evidence for this basis. There it is argued that genetic predisposition plays an important role in DM2. Elbein and Wolff (1997) for example argue that a history of type 2 diabetes in a first-degree relative doubles the risk of diabetes and that a child of two diabetic parents has an 80% lifetime risk to develop diabetes. This seems to point to the conclusion that DM2 is inheritable. The wide variation in incidence and prevalence of DM2 among different ethnic groups, such as among the Pima-Indians but also recently in an isolated population in the south-west of The Netherlands (Aulchenko et al. 2003) further seems to suggest a genetic aetiology (Elbein and Wolff 1997). However, it is only in a small number of cases that a direct genetic cause, in the form of an autosomal dominant, single gene form, could be established. Examples of these are mutations of glucokinase, the hepatocyte nuclear factors, and insulin promoter factor 1 (Elbein 2000). The little evidence for a clear genetic link to DM2 combined with the fact that the prevalence of the disease within a community can change dramatically over time (Jenkins and Campbell 2004) has lead to the consensus that diabetes is a disease involving both genetic and environmental factors (see also NHG Standaard, Rutten, Verhoeven et al. 2003).

Prevalence and populations

Diabetes is also a matter of numbers, prevalence and risks. Exact numbers on the prevalence of diabetes in The Netherlands as well as elsewhere are not available. The *NHG Standaard* reports a number of 260.000 diagnosed patients, but underlines that if the population would be screened this number should probably be doubled (Rutten, Verhoeven et al. 2003). Bakker and Bilo estimate that the total number of people with diabetes in the year 2000 had been 462.000 (Bakker and Bilo 2004). Both articles stress that the incidence of diabetes is increasing rapidly. Approximately 85% of Diabetes Mellitus patients have type 2 and especially in this group of patients the rapid growth is found. Moreover, not only the prevalence of diabetes is growing rapidly but also the average onset of the disease moves towards a younger age.

DM2 was classically described as a disease of the elderly; the onset of the disease occurs normally in people who are 50 years or older. Times are however changing. In the United States a new trend is observed. Children and especially adolescents are ever more diagnosed with DM2. According to estimates, 8-45% of the children with newly diagnosed diabetes have DM2 (Ramchandani 2004). But this trend is not restricted to the US; also in the UK research has been conducted on the prevalence of DM2 among children and adolescents. In the study of Ethisham, Hattersley and their colleagues a cross-sectional questionnaire survey of all paediatric diabetes centres in the UK during the year 2000 was conducted. They found that a considerable part of the group of children that were described as a-typical DM1 were actually type 2 diabetes patients (Ehtisham, Hattersley et al. 2004). Also Wabitsch and others have found in their study of a large cohort of 'Caucasian' children and adolescents with obesity living in

Europe that impaired glucose regulation and DM2 were present in a substantial portion of the studied patients. Their advice is therefore to screen for diabetes in severely obese children (Wabitsch, Hauner et al. 2004). In accord with these findings Wiegand, Maikowski and their colleagues conclude that insulin resistance DM2 is present in obese children. And they further claim that: "in Europe impaired glucose regulation and other signs of the insulin resistance syndrome in children and adolescents are far more common than so far believed and that they are not restricted to ethnic minority groups" (Wiegand, Maikowski et al. 2004). The prevalence of diabetes mellitus type 2 in children and adolescents in the Netherlands is unknown (Renders, Delemarre van et al. 2003) but diagnosing diabetes type 2 among this group is not uncommon (http://www.rivm.nl).
Next to the growing prevalence among younger persons, parts of the literature on DM2, also shows a higher prevalence among specific ethnic groups. We will go into that below.

Patients and managers
For patients diabetes type 2 often has an ambiguous character. On the one hand there are symptoms of the disease, which can vary in severity but are usually mild. On the other hand there are the complications that may occur in the future. Since these complications are serious and may even be life threatening, their prevention is the main aim of diabetes treatment. The disease implies the constant threat of disability in the form of blindness, kidney or heart failure, amputation of limbs or stroke. For the patient this means that although s/he may feel good or healthy at present, especially when the medication works well, s/he has to constantly keep the future in mind. This tension between present experience with the disease and future risks makes complying with the regime difficult for many patients.
The chronic character of the disease implies that behavioural adjustments that a patient has to make are directed towards the future and have to be chronic as well. A temporarily increase in physical activity or change in dietary habits does not take away the threat of severe complications. Since diabetes is a progressive disease, patients always have to be on their guard. Once diagnosed with diabetes, patients are enrolled in a medical regime and involved in the management of their disease. This consists of among others, the monitoring of their blood glucose level, compliance to prescribed dietary and medication, special care for their body (physical exercise, care for their feet, and the monitoring their sight). Management is also about control. Freeman and Loewe report on the importance of control, especially in the interaction with the clinician. In their study patients describe how frustrating it can be when the blood sugar is "out of control" due to stress or other emotions connected to their every day life (Freeman and Loewe 2000).
The progression of DM2 in a patient often implies a change of medication. Usually patients with DM2 start with oral medication. Switching from oral medication to insulin, as a result of the progression of the disease, is a large and difficult step for many patients to take. This has not only to do with fear for needles but also for the implication of having to use insulin,

which is irreversible. Patients may also be reluctant to start on insulin because they see this as a sign of deterioration and decline (Loewe and Freeman 2000).

By posing the question 'what is diabetes' we have shown that it can be many different things. We have seen that it is the absence and/or insensitivity to insulin, which are involved in a complex mechanism of blood and glucose level. It may also be described in genetic terms, but also these hold a relation with the environment in which genes and a DM2 patient are taking part. DM2 is also a matter of prevalence and populations at risk, something that is not stable but changes over time. And finally we have paid attention to DM2 as experienced by patients. Here symptoms, their severity and the patients' control of these came to the fore. It is tempting to view these different versions of DM2 as definitions that can be allocated to different persons or different contexts. But as we will see below such versions are not fixed somewhere but tend to travel across different settings.

2.2.2 Where can diabetes be found?

The fact that DM2 can be many things urges one to look at where it can be found and to trace what it is made to be in different settings. Here we confine ourselves to reporting about DM2's and their locales as described in the literature. However, even in the literature one hits upon many more locales than we describe below. We have chosen to pay attention to three locales: the clinic where DM2 presents itself to the professional; the population as an object of diagnosis and risk assessment; and the medical standards as a kind of policy guideline for good care.

In the clinic: the symptoms
The insidious onset of DM2 contributes to the under diagnoses of the disease (Ramchandani 2004). The risk of complications is already present before diabetes is diagnosed. The diagnosis of DM2 in the clinic (this may be at the hospital or at the GP) usually starts with the interpretation of the clinical presentation. However, over the last years there is increasing attention for screening and case-finding, which means that blood glucose is measured and diabetes diagnosed in persons with no symptoms of raised glucose levels. In the clinic, DM2 is a cluster of symptoms. The signs and symptoms of DM2 as described in the literature are:

- Polyuria
- Polydipsia
- Fatigue or lack of energy
- Obesity
- Hypertension
- Family history of diabetes
- Polycystic ovarian syndrome
- Acanthosis nigricans

- Fungal infections
- History of insulin resistance

Here we will address just those symptoms that are most commonly viewed as indications of DM2 in the clinic.
When a patient visits a clinician with complaints about frequent urination, excessive thirst and fatigue, DM is one of the first things to think about. These are classic symptoms that are always mentioned in leaflets or on websites. Information about these is aimed at patients (and others) as an urgent call to go and see a doctor. The already mentioned hereditary factor of diabetes makes inquiry by the medical professional about the prevalence of diabetes in a family a relevant diagnostic tool. However, family history by itself is not an indication for practitioners to screen for diabetes. In clinical practice, the first mentioned symptoms are, one could say, dominant over genes and family history.
Obesity in relation to DM2 is nowadays receiving much attention. Obesity is generally viewed as a sign of imbalance of the metabolic system. It is often accompanied by hypertension and blood fat disorders, such as high cholesterol. Insulin resistance and glucose intolerance are components of the metabolic syndrome (Ten and Maclaren 2004). According to epidemiological studies obesity by itself increases the risk of developing DM2 with a factor five (Gezondheidsraad 2004). The relationship between obesity and DM2 is nevertheless a complex one. It does not stand alone, but interacts with, for example, the occurrence of diabetes in the family. A person with obesity and a family history of diabetes has a 4 time higher chance of developing DM2 than a person who is not obese (Elbein and Wolff 1997). The degree of obesity is measured by calculating the relation between weight and height: the so-called body mass index (BMI). But Bhatia (2004) stresses that other factors than BMI play an important role, namely the way fat is dispersed over the body. Especially visceral fat that leads to truncal obesity is an indicator for diabetes (Bhatia 2004). The waist hip ratio of a patient is therefore clinically relevant. The Health Council of The Netherlands therefore advises screening for DM2 in cases of overweight and obesity. The council expects a considerable health gain due to early diagnosis (Gezondheidsraad 2004).
Another more straightforward symptom is Acanthosis nigricans. This is the hyper pigmentation of the skin in the neck or other skin folds and presents itself in the clinic as a 'neck that remains dirty no matter how hard it is scrubbed' (Ramchandani 2004). Although Acanthosis nigricans has been described as rare, Yamazaki et al stresses that:

> "This condition occurs in 56–92% of overweight children and adolescents with type 2 diabetes. Acanthosis nigricans has a dramatic ethnic predisposition and in any ethnic group, individuals with Acanthosis nigricans show a clinical surrogate for laboratory-determined hyperinsulinemia and insulin resistance indicating the highest risk for type 2 diabetes", p704 (Yamazaki, Ito et al. 2003).

The fact that it is clearly visible and that it is highly correlated to insulin resistance makes it a good screening tool for DM2 (Stuart, Gilkison et al. 1998; Kerem, Guttmann et al. 2001).
Fungal infections of for example the vagina but also the skin are symptoms that motivate testing for diabetes when the overall picture is not clear.
If symptoms point to DM laboratory test are conducted. The diagnosis of diabetes is based on the fasting blood glucose level; the blood glucose level is tested after an overnight fast. The NHG standard indicates that two times an elevated glucose level confirms the diagnosis.

In the population: the risks
The classical distinction between DM1 and DM2 in terms of juvenile and adult onset has become untenable. Cases of obese teenagers, as described in Pinhas Hamiel and Zeitler (1999), who at first sight were diagnosed with type 1, but who later on proved to be insulin resistant, an indication of type 2 diabetes, show that diabetes is a rather complex disease. This also indicates that a difference between populations at risk for DM2 based on age is changing (Pinhas Hamiel and Zeitler 1999). But there are more population differences that are more or less shifting. A considerable part of the literature focuses on ethnic differences and DM2.
According to Harris (1990) for example, patients with a different ethnic background tend to have different symptoms and even different insulin sensitivity and secretion. In her study Harris has found that “Blacks” have not only a higher prevalence of DM2 than “Whites”, but also more complications. The rates of loss of vision, amputations and renal disease are 1, 5-4 times higher in “Blacks” than in “Whites”. She concludes that:

> “Risk factors for DM2, including age, sex, obesity and family history of DM all operate within both race groups and probably interact with each other. The effect of gender and family history on rates of DM is similar in Blacks and Whites. Blacks have higher rates of DM at each obesity level, indicating that obesity alone cannot explain the differential in prevalence between races” (Harris 1990).

Ku, Gower and others go even further in their conclusions on racial difference in insulin resistance. Based on their research they argue that neither body composition, fat distribution, cardiovascular fitness, nor the amount of physical activity could explain the difference in insulin secretion (higher in African Americans) and sensitivity (which is lower in African Americans) (Ku, Gower et al. 2000). Other studies identify other ethnic populations at risk. The classic examples are the Pima-Indians, but also South East Asians, Hispanic Americans, and Pacific Islanders, (Dabelea, Pettitt et al. 1999, Krosnick 2000, Grinstein, Muzumdar et al. 2003; McNeely and Boyko 2004). In The Netherlands immigrants such as Hindustan-Surinam, Moroccan, Turkish and other people from the Surinam are according to the RIVM more at risk to develop diabetes than people of Dutch origin. Middelkoop and others have identified South-Asians as a risk group. In their study this group consists of inhabitants of The Hague who are immigrants from Surinam and have an Indian ethnic background. They have found that DM is extremely common among this group. They

conclude that the differences in prevalence of DM2 between socioeconomic groups within this South Asian population are smaller than the difference in DM prevalence between South Asians and the Dutch population (Middelkoop, Kesarlal Sadhoeram et al. 1999).
Given these studies, it is tempting to conclude that genetics play a part in the prevalence of DM2. Weijers, Bekedam and their colleagues however stress that the prevalence of diabetes among immigrants from Turkey, Morocco, Surinam and the Dutch Antilles is much higher than that of populations in their countries of origin, which indicates that lifestyle plays a major role in DM2 (Weijers, Bekedam et al. 1998). They further indicate that, although prevalence of DM2 is higher under certain minority groups, obese children of Caucasian origin also have much higher prevalence of impaired glucose tolerance then previously thought. Moreover Riste and others have found a surprisingly high prevalence of DM2 in "Europeans". They did their research in Manchester and suggest a 'history of impoverishment' as one of the reasons for the unexpected high prevalence of DM2 in all ethnic groups that they have studied (Riste, Khan et al. 2001).
Next to age, ethnicity and socio-economic status also sex-differences are a focus in DM2 clinical research. The prevalence among the sexes does not differ. However, since DM2 prevalence increases in higher age group, the absolute number of female DM2 patients in The Netherlands is also higher. This is because women simply live longer than men do. In addition, different studies suggest that DM2 does have other effect in female patients than in the male patients. Johansen and Birkeland describe that male DM patients have a 2- to 4-fold risk for cardiovascular death compared to non-diabetic individuals, while the risk in female patients is 3- to 5-fold higher(Johansen and Birkeland 2003). The American Diabetes Association warns on its website that the risk for cardiovascular diseases is more serious for diabetic women then for diabetic men (http://www.diabetes.org/diabetes-statistics/women.jsp). It seems that the protective effect of female sex before menopause on cardiovascular diseases and mortality is lost in diabetes patients. Also the risk of diabetic ketoacidosis (DKA) is 50% higher among women than among men. Another risk for younger women is that of developing fertility problems due to Polycystic ovarian syndrome (Solomon 1999).

In medical standards: good care

There is great diversity in care. Equally, there may be a great diversity in good care. Medical standards and guidelines are concerned with that and directed, i.e. to a certain extent, at standardising and co-ordinating care practices. By implementing protocols and other organisational means, care should become more routinized and transparent for practitioners, patients and others. There are different DM2 guidelines for GP's (NHG Standaard) and hospitals (NDF/CBO Richtlijnen (NDF 2000)) in the Netherlands. Different studies however, report on problems of implementing these guidelines in GP-practice (Konings 1995; Konings 1996) and others on the implementation in hospital-practices. In their study of barriers for implementing guidelines in Dutch hospitals, Dijkstra, Braspenning and their

colleagues argue that these problems are related to the specificities of diabetes care. Whereas the health care system is traditionally organised around acute and episodic illnesses, diabetes is a complex multi-systemic chronic disease (Hiss 1996, Dijkstra, Braspenning et al. 2000). The guideline of NDF/CBO requires for example that nine different disciplines become involved in providing good care and that their activities for each and every patient are co-ordinated. In other words DM2 requires a different organisation of care, across disciplines.

In the above-mentioned guidelines attention is paid to the diagnosis, the symptoms, the medication and to groups that may be at a higher risk. In addition to the symptoms that we referred to above, also pregnancy diabetes and ethnicity appear on this list. According to the NHG standard (directed to the GP) as well as the Health Council of The Netherlands (Gezondheidsraad 2004) only Hindustan-Surinamese may be at a higher risk ('etnisch belast'). The literature, so the guideline indicates, is not conclusive about a higher prevalence in other ethnic groups. "Turks and Moroccans are, just like the Dutch, member of the Caucasian race; there is no evidence that diabetes mellitus type 2 is found more or less often among them than among autochthones" (NHG Standaard 2003, footnote 4). By contrast in the NDF/CBO Richtlijn, Moroccans and Turkish people are added to the ethnically higher risk group.

But guidelines do more than identifying populations at risk. In guidelines criteria are noted that determine who is diabetic and who is not. In 1999 the World Health Organization published recommendations on diagnostic values for blood glucose concentration (WHO 1999). The major change was a lowering of the diagnostic value of the fasting plasma glucose. The implication of this change is that under the new criteria more people are diagnosed as diabetic. Riste, Khan and others describe that the number of DM patient they find depends on which guideline they follow (Riste, Khan et al. 2001). They also stress that using fasting criteria alone according to their data creates difficulties. The postprandial glucose level (the blood glucose level taken 1 to 2 hours after eating), they argue, is more closely associated with cardiovascular disease and death across all ethnic groups. Also Cieriello et al stresses the importance of controlling postprandial glucose levels, since the control of hyperglycemia is an essential part of good clinical practice in diabetes. They suggest that: "controlling postprandial glucose levels can help to optimize metabolic control in diabetic patients and may be particularly important for the prevention of vascular complications' (Ceriello, Hanefeld et al. 2004). By contrast Buse emphasizes that routinely measuring and treating postprandial glucose is not desirable because setting reasonable targets is very difficult and because of the risk of hypoglycaemia (Buse 2003).

Next to the attention paid to groups at risk and determining who has diabetes mellitus, the standards point to the various different complications that may occur, such as the diabetic foot, bad eyesight, blood pressure and cardiovascular diseases in relation to either the blood glucose level or to smoking habits and lack of physical exercise. Since the medical professionals cannot by themselves prevent these complications, the guidelines place special emphasis on education and on involving the

patients in caring for their disease. In terms of professional care, the guidelines recommend that patients on oral medication should visit their GP on a regular basis, that they are referred to the ophthalmologist the latest six months after diagnosis, and that they should undergo a full physical check up once a year. Patients that take insulin should visit the diabetes nurse on a regular basis and the internist once in three months. Also these patients have to undergo a full medical check up once a year.

While trying to locate DM2 in different places we have also come across other versions of what it is. In the clinic it is a cluster of symptoms. These symptoms however also refer to different sites where it occurs, such as in the family of the patient, in the life of the patient where complaints manifest themselves and in the clinic where symptoms are recognized and diagnosed based on their characteristics or, on laboratory results. DM2 could also be located in the population, or better in populations at risk. In our discussion of the literature it became clear that population boundaries are shifting. Populations at risk are constantly being redefined. With these changes also DM2 is changing: from adult onset to something that may manifest itself at a younger age; it has also become something of body weight and the health risks that come with that, it has to do with demographic differences between the sexes as well as differences in the implications of DM2 for the sexes. And finally DM2 is also a matter of ethnicity. Its prevalence may differ in different ethnic populations. Different literature place different emphases on what DM2 is in these populations, it has to do with genetics or environment, with wealth or poor socio-economic conditions, with the country of origin and with the country where one lives. The diagnosis of DM2 and care for patients has to do with the organization of healthcare. The medical guidelines refer to that with reference to the hospitals as well as the physicians. Depending on whether care is well organized DM2 patients or populations at risk can be diagnosed at an early stage and cared for in a good manner. As we have seen the specificities and complexities of diabetes are a challenge for a good organization of diabetes care.

2.2.3 How is it handled?

As has become clear DM2 may be said to be a variety of different things. Many different worlds enter the clinic in order to diagnose the disease and to provide or organize care for the patient. We will therefore turn to how DM2 is handled according to the literature and view this first in medical guidelines and secondly in medical practice where health professionals and patients interact.

In medical guidelines and literature

The complexity of the treatment regime and the amount of responsibility that is placed on the patient, - to keep control of his or her blood sugar in combination with (a prospect on) the various and severe long term complications, - makes diabetes a very difficult disease for patient and

physician (Loewe and Freeman 2000). In type 2 diabetes the ruling notion is that "bad" dietary habits, obesity and lack of physical exercise play a major role in the onset of diabetes. The first advice to patients is therefore to change their dietary habits and to do physical exercises. Loosing weight is the only way to really cure diabetes type 2. Hence the emphasis placed on diet. One of the recurring questions is how attainable living healthy is in a modern society. (Hirasing, Fredriks et al. 2001) describe the rising trend of obesity among children. And Kreijl, Knaap et al. (2004) show that almost 10% of the Dutch population is severely overweight and that on average people eat too much fat and about half the recommended amount of fruit, vegetables and fish (Kreijl, Knaap et al. 2004). This problem is even more urgent with reference to adolescents with DM2, since they are a difficult group to counsel and to influence in terms of diet or life-style issues. This is a matter of concern especially because this group would have the greatest benefit by loosing weight.

In the literature there is an increasing awareness that DM2 is part of a cluster of metabolic problems. The underlying insulin resistance causes increased blood sugar but also hypertension, high cholesterol, obesity, and other imbalances of the fat metabolism. This cluster of disorders is called Metabolic syndrome (Doelle 2004; Fletcher and Lamendola 2004). By concluding that DM2 is part of larger problem the focus of the treatment and control has shifted (NHG Standaard 2003). The Dutch NHG Standaard emphasis that DM2 can no longer be seen as a merely an imbalance in the blood glucose level, and identifies different interventions that need to be performed simultaneously, a treatment of the syndrome in total: medication for diabetes, hypertension, as well as high cholesterol. Even though medication is crucial all these factor of the syndrome are related to (an 'unhealthy') lifestyle. Although in the clinic practitioners are aware of the difficulties to change established lifestyles, in almost all the literature the importance of physical exercise and a balanced diet is stressed (Harris, Petrella et al. 2003). People who have DM2 should live healthier, and those at risk, especially (young) persons with obesity and those with diabetes in the family, should be protected against developing DM2 (Pinhas Hamiel and Zeitler 2000, Tuomilehto 2003)

An early start of the treatment and prevention has become the main goal in various guidelines and policy documents. Medical guidelines, but also insurance companies, have placed physical exercise and education of the patients centre stage. One could say that DM2 is becoming an exemplar case with potentials for prevention based on dietary habit and other lifestyle components. A case in which preventive lifestyle measures play at least as big a part as physiology or genetics (van Dam 2003, Desiere 2004)

In medical practice: patients and professionals

In the medical practise, in the interaction between patients and clinicians control is a central goal. As indicated above the management of DM2 is highly dependent on the involvement of patients. Clinicians provide the patient with tools in the form of education and information, oral medication, insulin and blood glucose meters to be able to control the

blood glucose level. From there on, the ball is in the court of the patient and out of the physician's hands. They should try to be punctual with their medication, adjust their diet and exercise more. Freeman and Loewe have conducted a qualitative study focusing on the interaction between clinicians and patients and show that poor communication results 'from different conceptions of the disease and different treatment goals' (Freeman & Loewe 2000: 507). They describe the frustration of clinicians in caring for DM2 patients. Although clinicians seem to know best how to manage the disease, they do not have the tools to make the patients stick to recommendations. Compliance is therefore a major topic in diabetes care, because finding an insulin regime and stabilising the blood glucose level is a determining factor in the course of the disease. Moreover, failing to keep the blood glucose level stable may result in Hyperglycaemia ('hyper') or a Hypoglycaemia ('hypo'). Hyperglycaemia is a blood glucose level that is too high. When left untreated it can sometimes result in a ketoacidosis or diabetic coma. A low blood glucose level or Hypoglycaemia is a side effect of the DM treatment (notably too much insulin) and can also result in loss of consciousness. For patients however a 'hyper' and a 'hypo' are not only threatening and scary, they are also indications of their disease being "out of control", or, an indication of failure to control it. They encounter these in the form of e.g. headaches, paleness, tiredness, hunger, bad vision in a case of a hyper, or sleepiness, dry tongue, tiredness, and thirst in cases of a hypo.

The only way for physician to control the behaviour of the patient is by communication. Physicians are focussed on possible complications in the future and prevention of these. For patients however, despite the recommendations of the clinician, diet or exercises may seem superfluous 'in the pursuit of vague hypothetical outcome (e.g. avoiding nephropathy)' (Freeman & Loewe 2000: 507). In their communication with the patients clinicians try to find a balance between informing a patient about his or her disease and possible risks and complications on the one hand, and overemphasising future risks and thereby frightening the patient on the other. Although one of the tasks of a physician is to take away anxiety and supply hope, anxiety can also be a tool to communicate risks (Brown, Harris et al. 2002). The initial anxiety experienced by a patient at the time of diagnosis can be a good motivator to trigger behavioural changes. Also the switch from oral medication to insulin there or knowing somebody who is experiencing complications can motivate lifestyle changes.

Moreover, and related to patient-doctor communication, Loewe and others (1998) have found that doctors tend to focus on complications that are localised inside the body, such as nephropathy, neuropathy, and problems with the vascular system (Loewe, Schwartzman et al. 1998). Patients on the other hand are more focussed on the surface of the body; retinopathy or loss of limbs are mentioned in this respect. In addition, for them other problems also play an important role, such as sexual problems (loss of libido in males), or problems with exercising in relation to problems with their feet. The study of Loewe and Freeman (2000) has shown that for patients control is not restricted to e.g. blood glucose values, but also control over their lives, such as their social functioning and not being seen

as a disabled person, the prevention of stress and the coping with other emotions (Loewe and Freeman 2000). Controlling the blood sugar level is not always in par with controlling other aspects of one's life.

From our discussion of how DM2 is handled it has become clear that diabetes has changed from a specific disease to a syndrome consisting of a cluster of diseases. This implies that care and intervention have become directed not only to DM2 but to a number of related diseases. Moreover managing these diseases is not merely a matter of the right medication but highly dependent on lifestyle changes and education of the patient. Lifestyle has become a core definition of health and illness, and changing that of MD2 patients is a central goal in medical guidelines. In medical practice the caring for DM2 patients is dependent on good communication between clinicians and patients. Both aim at controlling the disease, but as we have seen what is to be controlled may differ between them. Also for the patient the disease is just one aspect of life in which many other things may become the objects of control. Finally just as in the medical guidelines, establishing lifestyle changes is central in the management of the disease.

2.3 Diabetes and diversity in practice

Since this review is concerned with diversity in clinical practice, we have decided to go out and see for ourselves how diabetes care is done. We started by contacting an internist in a general hospital in Amsterdam. After having interviewed him, he invited us to observe consultations with his patients (we have observed consultations with 11 patients) and he also helped us to organize interviews with diabetes nurses and with a general practitioner. As indicated above all of them are involved in the Diabetes Care Team. SH is a big general hospital and one in which diabetes care is rather well organized. Our purpose here is not to present a representative case. Rather we contrast our material to the literature and view whether and how the differences that we have traced above enter such a practice. To do so, we will first introduce the hospital and the practice of the physician. Then we will address the issue of diversity and diabetes by raising the questions: first "what is good care", and second "does diversity matter and how". Finally in our discussion we will go into the kinds of differences that enter clinical practices and what we might learn from these about the relevance of diversity.

Introduction of the sites

The general hospital, which we will refer to as SH, is quite involved in the improvement of diabetes care. They develop special courses for doctor's assistants and other medical personnel on complications such as cardiovascular diseases and diabetes. They have also conducted studies on the perception of diabetes by patients from ethnic minority groups and on their appreciation of the information provided about it. Also seven educational videos aimed at different groups of patients have been

produced. And recently the diabetes team is setting up a research project on strategies to make patients do more physical exercise. Moreover, in this hospital there are a number of diabetes nurses. Among them one is of Moroccan origin, and recently a diabetes nurse of Turkish descent has been appointed. The reason for this is that 50% of the diabetes patients in SH are of non-Dutch origin, 'allochtonen'. The group of Moroccan patients is the highest and suggests that the presence of the Moroccan nurse is related to that (Interviews with Diabetic nurses).
The high number of patients form ethnic minorities has brought many changes for the diabetes team and the patients in the hospital. One example of this is a study on the perception of diabetes among Moroccan patients (Malki 2002). The Dutch patients of the hospital were included in this study as the control group. Their knowledge of the disease was taken to be exemplary. However upon starting the research, which was based on questionnaires that included questions regarding knowledge of the disease, the researcher was shocked to realize that the knowledge of the Dutch-speaking patients about diabetes was presupposed, but did not match with what actually was going on. The presupposition was that language was the main obstacle in the communication about diabetes and care. The Dutch speaking patients were thus supposed to be well informed. And as a result of this surprise, the researcher decided to leave out the so called 'control group'. Upon this finding, the 'control group' itself became an object of intervention. The team had therefore decided to design a training course for Dutch speaking patients.
The general practitioner that we have spoken to is one of the five GP's that work in this centre. It is an interdisciplinary health care centre with physiotherapists, psychologists, a dentist, doctors' assistants as well as a so-called 'praktijkondersteuner' (a nurse specifically trained and appointed to take over some of the tasks of the GP's). It is also a multicultural centre where doctors and other personnel are able to communicate in various languages with their clients. The GP that we interviewed has been working in this health centre over the last 20 years and has 74 DM2 patients from at least twelve different countries of birth. Moreover, the health care centre works closely with hospital SH to which patients with complications are referred and where the nurses and other personnel follow special courses on diabetes.

2.3.1 What is good care?

Since diabetes is a chronic disease with a high risk for complications care is delegated to different actors, patients, professionals, medication and instruments. We have also seen that the organization of good care is an objective of many people involved. Next to proper and early diagnosis and medication, life style changes are placed high on the agenda. This we also found in SH and the health care centre. However, depending on which practice we observed the emphasis on what good care is was placed elsewhere.

In our conversations with the GP, the internist and the diabetes nurses, communication came up as a topic of concern. The problem was initially framed in terms of mastering of the Dutch language, however in the course of the conversation it became rather diffuse, or maybe a matter of styles in different practices. “Mastering the language does not imply understanding the message. We had this woman a couple of years ago, whom we told that she did not need to check [her blood sugar] any more. We had established an insulin regime, so we told her to stop checking (‘controleren’) and then she stopped to inject insulin all together” (interviews diabetes nurses). This woman did master the Dutch language, but for her ‘pricking’ for blood to check her blood glucose level was the same as injecting a syringe to bring in the insulin. In the patient’s life pricking (‘prikken’) had become connected to measuring the blood sugar level, whereas in the hospital ‘prikken’ is reserved for the insulin injection. This example does not do away with language problems, which exist in multicultural societies such as Amsterdam. Specifically since the highest number of patients is above 50 years old and may include a high number of immigrants (especially from Turkey and Morocco) who did not learn the language properly. As indicated above the staff of the health care centre is multicultural and speaks various different languages and in specific cases an interpreter may be phoned (‘tolktelefoon’) to help out. In the hospital they have appointed an Arabic and a Turkish speaking diabetic nurse. But also since they advise all patients to bring along somebody who is close to them, especially when the insulin regime has to be established, this person may function as an interpreter for those who do not speak the Dutch language well enough. So communication is for various reasons, and some of these will be elaborated on below, important for providing good care. But there are many more.

Given the dependence of clinicians on measures that patients will have to take themselves, such as living healthy, eating regularly, no stress, etc. etc., clinicians are pleasantly ironic about the limits of achieving optimal care. They are aware that they are not in charge and that the patient may give his or her best, but life is erratic and so is an every day disease such as DM2. Also in their advice on physical exercises, their point of departure is not an ideal one, but rather the patient’s every day life. So they try to establish compatibility between the recommendation and the every day practice of the patient.

The GP’s role is one of a gatekeeper. The symptoms are sometimes straightforward to the physician (such as sudden loss of weight, thirst, fatigue) and indication to screen for DM2. More often however the complaints are a-specific, and incite the GP to screen for DM2 especially if the patient belongs to one of the high-risk populations (see below). Once diagnosed the patient is set for oral medication and is enrolled in a chain of care. The patient receives a full check up, for cardiovascular complications, eyesight, the functioning of the kidneys, etc. S/he will also be referred to a dietician for advice, the assistant will provide information on living healthy and physical exercise, she will check the feet and advice on how to care for these. In some cases the patient may also be referred to a psychologist. One of the central issues of good care is compliance. This

holds not only for the medication, eating habits and physical exercise, but especially for showing up at appointments including the annual full check ups. The assistant will in such cases try to trace the patient and make him or her visit the centre.

Whereas the GP is concerned with the initial diagnosis and usually with patients with less or no complications, in the hospital the patients' disease is in an advanced state. They often have several diverse complications. The care in the consultation room is therefore organised around a cluster of diseases. And most of the internist's time goes into checking and adjusting the package of medication. We have attended eleven consultations in which going over the list of medication were central. The internist checks whether the patient still has and takes all prescribed medications (usually a cluster of seven different medicines), inquires about changes in the patient's health or complains and changes in behaviour. Upon this a change in medication will be suggested. The patients that we have witnessed were all very well informed about their medication and their knowledge about what they were taking was crucial for the internist. Especially since the patients see different specialist, the internist may not be up to date about what they have prescribed. Even though the medication is a central theme in the consultation, there are always elaborate (colloquial) talks about the every day life of the patient. So here is short excerpt and an example of such a consultation (see box next page).

Patient R:

Patient R. is not doing well. He looks exhausted as he walks into the room. He appeared to suffer from pinching nerves in his neck and his back is hurting. His left hand is cold, while his right hand has a normal temperature. While the patient was talking about a business dinner in a good restaurant, which he could not enjoy because of the pain, the doctor was going over his list of medication, and checking whether patient R. still has all of it. He takes medicines against gastric acid and still has a supply of that and is authorised to pick up his insulin without a prescription. The doctor writes prescriptions for his other medication, but not without inquiring whether the patient still takes these as prescribed. In between there is a talk about the patient's back and about 'manuele therapy'. After this the doctor goes into the diabetes problem.

D: A blood glucose level of 6.1 is very well ('keurig'). Any problems with hypo's?

P: No.

D: What about physical exercise ('bewegen')?

P: Problematic.

D: Cycling?

P: Also difficult.

Patient R. explains that his legs are hurting and that he can not walk in a stable way ('zwalken').

D: Where does it hurt?

Patient R. points to his legs: from the knees to the ankles. He also asks about loss of hair on his legs. He even called his ex-wife to verify whether he used to have hair on that part of his body in former times.

The doctor explains that hair loss is common in diabetes. He further inquires about his eating habits. The patient explains that he takes his evening meals in three steps.

D: Have you seen a neurologist.

P: That's why I am here.

D: And what about your feet, are they ok?

P: Yes. Patient explains that he wears special shoes.

The doctor explains that he still wants him to see a revalidation doctor. Since the patient takes painkillers, and since he is developing a neuropathy, which is causing the problems with walking, he wants to be on the safe side. Not that the patient hurt himself without noticing.

D: What about your eyesight?

P: I did have an operation on my left eye, but that is not connected to my diabetes.

In the consultation room of the internist the medication is central. But also 'walking' and hypo's (a low blood glucose level) were recurring themes. Upon, posing the question about hypo's, patients start to talk about their food habits in relation to physical exercise. Physical exercise may cause hypo's and the way to prevent that is to drink e.g. a grape sugar beverage, or to eat something. However, the patients are instructed to inject less insulin than usual when planning to go for a walk. That is the preferred advice for preventing hypo's not by taking extra food. This was specifically important in overweight patients.

Whereas patients see the internist once in three month, they visit the diabetes nurse with shorter intervals. Together with the patient the

diabetes nurse establishes an insulin regime. As indicated above the patient has to check his blood glucose level in relation to the amount of the insulin taken. Once a good regime has been establishes the nurse makes sure that it remains stable by seeing the patient on a regular basis. The diabetes nurses are also responsible for the education of the patient and may refer him or her to the dietician or podiatrist. The may also, in urgent cases, discuss the state of the patient with the internist and suggest to refer her to other specialists. For the diabetes nurses knowledge about a patient's every day life is crucial.

> "During consultation hours, not only the results of the blood sugar or other results that we have in the computer are important. It can be a conversation about a divorce that a daughter is going through. And that's relevant knowledge. Because sometimes we cannot explain why the regime is not stable or why it cannot be stabilized. Stress in the family or abuse all have an effect on the regime. The situation at home, the environment, the social setting are all very important [...]. Sometimes we do feel like social workers. But we try to help the patient, sometimes we refer them to a psychologist or social workers."

So, for the diabetes nurse in SH good care is dependent on knowledge of the patient's life. The monitoring of the blood glucose level on a regular basis is crucial for preventing extra complications. This is one of the reasons why they see the patients on a regular basis. However the blood glucose level, as well as cholesterol or other values, are not independent from a patient's life. They therefore keep that into focus and provide a large package of care. Related to this issue we also learned from the diabetes nurses that compliance is no guarantee for a good management of the disease. Many patients who do not comply with their medication and instructions appear to be able to keep their regime stable. And also an often-heard complaint, namely that non-Dutch patients often do not comply with their treatment, appears not to be in par with the experience of the nurses in SH. "And it is also not the case that 'allochtone' patients comply less than 'autochtone' patients. It is often claimed, but is not true" (interview diabetes nurses).

2.3.2 Does diversity matter and how?

Both in the hospital and in the practice of the GP diversity matters in the form of an availability of care given by different specialists. DM2 is a chronic disease and good care for the patient may differ over the long periods of time of illness. In the interviews we came across an interesting difference between the hospital and the GP-practice. Clinicians in both practices defined diversity as differences between ethnic groups or between the sexes. Whereas the GP saw diversity as an important tool to provide good care, the clinicians in the hospital seemed wary of diversity. We will discuss this difference at the end of this section.
The GP reported on a variety of difference-making practices. The interdisciplinary team in the health centre was one of these. But while talking about her patient we learned that she distinguished between

patients' whose regime can be established easily and patients for whom this is, or remains a difficult task. The reason for this difference may have to do with the patient's compliance to medication or instructions, or it may remain vague. Since the regime is her main tool for intervention and monitoring, this difference is a crucial one in her practice. She also reported on a difference between men and women: "women are difficult to supervise". The physical exercise instructions that patients receive are usually closely linked to the patient's every day life. Men however tend to be more outdoor, to walk more. It appears to be difficult to get especially Moroccan women to do more walking.

Next to these differences in the caring for and supervising of the patients, in the GP-practice other differences seem to be crucial for the diagnosis. For the GP populations at risk are crucial categories for identifying the symptoms and diagnosing DM2, especially in cases of a-specific complaints. For example, by contrast to patients with a normal body weight, patients who are overweight or obese tend to be screened for DM2. But also ethnicity is such a diagnostic tool. The prevalence of diabetes in different ethnic populations as described in studies based on clinical research on the one hand and the experience with diabetes among different ethnic groups in her practice on the other, make the GP more attentive to DM2 in specific populations. "I think it is important to be aware of these differences, because it also affects the measures that you take. And those differences are helpful for your diagnosis, right" (interview GP)?

Whereas diversity functions as a diagnostic tool in the GP-practice, in the hospital it is viewed with suspicion. In the interviews with the internist and the two diabetes nurses, the individuality of the patient and the equality in care was emphasized. "From what we read in the news papers, diversity seems like a cage where groups of people are fitted into. This is problematic because if you do that, you would have to adjust your goal [for different populations], which is not good. Also you produce the idea that certain groups of patients should probably have the same goals [in terms of treatment] which is not tenable" (interview Internist). An emphasis on diversity is thus seen as a threat to good treatment and care. Again the internist:

> "There are so many cultural differences, you see. There are cases where I cannot touch the arms of one Moroccan woman, and the next I get to investigate with her scarf still on, but her breasts naked. Without any problem. There are great differences. Any presupposition about how a patient would like to be treated, will prove me wrong. A patient knows when you are judgmental and not approaching them in an open way. You start to make mistakes. [...] If you focus on practical problems of patients you see that diversity in the true sense of the word hardly exists. It is much more an interpersonal thing" (interview Internist).

Yet in practice the hospital does take differences and diversity into account. As indicated above, they have taken special measure to come around the language problems. They have also special advises for Muslim patients who commit themselves to the Ramadan (the fasting month). And given the importance of physical exercise and differences that may exist

among different groups to do these, they have designed a sports project specifically aimed at Moroccan women. It seems however that the cultural specific measures that are taken are seen as tools to provide better care for the individual patient.

So there is a difference between the relevance of diversity in the practice of the GP and that of the hospital. It is fore grounded in the GP-practice and back grounded in that of the Hospital. As indicated the GP is the gatekeeper of the healthcare system. The task of a GP is to provide adequate diagnosis based on more or less oblique complaints of patients. The GP therefore tends to take diversity in terms of populations at risk more seriously. A specific cluster of symptoms or vague symptoms in a risk population are therefore more often worth a test for DM2.
By contrast, at the diabetes out-patient clinic, most patients who appear there are already diagnosed with diabetes. They are sent there because of complications or because it is difficult to establish a regime for them. Diversity as means to determine between populations at risk is not relevant there. Rather the individuality of the patient is placed central. Treating each patient as an individual case is as we saw important for the treatment and the interaction between patient and clinician. While emphasizing individuality, as a means to provide equally good care for every body, diversity appeared to be important as well. However, not so much in the form of population at risk, but rather as an object of intervention as to provide equal care for every body. Populations, or ethnic groups, who do not or can not comply with the treatment because they do not master the Dutch language or because of specific lifestyle habits, become objects of intervention as to enrol them, as individuals, in the management of their disease.

2.3.3 Discussion

This review is part of a cluster of reviews aimed at investigating the relevance of diversity in clinical research. In common knowledge diversity is viewed as a cluster of variables (in practice ethnic, age and sex-differences) that independently or in conjunction may have an effect on clinical results. Differences so it seems lay there for the researcher to include these or leave them out in her study. In this review however, our aim was to question this taken for granted notion of diversity. Our example was diabetes mellitus type 2. We have investigated the literature and clinical practice in an ethnographic mode and asked: “which differences are made to matter and where in diabetes care?” Thus, instead of taking sex, age and ethnicity as natural and static categories, and instead of presuming their relevance in research, we have taken an open view towards our material and looked for the relevance of these as well as other categories of difference in it. Broadening the category of diversity on the one hand, and taking its specificities into account on the other, is, as we have shown, crucial for providing good care.

The dynamics of diversity

The largest part of our paper has focused on published material. Let us first try to outline the relevance of our findings for clinical research and diversity, and than turn to the clinical practice that we have investigated. Our analysis of the literature has shown the complexity of diversity. Diversity is in itself highly diverse and dynamic. Here we will address five types of shifts that diversity has undergone in research and practice of diabetes care.

Diversity is about similarities and differences and about clusters of variables to identify these. However, as we have seen identities of and the relation between variables are uncertain. The relation between genes and environment, between ethnicity and socio-economic status are not stable, and their outcome is a matter of context. While in some populations a genetic basis for DM2 may be determined, in many others this is not. And for example the prevalence of DM2 among people from Moroccan descent in the Netherlands may be higher than that of Moroccans who live in Morocco. In this case, what first seems to hint at ethnicity or genes, turns out to be a matter of environment and socio-economic status.

A second indication of the dynamics of diversity can be found in the fact that the relevance and meaning of variables changes over time and context. Age had long been a key-variable for understanding DM2 and for separating DM2 from DM1 patients. However, recent literature shows that this distinction does not hold any more, since more and more youngsters are diagnosed with DM2. Nowadays, physiological processes rather than age are central in understanding DM2. The relation between variables and context is clearly present in the examples of sex-differences and lifestyle. Although sex-differences are not relevant in terms of prevalence of DM2, they are relevant for the complications that come with the disease. The number of women that suffer from complications is higher because more women live longer than men. Moreover, in the treatment of DM2 patients, 'lifestyle' is more dominant than physiology, genetics or any other factor.

A third way of understanding the shifts in diversity that matters is by looking at new differences that have become relevant over time. Obesity is a case in place. It has become a variable that initially was not there. In the diagnosis it has effectively back grounded age. As we have shown it is recommended that also young people who are overweight are screened for DM2. But also obesity itself has been refined. The danger of obesity is not only a matter of body mass index but especially of the location of fat on the body. The reviewed literature thus shows a complex interaction between sex, age and ethnicity in relation to obesity and risk for DM2.

A fourth way of understanding the dynamics of diversity and the way differences matter is by looking at shifts in diabetes itself. As had become clear DM2 is not singular but rather part of a cluster of diseases. The fact that DM2 cannot be reduced to one symptom indicates that the care provided to patients is also delegated to different professionals and practices. From the everyday life of the patient and her family to the room of the practitioner to the laboratory where tests may be conducted, other emphasis may be placed and other differences may come to matter. Thus

the dynamics of diversity is also related to the complexity and the context of the disease itself.
The fifth hint at the dynamics of diversity had to do with differences between 'textbook medicines' and clinical practice. Taking clinical practice into account has for example shown that it matters for patients whether they are insulin dependent or not. This is related to how they perceive of and deal with their disease in every day life. Whether or not insulin dependent makes a difference for where a patient is treated and how DM2 affects her life. The clinical practice also hinted at a much-debated issue, namely compliance. As had become clear, compliance did not simply have to do with knowing the language of the doctor, but rather with understanding the disease and the treatments suggested. Moreover, compliance does not give a guarantee for a successful treatment, as the literature seems to suggest. Life events turned out to influence the glucose level and knowledge about theses in the clinic is crucial for providing good care. Thus differences in the lives of patients make crucial diagnostic tools.

Taking diversity in practices seriously
So what are the lessons that could be learned from this for a clinical research that takes diversity into account and that wants to be relevant for clinical practice?
Let us first consider clinical practice to then address the relation between practice and research.
Diversity is not stable and cannot be reduced to a number of fixed variables. The relevance of one version of diversity or the other varies depending on the context. In the hospital, as well as in the GP-practice the diversity in patients, how they experience and manage their disease and the problem of compliance are present and relevant. However, given the difference that we have found between hospital and GP-practice it seems important to attend to different versions of diversity. For example, diversity in the hospital has the individual patient as its starting point for diabetes care, whereas the starting point for the GP is the population (at risk). Clinical studies on DM2 and the prevalence of it in different populations provide important knowledge, or diagnostic tools, for GP's. They have a signal function. However these studies are not by themselves. Also the experience of a GP with DM2 and the clustering of its prevalence in populations in her practice are relevant tools in the diagnosis. Countries of origin are used as way to distinguish between populations at risk. Body weight is second tool of distinguishing between populations at risk.
The relevance of the notion of a population at risk is different for the practice of a GP and that of the hospital. Being a gatekeeper in medical healthcare, the GP will have to make the initial diagnosis of diabetes, and more often than not this is based on rather vague complaints of the patient. In this context knowledge about populations at risk is crucial and may suggest testing for DM2. By contrast at the diabetes out-patient clinics patients are usually already diagnosed with DM2 and this notion becomes less relevant. In the hospital it therefore seems more crucial to take the individuality of the patient into account in order to provide good care. Diabetes is placed in the complex life of the patient, and knowledge

about a patient's every life and changes in that are pivotal to do so. In addition, also cultural differences between groups of patients are crucial. Extra measures may have to be taken in order to provide care. Differences in the form of language spoken, education, country of origin, culture, difference between the sexes, eating habits (including fasting), lifestyle, may become relevant categories for intervention and for specific programs. This indicates that in medical practice other differences may matter than those that can be found in the literature. Based on 'practical knowledge' practitioners add relevant versions of diversity that do not receive much attention in clinical research. Moreover, this clinical knowledge is not clearly articulated. Yet, the struggle and endeavour of practitioners is critical for identifying 'differences that matter' and putting these on the research agenda. The question is how to make them part of clinical research. Our paper suggests that ethnographic research of clinical practice may help articulate the kinds of diversity that are relevant to clinicians and patients, and help to develop methodological strategies for making them part of clinical studies.

Despite the relevance of knowledge on population level, in clinical practice it is crucial to keep an eye on the uniqueness of every single patient. This indicates that diversity is endless and fluid by nature. Does this mean the end of clinical research? Does this mean that research based on population studies is no longer relevant? Our answer to these questions is negative. In patient care results from clinical research has to be targeted to individual patients. If the study populations are more similar to the patient in front of the caregiver, the application of the findings is easier. Therefore research in more diverse populations is needed. However, instead of fixing the variable of difference that should be taken into account for such research, we should opt for a higher variety of these. Diversifying clinical research itself, in terms of methods applied, categories compared and variables studied, makes the knowledge produced even more relevant for clinical practice and eventually for the care provided to patients. As we have suggested above, taking the knowledge produced in the clinic may help diversify clinical research.

In clinical practice both an individual and a population approach are important. Specifically for knowledge about 'populations at risk', practitioners rely on clinical research. Our examination of the literature showed different diagnostic tools may affect the prevalence of the disease in the population. Thus population at risk and the risks themselves are therefore dependent on technologies and research. This suggests that it is important to keep a close eye on how in clinical research populations and differences between them are produced. What differences are made to matter, and are these the relevant ones? To do this, ethnographic approach to clinical research may provide insight and help to establish links between research and clinical practice. The increasing attention in clinical research for ethnicity and social-economic status as diversity markers is a good start to bring clinical research closer to clinical care. However, the diverse patient population in the current Dutch society is characterized by more markers, which should also be recognized in clinical research. Attending to a more fluid notion of population prevents a naturalization of differences

and insures an individual approach to care and the treatment of the patient.

References

Black's Medical Dictionary (1999) 39th edition ed. by Macpherson G.

Bakker K and HJ Bilo (2004). "Diabetes care in The Netherlands now and in the future." Practical diabetes international : the journal for diabetes care teams worldwide 21(2): 88-91.

Bhatia V (2004). "IAP National Task Force for Childhood Prevention of Adult Diseases: insulin resistance and Type 2 diabetes mellitus in childhood." Indian Pediatr 41(5): 443-57.

Brown JB, SB Harris, et al. (2002). "The role of patient, physician and systemic factors in the management of type 2 diabetes mellitus." Fam Pract 19(4): 344-9.

Buse JB. (2003). "Should Postprandial Glucose Be Routinely Measured and Treated to a Particular Target? No!" Diabetes Care 26(5): 1615-1618.

Ceriello A, M Hanefeld, et al. (2004). "Postprandial Glucose Regulation and Diabetic Complications." Arch Intern Med 164(19): 2090-2095.

Dabelea D, DJ Pettitt, et al. (1999). "Birth weight, type 2 diabetes, and insulin resistance in Pima Indian children and young adults." Diabetes Care 22(6): 944-50.

Desiere F. (2004). "Towards a systems biology understanding of human health: Interplay between genotype, environment and nutrition." Biotechnol Annu Rev 10: 51-84.

Dijkstra RF, JCC Braspenning, et al. (2000). "Perceived barriers to the implementation of diabetes guidelines in hospitals in The Netherlands." N J Med 56(3): 80-85.

Doelle GC. (2004). "The clinical picture of metabolic syndrome." Postgrad Med 116(1): 30.

EhtishamS, AT Hattersley, et al. (2004). "First UK survey of paediatric type 2 diabetes and MODY." Arch Dis Child 89(6): 526-9.

Elbein SC. (2000). "Genetics of type 2 diabetes: an overview for the millennium." Diabetes Technol Ther 2(3): 391-400.

Elbein SC and GL Wolff (1997). "The genetics of human noninsulin-dependent (type 2) diabetes mellitus." J Nutr 127(9): 1891S.

Fletcher B and C Lamendola (2004). "Insulin Resistance Syndrome." J Cardiovasc Nurs 19(5): 339.

Freeman J and R Loewe (2000). “Barriers to Communication About Diabetes Mellitus.” J Fam Pract 49: 507.

Gezondheidsraad (2004). “Screening on type 2 diabetes.” Den Haag: Gezondheidsraad.

Grinstein G, R Muzumdar, et al. (2003). “Presentation and 5-Year Follow-Up of Type 2 Diabetes mellitus in African-American and Caribbean-Hispanic Adolescents.” Horm Res 60: 121-126.

Harris MI. (1990). "Noninsulin-dependent diabetes mellitus in black and white Americans." Diabetes Metab Rev 6(2): 71-90.

Harris SB, RJ Petrella, et al. (2003). "Lifestyle interventions for type 2 diabetes. Relevance for clinical practice." Can Fam Physician 49: 1618-25.

Hirasing RA, AM Fredriks, et al. (2001). "Toegenomen prevalentie van overgewicht en obesitas bij Nederlandse kinderen en signalering daarvan aan de hand van internationale normen en nieuwe referentiediagrammen." Ned Tijdschr Geneeskd 145(27): 1303-1307.

Hiss RG. (1996). "Barriers to care in non-insulin-dependent diabetes mellitus." Ann Intern Med 124(1): 146.

Jenkins AB and LV Campbell (2004). "The genetics and pathophysiology of diabetes mellitus type II." J Inherit Metab Dis 27: 331-347.

Johansen OE and KI Birkeland (2003). "Preventing macrovascular disease in patients with type 2 diabetes mellitus." Am J Cardiovasc Drugs 3(4): 283-97.

Kerem N, H Guttmann, et al. (2001). "The Autosomal Dominant Trait of Obesity, Acanthosis nigricans, Hypertension, Ischemic Heart Disease and Diabetes Type 2." Horm Res 55(6): 298-304.

Konings GPJM. (1995). "Waarom werken huisartsen niet volgens de NHG-Standaard Diabetes Mellitus Type II?" Huisarts Wet 38: 602-607.

Konings GPJM. (1996). "Implementatie van een standaard." Huisarts Wet 39: 560-564.

Kreijl CF, AGAC Knaap, et al. (2004). "Ons eten gemeten. Gezonde voeding en veilig voedsel in Nederland." Bilthoven, RIVM: 365.

Krosnick A. (2000). "The diabetes and obesity epidemic among the Pima Indians." N J Med 97(8): 31-7.

Ku CY, BA Gower, et al. (2000). "Racial differences in insulin secretion and sensitivity in prepubertal children: role of physical fitness and physical activity." Obes Res 8(7): 506-15.

Loewe R and J Freeman (2000). "Interpreting Diabetes Mellitus: Differences between Patient and Provider Models of Disease and their Implications for Clinical Practice." Cult Med Psychiatry 24: 379-401.

Loewe R, J Schwartzman, et al. (1998). "Doctor talk and diabetes: towards an analysis of the clinical construction of chronic illness." Soc Sci Med 47(9): 1267-1276.

Malki F (2002). "Marrokaanse Diabetespatiënten." Project Marokkaanse Diabetes Patiënten Slotervaartziekenhuis Amsterdam.

McNeely MJ and EJ Boyko (2004). "Type 2 diabetes prevalence in Asian Americans: results of a national health survey." Diabetes Care 27(1): 66-9.

Middelkoop BJC, SM Kesarlal Sadhoeram, et al. (1999). "Original Articles - Diabetes mellitus among South Asian inhabitants of The Hague: High prevalence and an age-specific socioeconomic gradient." Int J Epidemiol 28: 1119-1123.

NDF (2000). "Richtlijnen en Adviezen voor goede diabeteszorg." Utrecht.

NRC (2004). "Minister: hulp aan diabetici moet beter. NRC Handelsblad." Den Haag 6 Juli: p3 Binnenland.

Pinhas Hamiel O and P Zeitler (1999). "Clinical problem-solving. The importance of a name." N Engl J Med 340(18): 1418-21.

Pinhas Hamiel O and P Zeitler (2000). "Who Is the Wise Man?-The One Who Foresees Consequences:" Childhood Obesity, New Associated Comorbidity and Prevention." Prev Med 31(6): 702-705. Review.

Ramchandani N. (2004). "Type 2 diabetes in children. A burgeoning health problem among overweight young Americans." Am J Nurs 104(3): 65-8.

Renders CM, DWHA. Delemarre van, et al. (2003). "Insulin resistance and diabetes type 2 in overweight children." Ned Tijdschr Geneeskd 147(42): 2060-3.

Riste L, F Khan, et al. (2001). "High Prevalence of Type 2 Diabetes in All Ethnic Groups, Including Europeans, in a British Inner City: Relative poverty, history, inactivity, or 21st century Europe?" Diabetes Care 24(8): 1377-1383.

Rutten GEHM, S Verhoeven, et al. (2003). NHG-Standaard Diabetes mellitus type 2, NGH. 2004.

Solomon CG. (1999). "The Epidemiology of Polycystic Ovary Syndrome: Prevalence and Associated Disease Risk -." Endocrinol Metab Clin North Am 28(2): 247-264.

Stuart CA, CR Gilkison, et al. (1998). "Acanthosis Nigricans as a Risk Factor for Non-Insulin Dependent Diabetes Mellitus." Clin Pediatr 37(2): 73-80.

Ten S and N Maclaren (2004). "Insulin resistance syndrome in children." J Clin Endocrinol Metab 89(6): 2526-39.

Tuomilehto J. (2003). "Reducing coronary heart disease associated with type 2 diabetes: lifestyle intervention and treatment of dyslipidaemia." Diabetes Res Clin Pract : Suppl 61: S27-34. Review.

van Dam RM. (2003). "The epidemiology of lifestyle and risk for type 2 diabetes." Eur J Epidemiol. 18: 1115-1126. Review.

Wabitsch M, H Hauner, et al. (2004). "Type II diabetes mellitus and impaired glucose regulation in Caucasian children and adolescents with obesity living in Germany." Int J Obes Relat Metab Disord 28(2): 307-13.

Weijers RN, DJ. Bekedam, et al. (1998). "The prevalence of type 2 diabetes and gestational diabetes mellitus in an inner city multi-ethnic population." Eur J Epidemiol 14(7): 693-9.

WHO (1999). World Health Organization department of Noncummunicicable Disease Surveillance: Definition, Diagnosis and Classification of Diabetes Mellitus and Its Complications Part 1: Diagnosis and Classification of Diabetes Mellitus. Geneva, World Health Org.

Wiegand S, U Maikowski, et al. (2004). "Type 2 diabetes and impaired glucose tolerance in European children and adolescents with obesity -- a problem that is no longer restricted to minority groups." Eur J Endocrinol 151(2): 199-206.

Yamazaki H, S Ito, et al. (2003). "Acanthosis nigricans is a reliable cutaneous marker of insulin resistance in obese Japanese children." Pediatr Int 45(6): 701-5.

Review 3

Socio-cultural and political factors which facilitate and constrain representation and analysis of diversity in clinical research

Anita Hardon[3] and Eline van Haastrecht

[3] I thank the members of the research team, including Nicolien Wieringa, Amâde M'charek, Karien Stronks, Niek Klazinga, Patrick Bossuyt, Andrew Herxheimer and Martin Offringa for valuable comments on earlier drafts. It has been a pleasure to work in this multidisciplinary team. Trudy Dehue did an excellent review of the final draft at our Expert Conference, and sent me very interesting references which helped write the final version of this paper, and I thank Charles Medawar for many editorial suggestions and inspiration.

Contents

3

Socio-cultural and political factors which facilitate and constrain representation and analysis of diversity in clinical research

Summary

This paper deals with the way in which diversity in efficacy and safety of therapies is dealt with in randomised controlled trials, by scientists conducting studies, by the agencies that fund them, and by those involved in interpreting and presenting the data. The paper explores the socio-cultural assumptions and political-economic forces that constrain consideration of heterogeneity in clinical research.

Key points include:

- RCTs are a gold standard for evidence production in medicine. Their contribution to a better understanding of diversity in treatment outcomes is limited, because of an orientation towards comparing outcomes by artificially composed treatment groups.
- Legislation in The United States, resulting from an intensive lobby by women's health and minority groups, requires that women or members of minority groups be included as subjects in trials and that the trials be designed and carried out in a manner sufficient to provide for valid analysis of the variables studied.
- Trial populations (N = 51 recently published trials on hypertension, AIDS, Diabetes and Epilepsy) tend to be relatively heterogeneous in terms of age and sex, but less so for ethnicity, and sub-group *analysis* is rarely done. The exception is hypertension, for which around one-third of the trials reviewed included subgroup analysis for age and sex, and 18% for ethnicity.
- Standard treatment guidelines do not systematically present evidence on diversity in treatment outcome by age, sex and ethnicity, despite available data. As a result health care practitioners are not alerted to diversity issues that matter.

- Clinical researchers should take as point of departure that biological (including genetic), social, cultural, economic and environmental factors interact and co-produce efficacy and safety of medicines, and health of individuals. This conceptualisation broadens the diversity research agenda.
- A mix of research methodologies can be used to filter from etiological studies, from the realities of health care practice, as well as from ongoing RCTs and meta-analyses *diversity issues*, which need to be further explored. The diversity hypotheses, emerging from this range of studies, can be tested in future more focused RCTs.
- A diversity research agenda does not only require more funding, it requires a change in the kinds of questions addressed in clinical research.
- More attention for diversity in benefits and risks of medicines in clinical trials is constrained by political and economic factors: it is not in the interest of the pharmaceutical companies that fund most RCTs, as findings may limit markets for their products. Systematic consideration of diversity in efficacy and safety studies of medicines implies that regulators would need more time to study the registration dossier. Under the current fee-for-service payment structures, it is not in the interest of drug regulators to demand more complicated dossiers. Insurers pay for adverse health outcomes. They are therefore potential allies in the implementation of methodological reforms aimed at identifying diversity issues that matter.
- Health care practitioners and users have much to contribute and gain from more attention for diversity in clinical research. The mix of methods proposed allows for more systematic review of their experiences and views, providing evidence through which health outcomes are likely to improve.

3.1 Introduction

Twenty years ago, AH was confronted with the concept of *hiyang*, in the course of fieldwork on the misuse of medicine in self-medication in the Philippines. In an interview on childhood illnesses one mother reported on the asthma case of her son:

> "One week ago I went to the doctor with him, in the provincial hospital. The doctor prescribed Ventolin. I bought it in the pharmacy in town. It cost me 32 pesos. Ventolin is expensive. I gave it to him, but he did not get better. Probably it is not hiyang for him. It is hard to find a suitable drug for this small boy" (Whyte et al. 2002, p 28).

People in The Philippines use the concept of *hiyang* when a drug does not work. They attribute this to the relationship between the medicine and the individual: the medicine is not compatible with the patient. The informants in this study took individual differences in efficacy into consideration, in contrast to the tendency of biomedical practitioners to consider efficacies to be universal.

Nearly twenty years later, Dr. Allen Roses, a senior executive of GlaxoSmithKline (GSK), told a scientific meeting in London that the "vast majority of drugs only work in 30 or 50% of people." He cited therapeutic efficacy rates ranging from 25% in oncology to 60% in diabetes and asthma. These findings were reported on the front page of the *Independent* newspaper on December the 8 2003. The biggest drug company in Europe, GlaxoSmithKline (GSK), subsequently saw its share price fall last, despite unveiling details of its "product pipeline" for the next five years. In a subsequent GSK background paper, Roses explains that most drugs work in less than one in two patients mainly because the recipients carry genes that interfere in some way with the effects of the medicine. He reported:

> "Not every drug will work for everybody. This should come as news to no one. Most people have had the experience of going to the doctor and getting a medicine and having to go back and try another one" (GSK 2004, 51).

This paper deals with the way in which diversity in efficacy and safety of therapies is dealt with in clinical research, both by scientists conducting studies and by agencies funding them. Beyond describing the constructions of diversity the paper aims to explore the socio-cultural assumptions and political-economic forces that constrain and could facilitate consideration of heterogeneity in clinical research. To do so it first discusses the methodological characteristics of RCTs and their history. It then describes how in the late 1980s the need to include `women and minorities` and later children in trials became an issue of public debate. The debate led to legislation in the United States requiring women and member of minority groups to be included as subject in trials to allow for valid sub-group analysis. The paper assesses the implementation of these policy reforms by reviewing the design of 51 recently published randomised controlled trials, and assessing the attention for diversity in evidence-based guidelines for treatment.

3.2 The origins and key characteristics of RCTs – moving beyond individual judgements

Randomised controlled trials (RCTs) represent the gold standard in present day evidence-based medicine, and form the basis of modern medical practice. They are used not only to assess efficacy and safety of medical technologies, but also to evaluate health services and community health interventions, on the basis that RCTs produce better evidence than observational studies. Our discussion of the origins, characteristics and limitations of RCTs is therefore of importance to heath research in general.

How is diversity dealt with in RCTs? The development of the randomised controlled trial (RCT) in the 1940s and 1950s was, according to the official histories of this technique, a step that moved the art of healing from a pre-scientific to a scientific stage. It eliminated the subjective element in the evaluation of new treatments, and replaced it with

quantitative and objective data, radically separating the "hard" scientific aspect of healing from its "soft" social and cultural aspects. In order to establish RCTs as gold standard health professionals started working with statisticians to construct objective measures of the clinical status of patients (such as results of laboratory tests), to replace the anecdotal evidence based on individualized and embodied skills of the practitioner (Marks 1997).

Dehue (2001) dates the turn towards experimentation in the late 19th century, when the British statistician, philosopher and economist, Stanley Jevons, argued that crucial policy interventions, such as the legislation of the free trade of beer (to drive back the over-consumption of gin) should be tested before being brought into operation (Jevons 1880). She shows how subsequently in The United States, economists, psychologists, political scientists and sociologists developed methodological tools to provide governments with facts, which could replace interpretations and views in decision-making. Key to such experiments were comparisons of artificially composed intervention and control *groups*, an idea developed from the writings of the statistician Fisher. Fisher emphasized logical reasons for the randomisation needed for the reliable application of is statistical significance testing techniques. Objectivity increasingly became synonymous with *blindness* in experimental conditions, to exclude the possibility of free interpretation and personal judgement. Doing research for social policy purposes became equivalent to *standardized* methodological procedures. Individual interpretations and assessments were increasingly distrusted for their idiosyncrasy and likely pollution by self-interest (Dehue 1999, 2004).

The *British Medical Journal* recently celebrated the 50th Anniversary of RCTs in medicine. The journal claims to have published the first medical article, which explicitly described the RCT methodology, in 1948. It was a trial of streptomycin in pulmonary tuberculosis (*BMJ* 1998) that was designed, according to the historian Yoshioka (1998), to help clinicians decide who got access to the treatment when only a tiny amount of the drug was available. In the early 1960s the Thalidomide[i] disaster had a profound impact on the use of RCTs in government decisions on licensing and marketing of drugs. Safety of medicines became an issue of public concern. In The United States, Kefauver-Harris amendment imposed efficacy testing of new drugs using RCTs before granting of a marketing licence. Since 1962, the US Food and Drug Administration has been required to routinely evaluate the efficacy and safety of new drugs. Industrialised countries later followed The US example and also set up drug regulatory mechanisms, requiring pharmaceutical companies to provide scientific dossiers including pre-clinical studies and RCTs, to give reliable evidence of the safety and efficacy of their products. New drugs are not allowed to be marketed until such approval has been obtained. Moreover, after a drug comes to market, post-marketing surveillance is done to monitor any rare or long-term effects, and comparative efficacy trials may

also be done, to compare the benefits of a new compound with existing treatments.

Randomization, comparative testing and 'blind' assessment and are the three basic requirements in clinical trials:

- Randomization means that the test and control groups are selected in a manner which ensures an unbiased distribution of participants in the trial. Thus, a homogenous composition of participants in both groups can be expected.
- Comparison in trials involves one (test) group treated with the method or product being assessed, while another (control) group is treated with a competing product or placebo.
- In a 'single-blind' study, trial participants are unaware of the treatment they get; in a 'double-blind' study, neither participants nor their doctor know who receives what.

Criteria for inclusion are also typically used to determine which people can participate in the trial. Criteria for exclusion might, for example, identify pregnant women or someone with a liver disease. These criteria are important in establishing the clinical relevance of trial results, as are the reasons why participants withdraw from trials.

Trial design, written up in a formal 'protocol' is crucial. The protocol should define meaningful and appropriate outcome measures before the trial begins; researchers define these outcome measures and end points based on their preliminary understandings of the mechanism by which the compound works. The protocol should take account also of statistical factors, to ensure that any differences found between the control and test groups are real, and not due to chance.

Generally three phases of trials are done, following pre-clinical research (including long-term toxicity studies in animals). Phase one clinical trials are conducted to assess toxicity and the compound's minimum effective dosage in humans. Such trials usually include upwards of 20 volunteers and may last from one to two years. Then, the first check on efficacy and toxicity in humans is done with a few hundred patients in phase two trials. After these trials are completed, larger, phase three clinical trials are carried out, involving probable a few thousand patients. This larger study sample makes it possible to identify and begin to quantify the range of different responses to a drug, and often to compare results with other treatments.

3.3 Limitations of clinical trials

Challenges to this conventional approach to clinical trials have come from within the field of clinical medicine and beyond (Hansen and Launso,1989; Kaptchuk, 2001). The relevance of the results of controlled clinical trials for routine clinical practice is also increasingly questioned within the field

of clinical epidemiology. It is stressed that under routine clinical circumstances, inclusion and exclusion criteria cannot be applied for obvious ethical and therapeutic reasons. More pragmatic approaches are called for, incorporating the "heterogeneity, occasional or frequent ambiguity and other messy aspects of ordinary clinical practices" in the trial design (Feinstein 1983, 545). Outside the scientific arena, AIDS advocates have challenged the routine conduct of clinical trials most vocally. They have questioned the need for (and ethical basis of) randomisation procedures and have opposed the application of inclusion and exclusion criteria, defending the fundamental right of AIDS patients of access to treatments that could possibly save their lives (Epstein 1997). The list of limitations of RCTs is long. One important limitation of randomised controlled clinical trials is that they are done in "controlled" settings. This means that the methods' effects in less "controlled" settings cannot be predicted. The use of inclusion and exclusion criteria further has as a consequence that the results of clinical trials cannot be extrapolated to groups excluded from the trial. A further limitation of clinical trials is that the effects that are to be monitored in the trial are generally determined by the researchers, and may not be strictly relevant to actual clinical outcome, nor patient experience. Scientists devise the framework to evaluate safety and efficacy of medicines -- health effects that do not fall within this framework are generally not measured. As a result, once medicines are on the market, we are often confronted with unexpected adverse effects – effects not measured/observed in the initial trials.

Another limitation relates to the duration of trials: relatively few trials continue for more than a year, making it impossible to assess long-term effects[ii]. Lack of statistical 'power' due to limitations in number of subjects enrolled in the trial may further limit the significance of findings -- less common safety problems may go undetected. Moreover, because of the limited statistical power, analysis of differences in effects and safety within the intervention groups are rarely done, as this paper later describes. Another limitation is that RCTs are based on comparison of *artificially composed* groups. The underlying assumption is that medicine efficacy is best studied in unconnected individuals. Dehue has questioned this assumption in a case-study on a Dutch RCT testing the effects of free heroin provision to heroin abusers. She argues that the researchers had problems recruiting heroin users into the trial, because they were recruiting randomly assigned individuals to attend a clinic-like provision of heroin, thus underestimating the importance of social relations and the culture of heroin use (Dehue 2002, 2004). Finally, though developed as objective measures of efficacy, in practice the measurements used in RCTs are value-laden. Richards (1988) stresses that "judgements about experimental findings are necessarily contingent upon the professional values and interests of the adjudicating community, and may be structured by wider social interests such as consumer choice and market forces. We may infer that the very notion of efficacy is politically defined and defended, and the practical success of a therapy is asserted and sustained by the power of the interests and the sponsors that maintain it" (Richards 1988, 685).

The idea that therapeutic evaluation is inherently a social and political process becomes most apparent when researchers contest safety and efficacy claims, as Richards has shown in relation to the controversy over the efficacy of Vitamin C.

Despite this long list of limitations, findings from randomised controlled trials are at the core of 'evidence-based medicine' (EBM), a trend that emerged a little more than a decade ago. EBM encouraged physicians to apply the best available evidence in their medical practice, aiming to reduce the emphasis on unsystematic clinical experience (Guyatt et al. 2004). The emergence of evidence-based medicine has been facilitated by the systematic reviews reported by the Cochrane collaboration, and by the availability of many different databases and electronic tools to help searches for evidence. Not surprisingly, given the limitations addressed above, evidence-based medicine faces challenges in practice. Increasingly it is realized that, in patient care, clinicians have to balance the available evidence with the needs, circumstances and preferences of individual patients. Moreover, there are still many areas of practice, such as paediatrics, where evidence is limited (Offringa 2003). Evidence-based medicine is constrained by all the above limitations, including the lack of attention to diversity in clinical research. This makes it harder for clinicians to translate the generalized conclusions of trials and systematic reviews to individuals in need of care.

3.4 The trend to diversity in the conduct of clinical research

In the late 1980s, a marked shift towards diversity in clinical research took place in The United States (Cotton 1990, Hamilton 1996, Epstein 2003). For diseases like heart disease and HIV-AIDS, activists had emphasised that the lack of clinical data on women led to low standards of care. The US General Accounting Office, the research arm of the Congress, was called on to investigate to what extent the National Institutes of Health (NIH) was encouraging inclusion of women and minorities. The investigators found that, despite the existence of a 1986 NIH guidelines[iii], the policy had been poorly communicated to researchers and applied inconsistently "American women have been put at risk" the investigators said, presenting their study to the Congress subcommittee on Health and Environment. The GAO investigators cited an NIH funded physician's health study, begun in 1981, which had investigated the role of aspirin use in preventing heart attacks. The study had enrolled 22,000 male doctors; NIH officials had explained that including women would have increased the cost (Epstein 2003, page 180). In a 1990 JAMA article, addressing the issue of middle-aged white male bias, a female cardiologist complained that, in the absence of data, it comes down to "judgement call.... I'm uncomfortable putting all women with coronary heart disease on aspirin prophylactically because the data are not definitive. But I've seen clinical benefits with unstable angina, so in acute situations I feel I can extrapolate" (Cotton 1990, 1055)

In 1993, President Clinton, signed the NIH Revitilization Act, to require the National Institutes of Health, (the worlds largest public financier of medical research), to ensure that women and also members of racial and ethnic minority groups be included as subjects, in each clinical study funded by the agency from 1995 onwards. The NIH Revitalization Act of 1993, PL 103-43, specifies that: “In the case of any clinical trial in which women or members of minority groups will be included as subjects, the Director of NIH shall ensure that the trial is designed and carried out in a manner sufficient to provide for valid analysis of whether the variables being studied in the trial affect women or members of minority groups, as the case may be, differently than other subjects in the trial. 492B(c)” and that “the costs of such inclusion in the trial is (sic) not a permissible consideration in determining whether such inclusion is inappropriate. 492B(d)(2).” Specifically addressing the issue of minority groups, the statute states that “The term "minority group" includes subpopulations of minority groups. The Director of NIH shall, through the guidelines established… define the terms "minority group" and "subpopulation" for the purposes of the preceding sentence 492B(g)(2)”.

The NIH implemented these guidelines in 1994, by issuing a new subsidy guide. In this guide it explains to grantees “Since a primary aim of research is to provide scientific evidence leading to a change in health policy or a standard of care, it is imperative to determine whether the intervention or therapy being studied affects women or men or members of minority groups and their subpopulations differently. To this end, the guidelines published here are intended to ensure that all future NIH-supported biomedical and behavioural research involving human subjects will be carried out in a manner sufficient to elicit information about individuals of both genders and the diverse racial and ethnic groups and, in the case of clinical trials, to examine differential effects on such groups. Increased attention, therefore, must be given to gender, race, and ethnicity in earlier stages of research to allow for informed decisions at the Phase III clinical trial stage.”[iv]

For the purpose of this review it is interesting to see how minority and majority groups are defined in these 1994 guidelines. The box on the next page provides the definitions. These are based on categories used in US databases.

In 1993, in line with the NIH revitalization act, the US Food and Drug Administration also issued new guidelines governing the participation of women in clinical trials sponsored by drug companies. Since 1977, women of “childbearing potential” had been routinely excluded from many such trials, whether they were pregnant or not, or intended to be, ostensibly out of concern that an experimental drug might harm a foetus. In intent, this restriction applied only to Phase I and II trials of new drugs, whose potential for causing birth defects was still unknown. In practice, the broad

NIH defininitions of minority groups (NIH 1994)

The NIH specified that a minority group is a readily identifiable subset of the U.S. population which is distinguished by either racial, ethnic, and/or cultural heritage. The categories suggested are:

American Indian or Alaskan Native:

A person having origins in any of the original peoples of North America, and who maintains cultural identification through tribal affiliation or community recognition.

Asian or Pacific Islander:

A person having origins in any of the original peoples of the Far East, Southeast Asia, the Indian subcontinent, or the Pacific Islands. This area includes, for example, China, India, Japan, Korea, the Philippine Islands and Samoa.

Black, not of Hispanic Origin:

A person having origins in any of the black racial groups of Africa.

Hispanic:

A person of Mexican, Puerto Rican, Cuban, Central or South American or other Spanish culture or origin, regardless of race.

The majority group in the definition of NIH is: White, not of Hispanic Origin:

A person having origins in any of the original peoples of Europe, North Africa, or the Middle East.

NIH recognizes the diversity of the U.S. population and that changing demographics are reflected in the changing racial and ethnic composition of the population. The terms "minority groups" and "minority subpopulations" are meant to be inclusive, rather than exclusive, of differing racial and ethnic categories.

Each minority group contains subpopulations which are delimited by geographic origins, national origins and/or cultural differences. It is recognized that there are different ways of defining and reporting racial and ethnic subpopulation data. The subpopulation to which an individual is assigned depends on self-reporting of specific racial and ethnic origin. Attention to subpopulations also applies to individuals of mixed racial and/or ethnic parentage. Researchers should be cognizant of the possibility that these racial/ethnic combinations may have biomedical and/or cultural implications related to the scientific question under study.

and automatic exclusion of pre-menopausal women from new drug development had become commonplace. The 1993 guidelines permitted the inclusion of women, even in early clinical trials, provided the female subjects used some kind of birth control, and also called for drug companies to submit data on the effects of new drugs on both men and women (Epstein 2003).

The emphasis on the need for inclusion of women and minorities in trials was not unopposed. Some clinical researchers opposed the micro-management of their trials by NIH, arguing that sub-group analysis, if the results are to be of statistical significance, increases size of trial populations, and the cost of trials. They prefer to stress similarities between people:

> "...people have more biological similarities that differences. Penicillin will kill bacteria in blacks, whites, Cuban-Americans, Mexican-Americans, men, women, dogs, cats, birds, and petri-dishes" (Piantadosi and Wittes 1993: 565). They critique the undertaking of a second, larger and more costly trials in female nurses on the preventive effects of aspirin for cardio-vascular disease, in the absence of any plausible hypothesis that women respond differently than men, and question the "new law of studying each medical question within ´minority´ ...if we are studying sickle cell anemia, do we really have to include Mexican-Americans...In a study of diabetes, must we include Pima Indians whose insulin appears to be different from other people? Should every trial in Manhattan include Dominicans and Haitians?" (Piantadosi and Wittes 1993, 566).

Other supported the NIH Guidelines. They acknowledged that large sample-sizes (tens of thousands participants) would be needed for valid analysis, but suggested that decisions to do such large scale trials should be based on prior indications that differences matter. Appropriate representation of minority groups and women in the trial population is key in their view. Comparisons between racial and ethnic groups should be based on prior evidence that subpopulations are "unusually" affected by certain diseases (Freedman et al. 1995).

Did the 1994 guidelines make a difference? A report on the representation of African-Americans, Hispanics, and White in National Cancer Institute Cancer Treatment Trials (Tejeda et al. 1996), suggests that they did. The different populations groups are included in the trials, in a proportional way. However, proportional inclusion does not necessarily allow for subgroup analysis, and the researchers warn that, "Over sampling of these minority groups may be necessary to increase statistical power when there is reason to suspect differences (Tejeda et al 1996, Freedman et al. 1995). These researchers point with concern to the lower participation rate of Americans 50 years of age and earlier, suggesting that age might become an issue of concern (Tejeda et al. 1996).

While congressional and public attention initially focused on gender and racial diversification in biomedical research, attention to the needs of children came a few years later. The call for their inclusion in clinical trials came from paediatricians who pressed for policy change, pointing to evidence that the vast majority of medications used by children had never been tested on children. They argued that children are not simply miniature adults, and that it was crucial to study differences between adults and paediatric populations rather than simply extrapolating from the former to the latter (Epstein 2003). In 1998, four years after the publication on women and minorities, the NIH published guidelines on the inclusion of

children in trials. The FDA modernization act of 1997 encouraged pharmaceutical companies to study differences between adults and children in the safety and efficacy of drugs by offering a six month extension of patent protection[v].

With The United States leading the way in reforms in clinical research to include better representation of diverse populations, the International Conference on Harmonisation in the 1990s developed diversity guidelines. Based on these guidelines (ICH E8), the *General considerations for clinical research* published in 1997 (ICH E8) by the European Medicines Evaluation Agency (EMEA) identifies what is expected from companies, prior to registration of new drugs. The guidelines acknowledge a need to limit variation in Phase I and II of clinical research, by means of inclusion and exclusion criteria. The guidelines recommend, however, that Phase III trials should be more representative of the general population in which the drug is to be used. Women of reproductive age should be included, as long as they use a contraceptive. In addition there are guidelines on the inclusion of specific populations including *Ethnic factors in the acceptability of foreign clinical data* (ICH E5, 1998), *Studies in support of special populations: geriatrics* (ICH E7, 1995) and *Clinical investigation of medicinal products in the paediatric population* (ICH E11, 1998). The guidelines indicate that "some groups in the general population may require special study because they have unique risk/benefit considerations that need to be taken into account during drug development, or because they can be anticipated to need modification of use of the dose or schedule of a drug compared to general adult use. Pharmacokinetic studies in patients with renal and hepatic dysfunction are important to assess the impact of potentially altered drug metabolism or excretion."

The guidelines also call for special attention for ´vulnerable populations´. These include, "Individuals whose willingness to volunteer in a clinical trial may be unduly influenced by the expectation, whether justified or not, of benefits associated with participation, or of a retaliatory response from senior members of a hierarchy in case of refusal to participate. Examples are members of a group with a hierarchical structure, such as medical, pharmacy, dental and nursing students, subordinate hospital and laboratory personnel, employees of the pharmaceutical industry, members of the armed forces, and persons kept in detention. Other vulnerable subjects include patients with incurable diseases, persons in nursing homes, unemployed or impoverished persons, patients in emergency situations, ethnic minority groups, homeless persons, nomads, refugees, minors and those incapable of giving consent" (ICH E6, 1997). The main reason for identifying these groups is ethical: it is important to ensure rigorous informed consent procedures, to avoid undue influence or expectation. Ethnic factors are categorized into intrinsic (genetic, physiological) or external (cultural and environmental factors). The EU is especially concerned about the acceptability of using "foreign clinical data". For "ethnically insensitive" medicines, extrapolation of date from one region of

the world to another is allowed. For "ethnic sensitive" medicines this is not the case.

The ICH guidelines further advice that patients involved in clinical research should be reasonably representative of the population at large, and that drugs are studied in all age groups, including the elderly. Researchers are also required to take into account specific risks for elderly drug users, including the risks of drug interaction due to concomitant drug therapy, especially for the elderly. The guidelines suggest that differences in pharmacokinetics of drugs should be studied by age, to take account of the natural diminution of liver and kidney functions in elderly populations. The guidelines on clinical research among children identify five different age-groups between birth and 18 years of age. It is suggested that these age groups differ in kinds of pathologies suffered, physiological endpoints for clinical research, adverse effects and compliance.

3.5 Diversity in practice: a review of published reports of clinical trials

To what extent are the guidelines for diversity developed by funding agencies reflected in the practice of clinical research. To assess this, we reviewed published reports on clinical trials of four major diseases (Diabetes I and II, Epilepsy, Hypertension, and HIV-AIDS) in two major journals over the past four years[vi] (Table 3.1). It would seem safe to assume that, if research reports published in leading journals do not follow the guidelines, then trials reported in other medical journals are unlikely to do so.

This selected review suggests that, in the new millennium, trial populations do generally include men and women for the selected diseases. The extent, to which children and the elderly are included in studies, depends on the type of disease. All epilepsy RCTs included children, as did around 60% of Diabetes I studies, and half of the AIDS studies. The elderly are nearly always included in the Diabetes 2 and the hypertension studies, which again makes sense medically, because these diseases are more prevalent in the elderly. The main omission, in terms of representativeness of the trial population, is ethnicity. Remarkably, most of the trials for all disease categories, do not define the ethnicity of people included in the trial, with the exception of the hypertension trials, in which 70% did so.

Overall the findings suggest that diversity is considered most in trials on hypertension and AIDS. The public debate on the need to include women and minorities in The United States focused on these two illness conditions. For epilepsy and diabetes this is much less the case. This seems all the more remarkable because variations in aetiology and response to treatment by sex and age have been reported in the medical literature for both of conditions.

Table 3.1. Consideration of subgroups in RCTs on Diabetes 1 and 2, Hypertension, Epilepsy and AIDS published in two leading medical journals (Lancet, and JAMA), 2000-2004 (N = 51 studies).

	Diabetes 1	Diabetes 2	Hypertension	Epilepsy	AIDS
Number of studies	(5) [€]	(10) [€]	(17)	(6)	(13)
• Average percentage men in population	65	69	49	50	58/ 44[Ω]
• Percentage studies including children below 18 years	60	0	0	100	50[Σ]
• Percentage studies including elderly above 65 years	20	90	94	17[£]	8[Σ]
• Percentage ethnicity defined?	20	30	71	0[¥]	39
• Percentage studies with age used as inclusion/exclusion criteria (yes-no)	100	100	100	50	69
• Percentage studies with sex-related inclusion-exclusion criteria, such as pregnancy, menopause (yes-no)	80	10	24	0	85[Φ]
• Percentage studies with subgroup analysis for sex	20	10	35	17	8
• Percentage studies with subgroup analysis for age	20	10	35	17	8
• Percentage studies with subgroup analysis for ethnicity	0	0	18	0	0

[€] One study has been counted twice (for diabetes I AND II), because this study included both diabetes I and II.

[£] One study contained unclear information on maximum age of participants. Guideline states 12-65 yrs, but in the informative table it is stated that participants are 12-72 yrs --> Chadwick 1999)

[¥] 'Ethnicity' not explicitly mentioned in any of the trial papers, but 2 out of 6 studies have been carried out in rural India.

[Ω] 3 out of 13 studies are one-gender-only (only pregnant women). The first percentage given is therefore calculated for only mixed-gender studies, the second percentage is calculated for all studies (including the women-only studies).

[Σ] Of the 13 studies, three studies do not report on whether participants <18 yrs are included in the trial, of the studies that do report on this fact, 50% include children below 18. The same goes for adults >65 yrs old. Only one study explicitly states that adults over 65 yrs old are included. For 8 of the thirteen studies it is not possible to determine whether persons over 65 yrs old have been included. One study included only children (2-17 yrs old).

[Φ] Including 3 one-gender-only studies (only pregnant women)

There are many reasons for the general lack of representation of different ethnic groups in the trial populations. Clinical researchers often stress the operational problems of involving people from different ethnic backgrounds, including the need to translate interview questions (see also review 4 and 5 for more details). But, a biological universalistic idea that people generally are more similar than different may also play a role. Whatever the reason, the lack of representation is not in line with NIH guidelines. Our findings suggest that minorities are often excluded from trials. This is not only a problem for the production of medical knowledge, but it is also unethical as will be argued in review 4.

Even though the trial populations are relatively heterogeneous in terms of age and sex, the data suggest that also for these categories of difference, sub-group analysis is rarely done. The exception is hypertension, for which

around one-third of the trials included subgroup *analysis* for age and sex, and 18% for ethnicity. In all other diseases, sub-group analysis was rare for age and sex, and absent for ethnicity[vii].

A closer look at the studies in which subgroup analysis is done, suggests major limitations in the analyses. Firstly, it is unclear from the methodology sections of the papers, if sample-sizes were chosen to allow for sub-group analysis. Even so, statistical analysis of the difference between age, and sex groups (less so for ethnicity) are given in figures and tables in the results sections of the articles. Difference by age and sex are accounted for in published presentation of results, but they are rarely highlighted in conclusions, and never in the abstracts. Generally the trials report that the subgroups do not differ significantly in terms in terms of efficacy of the medicines studied. Below we present in more detail the kinds of results that are reported in the clinical trials in which some form of sub-group analysis was done. In the following analysis, we also mention the objectives of the trials and the funding sources.

The only Diabetes I trial to include subgroup analysis was designed to assess whether high dose of nicotinamide delays onset of diabetes. The study was funded by the EU and the Juvenile Diabetes Research foundation. It reports:

> " Data in table 3 show that there was no evidence of a treatment effects in groups divided on the basis of age, sex, oral tolerance status, antibody status, or first phase insulin response" (ENDIT 2004: 928).

The table that these researchers refer to typifies the presentation of sub-group data in this series of trials. (see Table 3.2.)

Table 3.2, Selection of data from Table 3 in the original article (Gale et al. 2004). Harard ratio's for developing diabetes within five years.

	Placebo (N = 275)	**Nicotinamide (N = 274)**	**Hazard ratio (95% CI)**	**P**
Overall	77 (28%)	82 (30%)	1.07 (0.78-1.45)	0,69
Sex				
• Male	47 (33%)	48 (33%)	0.99 (0.66-1.48)	0.97
• Female	30 (23%)	34 (26%)	1.17 (0.71-1.90)	0.53
Age at baseline				
• > 20 years	66 (40%)	64 (39%)	0.98 (0.69-1.39)	0.91
• ≥ 20 years	11 (10%)	18 (17%)	1.42 (0.70-2.90)	0.33

The main conclusion of the trial is that nicotinamide was ineffective at the dosage used. The data suggest that the drug is ineffective for all subgroups, but differences between groups are not tested.

The only Diabetes II trial found to conduct subgroup analysis was designed to assess the effects of acarbose on the development of diabetes. This study, funded by Bayer, reports that:

> "The beneficial effect of acarbose was consistent irrespective of ages, sex and body mass index (figure 2). Body mass index (as did weight change) significantly affected development of diabetes (p = 0,0066), whereas age and sex did not (p = 0,8643 and p = 0,4385 respectively)" (Chiasson et al. 2002, 2074).

The only Epilepsy trial that was found to conduct subgroup analysis was designed to compare efficacy and acceptability of phenobarbital and phenytoin as monotherapy for children in India (Pal et al. 1998). The researchers found no difference in efficacy, and no difference in the incidence of behavioural side effects between the drugs. The number of children included in the trial is limited (only 62 children completed the trial). The researchers report that:

> " Behavioural problems were more common among children with cerebral impairments, those under 5 years old, and girls. All these *patterns* (emphasis added) were stronger in children treated with phenytoin than in those who received phenobarbital" (Pal et al. 1998:23).

Though constrained by limitations of sample size, these authors appear to have carefully examined the influence of sex and age differentials.

The only AIDS trial that was found to conduct subgroup analysis compares fungicidal activity of combinations of drugs. The authors report that, "Adjustments for age, sex and other variables had no substantial effect on these estimates" (Brouwer et al. 2004, p 1766).

Subgroup analysis was most often done in the hypertension trials, as we have seen. Table 3.3 provides an overview of the seven trials in which this was done. It is apparent from this table that many of these studies are comparative trials, designed to show equivalence or superiority of a specific drug to other therapies. Six out of seven trials are (co)funded by pharmaceutical companies producing the drugs being tested.

These findings make it clear that, where differences in effects by age, sex or ethnicity are reported, these tend to be discounted by referring to the lack of statistical significance, or the lack of relevance in terms of endpoints as illustrated by the conclusion of the ALLHAT-LLT (2002) study:

> "ALLHAT-LLT included larger proportions of older participants, women, blacks and Hispanics than any other statin trial completed. However, subgroup analysis of ALLHAT-LLT…do not show age-, or sex-related differences in RRs for CHD event rates. The RR for prevastatin vs. usual care was significantly lower in blacks than non-blacks for CHD events, but was higher for strokes, with no overall difference for combined cardiovascular events .."(p 3007).

Table 3.3. Overview of objectives, main finding, and results of sub-group analysis of hypertension trials published in the Lancet and JAMA (2000-2004).

Author-year and funding	Objectives	Main finding	Subgroup-analysis results
ALLHAT 2000 National Heart, Lung and Blood Institute, Pfizer, and other comp.	Compare effects of doxazosin (an α-blocker, with chlorthalidone, a diuretic, on incidence of CVD in patients with hypertension.	Chorthalidone reduces the risk of combined CVD events in high-risk hypertensive patients.	The differential effect of treatment on CVD outcomes was consistently observed among patients younger than 65 years vs. those aged 65 years or older; black persons vs. non black persons (p 1971).
ALLHAT-LLT 2002 National Heart, Lung and Blood Institute, Pfizer, and other companies	To determine whether pravastatin compared with usual care reduces all-cause mortality in older, moderately hypercholesterolemic, hypertensive participants with at least 1 additional CHD risk factor	Prevastatin did not reduce all cause mortality	ALLHAT-LLT included larger proportions of older participants, women, blacks and Hispanics than any other statin trial completed. However, subgroup analysis of ALLHAT-LLT...do not show age-, or sex-related differences in RRs for CHD event rates. The RR for prevastatin vs. usual care was significantly lower in blacks than non-blacks for CHD events, but was higher for strokes, with no overall difference for combined cardiovascular events .(p 3007)
Black 2003 Searle and Pharmacia	To determine whether initial therapy with vipramil is equivalent to a physicians´s choice of atenolol or hydrochlorthiazide in preventing cardiovascular disease	The trial did not demonstrate equivalence between vipramil and atenolol or hydrochlorthiazide.	When participants were grouped according to baseline characteristics, there was no evidence of treatment by subgroup interaction for any of the predefined baseline subgroups (sex, age, geographical origin), p. 2079.
Brown 2000 Bayer	To compare effects of calcium channel blocker nifedipine with the diuretic combination of co-amilozide on cardiovascular mortality and morbidity in high-risk patients with hypertension		Age and sex are defined as ´risk factors´. And the author conclude: Event rates, as expected, were higher in some risk groups than others, but with no apparent difference between the two treatment groups....(p 370).
Pepine et al. 2003 Basf and Abbott Laboratories	To compare mortality and morbidity outcomes in patients with hypertension and CAD treated with a calcium antagonist strategy (verapamil) or a non-calcium antagonist strategy.	Verapamil strategy is as clinically effective as the non calcium-antagonist strategy in hypertensive CAD patients.	This authors provide a table with the effects of the treatment strategies in primary outcomes in subgroups including ages, sex and race-ethnicity, but don´t discuss the results for these subgroups. The focus on differential effects by prior health conditions, such as prior health failure.
Sever et al. 2003 The Ascot-lla study, funded by Pfizer	To assess the benefits of cholesterol lowering in the primary prevention of coronary heart disease.	Cholesterol lowering with atorvastatin conferred a 36% reduction in fatal CHD.	The proportional effect of atrovastatin on the primary endpoint did not differ significantly in any prespecified subgroup, although the benefit was not significant in six subgroups, including patients with diabetes, and -no benefit was apparent among women. p 1153 The apparent lack of significant benefit of atorvastatin on the primary endpoint among women may reflect the small number of events they experiences...these results highlight a potential shortcoming of ASCOT, which..included mainly white male participants. p. 1155.
Taubert 2003 Not given	Does dark chocolate lower BP in individuals with mild isolated hypertension	A calorie-balanced increase in the consumption of dark chocolate may favourably affect BP in previously un-treated elderly hypertensive individuals.	There were no sex differences in the effects of chocolate on BP, P 1029

In the results of the statin trial conducted by Sever et al. (2003):

> "The apparent lack of significant benefit of atorvastatin on the primary endpoint among women may reflect the small number of events they experiences…these results highlight a potential shortcoming of ASCOT, which….included mainly white male participants". p. 1155.

None of the above studies make explicit if the sample-sizes included in the studies are large enough to conduct subgroup analysis. The main objective is to demonstrate relative efficacy and to show that different populations groups were represented in the trial, not to demonstrate differences between population groups.

This is also apparent in a lack of attention for differentials in experiences of adverse drug effects. Such data, if only descriptive, would be relevant both for clinical practitioners and patients. As the trials cited above were not designed to assess safety of medicines, no subgroup safety analysis was done. An exception is the epilepsy trial in India, where the investigators point to a *pattern*, suggesting that behavioural adverse events were more common among children with cerebral impairments, in those aged under 5 years old, and in girls.

One can argue that the lack of attention to safety is related to the specific objective of the trials that we reviewed. Pharmaceutical companies have to undertake many (pre)clinical studies, including tests for carcinogenicity, mutagenicity and teratogenicity before registering drugs.

The failure of trials to analyse differential efficacy is also a problem for patients. Medicines tend to be put on the market in dosages that are believed to be most suitable for most patients – and even in 'one-size-fits-all' formulations – though, in practice, individual response to different dosages of medicines may vary widely. For children, the need to adjust dose is quite well recognized, but the lack of sub-group analysis on efficacy, means that trials may fail to point to differences in drug response that are attributable to lack of dosage adjustment.

Earlier studies in The US that are designed to evaluate the implementation of the Revitilization Act confirm our findings that sub-group analysis is rarely done. A study by Vidaver et al (2000), looking at original articles in NEJM, JAMA, JNCI and Circulation, 1993/1995/1997/1998, showed that although the proportion of trials including women showed an increase from 72% in 1993 to 91% in 1998, very few of the clinical trials make any mention of subgroup analysis of outcomes in relation to sex. Moreover, in most cases, sample sizes are too small to do adequate subgroup analysis by sex. (On request, some authors reported that in cases where sex-specific analyses had been done, they found no differences and thus did not report this information.) Woosley et al. (2000) points out that NIH legislation only requires valid analyses of the role of sex in phase III clinical trials. However, in order to make data available for Phase III clinical trials, research should be designed to look for possible sex differences in all

phases of clinical research. Lack of proper anticipation may lead to avoidable toxicity or failure to observe beneficial effects of drug.

3.6 The relation between clinical trial data and practice

Given this lack of evidence on differences in effects and safety of drug by sex, age and ethnicity, how do clinical guidelines, which are supposed to follow the principles of evidence based medicine, confront the issues? A recent review of clinical practice guidelines, found that most ethnic specific statements were not backed by clinical trial data, but by results of descriptive studies or narrative reviews (Manna et al. 2003a). Given the lack of sub-group analysis by ethnicity this should not come as a surprise. The authors reviewed the clinical practice guidelines from the USA, Canada, the UK and the Netherlands. The USA guidelines contained the most ethnic specific statements and the Dutch guidelines the least.

Manna et al (2003b) investigated specifically whether differences related to patients' ethnic background are mentioned in the scientific evidence used by The Netherlands association for general practitioners, the NHG (for diabetes type 2, hypertension, asthma). They concluded that available scientific evidence on differences by ethnicity was not routinely presented in the standards, partly because the experts making the standards feared that highlighting differences would perpetuate inequality in health care. Commenting on the review of Manna et al, Assendelft (2003) suggests that the NHG standards are successful because they are short, clear, and practice oriented. Important subgroups of patients are described separately only if the standard advice would not provide adequate care. He argues that guidelines should highlight differences by ethnicity or sex, only when these were found to be statistically significant differences, and when differences were clinically relevant. He also questions whether ethnicity findings from The US cannot simply be extrapolated to The Netherlands. Possibly, differences in health and health care vary between ethnic groups in different countries. Only in instances where the NHG standard would cause harm to specific subgroups should these issues be reported on separately.

Keuken et al. (2004) reviewed the extent to which Dutch guidelines for Hypertension, Rheumatoid Arthritis, Osteoporosis, and Depression address sex-related diversity. They found that the guidelines for Osteoporosis paid most attention to sex-related diversity and those for depressions least. These reviewers suggest that for sex-differentials to be considered more systematically, the agencies preparing guidelines should develop committees that include men and women, they should include sex-differentials in the terms of reference of the committees, in the searches for relevant literature and in the recommendations. Lack of evidence is not cited as a reason for the lack of systematic attention for sex-differentials.

3.7 Towards methodological reforms

The diversity agenda has been pushed by powerful advocacy groups. We have seen that the focus on inclusion of women, blacks, children and the elderly in the trials emerged out of explicit and tacit alliances among an array of diverse actors, including women's health groups, paediatricians and organizations of racial and ethnical minorities (Esptein 2003). The need to consider differences between these population groups was accepted by medical researchers, and medical institutions because they reinforce a tenacious biological, reductionist assumption that medicines either are or are not 'effective' by biologically defined group. The rationale for including diverse populations in research is that their biological make-up differs from that of white males, the preferred trial population prior to the reforms.

We have also seen that in practice subgroup analysis is hardly done. The emphasis has been on making sure that differences are *accounted* for in the trial population. Differences are not *analysed* adequately, and as a result clinical trials continue to report their main findings on efficacy of medicines by treatment and control group, not a differentiated picture by age, sex and ethnicity. The major contribution of the Revitilization Act is that groups that were first excluded from trial populations now tend to be included, though our analysis suggests that the act is not successful in encouraging more heterogeneous trial populations in terms of ethnicity.

The rates of subgroup analysis that we observed in our review of 51 RCT recently published in leading medical journals suggests that sub-group analysis is done in the *minority* of trials, mainly for the categories sex, and age. Ethnicity appears to be analysed only if there is some indication of (or reasons to suspect) differences in treatment outcomes. Randomised clinical trials at present are oriented towards the measurement of pre-determined outcome measures, calculated as averages for treatment and control groups. By not routinely addressing diversity, they have a *homogenizing* effect on health care.

One reason for the lack of attention for diversity in treatment outcomes may be the complexities involved in sub-group analysis. Common diseases and treatment outcomes are influenced by complex interplay of factors, including genetic factors and a significant environmental/lifestyle component (Pierce et al. 2004, Cooper 2003). Differences by ethnic group, age or sex can thus be a marker for socio-economic, cultural, biological differences. Negative outcomes of AIDS medication in a particular ethnic group may for example be caused by biological differences, or by sub-optimal adherence to the medication regime, not the pharmaceutical properties of the drugs. The low adherence could be caused by specific cultural notions of efficacy and safety of the medicines, or by socio-economic constraints in accessing the medications. And dietary practices influencing uptake of the medicines could have an influence. This complexity is a conceptual constraint to diversity research, as it implies that sub-group analysis based only on hypothesis concerning differences in

biology alone will not suffice in any attempt to take diversity serious in clinical research.

ZonMw, The Netherlands health research agency that funded our work, includes not only biological, but also cultural and socio-economic differences in their policy framework. The agency appears to go beyond defining diversity necessarily by *group*. Its policy document on diversity (Kleurstof) refers to "individual" differences, and defines three causes of differences in health: socio-economic, biological, and cultural/behavioural. ZonMw further acknowledges that health care perpetuates these differences.

Given the limited successes of the past, how can public agencies involved in funding of medical research, such as ZonMw and the NIH encourage clinical researchers to incorporate such a comprehensive conceptualisation of diversity in their study designs? In addition to recommending studies with enough statistical power to conduct subgroup analysis, such a conceptualisation of diversity requires a *paradigm-shift* in clinical research. Researchers should not only account for differences by predefined biologically categories, but also keep an eye open for more complex interplays of factors that cause diversity. More attention is needed for the generation of *diversity hypotheses* that can guide more focused studies on diversity issues. We propose six specific methodological reforms to generate and test such hypotheses.

In theory, biomedicine should be able to develop tests to establish how an individual's genetic make-up might affect response to drugs, and this could lead to tailor-made treatments. Genetic tests are increasingly used to define new kinds of groups: carriers of genes, who are defined as ´at risk populations´. One area of methodological reform could involve more routine attention for possible genetic variation in effects of treatments in the design and analysis of clinical trials.

Secondly, we propose that when conducting RCTs, researchers should be more alert to phenomena that were not included in the study design (Offringa 2003). They should pay special attention to outliers, and other unexplained variance. Such analysis can lead to hypotheses on why variance occurs.

Thirdly clinical trials can be used more systematically to *test* diversity hypotheses derived from observations within RCTs, as well as from etiological and observational studies, case-reports and user-studies, which are generally seen to have less scientific value than RCTs (Glasziou et al. 2004).

A fourth suggestion for reform is a topic for medical scientists and statisticians who do meta-analyses of clinical trials. Meta-analysis is generally done with a view to providing sound medical evidence, as a basis for guidelines on medical practice. Meta-analysists tend to focus on

similarities in outcomes, rather than differences, thus contributing to a homogenizing effect on medical practice. We propose that meta-analysts consider differences between the study populations included in the various RCT as data from which both similarities in outcome, and potential differences in treatment outcomes can be derived. In an ideal world, the meta-analysists would have access not only to the study protocols and published reports, but also to the raw data. This would allow for re-analysis to test certain diversity hypotheses.

Fifth, we suggest that data from routine health care practice can be used to explore much more systematically differences that matter in practice. Health practitioners are confronted with individual clinical realities that do not conform to the evidence from clinical trials. Observations from individual practitioners can lead to valuable hypothesis. Delphi methods could be used to generate hypothesis from a multitude of individual practitioners, allowing for more systematic feedback from health workers´ experiences. Computerised health practitioner data-bases can likewise be used to explore and test diversity issues.

Finally, user's experiences can also be explored much more systematically. There is already some volume of information available on the Internet where users of medicines report their experiences. There are also various systems of pharmacovigilance. This body of data can be mined much more effectively for hypotheses on diversity in drug safety and efficacy.

3.8 Factors that facilitate and constrain diversity in practice

These six suggestions for methodological reform, clearly present a challenge for the scientific research community and to other actors in health care. Reform requires researchers to integrate understandings that come from evidence that is 'objective', but also incomplete, and evidence derived from individual health practitioners´ and patients´ experiences, values and views. The methods suggested above can lead to large number of diversity hypothesis. Before large-scale trials are done to test these hypotheses, a prioritisation would need to be made --- methods need to be developed to identify the *diversities that matter*. Qualitative and participatory methods involving clinical practitioners, general practitioners, insurers and clients of health care as stakeholders, need to be developed.

Diversity knowledge needs to find its way into treatment guidelines and health care policy. We have seen that clinical research has a homogenizing effect on medical practice. Not surprising, the professional committees developing treatment guidelines do not consider diversity in a systematic way - they lack hard evidence from RCTs. The assumption in medical practice is that generally people will respond the same to treatments. Differences are an exception, rather than the rule.

Pharmaceutical companies commission the bulk of pre-marketing research and are responsible for the initial definition of safety and efficacy assessment of their products by means of clinical trials. They fund most RCTs, and can thus be held responsible for the lack of attention for diversity. They are not likely to be interested in diversity in treatment outcomes. Drug development costs are said to be as high as US$800 million per new chemical entity[viii]. For a drug to be profitable, and allow for recovery of development costs, companies generally need big markets. Research results which point out that a treatment is less effective, or even harmful, for some subgroups are not welcomed by company marketing directors. Once a drug is on the market, it is not in their commercial interest to assess safety issues in depth, in different populations. If safety problems were to emerge in some groups, this would lead to a reduction of the market for the drug; accordingly, comparative efficacy trails are designed to increase markets for drugs, not reduce them. Unless a regulatory agency demands subgroup analysis -which is generally only when there is a reason to believe that drugs work differently in specific group - companies are unlikely to want to conduct such expanded trials. The exception is the conduct of trials among children, for which regulatory agencies have developed an incentive: a patent extension of six months which can mean a significant increase in profitably, especially when generic competition is threatened.

Drug regulatory agencies are currently dependent on a fee for service in both The United States and Europe. They are under pressure to shorten the time needed for registration of drugs. Drug companies need government agencies to evaluate their drug fast and efficiently, because they suffer economically from delays in bringing new drugs to market. Systematic consideration of diversity in efficacy and safety of medicines implies that regulators would need more time to study the registration dossier. Under the current fee-for-service payment structures, it is not in the interest of drug regulators to demand more complicated dossiers. Resistance to diversity becomes apparent in controversies, for example, in relation a possible increase in suicide risks, in children taking SRRIs. Only after media coverage, and public concern, did the regulators re-analyse the clinical data, to find that the data confirmed the increased risk. If trials in children (major users) had been required, and systematic subgroup analyses had then been done, such drugs would have been contraindicated for children, years ago.

Insurers base their decisions to reimburse treatments on evidence from clinical trails and increasingly on assessment of cost-effectiveness under routine health care conditions. Homogeneity in treatment is easier to manage, than diversified medical practice, and therefore can be seen to be more cost-effective. On the other hand, if patients do not become better, or suffer adverse effects of treatment, the insurers will have to pay for follow-up health care interventions. Ultimately insurers have an interest in better health outcomes for their clients. This makes them a potential ally in the implementation of above methodological reforms, which aim at

identifying diversity issues that matter for clinical practice and patient health.

3.9 In conclusion

This paper has focused on factors, which facilitate and constrain representation and analysis of *diversity* in clinical research – i.e. the extent to which, and ways in which, the response to health interventions differs in identifiable subgroups, from those of the whole treat population. We have shown that randomised controlled clinical trials have their limitations in addressing diversity issues. Clinical researchers generally suggest that the main obstacle they face is to do with organising (and funding) studies of an appropriate size, to allow for subgroup analyses with sufficient statistical power. But, there seem to be many other limitations besides.

We have seen that, in practice, subgroup analysis is rarely done and that, when it is, it is generally confined to diversity categories which can be defined in biological terms, i.e. age and sex.

We argue the need for a paradigm shift in clinical research that allows it to go beyond biological defined categories of difference, and beyond the use of randomised controlled trials (RCTs) as some gold standard. We propose some broader definition, and recognition, of diversity. It may include biological, socio-cultural, psychological and economic factors, as well as living conditions and diet, which are underlying causes of differences by age, sex and ethnicity. We need, as a point of departure, to accept that many features of diversity will interact and co-produce the efficacy and safety profiles of medicines, and health of *individuals*. Such a conceptualisation of diversity is challenging as it results in an endless list of diversity hypotheses, impossible to incorporate in clinical research. We have argued that a mix of research methodologies can be used to filter from etiological studies, from the realities of health care practice, as well as from ongoing RCTs and meta-analyses *diversity issues that matter*, which need to be further explored. Diversity hypotheses, emerging from this range of studies, can be tested in future more focused RCTs. This does not only require more funding, it requires a change in the kinds of questions addressed in RCTs.

References

Agodoa LY, Appel L, Bakris GL, Beck G, Bourgoignie J, Briggs JP et al. (2001). Effect of Ramipril vs Amlodipine on renal outcomes in hypertensive nephrosclerosis. A randomized controlled trial. JAMA, 285(21), 2719-2728.*

ALLHAT. (2000) Major cardiovascular events in hypertensive patients randomized to Doxazosin vs Chlorthalidone. The Antihypertensive and Lipid-Lowering Treatment to Prevent Heart Attack Trial (ALLHAT-LLT). JAMA, 283(15), 1967-1975.*

ALLHAT. (2002). Major outcomes in moderately hypercholesterolemic, hypertensive patients randomized to Pravastatin vs usual care. The Antihypertensive and Lipid-Lowering Treatment to Prevent Heart Attack Trial (ALLHAT-LLT). JAMA, 288(23), 2998-3007.*

Angell M. The truth about the drug companies. New York: The Random House, 2004. p. 37-51.

Apfel SC, Schwartz S, Adornato BT, Freeman R, Biton V, Rendell M, et al. (2000). Efficacy and safety of recombinant human nerve growth factor in patients with diabetic polyneuropathy. A randomized controlled trial. JAMA, 284(17), 2215-2221.*

Assendelft WJJ (2003). Commentaren - Etniciteit in standaarden van het Nederlands Huisartsen Genootschap. Nederlandsch Tijdschrift voor Geneeskunde, 147(35), 1668-1671.

Black HR, Elliott WJ, Grandits G, Grambsch P, Lucente T, White WB, et al. (2003). Principal results of the Controlled Onset Verapamil Investigation of Cardiovascular End Points (CONVINCE) Trial. JAMA, 289(16), 2073-2082.*

Brouwer AE, Rajanuwong A, Chierakul W, Griffin GE, Larsen RA, White NJ, et al. (2004). Combination antifungal therapies for HIV-associated cryptococcal meningitis: a randomised trial. Lancet 363(9423), 1764-1767.*

British Medical Journal. (1998) Editorial. Fifty years of randomised controlled trials. BMJ 1169.

Brown MJ, Palmer CR, Castaigne A, Leeuw PW de, Mancia G, Rosenthal T, et al. (2000). Morbidity and mortality in patients randomised to double-blind treatment with long-acting calcium-channel blocker or diuretic in the International Nifedipine GITS Study: Intervention as a Goal in Hypertension Treatment (INSIGHT). Lancet 356(9227), 366-372.*

Chadwick D. (1999). Safety and efficacy of vigabatrin and carbamazepine in newly diagnosed epilepsy: a multicentre randomised double-blind study. Lancet, 354(9172), 13-19.*

Chaillous L, Lefèvre H, Thivolet C, Boitard C, Lahlou N, Atlan-Gepner C, et al. (2000). Oral insulin administration and residual β-cell function in recent-onset type 1 diabetes: a multicentre randomised controlled trial. Lancet 356(9229), 545–49.*

Chiasson JL, Josse RG, Gomis R, Karasik A, & Laakso M. (2002) Acarbose for prevention of type 2 diabetes mellitus: the STOP-NIDDM randomised trial. Lancet 359(9323), 2072-2077.*

Chiasson JL, Josse RG, Gomis R, Hanefeld M, Karasik A. & Laakso M. (2003). Acarbose treatment and the risk of cardiovascular disease and hypertension in patients with impaired glucose tolerance. The STOP-NIDDM trial. JAMA, 290(4), 486-494.*

Chiron C, Marchand MC, Tran A, Rey E, d'Athis P, Vincent J, et al. (2000). Stiripentol in severe myoclonic epilepsy in infancy: a randomised placebo-controlled syndrome-dedicated trial. Lancet, 356(9242), 1638–1642.*

Collier AC, Kalish LA, Busch MP, Gernsheimer T, Assmann SF, Lane TA, et al. (2001). Leukocyte-reduced red blood cell transfusions in patients with anemia and human immunodeficiency virus infection. The viral activation transfusion study: a randomized controlled trial. JAMA 285(12), 1592-1601.*

Cotton P. (1990). Is there still too much extrapolation from data on middle-aged white men. Medical news and perspectives. JAMA 263(8), 1049-1050.

Cotton P. (1990). Examples abound of gaps in medical knowledge because of groups excluded from scientific study. JAMA 263(8), 1051-1055.

Colhoun HM, Betteridge DJ, Durrington PN, Hitman GA, Neil HAW, Livingstone SJ, et al. (2004). Primary prevention of cardiovascular disease with atorvastatin in type 2 diabetes in the Collaborative Atorvastatin Diabetes Study (CARDS): multicentre randomised placebo-controlled trial. Lancet 364(9435), 685–96.*

Cooper RS. (2003). Race, genes and health – new wine in old bottles? International Journal of Epidemiology 32, 23-25.

Dahlöf B, Devereux RB, Kjeldsen SE, Julius S, Beevers G, Faire U, de, et al. (2002) Cardiovascular morbidity and mortality in the Losartan Intervention For Endpoint reduction in hypertension study (LIFE): a randomised trial against atenolol. Lancet 359(9311), 995–1003.*

Dehue T. (1999). Testing treatments, managing life: on the history of randomized clinical trials. History of the Human Sciences, 12(1), 115-124.

Dehue T. (2001). Establishing the experimenting society: The historical origin of social experimentation according to the randomized controlled design. American Journal of Psychology, 114(2), 283-302.

Dehue T. (2002). A Dutch treat: randomized controlled experimentation and the case of heroin-maintenance in the Netherlands. History of the Human Sciences, 15(2), 75-98.

Dehue T. (2004) Historiography taking issue: analysing an experiment with heroin abusers. Journal of the history of the behavioral sciences, 40(3), 247-264.

DIAS. (2001). Effect of fenofibrate on progression of coronary-artery disease in type 2 diabetes: the Diabetes Atherosclerosis Intervention Study, a randomised study. Lancet, 357(9260), 905–910.*

Elam MB, Hunninghake DB, Davis KB, Garg R, Johnson C, Egan D, et al. (2000). Effect of Niacin on lipid and lipoprotein levels and glycemic control in patients with diabetes and peripheral arterial disease. The ADMIT study: a randomized trial. JAMA, 284(10), 1263-1270.*

EMEA (European Agency for the Evaluation of Medicinal Products). (1998). ICH Topic E 5. Ethnic factors in the acceptability of foreign clinical data. Step 4, Consensus Guideline, 5 February.

EMEA. (1997). ICH Topic E 6. Guideline for Good Clinical Practice. 1997.

EMEA. (1995). ICH Topic E 7. Studies in support of special populations: Geriatrics.

EMEA. (1997). ICH Topic E 8. General considerations for clinical trials. Step 4, Consensus Guideline, 17 July.

EMEA. (1998). ICH Topic E 11. Clinical investigation of medicinal products in the paediatric population.
ENDIT. (2004). European Nicotinamide Diabetes Intervention Trial (ENDIT): a randomised controlled trial of intervention trial of intervention before the onset of type 1 diabetes. Lancet, 363(9413), 925-931.*

Epstein S. (1997). Activism, drug regulation and the politics of therapeutic evaluation in the AIDS era: a case-study of ddC and the "Surrogate Markers" debate. Social studies of science. 27(5): 691-726.

Epstein S. (2003). Inclusion, diversity, and biomedical knowledge making: the multiple politics of representation. In: Oudshoorn, N, & Pinch T, (Ed). How Users Matter. The Co-Construction of Users and Technologies (pp.173-193). Cambridge: MIT Press.

Feinstein AR (1983). An additional basic science for clinical medicine: the constraining fundamental paradigms. Annals of Internal Medicine. 99: 3930397

Fitzgerald DW, Desvarieux M, Severe P, Joseph P, Johnson WD, Pape JW, et al. (2000). Effect of post-treatment isoniazid on prevention of recurrent tuberculosis in HIV-1-infected individuals: a randomised trial. Lancet, 356(9240), 1470-1474.*

Fonseca V, Rosenstock J, Patwardhan R, Salzman A. (2000). Effect of Metformin and Rosiglitazone combination therapy in patients with type 2 diabetes mellitus. A randomized controlled trial. JAMA, 283(13), 1695-1702.*

Freedman L, et al. (1995). Inclusion of women and minorities in clinical trials and the NIH Revitalization Act of 1993 –the perspective of NIH clinical trialists. Controlled Clinical trials, 16(...), 277-285.

French N, Nakiyingi J, Carpenter LM, Lugada E, Watera C, Moi K, et al. (2000). 23-valent pneumococcal polysaccharide vaccine in HIV-1-infected Ugandan adults: double blind, randomised and placebo controlled trial. Lancet, 355(9221), 2106-2111.*

Glasziou P, Vandenbroucke J, Chalmers I. (2004). Assessing the quality of research. British Medical Journal, 328(7430), 39-41.

Gordin F, Chaisson RE, Matts JP, Miller C, Lourdes Garcia M de, Hafner R, et al. (2000). Rifampin and Pyrazinamide vs Isoniazid for prevention of tuberculosis in HIV-infected persons. An international randomized trial. JAMA, 283(11), 1445-1450.*

GSK (2004), The drugs don´t work. Life Sciences, Enlargement Europe, 50-53.

Guyatt G, Cook D, Haynes B. (2004). Evidence based medicine has come a long way. British Medical Journal 329(7473), 990-991.

Hamilton JA. (1996). Women and health policy: on the inclusion of women in clinical trials. In C.F. Sargent and C.B. Brettel (Ed), Gender and Health: an international perspective.Upper Saddle River, NJ: Prentice Hall.

Hammer SM, Vaida F, Bennett KK, Holohan MK, Sheiner L, Eron JJ, et al. (2002). Dual vs single protease inhibitor therapy following antiretroviral treatment failure. JAMA, 288(2), 169-180.*

Hansen EH & Lausno L. (1989). Is the controlled clinical trial sufficient as a drug technology assessment. Journal of social and administrative pharmacy, 6(3), 117-126.

Hansson L, Hedner T, Lund-Johansen P, Kjeldsen SE, Lindholm LH, Syvertsen JO, et al. (2000). Randomised trial of effects of calcium antagonists compared with diuretics and ?-blockers on cardiovascular morbidity and mortality in hypertension: the Nordic Diltiazem (NORDIL) study. Lancet, 356(9227), 359–365.*

HOPE. (2000). Effects of ramipril on cardiovascular and microvascular outcomes in people with diabetes mellitus: results of the HOPE study and MICRO-HOPE substudy. Lancet, 355(9200), 253–259.*

Julius S, Kjeldsen SE, Weber M, Brunner HR, Ekman S, Hansson L, et al. (2004). Outcomes in hypertensive patients at high cardiovascular risk treated with regimens based on valsartan or amlodipine: the VALUE randomised trial. Lancet, 363(9426), 2022–2031.*

Jevons SW. (1880). Experimental legislation and the drink traffic. Contemporary Review, 37(jan./june), 177-192.

Kahn JO, Weng Cherng D, Mayer K, Murray H, Lagakos S. (2000). Evaluation of HIV-1 Immunogen, an immunologic modifier, administered to patients infected with HIV having 300 to 549 x 10-6/L CD4 cell counts. JAMA, 284(17), 2193-2202.*

Kaptchuk TJ. (2001). The double-blind, randomized, placebo-controlled trial: gold standard or gold calf? Journal of Clinical Epidemiology, 54(6), 541-549.

Keuken DG, Moerman CJ, Haafkens JA, Klazinga NS, Riet ter G. (2004). Aandacht voor sekseverschillen in wetenschappelijk onderzoek en richtlijnontwikkeling in de Nederlandse Gezondheidszorg. Rapportage aan ZON-Mw, 8 April 2004.

Kjeldsen SE, Dahlöf B, Devereux RB, Julius S, Aurup P, Edelman J, et al. (2002). Effects of Losartan on cardiovascular morbidity and mortality in patients with isolated systolic hypertension and left ventricular hypertrophy. A Losartan Intervention for Endpoint Reduction (LIFE) substudy. JAMA, 288(12), 1491-1498.*

Laurent C, Kouanfack C, Koulla-Shiro S, Nkoué N, Bourgeois A, Calmy A, et al. (2004). Effectiveness and safety of a generic fixed-dose combination of nevirapine, stavudine, and lamivudine in HIV-1-infected adults in Cameroon: open-label multicentre trial. Lancet, 364(9428), 29-34.*

Lindholm LH, Anderson H, Ekbom T, Hansson L, Lanke J, Dahlöf B, et al. (2001). Relation between drug treatment and cancer in hypertensives in the Swedish Trial in Old Patients with Hypertension 2: a 5-year, prospective, randomised, controlled trial. Lancet, 358(9281), 539–544.*

Mandelbrot L, Landreau-Mascaro A, Rekacewicz C, Berrebi A, Bénifla JL, Burgard M, et al. (2001).Lamivudine-Zidovudine combination for prevention of maternal-infant transmission of HIV-1. JAMA, 285(16), 2083-2093.*

Mani KS, Rangan G, Srinivas HV, Srindharan VS, Subbakrishna DK. (2001). Epilepsy control with phenobarbital or phenytoin in rural south India: the Yelandur study. Lancet, 357(9265), 1316-1320.*

Manna DR, Bruijnzeels MA, Mokkink HGA, Berg M. (2003a). Ethnic specific recommendations in clinical practice guidelines: a first exploratory comparison between guidelines from the USA, Canada, the UK, and the Netherlands. Qual Saf Health Care, 12, 353-358.

Manna DR, Bruijnzeels MA, Mokkink HGA, Berg M. (2003). Minder etnische kennis in de standaarden van het Nederlands Huisartsen Genootschap over diabetes mellitus 2, hypertensie en over astma bij volwassenen dan in de onderbouwende literatuur. Nederlandsch Tijdschrift voor Geneeskunde 147(35), 1691-1695.
Marks HM. (1997) The progress of experiment: science and therapeutic reform in the United States, 1900-1990. New York: Cambridge University Press.

Mwinga A, Nunn A, Ngwira B, Chintu C, Warndorff D, Fine P, et al. (2002). Mycobacterium vaccae (SRL172) immunotherapy as an adjunct to standard antituberculosis treatment in HIV-infected adults with pulmonary tuberculosis: a randomised placebo-controlled trial. Lancet, 360(9339), 1050-1055.*

Nachman SA, Stanley K, Yogev R, Pelton S, Wiznia A, Lee S, et al. (2000). Nucleoside analogs plus ritonavir in stable antiretroviral therapy-experienced HIV-infected children. A randomized controlled trial. JAMA, 283(4), 492-498.*

Nduati R, John G, Mbori-Ngacha D, Richardson B, Overbaugh J, Mwatha A, et al. (2000). Effect of breastfeeding and formula feeding on transmission of HIV-1. A randomized clinical trial. JAMA, 283(9), 1167-1174.*

AUTEUR. (1994). NIH Guidelines On The Inclusion Of Women And Minorities As Subjects In: Clinical Research, NIH GUIDE, Volume 23, Number 11, March 18, 1994

NIH Guidelines on The Inclusion Of Women And Minorities As Subjects In Clinical Research - Updated August 1, 2000.

section 492B of the Public Health Service Act, added by the NIH Revitalization Act of 1993, Public Law (PL) 103-43, supersede The NIH Revitalization Act of 1993, PL 103-43, signed by President Clinton on June 10, 1993

NIH. (1998) NIH Policy and guidelines on the inclusion of children as participants in research involving human subjects. NIH Guide for grants and contracts.

Offringa M. (2003) Meten en Wegen. Amsterdam: Vossiuspers UvA.

Offringa M, Assendelft WJJ, Scholten RJPM. (2003). Inleiding in evidence-based medicine: klinisch handelen gebaseerd op bewijsmateriaal. Houten: Bohn Stafleu Van Loghum.

Pal DK, Das T, Chaudhury G, Johnson AL, Neville BGR. (1998). Randomised controlled trial to assess acceptability of phenobarbital for childhood epilepsy in rural India. Lancet, 351(9095), 19-23.*

Pepine CJ, Handberg EM, Cooper-DeHoff RM, Marks RG, Kowey P, Messerli FH, et al. (2003). A calcium antagonist vs a non-calcium antagonist hypertension treatment strategy for patients with coronary artery disease. The International Verapamil-Trandolapril Study (INVEST): A randomized controlled trial. JAMA, 290(21): 2805-2816.*

Perry HM, Davis BR, Price TR, Applegate WB, Fields WS, Guralnik JM, et al. (2000). Effect of treating isolated systolic hypertension on the risk of developing various types and subtypes of stroke. The Systolic Hypertension in the Elderly Program (SHEP). JAMA, 284(4), 465-471.*

Peto R, & Baigent C. (1998) Trials: the next 50 years. Large scale randomised evidence of moderate benefits. British Medical Journal, 317(7167), 1170-1171.

Piantadosi S, & Wittes J. (1993). Politically correct trials. Controlled Clinical trials, 14(6), 562-567.

Pierce N, Foliaki S, Sporle A, Cunningham C. (2004). Genetics, race, ethnicity and health. BMJ, 328, 1070-1072.

Pradhan AD, Manson JE, Rifai N, Nuring JE, Ridker PM. (2001). C-reactive protein, Interleukin 6, and risk of developing type 2 diabetes mellitus. JAMA, 286(3), 327-334.*
PREMIER. (2003). Effects of comprehensive lifestyle modification on blood pressure control. Main results of the PREMIER clinical trial. JAMA, 289(16), 2083-2093.*

Raz I, Elias D, Avron A, Tamir M, Metzger M, Cohen IR. (2001). β-cell function in new-onset type 1 diabetes and immunomodulation with a heat-shock protein peptide (DiaPep277): a randomised, double-blind, phase II trial. Lancet, 358(9295), 1749–1753.*

Reiber GE, Smith DG, Wallace C, Sullivan K, Hayes S, Vath C, et al. (2002). Effect of therapeutic footwear on foot reulceration in patients with diabetes. A randomized controlled trial. JAMA, 287(19), 2552-2558.*

Richards E. (1988) The politics of therapeutic evaluation: the vitamin C and cancer controversy. Social Studies of Science 18: 653-701

Scott RC, Besag FMC, Neville BGR. (1999). Buccal midazolam and rectal diazepam for treatment of prolonged seizures in childhood and adolescence: a randomised trial. Lancet, 353(9153), 623-626.*

Sever PS, Dahlöf B, Poulter NR, Wedel H, Beevers G, Caulfield M, et al. (2003). Prevention of coronary and stroke events with atorvastatin in hypertensive patients who have average or lower-than-average cholesterol concentrations, in the Anglo-Scandinavian Cardiac Outcomes Trial – Lipid Lowering Arm (ASCOT-LLA): a multicentre randomised controlled trial. Lancet, 361(9364), 1149-1158.*

Silva M de, MacArdle B, McGowan M, Hughes E, Stewart J, Neville BGR, et al. (1996). Randomised comparative monotherapy trial of phenobarbitone, phenytoin, carbamazepine, or sodium valproate for newly diagnosed childhood epilepsy. Lancet, 347(9003), 709-714.*

Skyler JS, Cefalu WT, Kourides IA, Landschulz WH, Balagtas CC, Cheng S, et al. (2001). Efficacy of inhaled human insulin in type 1 diabetes mellitus: a randomised proof-of-concept study. Lancet, 357(9253), 331–335.*

Staessen JA, Den Hond E, Celis H, Fagard R, Keary L, Vandenhoven G, et al. (2004). Antihypertensive treatment based on blood pressure measurement at home or in the physician's office. A randomized controlled trial. JAMA, 291(8), 955-964.*

Taha TE, Kumwenda NI, Hoover DR, Fiscus SA, Kafulafula G, Nkhoma C, et al. (2004). Nevirapine and Zidovudine at birth to reduce perinatal transmission of HIV in an African setting. A randomized controlled trial. JAMA, 292(2), 202-209.*

Taubert D, Berkels R, Roesen R, Klaus W. (2003). Chocolate and blood pressure in elderly individuals with isolated systolic hypertension. JAMA, 290(8), 1029-1030.*

Tejeda HA, Green SB, Trimble EL, Ford L, High JL, Ungerleider RS, et al. (1996). Representation of African-Americans, Hispanics, and Whites in National Cancer Institute Cancer Treatment Trials. Journal of the National Cancer Institute, 88(12), 812-816.

Vidaver RM, Lafleur B, Tong C, Marts SA (2000). Women subjects in NIH funded research literature: lack of progress in both representation and analysis by sex. Journal of women's health and gender-based medicine. 9: 1041-1047

Whyte SR, Geest S van der, Hardon A. (2002). Social lives of medicines. Cambridge: Cambridge University Press.

Yoshioka A. (1998). Use of randomisation in the Medical Research Council's clinical trial of streptomycin in pulmonary tuberculosis in the 1940s. British Medical Journal, 317(7167), 1220 - 1223.

Yusuf S, Gerstein H, Hoogwerf B, Pogue J, Bosch J, Wolffenbuttel BHR, et al. (2001). Ramipril and the development of diabetes. JAMA, 286(15), 1882-1885.*

---

[i]. A sleeping pill prescribed to pregnant women, which had serious adverse effects on the unborn child. Many babies were born with serious disabilities.

[ii]. Long-term animal studies have been done on hormonal contraceptives, but the general consensus is that the animal models have very little predictive value for humans.

[iii]. This was a predecessor to the 1994 guidelines which more explicitly demand attention for women and minorities.

[iv]. In a 2000 update on the Guidelines NIH has strengthened its reporting requirements (NIH 2000). "The Research Plan in the application or proposal must include a description of plans to conduct the valid analyses of the intervention effect in subgroups. The final protocol approved by the IRB must include these plans for analysis. The award will reguire the results of subset analyses must be reported to NIH in Progress Reports, Competitive Renewal Applications (or Contract Renewals/Extensions), and in the required Final Progress Report. Inclusion of the results of subset analyses is strongly encouraged in all publication submissions. If the analysis reveals no subset differences, a brief statement to that effect, indicating the subsets analyzed, will suffice".

[v]. The European Commission announced similar patent protection for companies who undertake clinical research on medicines in children in September 2004.

[vi]. For Epilepsy we had to extend the period to eight years, because we only found two trials in the reference period.

[vii]. To assess if your conclusions hold for more specialized medical journals we also reviewed a systematic sample (everye third RCT) published in the journals Diabetes Care and AIDS in the period 2000-2004. This resulted in a sample of 31 Diabetes RCTs and 29 AIDS RCTs. We found that subgroup analysis is less common in these journals than in the JAMA and Lancet. Only 13% of the Diabetes studies included subgroup analysis for sec, 10 % for age, and non for ethnicity. Of the 29 AIDS RCTs, only 3% had done subgroup analysis for sex, 3% for age and non for ethnicity. The lower percentages of subgroup analysis in these specialized medical journals may reflect that RCTs, with more power for subgroup analysis, are more likely to be published in the major medical journals.

[viii]. In the rhetoric about research and development cost of a new drug the price of $800 million is often mentioned, but the real cost may well be under $100 million (Angell 2004).

Review 4

Ethics and diversity in clinical research

Dick Willems, Tsjalling Swierstra

Contents

4

Ethics and diversity in clinical research

Summary

This chapter addresses normative questions related to the increasing attention for diversity in clinical research. It raises three questions: what normative concepts, especially concepts of justice, are used to argue for or against attention for different types of diversity in clinical research? What are normative issues related to new forms of diversity (other than the traditional age, gender, and ethnicity)? Should 'diversity' become a new criterion in the assessment of research protocols by research ethics committees and if so, what would that imply? These three questions are answered, as much as possible, on the basis of an electronic literature search.

The point of this article is that, in the ethics of clinical research, we need to move from a singular perspective on a right to protection to a mixed perspective, where the right to protection needs to be balanced against a right to representation or participation. Increasing attention for diversity, together with increasing patient pressure on participation, may thus reorient the discussion on the ethics of medical research with patients from a predominant focus on protection of subjects to a more balanced mix of focuses, in which the right to participation in research will be another important value.

4.1 Introduction

Ethical theory about clinical research was developed in the sixties, when it became clear that the abuse of patients for clinical research had not been limited to Nazi Germany, but also occurred in government-sponsored research in the US and Europe. It was argued, moreover, that the risk that

patients might be inappropriately subjected to research protocols was intrinsic to the whole enterprise, because of the 'double agency' of physicians as both caregivers and researchers. It was suggested that even the most benevolent doctor, put before the choice between the interest of his beloved research project and his patient, might sacrifice the latter's interests. This lead to the idea that patients had a right to protection from this risk, a protection that was realised, among others, by an ethical assessment of protocols.

However, times have changed. Starting with the AIDS community (Gifford 2002) and patients with rare diseases (Rai 2002), patients have increasingly argued that they not only wanted a right to protection from abuse in clinical research (some argued that the focus on protection was entirely paternalistic (Edwards 2004)), but also a right to be represented, as a group, in research and even an individual right to actually participate. Times have changed in another sense as well: it has become increasingly clear that there are important differences between patients with regard to the outcomes of clinical research (effect modification): differences between patients and patient groups are relevant for medical research and there are increasing technical possibilities to construct and assess such differences. Thus, new forms of diversity emerge, especially from new genomics technologies. These two developments, the demand from patient groups to be included and the emerging importance of effect modification, point in the same direction: diversity matters and should be an element in clinical research.

A third development that contributes to challenging the protection paradigm in the ethics of clinical research is the very expansion of research. Whereas, in the sixties, participating in a clinical study was rare for patients, and the large majority of patients were exclusively treated within the context of good care, nowadays, even the smallest rural hospital tends to participate in studies. Consequently, the ethical attention for clinical research can no longer be based on the idea that it is somehow alien to care - this makes the need for protection less obvious without, of course, abolishing it. Additionally, some patients (especially cancer patients) can only receive treatment within the context of a trial. In other words: clinical studies are no exception anymore, but tend to become a standard part of care and non-participation in trials could increasingly become non-participation in care. Therefore, a normative framework built on the idea that research is a corpus alienum in care hardly reflects the current situation, at least in hospital care. A final problem for the protection paradigm might be that it increasingly becomes clear that the bulk of clinical research with patients is quite innocent, often involving no more risk than a blood draw; so, much of the current clinical research hardly raises ethical concerns, and could therefore, maybe, be spared an elaborate ethical assessment.

The usual approach in such assessments, as specified in the various versions of The Helsinki Declaration, employs concepts of justice that

relates to the protection of those who, because of illness and/or a double bind, are considered vulnerable and, to a large extent, incapable of deciding freely for themselves. Those who argue for diversity in clinical research on the other hand, bring in concepts that relate to distributive justice: fairness and equality of chances, simple equity, or equity related to needs. From a right to protection to a right to representation and even a right to participation.

This chapter investigates the arguments and consequences of this shift, and proceeds as follows: first, we will outline the basics of the justice-as-protection paradigm, showing that it hinges on a contradiction between the physician as caregiver and the physician as researcher (viz. the hospital as a caring institution versus a scientific laboratory). The second section will describe the results of an electronic search in MEDLINE and Philosopher's Index, to find answers in the literature to the following questions:

1. In what terms is increasing attention for diversity justified or criticized? Is that in terms of justice, egalitarianism, or political correctness (for instance, fear of accusations of racism or ageism)? Are (emerging) new forms of clustering and diversity within clinical research subject of normative debate?
2. On what arguments did patients and patient groups demand inclusion in instead of protection from trials?

In the discussion section, we will briefly address the question how a right to protection and a right to representation or participation can be balanced.

4.2 Protecting patients from inappropriate research

Historically, issues related to ethnicity, one of the key forms of diversity, were at the core of the ethics of clinical research with human subjects. The Tuskegee syphilis study, conducted in The United States from 1932 to 1972, has become an eponym for morally bad clinical research – almost on a par with the Nazi Germany experiments in concentration camps (Fairchild 1999; Bowman 1999). In that study, which was sponsored by the US Public Health Service, African American crop workers from Alabama were denied treatment for tertiary syphilis, so that its natural course could be studied. As a part of the study, the men were told they had 'bad blood' instead of their true diagnosis – a clear example of racist use of so-called culturally sensitive concepts. In 1997, President Clinton offered apologies for the study to the African American population; some say these apologies were at least partly intended to diminish what has been called a 'Tuskegee effect' (Bates 2004): the distrust of black Americans toward the medical establishment, and clinical research in particular (White 1997). Even though the experiment is always regarded as the starting point of ethical concerns around clinical research, which were later translated into a protection-based set of rules, the initial opposition against the experiment was inspired by distributive concerns to the effect that the pains of medical research should not exclusively be inflicted on the disadvantaged.

The foundation of what we call the protection paradigm lies in the principle of non-malfeasance ('do no harm'). Non-malfeasance is a core element of medical ethics. And so it has been, it seems, since the early days of the profession. The Hippocratic Oath contains the principle that the practitioner should do everything possible to help the patient, and to refrain from actions that might harm him or her. And to take a recent example: in Beauchamp and Childress's influential Principles of Biomedical Ethics it is one of the four key principles of medical ethics.

However, this continuity is not as straightforward as it seems. Instead, it has been fought over, especially in modern times. Let us try to return to the original 'context of invention' of this principle and ask ourselves why the medical profession in Antiquity came up with a professional oath in the first place. This is not self-evident. Ancient bakers and butchers did not develop professional codes. So why did the doctors? It can be assumed that the oath served as a countermeasure/answer against the basic, latent distrust of the patient towards the medical profession. Such distrust is understandable, because the doctor wields great power over life and death whereas the patient is relatively wanting, and power- and defenceless. So, the promise not to harm the patient can be understood as a trust-generating device. And a quite necessary device as well, because without it, who would be willing to invite a doctor to the side of one's bed?

The principle of non-malfeasance belongs to a practice that can be characterized by three features. In the first place, the relation between the doctor and the patient was highly individual. Secondly, the patient had at least some power over the doctor in so far as (s)he could decide whether to pay him or not for his services. And thirdly, the doctor had no other interests than to cure the patient. These three features may have characterized the doctor-patient relationship for thousands of years, but for many patients they changed in the course of the nineteenth century when the methods and goals of the experimental sciences intruded into the domain of medicine. Scientists like Pasteur, Koch and Eijkman rapidly progressed in acquiring knowledge about the causes of infection diseases (ten Have 2002). But this scientific progress demanded more and more subjects for experimental research.

This intrusion itself was made possible by the fact that more and more poor people were concentrated in hospitals, barracks, and prisons. These new spatial configurations made them more easily available as objects for medical research. In the relation between these groups of patients and their doctors, the three features mentioned above were often absent. The doctors treated populations, not individuals. Furthermore, these patients were relatively powerless – more object than subject, as it were – because they did not pay the doctors themselves – being a prisoner, a soldier, or simply too poor to afford medical care. Finally, the agenda of the doctor had often changed: not simply helping the patient, but using the patient to further medical knowledge. This change occurred against the ideological background of nationalist discussions about the need for a healthy

population. The population as a whole became the object of medical intervention. As a result, the traditional 'hippocratic' relation between the doctor and the patient changed into a ménage à trois: the doctor-researcher saw himself confronted with a conflict of loyalties. The answer 'don't do any harm' would often turn into a question: who not to harm, the individual patient/research object, or the population (or even mankind) as a whole? From a consequentialist perspective the answer often was all too clear: some individuals have to be harmed, to avoid much larger harm to, or – later - to promote the benefit of the collective. And this answer was all the more easily given, when these individuals were prisoners, prostitutes or ethnic/racial minorities.

As we all know, especially the totalitarian systems of the twentieth century have put this utilitarian calculus to a horrendous use before, during and after the Second World War. But this calculus was not confined to those regimes. As late as 1966, the American Harvard professor Henry Beecher could publish a famous article in which he listed twenty two examples of medical research that sacrificed individuals to the long term good of the collective. This article lead to an uproar around the western world, and finally resulted in American and European laws that clearly forbid these scientific practices. A major explanation for this tremendous effect of the article can be found in the emancipation of the 'usual subjects' of these types of medical research: the poor, blacks, prisoners, prostitutes, psychiatric patients, etcetera. As a result, today the principle of non-malfeasance is usually seen as a cornerstone of medical ethics, as central as in the days of Hippocrates. And again, it can be understood as a device to solicit trust and cooperation of the patient.

In this new context, the ancient non-malfeasance principle acquires a new, modern thrust: it stresses that the patient should be protected against the medical scientist, and that (s)he is first and foremost an individual with moral rights instead of member or representative of a group whose rights can be in any way subsumed under the welfare of the collective. However, this moral paradigm that has been dominant for the last forty or fifty years, is put under a strain by two recent developments. In the first place, instead of demanding to be protected from medical experiments, in a new development certain patient groups are now demanding to be included in them. Secondly, instead of treating an individual singularly as an individual, recently patients themselves have demanded to be treated as a member or representative of a group, of a collective. In the next section we will turn our attention to these two developments that undermine the prevailing moral paradigm concerning the use of human beings as subjects for medical experimentation.

The Helsinki Declaration (latest version 2000) states that 'In medical research on human subjects, considerations related to the well-being of the human subject should take precedence over the interests of science and society.' The central focus of the ethical assessment of research, the protection of the study subjects against the sacrifice of their well-being to

scientific interests, could hardly be more clearly expressed. The reason for this focus on protection lies in two vulnerabilities: firstly, the essentially vulnerable position of the patient who may feel dependent for her care on the same physician who asks her to participate in a research project. And secondly, the physical vulnerability of a sick person who may be less resistant to experimental procedures and treatments than healthy subjects. Part of this protection task is to ensure that patients are only asked to participate in valid forms of research, since bad research is in itself considered unethical.

The protection argument is a form of justice argument, in which fairness means that the strongest shoulders carry the largest weights (and inversely, the weakest should have most benefits): possibly harmful or burdensome research with no expected direct benefit should be done on healthy volunteers, probably beneficial and hardly burdensome studies on sick people. That is why research that can only be harmful, like phase-I trials, should be conducted on healthy volunteers, not sick people. Similar reasoning would imply that studies that can be done on adults should not be done in children, not only because of consent issues, but above all because children are considered more vulnerable. Only if the other form of equity argument, fairness of results, would override this form of equity, would it be acceptable to include children.

4.3 Challenging patient protection as the exclusive ethical problem

If attention for diversity is argued for in terms, not of protective, but of distributive justice, this leads to the idea that every patient has a prima facie right to participate in clinical research, and that the exclusion of patients or healthy volunteers should be based on the research questions, not on discrimination. The right to be represented (that is usually argued for in terms of effect modification) and the claim to equal access for all, both point in the direction of greater diversity.

Representation
Representation as an argument states that the participants in trials should be a faithful representation of the groups that will potentially benefit from the results of the trial. One of the main reasons for this argument is that otherwise, results will be applied in patient groups for which the balance of benefits and harms may be substantially different than in the research population (effect modification).

The shift from protection to representation can be clearly shown in the discussion on the involvement of children in medical research. An example is the way in which attention for children in research is argued for in terms of the risks of applying knowledge obtained in adults without proper research in children. In the older literature, the inclusion of children is sometimes warned against because of the (supposed?) risks involved in participation. This seems to be a waning type of argumentation against

participation; nowadays, the argument is more often couched in terms of balancing recruitment and protection (Meaux 2001). Some papers argue in terms of a right to access of children to research (Arnold 1995; Sugarman 2004).

The problems surrounding informed consent for research with children are often mentioned in the literature, but rarely as an argument against including them (Nelson 2003; Kodish 2003). Kodish, in a recent editorial, argues that the time that parents decided about the participation of children without consulting them ("previous history when children were considered chattel"), is over: efforts should be made, he says, to give even very young would-be participants a voice in this decision (Kodish 2003). Obviously, it is not clear whether this will lead to increasing or decreasing participation of children in (some) trials.

Interestingly, one publication argued for increasing attention to older adolescents who, above all for reasons of compliance, are under-researched in paediatric oncology (Bayer 2002). From the paper, it is not entirely clear whether this is bad because of effect modification (compliance!), or because it is considered bad in itself to neglect certain groups in research, for reasons of 'simple equity'. Moreover, the example shows a point that we will have to come back to: increasing the number of participants from diverse groups means more than just widening inclusion criteria, it also means making participation in research 'do-able' for these other groups. Adolescent cancer patients might ask for different study designs or visit schedules if non-compliance is the main reason why they are not studied.

With regard to the elderly, a similar shift in ethical attention may be noted as concerning children. Arguments in favour of the inclusion of the elderly population in research are usually couched in terms of the increasing number of elderly in Western societies, but also in terms of the specific risks elderly people, who often have extensive co-morbidity, run when prescribed treatments have only been tested on younger adults with one disease at a time (Wenger 1992, High 1995, Bayer 2000). One author explicitly couches the justification of attention to diversity in political terms (adequate representation): all groups should as much as possible, be represented in clinical trials (Weijer 1999). This would apply for race/ethnicity, age, gender, but also for mental illness, HIV, substance abuse, and the like. Other issues that are raised are, of course, the question of mental competence and proxy representation in older people (Olde Rikkert 1995, Merson 1994, Kapp 1994, Barron 2004), but also the financial, logistic and social accessibility of clinical research. As Bayer (2003) notes, many elderly people have trouble organising and making the extra expenses for their participation if that involves more frequent hospital visits, regular questionnaires, etc. Increasing age diversity in trials, then, would involve creating favourable conditions for elderly, which on the other hand may make trials less attractive to payers.

Participation
An individual right, not only to representation, but to actual participation can be claimed by pointing to the intrinsic advantage of participating in trials: either because participants had at least a chance of being the first to receive a possibly beneficial treatment (this was the reason why AIDS patients asked for access), or because of the discovery of the so-called inclusion benefit: there is increasing evidence that participating in a trial is beneficial, no matter in what trial arm the individual ends up. Even those who get placebo seem to be better off than patients receiving regular care – therefore, it may be argued that all patients should have a fair chance to participate in trials.

Patients could start claiming a right to this form of health care benefit. On the other hand, the risks inherent in the participation in trials (which may be smaller than usually thought) may also lead to attention for diversity. Both benefits and risks would need to be justly distributed.
The call for attention to diversity in terms of gender may be taken as an example where both effect modification and simple equity play a role. It has been widely discussed in the literature. Since female patients represent more than half of the potential users of drugs, it is hardly surprising that this is the only form of diversity that enjoys the favourable interest of the pharmaceutical industry (Bush 1994). Originally, exclusion of women from early-phase trials was motivated, if at all, with protection arguments, especially for women in the childbearing age. More than ten years ago, the FDA has modified its policies on this issue, stating that sex-specific studies of drugs are required as a part of the registration of new drugs, and that 'being of childbearing age' no longer counts as an exclusion criterion per se (Merkatz 1994). In a paper from the same period, the exclusion rate of women from clinical trials is presented as an injustice, not only because it may lead to inadequate health care, but because it is linked to women's oppression (DeBruin 1994).

4.4 What type of diversity counts?

One paper investigating racial categories used in cancer research, addressed this issue (Figgs 2003). Its conclusion is that the use of racial category of 'African-American' should be abandoned for the purposes of research into breast cancer, not so much because of political correctness or normative reasons, but because they are incorrect and hardly useable criteria.

An area where new categories of respondents (and therefore, new forms of diversity), that may claim equal representation in research, are constantly formed is pharmacogenetics, the study of diversity in response to medication. Even though the use of standard ethnic or racial classifications is common in this field of research, several authors have urged for the radical abandonment of such categories, and for the

development of new ones based on commonalities in drug response (Keville 1994, Soskolne 1997).

The first pharmacogenetic studies in the Fifties concerned ethnic differences in the response to medication (Bayer 2000), and thus classical categories of diversity. Evans, one of the 'big shots' in pharmacogenetics, argues, in a review in Science, that racial distinctions are crucial for research into the risk of developing disease, but also into differences in response to medication (Evans 2002). Evans even expects an increasing importance of race and ethnicity in the selection of patients for research and treatment. However, some authors hope and expect that the traditional ethnic differences will loose their relevance as a consequence of genomics (Foster 2001).

Where authors like Foster see the possibility that pharmacogenetic information will create new categories and distinction between people as a danger, there could just as well be the possibility that a proliferation of categories would lead to their disappearance as categories. Foster, however, argues that this will not happen all on its own, since racial categories are usually taken as the starting point for the selection of research subjects.

A recent study in Nature Genetics seems to confirm the expectation that new forms of diversity will emerge as a consequence of – mainly - pharmacogenetics: the existing phenotypical criterions of ethnicity have proven to be less fruitful than totally new pharmacogenetic forms of clustering.(McLeod 2001) The authors of that study argue, contra Evans, that standard ethnicity categories should be left aside in clinical research. This is a discussion about which categorization of humans is acceptable and desirable in clinical research. Would a categorization along (pharmaco)genetic lines be politically and ethically preferable to a categorization along phenotypic or phenomenological lines? One of the advantages could be that it would make crude generalizations and even discrimination harder or even impossible. All of us may be fast metabolizers of some drugs or foodstuffs, all of us may be part of the ethnic minority of those how develop severe side-effects to – say – chloramphenicol. Without falling in the trap of yet another form of positivism (the one that says that biology will tell us what distinctions we need to make), one could say that the non-traditional forms of diversity would make distinctions more open than the classical triad of gender, age and race allows for.

New forms of diversity that might result from the use of pharmacogenetics techniques will probably be less visible than the forms of diversity this project is about (age, gender, ethnicity). Herceptin is a case in point: the chances of responding to the drug cannot be known without a specific test.

A consequence of the increase of using genetic inclusion criteria could be that inclusion would become more instead of less restrictive. Foster, for instance, says that pharmacogenetics may lead to two societal dangers: the widening of the difference between those for whom a treatment will be available on the one hand, and increasing possibilities for discrimination because the knowledge about the expected response to medications may come to function as a selection criterion for insurance and jobs.

4.5 The role of Research Ethics Committees

In a limited number of commentaries and editorials, the question is raised if, and to what extent, diversity should be an issue for the ethical assessment of protocols by Research Ethics Committees (RECs). As early as 1994, the US National Institutes of Health issued a guideline stating that 'the Inclusion of Women and Minorities as Subjects in Clinical Research' should be a concern for those who perform ethical assessments. REC's should see to it, that women and minorities will be included in trials in a manner that makes valid sub-analyses possible. However, the NIH Guidelines stipulate that not only RECs, but also (and foremost) funding agencies should take diversity into account when making funding decisions (NIH 1994). The Dutch Manual for RECs, however, does not address the issue of diversity (CCMO 2002).

In most Western countries, clinical research projects are assessed for ethical aspects by RECs (Institutional Review Boards). A second, related, aim of the ethical assessment is to ensure that patients are fully aware of and agree with the fact that their treatment is wholly or partly experimental. This, clearly, stems from the indignation about experiments that patients were subjected to unknowingly, such as the Tuskegee study. The question is whether attention for justice as equal representation and/or a right to participation, for instance because of relevant diversity, could be an additional aim of ethics assessment. As far as attention for diversity increases the scientific validity of the study, it could easily be shown to be part of the requirement that patients are only subjected to quality research. However, if the demand to pay attention to diversity stems from notions of a right to participation, this seems to imply a different perspective that may be more difficult to reconcile with a protection task.

In a 2001 paper, philosopher Lisa Eckewiler argues that attention for diversity as a part of the inclusion criteria is not enough for the ethical assessment of protocols. She says it should lead to the involvement of relevant populations groups in the assessment of protocols, and to effective consultation, by RECs, with the various communities affected by the research project.

A further question is whether it is the RECs that have a task in safeguarding the interests of patient groups by looking at the conditions that favour or disfavour there representation. Attributing such a task to

RECs would demand that they are or become competent to assess these matters. Even though it is not difficult for RECs to ask questions about the relevance of diversity for a clinical study, they may not have the competence to assess the answers given by researchers. This would lead to the hardly satisfactory situation that any answer would have to be accepted.

If attention for diversity should really become a guiding principle in the design and the performance of studies, then these should not only allow for the participation of children, women, ethnic minorities and aged persons, they should indeed create the opportunities to participate for these groups. For instance, the Tuskegee scandal has made it unattractive for Afro-Americans to participate in trials, and so might all sorts of organisational, social, and physical obstacles make participation unattractive to, for instance, the elderly. If elderly patients, and especially those with functional limitations, are to be enrolled in trials, it will hardly be sufficient to modify inclusion criteria; it might be necessary to accept concessions in the number of control visits, or in the intrusiveness of diagnostic procedures. It could, then, be advisable that those who assess research proposals, be it financiers or ethics committees, would take the actual accessibility of trials as an item of assessment.

An even further question would be whether some form of affirmative action would be necessary. If it is true that most trials up to now have been for white, relatively young males with no more than one condition, the exclusion of some groups from trials during the last decades might make it necessary not only to facilitate their inclusion, but even to redress the balance; this has been argued for with regard to children.

This would mean that studies should be designed and organised to make participation attractive for these groups, they should also test procedures and treatments that are relevant to them and that would be of value to these groups if the study shows them to be efficacious. It would mean that RECs would have to assess whether these opportunities are sufficiently guaranteed in the project.

4.6 Conclusion/recommendation

We have argued that there may be various forms of fairness arguments involved in the discussion about the ethical aspects of diversity in clinical research: equitable distribution of benefits and harms associated with research itself, and equitable distribution of beneficial and harmful results. We would recommend that the former, that most clearly has to do with distributive justice, needs to be a primary concern for the financiers and policymakers in the area of clinical research, whereas the second, that has to do with effect modification, also needs to be a part of the ethical assessment of protocols by RECs. Individual RECs will have the task to balance the right to protection and the right to representation or

participation: they will have to sort out in what situation will protection need to remain central, and when is representation or participation as, or even more, important.

Quite some attention has been given, in this review, to a different level in which ethics and diversity may come together: the ethical issues surrounding new forms of diversity that result from medical research. This has little to do, probably, with the assessment of research protocols, but the more with the politics of diversity as they develop in the genomics era. The third section of this paper is essentially a recommendation for normative studies of the kind of categorizations and classifications that may be encouraged by clinical research. Such studies should look at clinical research as a way to, to use Austin's term, perform new forms of categorization.

References

Arnold LE, Stoff DM, Cook E Jr, Cohen DJ, Kruesi M, Wright C et al. Ethical issues in biological psychiatric research with children and adolescents. J.Am.Acad.Child Adolesc.Psychiatry 1995;34(7):929-39.

Barron JS, Duffey PL, Byrd LJ, Campbell R, Ferrucci L. Informed consent for research participation in frail older persons. Aging Clin Exp Res 2004;16(1):79-85.

Bates BR, Harris TM. The Tuskegee Study of Untreated Syphilis and public perceptions of biomedical research: a focus group study. J Natl Med Assoc 2004;96(8):1051-64.

Bayer A, Fish M. The doctor's duty to the elderly patient in clinical trials. Drugs Aging 2003;20(15):1087-97.

Bayer A, Tadd W. Unjustified exclusion of elderly people from studies submitted to research ethics committee for approval: descriptive study. BMJ 2000;321(7267):992-3.

Bhopal R. Is research into ethnicity and health racist, unsound, or important science? BMJ 1997;314(7096):1751-6.

Bleyer A. Older adolescents with cancer in North America deficits in outcome and research. Pediatr Clin North Am 2002;49(5):1027-42.

Bowman JE. Tuskegee as a metaphor. Science 1999;285(5424):47-50.

Bush JK. The industry perspective on the inclusion of women in clinical trials. Acad Med 1994;69(9):708-15.

CCMO. Manual for the review of medical research involving human subjects. 2002. Den Haag, CCMO.

DeBruin DA. Justice and the inclusion of women in clinical studies: an argument for further reform. Kennedy Inst Ethics J 1994;4(2):117-46.

Eckenwiler L. Moral reasoning and the review of research involving human subjects. Kennedy Inst Ethics J 2001;11(1):37-69.

Edwards SJ, Kirchin S, Huxtable R. Research ethics committees and paternalism. J Med Ethics. 2004;30(1):88-91.

Fairchild AL, Bayer R. Uses and abuses of Tuskegee. Science 1999;284(5416):919-21.

Figgs LW. Breast cancer research among African-American women: accurate racial categories? Cancer 2003;97(1 Suppl):335-41.

Gifford AL, Cunningham WE, Heslin KC, et al. Participation in research and access to experimental treatments by HIV-infected patients. N Engl J Med. 2002;346(18):1373-82.

High DM, Doole MM. Ethical and legal issues in conducting research involving elderly subjects. Behav Sci Law 1995;13(3):319-35.

Horner JS. Retreat from Nuremberg: can we prevent unethical medical research? Public Health 1999;113(5):205-10.

Kapp MB. Proxy decision making in Alzheimer disease research: durable powers of attorney, guardianship, and other alternatives. Alzheimer Dis Assoc Disord 1994;8 Suppl 4:28-37.:28-37.

Keville TD. The invisible woman: gender bias in medical research. Women's Rights Law Report 1994;15:123-42.:123-42.

Kodish E. Informed consent for pediatric research: is it really possible? J Pediatr 2003;142(2):89-90.

Marson DC, Schmitt FA, Ingram KK, Harrell LE. Determining the competency of Alzheimer patients to consent to treatment and research. Alzheimer Dis Assoc Disord 1994;8 Suppl 4:5-18.:5-18.

Masoudi FA, Havranek EP, Wolfe P, Gross CP, Rathore SS, Steiner JF et al. Most hospitalized older persons do not meet the enrollment criteria for clinical trials in heart failure. Am Heart J 2003;146(2):250-7.

McLeod HL. Pharmacogenetics: more than skin deep. Nat Genet 2001;29(3):247-8

Meaux JB, Bell PL. Balancing recruitment and protection: children as research subjects. Issues Compr Pediatr Nurs 2001;24(4):241-51.

Merkatz RB, Junod SW. Historical background of changes in FDA policy on the study and evaluation of drugs in women. Acad Med 1994;69(9):703-7.

National Institutes of Health. NIH Guidelines on the Inclusion of Women and Minorities as Subjects in Clinical Research. Fed Regist 1994;59:14507-13.

Nelson RM, Reynolds WW. Child assent and parental permission: a comment on Tait's "Do they understand?" Anesthesiology 2003;98(3):597-8.

Olde Rikkert MG, Verweij MF, Hoefnagels WH. [Informed consent and mental competence of the elderly in medical-scientific studies]. Tijdschr Gerontol Geriatr 1995;26(4):152-62.

Rai AK. Pharmacogenetic interventions, orphan drugs, and distributive justice: the role of cost-benefit analysis. Soc Philos Policy 2002;19(2):246-70.

Soskolne CL. Ethical, social, and legal issues surrounding studies of susceptible populations and individuals. Environ Health Perspect 1997;105 Suppl 4:837-41.

Sugarman J. Determining the appropriateness of including children in clinical research: how thick is the ice? JAMA 2004;291(4):494-6.

Weijer C. Selecting subjects for participation in clinical research: one sphere of justice. J Med Ethics 1999;25(1):31-6.

Wenger NK. Exclusion of the elderly and women from coronary trials. Is their quality of care compromised? JAMA 1992;268(11):1460-1.

White RM. Grand Dragon or windmill: why I opposed the Presidential apology for the Tuskegee Study. J Natl Med Assoc 1997;89(11):719-20.

Witzig R. The medicalization of race: scientific legitimization of a flawed social construct. Ann Intern Med 1996;125(8):675-9.

Appendix

Methods

This review is, as much as possible and relevant, based on data from a literature review. An electronic search has been performed in MEDLINE and Philosopher's Index.

1. Medline. Using Pubmed, searches have been performed using the following search terms, always in combination with 'clinical research': ethnic, ethnicity, age, elderly, children, gender, women, and diversity. As 'limits' I used: Bioethics, Review, Editorial, History of Medicine. The various combinations of these search phrases and their limits led to a total of 83 non-duplicating hits. From these, I first selected editorials and historical studies and then, on the basis of the abstracts, I further selected papers that had normative issues as a core subject. Moreover, papers in ethics journals were retained for analysis. Thus, for example, descriptive studies in medical journals about the extent to which diversity is addressed in research were eliminated from the analysis, just as studies addressing the study of ethnicity as a subject on its own. The total number of publications retained for analysis was 45.

2. Philosopher's Index. A search using the same terms resulted in a limited number of hits, none of which were new compared to the Medline search.

Review 5

Methodological implications of focussing on diversity in clinical research

Madelon van Wely, Ronald Stolk, Patrick Bossuyt, Martin Offringa, Rick Grobbee, Karien Stronks

Contents

5

Methodological implications of focussing on diversity in clinical research

Summary

When characteristics like age, gender and ethnicity are modifiers of a treatment effect this might have implications for the generalisability of the results of medical intervention studies. On the other hand, the variability between diversity groups is usually smaller than the variability within subgroups. Therefore, a proportional representation of gender, age and ethnicity subgroups is, by itself, not a prerequisite for generalisation of the results.

We should first investigate whether there is a clinical relevant modification of effect of the intervention. In the presence of effect modification one may choose to design new trials that specifically address the effect of the intervention in particular subgroups. Another option is to enlarge the original trial with sufficient members from each subgroup such that subgroup analysis can be performed. Where subgroup analysis is performed the results may suggest differences between the subgroups included, but further research will be necessary to confirm these. A decision must be made whether to power the original research sufficiently to do subgroup analysis in the first place.

When aiming at including a more heterogeneous population one may encounter problems with adequate trial enrolment, self report and completeness of follow-up in different age, gender and ethnic groups. A systematic literature search was performed to elucidate these problems and to suggest solutions.

The analysis of a study should consider possible effect modification. With pre-planned subgroup analyses in an adequately powered study it would be most elegant to first perform statistical tests of interaction, followed by subgroup analyses to estimate the differences. A cost-benefit analysis may be valuable in deciding on whether adequate representation of subgroups is worthwhile.

5.1 Introduction

The validity of the findings generated by a randomised trial is an important dimension of quality. Internal validity of the trial implies that the differences observed between groups of patients allocated to different interventions may, apart from random error, be attributed to the treatment under investigation. In general in clinical trials strict inclusion and exclusion criteria are set to create a homogeneous study population such that the chance of side effects, co-morbidity and early drop-out becomes as small as possible[1-3]. This will increase the internal validity of the trial. In contrast, external validity, or generalisability, is the extent to which the results of a study provide a correct basis for generalisation to patients that were not included in the study.

Internal validity is a prerequisite for external validity: the results of a flawed trial are invalid, and the question of its external validity then becomes redundant. There is no external validity per se; the term is meaningful only with regard to clearly specified disease features that were not directly examined in the trial. Can results for instance be generalized to a given individual patient or to groups that differ from those enrolled in the trial with regard to age, sex, ethnicity, and co-morbid conditions?

The generalisability of results obtained from randomised clinical trials is often questioned as the patient population selected for inclusion is in general quite homogeneous. Relevant differences in patient characteristics between the study population and the total patient population to whom the intervention of interest may be applied do only affect the generalisability of the results when there is a significant difference in treatment effect between subgroups of patients. Outcomes can change with age: the probability to become pregnant following IVF, for example, decreases with age. Results from an IVF trial in women aged below 30 are therefore not being necessarily generalisable to women above 35 years of age. Furthermore, due to differences in genetic predisposition, certain ethnic groups may respond differently to a given medical intervention than others. In these examples the factors involved in an interdependent relationship can be regarded as having their effects modified by each other, which gives rise to the terms "effect modification" and "effect modifier". Since very few exposures cause disease entirely by themselves, nearly every causal factor must modify the effect of other causal factors and have its effect modified by them. When these other factors are unidentified, they are generally regarded as part of the background environment, assumed to be uniformly distributed between the intervention and comparison groups and hence disregarded. Part of the challenge of epidemiologic research is to identify major modifying factors that are not uniformly distributed, so that differences in findings across studies and between subgroups can be understood.

Despite many theoretical reasons and some empirical examples of substantial effect modification by characteristics of patients, systematic comparisons of randomised and non-randomised studies have so far failed to show many differences. It appears that differences between subgroups are not very frequent[4].

If characteristics like gender, age and ethnicity do modify the intervention effect this will have implications for the generalisability of the reported trial results. In this case one may choose to design new trials that specifically address the subgroups that appear to have altered treatment effects. Another option is to enlarge the original trial with members from these subgroups, i.e. to include a more heterogeneous population. This review aims to describe analytical and interpretational problems that may arise when including different gender, age, co-morbidity and ethnic groups in randomised intervention studies. An attempt is made to guide investigators and policy makers in dealing with these problems.

5.2 Material and Methods

Studies included in this review

We will limit this review to explanatory randomised trials only. These studies aim to estimate the existence and size of the effect of an intervention in an ideal situation with maximal internal validity. In contrast, pragmatic randomized trials are set up to study effectiveness in real clinical circumstances. The potential differences in effect by age, gender and ethnicity are especially an issue in explanatory trials. In pragmatic trials the study is conducted on patients who represent the full spectrum of the population to which the treatment might be applied[5].

The danger of pragmatic trials is that internal validity may be overly compromised in the effort to ensure generalisability. Therefore, in an evidence-based medical approach explanatory randomised controlled trials (RCTs) are generally viewed as a paradigm for research on the benefits of interventions.

We excluded non-randomised studies from this review as results coming from RCTs tend to have a larger impact on decision-making than the results of non-randomised research. In addition, the inclusion of patients in explanatory RCTs is often more selective than in non-randomised trials, which gives rise to questions on generalisability.

Focus of this review

The factors involved in an interdependent relationship can be regarded as having their effects modified by each other, which gives rise to the terms "effect modification" and "effect modifier". If gender, age and ethnicity modify the intervention effect this will have implications for the generalisability of the outcomes.

We will focus on how to deal with effect modification in large RCTs and on alternative methodological approaches to deal with problems that may occur when introducing heterogeneity in gender, age, co-morbidity and ethnicity in randomised trials. Such approaches will pertain to all stages of a trial, i.e. the preparation of the trial, recruitment/enrolment, randomization, measurements, follow-up, and analyses. The consequences of these approaches will be discussed.

We will also focus on issues that may interfere with the inclusion of heterogeneous populations or with obtaining specific data. For instance

when dealing with questionnaires at the measurement stage, these questionnaires may not be validated for different patient groups. This problem often arises when different ethnic groups are included in a study. Furthermore, language problems may limit the inclusion of certain ethnic groups.

Methodological issues

A separate literature search was performed to identify methodological studies concerning all stages in the design and execution of a RCT. The basic search strategy consisted of a MedLine and Embase search (from 1990 until December 2004) using the MeSH words randomised clinical trial, controlled clinical trials, random allocation, population characteristics, reproducibility of results, age factor, sex, ethnicity and the text words generalisability, external validity, heterogeneity, age, gender, minorities. Subsequently, a specific search was performed at the separate stages of the trial using the following MeSH and text words [tw]:

At stage 0 - preparation of the trial: effect modifiers, effect modification [tw], generalisability [tw], trial design.
At stage 1 - enrolment of a heterogonous population: patient selection, patient recruitment, patient selection [tw], recruitment [tw], inclusion criteria [tw], eligibility.mp, informed consent, disclosure, informed consent[tw], participation[tw].
At stage 2 - allocation of interventions: stratification, research design
At stage 3 - measurements: quality of life, questionnaires, validation.mp, illiteracy, language.mp
At stage 4 - MeSH patient dropouts
At stage 5 - analysis: statistics, effect modifiers, effect modification [tw], interaction [tw], subgroup analysis [tw]

For all stages of the design and execution of a randomised controlled trial we formulated the following questions:

1. In situation with diverse populations what is the current practice and how does this practice consider diversity?
2. Is including a diverse or heterogenic study population in a given field necessary? If the answer to the second question is definitely yes, two more questions follow:
3. Which factors complicate the implementation of the ambition described in question 2?
4. What can be done to resolve these problems?

5.3 Stage 0 - Preparation of the trial

One of the central issues concerning the design of any RCT is what type of patients should be selected to participate in the trial. Homogeneity of the population to be studied is considered to be important as homogeneity improves the internal validity and precision of the results. On the other hand, some scientists point towards the need for inclusion of a

representative presentation of the total population: the correct case-mix must be included to be able to generalise the results[6].
Representation of subgroups that differ in gender, age and ethnicity is, by itself, not a necessity for generalisation of the results[7]. As explained before we first have to determine whether there is a potential clinical relevant modification of effect of the intervention, i.e. effect modification by certain patient characteristics. Therefore, previous research evidence regarding potential differences among the studied intervention effect in different gender, age or ethnic groups should be gathered and studied when designing a trial[8]. In any trial, there are always several potential effect modifying factors. The following table describes potential sources of effect modification:

Table 5.1 Potential sources of effect modification.

Potential effect modifying factors	Examples
Study design characteristics	Setting, patients, co-medication, study duration, measurements and method used for measurements, completeness of follow-up, methodological quality of the study
Patient characteristics	Age, gender, ethnicity, biochemical markers, genetic polymorphism
Disease characteristics	Method and sensitivity of diagnosis, severity of disease, staging and biological response
Intervention characteristics	Application, route, dosing, intensity, time point of treatment, duration and compliance

Performing a multi centre RCT will already account for some potential effect modifying factors, like setting and patient characteristics. A multi centre study not only captures all kind of differences in care between medical centres but is also more likely to include a diverse population. This will eventually improve the generalisability of the results[9]. Depending on the object of the study it may be worthwhile to select patients not only at second and third line centres but also at first line to enable recruitment of participants from the general patient population.

What would be the ideally situation?
In the ideal situation presence of effect modification is identified or at least suspected before starting the trial. Thus, the chance of identifying actual differences in treatment effect can be maximised at the trial design stage.

How can potential effect modification be detected?
A frequently used method for the identification of effect modifiers is subgroup analysis. Subgroup analyses are usually retrieved with the help of statistical analyses from population data. The power of the evidence coming from subgroup analyses is generally low[10,11]. Evidence from other study types like fundamental research, pharmacokinetic and pharmacodynamic studies should be taken into account. Furthermore,

potential effect modifiers can be identified on basis of the mechanism of action of the intervention combined with knowledge of the pharmacology, pathophysiology, etc[12,13].
Small studies may provide an impression of the size of the effect modification, for example by using intermediate endpoints.
Next to biological factors, cultural factors may also act as modifier by affecting compliance and completeness of follow-up. This is quite a different phenomenon that, unlike the presence of an intrinsic biological factor, is in essence preventable. This is further discussed later in this paper (at trial stage 4).

How to deal with potential effect modification of gender, age or ethnicity?
If preliminary evidence supports the existence of subgroup differences (modification) in intervention effect there are two options. A separate RCT can be performed in the specific subgroup or members from different subgroups are included into the trial[14,15]. For instance, consider the situation where women respond differently than men to a certain drug but until now trials have only been performed in men. In this case it would be most sensible to perform a separate study with women only. In another situation, when a new drug is tested and age is expected to be an effect-modifier on the basis of known biological mechanisms it would be most effective to include an adequate number of members from different age groups.
For instance, children are a very heterogeneous group in which age can be expected to be an important effect modificator. Large differences exist between a newborn infant and a 16 year old teenager in terms of body size, composition, surface area, physiology, pharmacokinetics and pharmacodynamics. Division of children into age groups in order to create more homogeneous groups is widely accepted in research: Medline separately indexes papers concerning neonates (0-1 month), infants (1 month to 2 years), children (2 years to 12 years), and adolescents (12 to 16 years). Within these groups, the development and physiology are considered to be comparable, which adds to the statistical power of studies and makes the research more efficient[16]. The other approach is to increase the generalisability of a large RCTs results by including both children and adolescents.
When including a more heterogeneous population into the trial to be able to identify effect modification the distribution of trial participants should not necessarily mirror the distribution of a subgroup in the general population. Subgroups may be too small to carry out adequately powered subgroup analyses. For that reason, it will frequently be necessary to over sample certain subgroups, and, in order to make sure that treatment allocation is balanced in the subgroup, to stratify the randomisation by subgroup status. Optimally, the trial is enlarged such that enough subjects from each subgroup are included to allow for an adequate and preferably statistically powerful assessment of the treatment effect. Subsequently, when 1) it is likely that age, gender and/or ethnicity modify the intervention effect and 2) the study aims to included enough subjects from each subgroup, then

predefined secondary objectives are to be formulated related to subgroup analyses and a predefined data analysis plan should be present[13].

Problems can be encountered when aiming at such over sampling of subgroups, i.e. when a relative large sample of the subgroup population is included. Over sampling is not always feasible due to limited availability of the specific population, e.g. in diseases that are uncommon, and due to limitations in available financial means. If over sampling is not possible within the realm of a RCT it is advisable to try and gather evidence on the comparability of the treatment effects within subgroups. To this aim, data obtained from post-marketing studies can help to answer some of the research questions in minority groups.

When there is no evidence for effect modification of gender, age or ethnicity?

Preliminary evidence from other sources may also support that there is no substantial difference between the intervention effects within subgroups, hence that the results are likely to be generalisable across various different patient populations. If there are differences these are probably small and very large subpopulations would be needed to detect them. In this case there is a no scientific argument to include different subgroups. To confirm the presumed generalisability of the results from RCTs, again data obtained from post-marketing studies might be used.

An example of potential effect modification

When studying heart failure or another cardiovascular event, gender is often taken into account as a potential effect modifier. Men have a higher incidence of heart failure, but the overall prevalence rate is similar in both sexes, since women survive longer after the onset of heart failure[17]. Hypertension, diastolic dysfunction, diabetes, obesity, and inactivity are more important risk factors for heart failure in women, whereas ischemic heart disease and systolic dysfunction are more important risk factors in men[18]. Regarding diagnostics, there are gender differences in the specificity and sensitivity of some non-invasive diagnostic tests for cardiovascular events. Given these biological differences in pathophysiology and clinical presentation of the condition there is every reason to take gender into account as a potential effect modifier in trials that study interventions of hart failure or other cardiovascular events. This does not necessarily mean that gender is an effect modifier in these trials - it needs to be considered as one. As a matter of fact, large studies on lifestyle modification and the use of statins[19,20] or nicorandril[21] found no evidence for modification by gender on all-cause mortality and heart diseases.

5.4 Stage 1 - Enrolment of a heterogeneous population

What is the current practice?

Trial guidelines and eligibility requirements are developed by the researchers and usually include criteria for indications and contra-indications like age, sex, type and stage of disease, previous treatment history, and other medical conditions. Inclusion and exclusion criteria -- medical or social standards used to determine whether a person may or may not be allowed to enter a clinical trial -- help to identify appropriate participants and help to exclude those who may be put at risk by participating in a trial.

In addition, in general it seems that inclusion/exclusion criteria are set to make the study population more homogeneous, such that the chance of side effects, co-morbidity and early drop-out becomes as small as possible[1-3]. Financial considerations can also be of importance. For instance, the costs of recruitment may be higher in certain groups. Many studies do not adequately report on trial design and underlying reasons for the choice of the particular inclusion and exclusion criteria[22].

Decades of clinical research have excluded women and minority groups[14,23-29]. Excluding women has long been defended as a way to protect women. However, representation of women in trials has increased between 1966 and 1990[14]. According to some, gender bias in clinical trials is now not more than a perceived bias[30]. Indeed in Review 3 it was found that most trials nowadays appear to include both men and women.

Exclusion or non-inclusion of ethnic minorities from randomised trials minorities still is a major problem. The percentage of minority patient enrolments appears to have decreased between 1996 and 2002[14].

As for age, most papers refer to the exclusion of person of older age. There is ample specific information on inclusion or exclusion of certain age groups within randomised trials in children.

The exclusion of specific groups has been frequently shown to occur through other, more indirect criteria. For instance, exclusion of people that do not speak Dutch or English in Dutch trials may result in the exclusion of certain immigrant groups. Sometimes subgroups are not included into a trial due to perceived problems (e.g. in compliance) by the physician. Informed consent may present problems in individuals of other ethnic origins due to differences in language and cultural background. Furthermore, the willingness to participate may also differ between groups.

What would we ideally want when aiming for more diversity?

Let us first assume that age, gender and/or ethnicity are treatment effect modifiers and that the aim is to enrol individuals from those different groups. Then, in the optimal scenario there would be no barriers at the enrolment stage.

Which factors complicate the implementation of this ambition?

1. The inclusion criteria are too narrow. As a results persons of different ethnic origin, gender, age, etc. are not included in the trial

2. Persons of different ethnic origin, gender, age, etc. are more likely to refuse to participate in the trial

5.4.1 The inclusion criteria are too narrow. As a result persons of different ethnic origin, gender, age etc are not included in the trial

Many reasons for exclusion of certain subgroups from a trial have been described in the literature. In general these can be grouped into four categories, i.e. 1) exclusion on the basis of medical reasons, 2) exclusion on basis of practical reasons, 3) exclusion due to perceived problems with responsible behaviour and 4) indirect exclusion.

Exclusion on basis of medical reasons is mainly important in elderly and children. Older people have an increased chance on a history of prior malignancy[31]. The chances on an advanced-stage disease, co-morbid condition or concomitant medication increase with age[27,31-33]. The elderly may also experience and increased toxicity and a slower recovery of elderly from curative procedures like surgery or radiotherapy[33].

In children there are different considerations that are taken into account in RCTs that may interfere with their inclusion, like discomfort, pain, treatment effects on growth of developing organs, and the size and volume of bio samples[34].

Exclusion on basis of practical reasons is important in children and in ethnic minorities. For instance, including children in chemotherapy trials might be problematic because many conditions are uncommon in children. Research priorities are being adult-focused because of greater burden of disease in adults[34].

Minorities are often excluded simply because they speak a different language. There may be no explicit exclusion criteria directly related to ethnic origin, but all consent forms are usually in local language only[6]. Furthermore, strategies for recruitment (e.g. translators and translation of info-sheets) are often unavailable.

Exclusion due to perceived problems with responsible behaviour plays a role in different age and minority groups. As for older people the perception that older patients are less likely to benefit from trial (on the part of the patients themselves, their families, or their physicians) can be a reason not to include them[15,31,33]. The assumption that treatment is too hard on older people and that the risk and discomfort will not be worth the benefits has a negative influence on the recruitment of the elderly. One recent study showed that physicians asked 51 percent of their eligible patients under 65 to participate in clinical trials but only 35 percent of those over 65[35]. Furthermore because of the perception that elderly are less compliant to trials physicians may therefore not attempt to recruit these individuals[36].

Similarly in adolescents and minority groups have been excluded due to perceived lack of compliance[37;38].

Indirect exclusion happens when there is, for instance, a link between gender, race and illicit drug use. Most HIV infected women tend to be also coloured and drug users [36]. Exclusion on basis of drug use in HIV trials

would therefore result in the exclusion of a large part of the coloured female population.

What can be done to resolve these problems?

As for exclusion due to medical reasons studies show that older people in otherwise good general health -- those having what physicians call "good performance status" -- tolerate standard chemotherapy regimens almost as well as younger persons. Data have also shown that the mortality from most operations, even such major surgeries as liver resections, is no different for the fit elderly than for younger patients[35]. Unless exclusion of elderly persons is explicitly needed, advanced age alone is not a sufficient reason to limit participation[33,33,39]. When barriers like co-morbidities and psychological and socioeconomic problems are addressed, the number of older people in trials is likely to increase (National Cancer Institute. Cancer trial barriers falling for people over 65. Available at: www.nci.nih.gov/clinicaltrials/developments/barriers-falling0101).

The involvement of children in clinical trials represents a dilemma. An important medical argument in favour of including children in a trial is that the present practice of off-label use and/or unlicensed drug use may harm children more than studying new interventions. In this context it is important that physicians and parents realise that the surroundings of clinical trials are much more controlled than they will ever be in case of unlicensed or off-label drug use[34,40].

As to the exclusion due to practical reasons, more effort should be put into testing medication in children. Caldwell states that licensing and funding regulatory bodies in individual countries must demand trial-based data in children for pharmaceutical and non-pharmaceutical interventions of clinical value to paediatric patients before the necessary approvals are given[34]. A systematic co-ordinated process needs to be established world-wide to ensure that the most important or essential drugs are prioritised for paediatric development[41].

As to the exclusion due to perceived problems with responsible behaviour more attention for diversity in policy guidelines may be helpful. The NIH-guidelines have made a difference, at least as to more awareness for the importance of diversity.

Reasons not to include certain subgroups are sometimes based on expected non-compliance. Indeed compliance in minority groups can be a problem. It appears that this is mainly due to a less effective communication between the physician and persons from ethnic minorities. Mutual understanding between physician and patient has been shown to be a strong predictor for patient compliance[42]. When communication and quality of giving information are improved, which is necessary anyhow in a RCT, the compliance is likely to improve as well.

Indirect exclusion of certain subgroups because these are linked to a specific exclusion criterion can not be easily prevented. At least awareness of this possible exclusion is of importance.

In summary, it can be stated that it is preferable to keep entry criteria simple and wide. Such a strategy can be a positive virtue by helping to

attain the large numbers of patients that are usually needed to reliably detect the sorts of moderate treatment benefits that are plausible[1].

5.4.2 Persons from certain groups may be more likely to refuse to participate in the trial

Recently a systematic literature search documented factors may influence a patient consent to participate in a clinical trial[43]. The most frequent reason mentioned by patients for choosing not to participate in a trial were fear of side effects, the phenomenon of randomisation, and, especially, the use of a placebo arm. Reasons to withdraw consent were interference with work, complicated record-keeping requirements, difficulty rescheduling appointments due to lack of flexibility on the part of the study personnel[44,45].

Overall, patients from non-western origin are generally more reluctant to participate in trials due to a lower level of trust of medical research and fear of losing their autonomy[46]. A fatalistic attitude towards diseases such as cancer may also play a role. It is well known that language problems and cultural differences complicate the informed consent procedure[14]. Still, strategies used for recruitment have been noticed to be inappropriate, i.e. lack of translation of information and informed consent forms and absence of cultural sensitive appropriate educational materials[6,46]. Furthermore, socioeconomic factors, such as low income causes lack of access to health care and as a result also to clinical trials[6,46].

Older people, especially older women are also less likely to enrol into randomised cancer trials[14]. Here the barriers include physician bias - the assumption that treatment is too hard on older people and that the risk and discomfort will not be worth the benefits[33]. In this case it is the physician that discourages the person to participate in the trial. Other barriers can be psychological or socioeconomic, as older patients may lack financial, social and logistic support for participation in clinical trials[31,33] (also see: http://www.cancer.gov/clinicaltrials/developments/barriers-falling0101). Furthermore complicated and lengthy administrative procedure and informed consent process may present an obstacle for elderly[27].

In children the threshold for gaining consent is higher and the procedure is even more complex, because parents have to make decisions on behalf of their child. Furthermore, in contrast to adults, it is very uncommon for healthy child-subjects to participate in RCTs. Parents are very reluctant to give consent and expose their child to health risks associated with the study[47]. Payment for a child's participation in a RCT is illegal in many countries, yet it is allowed in the United States. Non-reimbursement for additional (e.g. travel-) costs may create another barrier to participation[34].

What can be done to meet these problems?

The informed consent taker has a large impact on the decision of an individual to participate in a trial. The ability of the person that asks for informed consent to present adequate information on all aspects of the trial is an important factor. Furthermore a good relation with and trust in the consent taker will have a positive affect on the participation[43].

To improve the participation of recruited individuals into a trial several suggestions have been put forward.
In general, it should be anticipated that many factors that influence the decision of individuals to participate into a trial depend upon the communication and quality of giving information. The necessity to present extra information on randomization has clearly been shown in the literature. Explaining why randomization is necessary and what it means for the participant has been shown to improve the participation rate[48-50]. For instance in HIV-infected patients it was found that providing all patients with information about the meaning, role, and availability of AIDS clinical trials at the initiation of HIV primary care reduces differences in participation rates by gender, race, and history of drug use[51]. Furthermore, using different methods for the distribution of patient oriented information on the trial may result in more consideration. Human contact tends to be most successful in improving understanding[52]. Giving training in communication and giving information to the persons that recruit patients into trials and that ask informed consent is likely to result in an overall improved participation[53]. It has also been shown that flexibility in scheduling and rescheduling appointments helps to improve participation in general[53].
In immigrant groups it is advisable to involve a translator and to present written information in the patients' mother tongue. However, one may need to assess the limitations of translated material. In some communities many people who cannot speak or read English, can only speak and not read their own language either. Also be sensitive to cultural traditions and gender related issues[43,54,55]. A multifaceted strategy that targets the cultural, perceptive and cognitive characteristics of specific populations has been shown to be effective for increasing the enrolment of older African-American women in a cancer prevention trial[56].
To furthermore enlarge participation of elderly and minority groups input from residents of the study community can be effective[39,53,57]. Because most patients appreciate the opinion of their social environment it may also be useful to involve the community leaders in the informed consent procedure[39,53,57].
Both in the elderly and in minority groups, the use of a standard consent process and an extra meeting with a qualified person has been shown to be a reliable as well as practical approach to improve understanding. There is no evidence that this qualified person needs to be a physician. Rather, it might be a nurse, ideally from the same ethnic or social community as the patients to be recruited[52].
Because of the difficulties encountered in recruiting children to clinical trials, researchers need to take into account the risk-benefit analysis parents make when considering their child's participation in trials and, accordingly, modify risk factors and costs whenever possible to enhance participation–– e.g. by keeping blood tests and hospital visits to a minimum, and by reimbursing travel and other costs. Researchers must try to build better relationships with paediatricians and parents by communicating more clearly and openly[47]. They need to address key issues such as the parents' emotional response to their child's involvement in a

trial, and the physician's concerns about trial participation disrupting their doctor-patient relationship[34].
Information for patients on trial participation has been published by Consumers for Ethics in Research (CERES) and can be downloaded at http://www.ceres.org.uk.

5.5 Stage 2 - Allocation to interventions

Randomisation procedures for clinical trials are intended to create groups of patients with nothing but random differences in both known and unknown baseline characteristics that possibly influence prognosis and modify the intervention effect other than the intervention being considered. Statistically significant differences between groups in endpoint rates may then be attributed to the difference in treatment, rather than to other prognostic features.
Stratified randomisation is intended to prevent imbalance in known prognostic features and to assure sufficient participants of the predefined subgroups. Stratified randomisation is accomplished by identifying stratification factors before research is begun. Each factor may have two or more levels. The total number of strata needed is equal to the product of the number of levels of each factor. For the purpose of this review the main advantage of stratification lies in the facilitation of subgroup analyses.

What do we ideally want?
If age, gender and/or ethnicity are identified as expected modifiers of the response then balanced treatment allocation within these groups is advisable. By stratifying one can ensure good balance of these factors across intervention groups and sufficient power to study the effect of the intervention in each subgroup.
Stratification is especially important in small populations. The chance of an imbalance in covariates increases as the population size decreases. On the other hand, the number of strata and the levels that define strata should be limited. With fewer strata, there will be more patients in each stratum making subgroup analyses more feasible. These issues should be resolved at stage 0, when designing the trial. As already mentioned before, it is important to ensure that the power of the study is sufficient to draw valid conclusions for each subgroup.
The analysis of a study with a stratified inclusion scheme should take the stratification into account.

5.6 Stage 3 - Measurements

What is the current practice and does this interfere with attention for diversity?
In scientific research one has to standardize all measurements procedures and instruments. For clinical measurements this usually does usually not

present any problems. The validity of a new diagnostic measurement technique is nowadays often being tested in different gender, age and minority groups[58,59]. This is important as the accuracy of particularly non-invasive tests can differ among subgroups.
Data from self report is gathered with the help of questionnaires or interviews. In randomised trials self reporting is used as an adjunct to clinical measurements, for example the assessment of quality of life and patient satisfaction. It can also be performed to gather data on health related behaviour. Application of the same measurement instrument in a heterogeneous group is usually not valid[54,60]. Therefore, several subgroup specific instruments have been developed that are validated within those subgroups of different gender, age and ethnicity.

For this section it is assumed that people from different diversity groups (ethnic minorities, elderly and women) are successfully enrolled into the trial.

What would we ideally want when aiming for more diversity?
In the ideal situation data of similar quality is obtained from all participants.

Which factors complicate the implementation of this ambition?
What are the barriers to collect reliable information in different groups?
1. The instruments used for measurement (i.e. questionnaires) are not validated in different groups
2. The participant suffers poor literacy levels

5.6.1 The instruments used for measurement (i.e. questionnaires) are not validated in different groups

The formal validation procedure contains a lengthy process of testing and may therefore not be feasible. Firstly, it may be too difficult to perform validation studies in small minority populations, secondly it would postpone the initiation of a new randomised trial, and thirdly associated costs will bear too much on the available funds.
Often, when interviewing non-native speaking ethnic minority communities, the problem with translation is not the respondents' understanding of the words themselves, but their lack of understanding of concepts which are taken for granted. For instance, a question like “do you have family and heart problem" is not clearly understood. The fact whether the language one wishes to translate into is a written language (e.g. Chinese) or a verbal language (e.g. Caribbean patois or Moroccan barbers) is relevant. If it is a verbal language, clearly there is little merit in trying to produce a written translated questionnaire[54].

What can be done to resolve this problem?
Firstly, it will be necessary to invest in the development of certain basic instruments in minority groups. Secondly, rather than pursuing cross cultural equivalence an alternative is to search for issues that are

meaningful only within a particular culture and for issues that are of concern to everybody. The end result would be a set of questions that would share common items supplemented by culture specific information[54]. Alternatively one can focus on the similarity of concepts rather than on the equivalence of items. All people have a notion of physical wellness but the implication of that notion may be different among culture. For some people the ability to kneel may be essential, for others the ability to walk to a shop or to play soccer[54].

5.6.2 The participant suffers poor literacy levels

Some non-native speakers also suffer poor literacy levels in relation to their own language. Therefore, it should not be assumed that a non-native speaking respondent will be able to fill in a translated self completion questionnaire or read translated show cards.

What can be done to resolve this problem?
One solution may be to use transliteration, whereby the words are translated into mother tongue, but written out in Dutch or English. A last resort solution would be to use interviewers who are fluent in both Dutch (or English) and the respondents' mother tongue and allow them some flexibility to adapt translations in line with cultural differences. However, if this route is chosen, it is vital that interviewers are briefed face to face to ensure that all are clear on the meaning of the questions and the terminology used, in order to minimise differences in translation and interpretation of concepts. (http://www.mrs.org.uk/networking/ern/faqs.htm#20)

5.7 Stage 4 - Follow-up and drop-out

What is the current practice?
Is there a higher loss-to-follow up in certain groups? Literature on retention of subjects to the end of the study is scarce. Ethnicity has been associated with the likelihood of study completion in several HIV trials[61,62]. Overrepresentation of ethnic minority groups among withdrawn subjects was also found in a large asthma trial[45]. This and another study found that especially stress, heavy work, family time demand, and decreased social support were predictive of study retention problems (i.e. attrition)[45,63]. Another asthma trial including children and adults found similar attrition rates in both groups[64]. In a large cohort of rheumatoid arthritis in adults, younger age, lower levels of education, and non-Caucasian race predicted attrition[65]. Gender, level of disability and disease duration were not associated with attrition.
Adolescents with behavioural problems and alcohol/drug users were shown to be difficult to retain in trials[66]. It seems that young children, when participating in a trial, generally have good study retention[67,68].

What would we ideally want when aiming for more diversity?
In the ideal situation all participant would complete the study.

How can study retention be improved?
Investigators planning large trials involving the recruitment of diverse samples of the population should give attention to retention when designing the study. In general, when subjects feel involved and valued by the research team, they are more likely to continue participation to the end of the trial despite logistical difficulties[45]. Retention strategies that warrant careful consideration include limiting the number of required visits to the study site[53]. Furthermore maximum flexibility in the scheduling of appointment times without jeopardising the protocol integrity was proven to be effective. Nearly complete follow-up was achieved in an asthma study that included children and caretakers from different ethnic groups by having a flexible staff, computer tracking, and face-to-face recruitment[63]. Lastly, training staff in communication strategies for engaging research subjects, and providing clinical feedback to subjects about their condition whenever possible has been described to help improve the retention[45].

5.8 Stage 5 - Analysis

The usual step to study effect modification is to perform subgroup analyses. Subgroup analyses involve splitting all the participant data into subgroups, often to make comparisons between them. For instance, a subgroup analysis may be performed to study the effect of a new pain killer in men and women separately.
Subgroup analysis may have two well-known problems. The first problem occurs when multiple subgroup analyses are performed. In this situation it is likely that a difference from the overall result will be found in one or more comparisons even if none exists. This phenomenon is referred to as a type 1 error. The second problem occurs if the study is not adequately powered to detect differences in subgroups. Though a randomised controlled trial may be powered to find a difference between the treatment and control groups, it is usually not powered to find differences between smaller subgroups so that a real difference in a subgroup may not be detected. This phenomenon is referred to as a type 2 error. To resolve this a multiplication of the required populations is necessary depending on the number of strata.
The credibility of subgroup analyses is improved if the analyses are confined to the primary outcome measure and to a few predefined subgroups, on the basis of biologically plausible hypotheses. This might include factors used to stratify randomisation[10,11,13].
It is useful to distinguish between the notions of qualitative effect modification and quantitative effect modification[69]. Qualitative effect modification exists if the direction of effect is reversed, that is if an intervention is beneficial in one subgroup but is harmful in another. This type of effect modification is considered to be rare. Quantitative effect modification exists when the size of the effect varies but not the direction,

thus the intervention is beneficial to different degrees in different subgroups. For instance there is quantitative effect modification by gender when a new pain killer is twice as effective compared to placebo in reducing pain in men as in women.
Effect modification is closely related to statistical interaction in regression models. In relative risk regression models a significant interaction between a treatment and a second variable implies that the second variable is an effect modifier. Hence, statistical tests of interaction more directly assess the evidence for an effect modifier, and are to be preferred above the use of subgroup P values or confidence intervals.
In clinical trials subgroup analysis is often performed when effect modification is suspected. Statistical tests of interaction however, are insufficiently used[13].

How can information optimally be extracted in the presence of possible effect modification?
Let us assume that preliminary evidence supported the existence of subgroup differences in intervention effect, that subgroups in the RCT were prespecified, and that we managed to include a sufficient number of participants of each subgroup into the trial. In this case it would be most elegant to first perform statistical tests of interaction followed by subgroup analyses to identify and quantitate these effect differences.

Statistical tests of interaction
There are different risk models that can be used to study interaction. Multiplicative or additive risk models are applied most frequently. Under an additive model, the increase in rate or risk from a combination of factors equals the sum of the increases from each factor by itself. The multiplicative model assumes that the relative risk (risk ratio, rate ratio) for the factors operating together equals the product of their relative risks. The decision on how to assess statistical interaction depends upon what model we employ to arrive at an expected joint effect to compare with the observed joint effect (or equivalently, upon the scale of measurement, hence the term "effect measure modification").
The additive model has been put forth by Rothman as the "natural" scaling[44]. Risks are probabilities, and the probability that either of two independent and mutually exclusive events will take place is the sum of the probabilities for each. For instance, smoking causes myocardial infarction or oral contraceptive use causes myocardial infarction. Therefore if the risk (probability of disease) in people with both exposures exceeds the sum of the risks for each exposure separately, then some non-independence or interaction must exist between these two disease events. Rothman's proposition appears to have become the consensus in terms of evaluating impact on public health and/or individual risk.
Biological interaction may or may not manifest itself as a statistical interaction on either the additive or multiplicative scales[45]. Biological interaction refers to interdependencies in causal pathways. Such interdependencies – situations where one factor may potentiate or inhibit

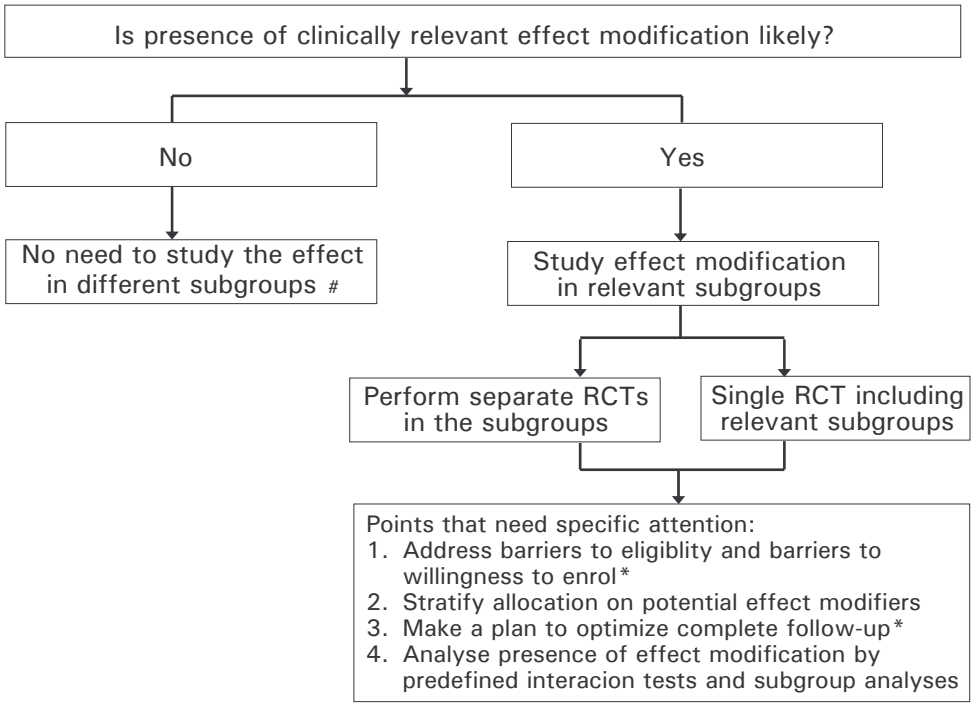
the effect of another – have implications for understanding of disease aetiology or effectiveness of treatment interventions.
Statistical interaction - or effect measure modification - and biological interaction clearly are two different phenomena.
An example is given for a heart attack prevention study that compared pravastatin with usual care in 10.355 older (>55 yrs), moderately hypercholesterolaemic, hypertensive participants with at least one additional coronary heart disease risk factor. In this study, conducted between 1994 and 2002, pravastatin did not reduce mortality or coronary heart disease when compared with usual care.
Of the enrolled population 49% was female, 55% was older than 65 years, 41% was white non Hispanic and 38% was black. Heterogeneity of effect in these and other subgroups was examined by testing for treatment-covariate interaction (with Cox proportional hazards regression). Age or sex-related differences for mortality and CHD event rates were not found, as also seen in previous statin trials. However, for pravastatin versus usual care the risk of CHD events was significantly lower in blacks than in non blacks for CHD events (RR 0.73, 95% CI 0.58 to 0.92).
In summary, the following steps in the statistical analyses are recommended[13]. The subgroups under study should be predefined on the basis of biologically plausible hypotheses. Investigators should recognise whether their trial is large enough to detect realistic subgroup effects. Unless the evidence is statistically convincing and clinically sensible, claiming a treatment difference in a subgroup when the overall treatment comparison is not significant is not justified. Statistical tests of interaction that assess whether a treatment effect differs between subgroups should be used rather than inspection of subgroup p values. Even when effect modification can not be proven, an indication of the size of the modification can be obtained and used in new trials.

5.9 Cost issues

In the real world financial issues often are another argument to limit the attention to diversity in clinical trials. Ensuring the representation of both males and females, different age groups and of minorities will have impact on costs. A cost-benefit analysis may be worthwhile in deciding on whether such representation is worthwhile[12].
Quantitative trials are usually powered to yield results that are statistically significant for the population as a whole. Where subgroup analysis is performed the results may suggest differences between the main population and the subgroup, but further research will be necessary to confirm these suggestions. Thus a decision must be made whether to power the original research sufficiently to do subgroup analysis in the first place[12]. The potential number of groups on whom subgroup analysis could be performed can be enormous. Obtaining the representation and performing the analysis will, however, cost additional research money. Simply translating an information sheet into another language will be costly and having an interpreter much more so. It can be questioned whether

these costs should be imposed unless there is a plausible and worthwhile benefit. As an extreme example, seeking Inuit representation in Dutch trials would seem to give little benefit for great cost. Conversely, seeking the representation of Moroccan or Antillean populations would seem worthwhile, but only where there is a plausible expectation of treatment differences. More difficult decisions would lie with groups such as those of Chinese or Japanese origin who are not greatly represented in The Netherlands. These issues also depend upon the incidence of the disease under study in a certain subgroup.
A recent RCT in the United Kingdom was designed to evaluate whether enhanced care for diabetes, tailored to the needs of the South Asian community with Type 2 diabetes, would improve risk factors for diabetic vascular complications such as hypertension compared to standard care[70]. Enhanced care involved Asian link workers and extra community diabetes specialist nurse sessions during one year. After adjusting for baseline measurement and age, only a small differential reduction in diastolic blood pressure in the enhanced care group was seen as compared to the standard care group. In this particular study, the benefit did not seem to weigh up to the additional cost of enhanced care.

Figure 5.1 Accounting for effect modification in RCTs



* Points 1 and 3 have to be addressed in every study, whether effect modification is suspected or not.

Confirmation of presumed generalisability remains necessary.

5.10 Discussion

An important question for randomised controlled trials is whether the results can be generalised to populations other than studied in the trial.

Currently, the generalisability of results obtained from randomised clinical trials is often questioned as the included patient population is in usually homogeneous. Relevant differences in patient characteristics between the study population and the total patient population that may benefit from the intervention at issue will only affect the generalisability of the results if there are significant differences between subgroups of patients. It is therefore important to investigate whether there is a clinical relevant modification of effect of the intervention by the presence of certain patient characteristics.

This review is part of a series of reviews on political, ethical and social aspects of diversity in clinical research. The current review focuses on the potential effect modifiers age, gender and ethnicity. However, age, gender and ethnicity are not the only variables that have to be taken into account in research on diversity. Co-morbidity and genetic markers, such as medication sensitivity, should be considered as well. Effect modification may be based on biological principles, for which ethnicity is only a proxy. Ethnicity is in essence a mix of genetic factors, illness, social factors, and behavioural and clinical characteristics. At this moment there are no workable alternatives for this proxy.

Evidence of effect modification by patient characteristics is scarce[4]. When aiming at addressing diversity, hypothesis generating evidence regarding potential differences among the studied intervention effect in different gender, age or ethnic groups should be gathered and studied when designing a trial[8]. This can largely be done by reviewing the literature as well as by studying the biological plausibility of potential differences[12,13].

In the presence of effect modifiers one may choose to design new trials that specifically address the treatment effect in relevant subgroups. For instance, age is a potential effect modifier in cancer trials. The majority of the information on cancer treatments however, has been obtained from data of patients below 70 years of age in spite of the fact that the older group is most likely to have the disease. Therefore, at present many new cancer trials are being performed in older men and women.

Another option is to enlarge the original trial with members from the other subgroups, such that subgroup analysis can be performed. When aiming at including a more heterogeneous population one may encounter problems with adequate trial enrolment, self report and compliance and retention in different age, gender and ethnic groups. It will be more complicated to deal with these problems in a more diverse population than in a homogeneous sample. Furthermore, subgroup analysis will be impossible or useless if in certain groups the participation was inadequate or if there was attrition bias.

The majority of these problems can be overcome, as has been shown is this review. Keeping entry criteria simple and wide is usually preferable. Such a strategy can help to include the large numbers of patients that are usually needed to reliably detect the sorts of moderate benefits that are plausible at a reasonable cost. Further success can be attained by improved communication, anticipating to cultural, perceptive, and cognitive characteristics of specific populations, and by involving the community. Thus, information can be generated that is relevant to many different

categories of patients with a particular condition[1]. The majority of these problems can be overcome. Financiering and governmental organisations can put diversity issues high on their agendas and invite investigators to argue whether and why subgroups should not be included into a new clinical trial.

At the analysis phase, the presence of differences in effect between groups, i.e. effect modification, is usually studied by subgroup analyses. Yet, findings from multiple subgroup analyses may be misleading since subgroup analyses are observational by nature and are not based on randomized comparisons. False negative and false positive significance tests increase in likelihood as more subgroup analyses are performed. If subgroup findings are presented as definitive conclusions there is clearly a risk of patients being denied an effective intervention or, on the other hand, treated with an ineffective or even harmful intervention. Shallow subgroup analyses can also generate misleading recommendations about directions for future research that, if followed, would waste scarce resources. Therefore, when differences in subgroups are found, research should focus on the biological plausibility of these differences. Misleading subgroups results can be largely prevented by pre-specifying which subgroups will be compared in the study protocol on basis of acknowledged potential effect modifiers.

Alternatively, statistical tests of interaction assess more directly the evidence for effect modification, as compared with subgroup P values or confidence intervals. In the most favourable situation power considerations for interaction tests will drive sample size planning in clinical trials. This, again, will have additional costs. On the other hand, if the study is underpowered for effect modification, the results may suggest differences between the main population and the subgroup, but further research will be necessary to confirm these. When high costs will be made anyhow by performing separate studies in subgroups, the cost-benefit can generally be expected to be in favour of one large adequately powered trial that includes those subgroups.

Still, lack of power is likely to remain an important consideration in interaction tests. There are a couple of potential ways around this problem, none of them completely satisfying. For instance, the interaction tests could be run using a higher type I error rate, leading to more power but also to more false positive results.

Including a diverse population in which subgroups are adequately represented will have costs. Easy low cost strategies that may improve the generalisability of the results include the performance of a multi centre trial and the use of less restrictive inclusion criteria. A multi centre study integrates the differences between medical centres and is more likely to include a diverse population[9]. Using wide and simple inclusion criteria will improve and facilitate the inclusion of a diverse population.

Because of the controlled setting, RCTs are best designed to test diversity hypotheses emerging from existing studies, which is one of the reasons why we limited this review to RCTs. However, methods on how to address heterogeneity as advised for RCTs are likely to be useful for non-randomised controlled studies as well. Also for uncontrolled studies the

suggestions to improve the inclusion and participation, the validity of measurements and completeness of follow-up will be valuable. Although largely similar analyses techniques one should be particularly aware of the problems with subgroup analysis in these studies. An important advantage of the controlled study setting is that the relative effectiveness of the intervention compared to the control can be studied in separate subgroups or in one large heterogeneous study population. When studying separate subgroups in uncontrolled studies one may end up with results that are not so easy to interpret due to the lack of a reference group.

5.11 Recommendation

Focussing on diversity is important for making optimal clinical decisions for individual patients and methodological implications can be addressed to in RCTs. We recommend that all relevant issues be considered in the RCT design phase. We further recommend that focussed methodological research in this field is stimulated by a separate research program from ZonMw.

References

1. Yusuf S, Held P, Teo KK and Toretsky ER. 1990. Selection of patients for randomized controlled trials: implications of wide or narrow eligibility criteria. Stat. Med. 9:73-83.

2. Strader DB. 2002. Understudied populations with hepatitis C. Hepatology 36:S226-S236.

3. Lewis JH, Kilgore ML, Goldman DP, Trimble EL, Kaplan R, Montello MJ, Housman MG, and Escarce JJ. 2003. Participation of patients 65 years of age or older in cancer clinical trials. J. Clin. Oncol. 21:1383-1389.

4. Zielhuis GA. 2000. [Roaming through methodology. XXIV. Problems in generalizing from randomized intervention studies with selected patient populations]. Ned. Tijdschr. Geneeskd. 144:1860-1864.

5. Godwin M, Ruhland L, Casson I, MacDonald S, Delva D, Birtwhistle R, Lam M, and Seguin R. 2003. Pragmatic controlled clinical trials in primary care: the struggle between external and internal validity. BMC. Med. Res. Methodol. 3:28.

6. Mason S, Hussain-Gambles M, Leese B, Atkin K, and Brown J. 2003. Representation of South Asian people in randomised clinical trials: analysis of trials' data. BMJ 326:1244-1245.

7. Bouter LM, and van Dongen MCJM. 1995. Epidemiologisch onderzoek: opzet en interpretatie.

8. Freedman LS, Simon R, Foulkes MA, Friedman L, Geller NL, Gordon DJ, and Mowery R. 1995. Inclusion of women and minorities in clinical trials and the NIH Revitalization Act of 1993--the perspective of NIH clinical trialists. Control Clin. Trials 16:277-285.

9. Petrucci N. 2003. Generalisability of randomised controlled trials: a matter of discussion. Intensive Care Med. 29:858.

10. Pocock SJ, Assmann SE, Enos LE, and Kasten LE. 2002. Subgroup analysis, covariate adjustment and baseline comparisons in clinical trial reporting: current practice and problems. Stat. Med. 21:2917-2930.

11. Pocock SJ, Collier TJ, Dandreo KJ, de Stavola BL, Goldman MB, Kalish LA, Kasten LE and McCormack VA. 2004. Issues in the reporting of epidemiological studies: a survey of recent practice. BMJ 329:883.

12. Allmark P. 2004. Should research samples reflect the diversity of the population? J. Med. Ethics 30:185-189.

13. Assmann SF, Pocock SJ, Enos LE and Kasten LE. 2000. Subgroup analysis and other (mis)uses of baseline data in clinical trials. Lancet 355:1064-1069.

14. Murthy VH, Krumholz HM and Gross CP. 2004. Participation in cancer clinical trials: race-, sex-, and age-based disparities. JAMA 291:2720-2726.

15. Hutchins LF, Unger JM, Crowley JJ, Coltman CA Jr. and Albain KS. 1999. Underrepresentation of patients 65 years of age or older in cancer-treatment trials. N. Engl. J. Med. 341:2061-2067.

16. Tan E, Cranswick NE, Rayner CR and Chapman CB. 2003. Dosing information for paediatric patients: Are they really 'therapeutic orphans'? Medical Journal of Australia. 179:18.

17. Stromberg A. and Martensson J. 2003. Gender differences in patients with heart failure. Eur. J. Cardiovasc. Nurs. 2:7-18.

18. Lund LH, and Mancini D. 2004. Heart failure in women. Med. Clin. North Am. 88:1321-45, xii.

19. Papademetriou V, Piller LB, Ford CE, Gordon D, Hartney TJ, Geraci TS, Reisin E, Sumner BM, Wong ND, Nwachuku C. et al 2003. Characteristics and lipid distribution of a large, high-risk, hypertensive population: the lipid-lowering component of the Antihypertensive and Lipid-Lowering Treatment to Prevent Heart Attack Trial (ALLHAT). J. Clin. Hypertens. (Greenwich.) 5:377-384.

20. Cheung BM, Lauder IJ, Lau CP and Kumana CR. 2004. Meta-analysis of large randomized controlled trials to evaluate the impact of statins on cardiovascular outcomes. Br. J. Clin. Pharmacol. 57:640-651.

21. IONA study group 2004. Impact of nicorandil in angina: subgroup analyses. Heart 90:1427-1430.

22. Sheikh A, Panesar SS, Lasserson T, and Netuveli G. 2004. Recruitment of ethnic minorities to asthma studies. Thorax 59:634.

23. Meinert CL and Gilpin AK. 2000. Enrollment of women in cardiovascular clinical trials. N. Engl. J. Med. 343:1972-1973.

24. Heiat A, Gross CP and Krumholz HM. 2002. Representation of the elderly, women, and minorities in heart failure clinical trials. Arch. Intern. Med. 162:1682-1688.

25. Flaskerud JH and Nyamathi AM. 2000. Attaining gender and ethnic diversity in health intervention research: cultural responsiveness versus resource provision. ANS Adv. Nurs. Sci. 22:1-15.

26. Moerman CJ and Van Mens-Verhulst J. 2004. Gender-sensitive epidemiological research: Suggestions for a gender-sensitive approach towards problem definition, data collection and analysis in epidemiological research. Psychology Health & Medicine . 9.

27. Bjarnason NH. 2004. Disparities in participation in cancer clinical trials. JAMA 292:922-923.

28. Hussain-Gambles M. 2003. Ethnic minority under-representation in clinical trials. Whose responsibility is it anyway? J. Health Organ Manag. 17:138-143.

29. Hussain-Gambles M, Atkin K, and Leese B. 2004. Why ethnic minority groups are under-represented in clinical trials: a review of the literature. Health Soc. Care Community 12:382-388.

30. Meinert CL, Gilpin AK, Unalp A and Dawson C. 2000. Gender representation in trials. Control Clin. Trials 21:462-475.

31. Trimble EL, Carter CL, Cain D, Freidlin B, Ungerleider RS and Friedman MA. 1994. Representation of older patients in cancer treatment trials. Cancer 74:2208-2214.

32. Lee PY, Alexander KP, Hammill BG, Pasquali SK and Peterson ED. 2001. Representation of elderly persons and women in published randomized trials of acute coronary syndromes. JAMA 286:708-713.

33. Yee KW, Pater JL, Pho L, Zee B and Siu LL. 2003. Enrollment of older patients in cancer treatment trials in Canada: why is age a barrier? J. Clin. Oncol. 21:1618-1623.

34. Caldwell PH, Murphy SB, Butow PN and Craig JC. 2004. Clinical trials in children. Lancet 364:803-811.

35. Kemeny MM. 2004. Surgery in older patients. Semin. Oncol. 31:175-184.

36. Diaz T. 1995. Women, hiv and condom use. Fam. Plann. Perspect. 27:226-227.

37. Bleyer WA, Tejeda H, Murphy SB, Robison LL, Ross JA, Pollock BH, Severson RK, Brawley OW, Smith MA and Ungerleider RS. 1997. National cancer clinical trials: children have equal access; adolescents do not. J. Adolesc. Health 21:366-373.

38. Bleyer A. 2002. Older adolescents with cancer in North America deficits in outcome and research. Pediatr. Clin. North Am. 49:1027-1042.

39. Unson CG, Ohannessian C, Kenyon L, Fenster J, Reisine S and Prestwood K. 2004. Strategies for enrolling diverse older women in an osteoporosis trial. J. Aging Health 16:669-687.

40. Steinbrook R. 2002. Testing medications in children. N. Engl. J. Med. 347:1462-1470.

41. Kauffman RE. 1999. Essential drugs for infants and children: North American perspective. Pediatrics 104:603-605.

42. Harmsen H, Meeuwesen L, van Wieringen J, Bernsen R and Bruijnzeels M. 2003. When cultures meet in general practice: intercultural differences between GPs and parents of child patients. Patient. Educ. Couns. 51:99-106.

43. Nievaard MA, de Vos R, de Haes JC and Levi M. 2004. [Reasons why patients do or do not participate in clinical trials; a systemic review of the literature]. Ned. Tijdschr. Geneeskd. 148:186-190.

44. Sateren WB, Trimble EL, Abrams J, Brawley O, Breen N, Ford L, McCabe M, Kaplan R, Smith M, Ungerleider R. et al 2002. How sociodemographics, presence of oncology specialists, and hospital cancer programs affect accrual to cancer treatment trials. J. Clin. Oncol. 20:2109-2117.

45. Janson SL, Alioto ME and Boushey HA. 2001. Attrition and retention of ethnically diverse subjects in a multicenter randomized controlled research trial. Control Clin. Trials 22:236S-243S.

46. Shavers VL, Lynch CF and Burmeister LF. 2002. Racial differences in factors that influence the willingness to participate in medical research studies. Ann. Epidemiol. 12:248-256.

47. Hoppu K. 1999. Patient recruitment--European perspective. Pediatrics 104:623-626.

48. Fallowfield LJ, Jenkins V, Brennan C, Sawtell M, Moynihan C and Souhami RL. 1998. Attitudes of patients to randomised clinical trials of cancer therapy. Eur. J. Cancer 34:1554-1559.

49. Ellis PM, Butow PN and Tattersall MH. 2002. Informing breast cancer patients about clinical trials: a randomized clinical trial of an educational booklet. Ann. Oncol. 13:1414-1423.

50. Wright JR, Crooks D, Ellis PM, Mings D and Whelan TJ. 2002. Factors that influence the recruitment of patients to Phase III studies in oncology: the perspective of the clinical research associate. Cancer 95:1584-1591.

51. Freedberg KA, Sullivan L, Georgakis A, Savetsky J, Stone V and Samet JH. 2001. Improving participation in HIV clinical trials: impact of a brief intervention. HIV. Clin. Trials 2:205-212.

52. Flory J and Emanuel E. 2004. Interventions to improve research participants' understanding in informed consent for research: a systematic review. JAMA 292:1593-1601.

53. Dilworth-Anderson P and Williams SW. 2004. Recruitment and retention strategies for longitudinal African American caregiving research: the Family Caregiving Project. J. Aging Health 16:137S-156S.

54. Hunt SM and Bhopal R. 2004. Self report in clinical and epidemiological studies with non-English speakers: the challenge of language and culture. J. Epidemiol. Community Health 58:618-622.

55. Newman LA, Hurd T, Leitch M, Kuerer HM, Diehl K, Lucci A, Giuliano A, Hunt KK, Putnam W and Wells SA. 2004. A report on accrual rates for elderly and minority-ethnicity cancer patients to clinical trials of the american college of surgeons oncology group. J. Am. Coll. Surg. 199:644-651.

56. Zhu K, Hunter S, Bernard LJ, Payne-Wilks K, Roland CL and Levine RS. 2000. Recruiting elderly African-American women in cancer prevention and control studies: a multifaceted approach and its effectiveness. J. Natl. Med. Assoc. 92:169-175.

57. Moreno-John G, Gachie A, Fleming CM, Napoles-Springer A, Mutran E, Manson SM and Perez-Stable EJ 2004. Ethnic minority older adults participating in clinical research: developing trust. J. Aging Health 16:93S-123S.

58. Dinis-Ribeiro M, da Costa-Pereira A, Lopes C, Barbosa J, Guilherme M, Moreira-Dias L, Lomba-Viana H, Silva R, Abreu N, and Lomba-Viana R. 2004. Validity of serum pepsinogen I/II ratio for the diagnosis of gastric epithelial dysplasia and intestinal metaplasia during the follow-up of patients at risk for intestinal-type gastric adenocarcinoma. Neoplasia. 6:449-456.

59. Yang KN, Schaller JL and Parker R. 2000. Factor structures of Taiwanese teachers' ratings of ADHD: a comparison with U.S. studies. J. Learn. Disabil. 33:72-82.

60. Hunt SM. 2004. International development of the Quality of Life in Depression Scale (Qlds). J. Affect. Disord. 80:95-96.

61. Wagner GJ, Maguen S and Rabkin JG. 1998. Ethnic differences in response to fluoxetine in a controlled trial with depressed HIV-positive patients. Psychiatr. Serv. 49:239-240.

62. Gifford AL, Cunningham WE, Heslin KC, Andersen RM, Nakazono T, Lieu DK, Shapiro MF and Bozzette SA. 2002. Participation in research and access to experimental treatments by HIV-infected patients. N. Engl. J. Med. 346:1373-1382.

63. Senturia YD, McNiff MK, Baker D, Gergen P, Mitchell H, Joseph C and Wedner HJ. 1998. Successful techniques for retention of study participants in an inner-city population. Control Clin. Trials 19:544-554.

64. Bender BG, Ikle DN, DuHamel T and Tinkelman D. 1997. Retention of asthmatic patients in a longitudinal clinical trial. J. Allergy Clin. Immunol. 99:197-203.

65. Krishnan E, Murtagh K, Bruce B, Cline D, Singh G and Fries JF. 2004. Attrition bias in rheumatoid arthritis databanks: a case study of 6346 patients in 11 databanks and 65,649 administrations of the Health Assessment Questionnaire. J. Rheumatol. 31:1320-1326.

66. Meyers K, Webb A, Frantz J and Randall M. 2003. What does it take to retain substance-abusing adolescents in research protocols? Delineation of effort required, strategies undertaken, costs incurred, and 6-month post-treatment differences by retention difficulty. Drug Alcohol Depend. 69:73-85.

67. Sexton K. 2004. Comparison of recruitment, retention, and compliance results for three children's exposure monitoring studies. J. Expo. Anal. Environ. Epidemiol.

68. Bisgaard H. 2004. The Copenhagen Prospective Study on Asthma in Childhood (COPSAC): design, rationale, and baseline data from a longitudinal birth cohort study. Ann. Allergy Asthma Immunol. 93:381-389.

69. Yusuf S, Wittes J, Probstfield J and Tyroler HA. 1991. Analysis and interpretation of treatment effects in subgroups of patients in randomized clinical trials. JAMA 266:93-98.

70. O'hare JP, Raymond NT, Mughal S, Dodd L, Hanif W, Ahmad Y, Mishra K, Jones A, Kumar S, Szczepura A. et al 2004. Evaluation of delivery of enhanced diabetes care to patients of South Asian ethnicity: the United Kingdom Asian Diabetes Study (UKADS). Diabet. Med. 21:1357-1365.

Review 6

Novel policy strategies to diversity in clinical research

Nicolien Wieringa[4], Andrew Herxheimer, Eline van Haastrecht,
Anita Hardon, Amâde M'charek, Niek Klazinga

[4] I thank John Grin for sharing his thoughts on research policy with me.

Contents

6

Novel policy strategies to diversity in clinical research

Summary

This review builds on the outcomes of the preceding reviews by identifying three main barriers to diversity in clinical research which have implications for policy development that are particularly relevant to the aims of this project. The first barrier concerns the ways in which diversity is being conceptualised in clinical research; the second is that research agendas do not adequately address diversity issues, and the third is that the study of diversity in clinical research is now largely limited to RCTs. Based on these barriers three policy strategies for broadening the conduct of research are considered. An analysis of policy strategies used by other research funding agencies, literature and our own experiences with the development and evaluation of research policies was used to identify and discuss strategies for ZonMw.

The first strategy involves new methods for the conceptualisation of diversity. Linking the patient categories age, sex/gender and ethnicity to underlying mechanisms of differences in disease or health outcomes is a way to increase the relevance of research questions in terms of medical practice and the people involved. Another strategy is to search for dimensions of diversity beyond age, sex/gender and ethnicity. Involvement of patients/consumers and practitioners into the research process is useful to broaden the perspective on diversity, what differences matter to whom and when.

The second strategy is to incorporate diversity into the research agenda. New methods are needed to develop alternative research hypotheses that focus on diversity, and to identify under-researched diversity issues. Systematic consultation of various stakeholders is therefore relevant.

Developing expertise on diversity is essential among researchers as well as in the reviewing processes.

In the third strategy a key element for ZonMw is to stimulate a diversity of research methods. In addition to using randomised controlled trials, a range of qualitative methods can be used to explore diversity issues to broaden the biomedical perspective and generate new research hypotheses. Also, participatory strategies involving patients/consumers and practitioners at all stages of research planning, decision-making and conduct can help to increase the relevance of outcomes for patients and practitioners. Finally, reviewing ZonMw's research procedures, for example concerning patient selection criteria, can identify barriers for change to increase participation of specific patient groups.

6.1 Introduction

The aim of this review is to analyse and discuss policy strategies to focus on diversity in clinical research. The preceding reviews concern the conceptualisation of diversity and its practical consequences for research on diversity. In this review, we aim to identify barriers to diversity in clinical research, discuss strategies for overcoming these barriers and to consider their feasibility and implementation.

The Netherlands Organisation for Health Research and Development (ZonMw) initiated this project, and the policy strategies are primarily intended for it. ZonMw functions in the wider context of health care, including governmental organisations, regulatory agencies, other research funding agencies, and organisations of researchers, professionals and patient/consumer. Its main themes for agenda setting are developed at a policy level in co-operation with the government and its advisory agencies. ZonMw translates these public health needs into research programmes and then decides on the funding of project proposals. ZonMw aims to further develop policy strategies to focus on diversity issues.

This review builds on the main outcomes of the preceding reviews and the discussions with experts during the project's international meeting in November 2004. Based on these findings, we identify issues for developing policy strategies, relevant to ZonMw. These concern the basic processes of ZonMw and involve mainstream developments in publicly funded clinical research. From the results presented in this review, we infer and discuss specific strategies for ZonMw and their feasibility.

6.2 Building on the lessons from our five reviews

The analysis of biomedical literature, as presented in review 1, revealed that in many diseases the aetiology, prognosis, disease perception and/or effects of interventions are modified by age, sex/gender and/or ethnicity

and may affect health outcomes. To provide optimal care for every individual, it is therefore necessary to take these factors into account. Present clinical research does not specifically focus on diversity issues. A number of research areas and methods were highlighted that would allow for better consideration of diversity. They included differentiation of treatment guidelines, research into specific populations and under-researched areas, possible effect modification, and other clinically relevant patient categorisations uncovering diversity issues.

Diversity issues reach beyond the classifications of sex/gender, age and ethnicity, as was shown in review 2. Many different kinds of diversity may be relevant to patients and their treatments, and their meaning and relevance may change over time and in various contexts. It is not only important to articulate the 'local knowledge' of patients and clinicians in order to learn more about diversity. This knowledge can also help to diversify clinical research. Instead of focusing on a preferred cluster of differences, clinical research should keep a sensitive eye for other differences that matter for patients and practitioners. As was shown in review 2, clinical research is an important means of identifying populations at risk and this knowledge is on its own terms crucial for general practitioners in order to diagnose patients. This makes the diversification of clinical research even more important since there is no stable category of 'population at risk'. Moreover, in determining relevant differences, clinical research should include the question: relevant to whom and when? Therefore, research into the dynamics of clinical research is important to identify which diversity issues are put on political, social and scientific agenda's, and which stakeholders are involved.

The analysis of assumptions underlying randomised controlled trials (RCTs) in review 3 shows that the RCT is not a neutral research design. The historical development of RCTs illustrates the paradigm shift from seeing the human body as "all unique" to "all similar". The design and conduct of RCTs needs to include the examination of diversity issues. However, an analysis of recent trials in a number of diseases in which diversity issues are known to be relevant revealed that very few subgroup analyses were performed and reported. The social context of clinical research, for instance drug research, should support the study and publication of diversity issues. The implications of under-representation of specific populations in clinical research have fuelled the debate on the limited generalisability of research findings.

Historically, ethical debates in clinical research have focused on patient protection, expressed mainly as doing no harm and autonomy in decision making. Review 4 showed that it is also beginning to be acknowledged that participation and representation in clinical research are becoming increasingly important in the ethics of research. An important issue is therefore to consider barriers to participation, how these may differ between patient groups, and how access to research can be achieved. Another ethical issue concerns personal implications of knowledge about

diversity, for example about genetic predisposition to diseases. Research into the individual, familial and societal implications of genetic screening as a preliminary for inclusion in clinical trials was found relevant.

Review 5, which deals with methodological implications of focussing on diversity in research, showed that RCTs form a good instrument to study effect modification. This term denotes differences in the outcomes of interventions between populations, caused by biological and/or socio-cultural differences. To study effect modification, current practices in the application of the method need to be considered at all stages, i.e. trial planning, enrolment of a heterogeneous population, allocation of interventions, measurements, follow-up and drop-out, analysis and cost issues. Since present research practices generally aim at including homogeneous populations, many barriers were identified that need to be addressed. Where relevant because of underlying differences in the aetiology of diseases, or differing mechanisms of action of diagnostic, preventive or treatment interventions, trials should study possible effect modification.

These research findings identify three main barriers to diversity in clinical research which have implications for policy development that are particularly relevant to the aims of this project. The first barrier concerns the ways in which diversity is being conceptualised in clinical research. To classify patients by sex, age and ethnicity may be relevant because of underlying biomedical and/or socio-cultural causes. What is considered important may differ between various stakeholders, in different contexts and times. The conceptualisation of diversity is therefore necessary for developing research strategies. The second barrier is that research agendas do not adequately address diversity issues. Thus, strategies for the process of research planning need to be developed. Which strategies help clinical research agendas to focus on diversity issues? The third barrier is that the study of diversity in clinical research is now largely limited to RCTs. In addition to studying effect modification in RCTs, other methods are also needed to include the perspectives of other stakeholders on diversity and their translation into research. For this purpose, a range of methods can be used. Therefore, the third subject for policy development involves strategies for broadening the conduct of research on diversity issues.
The next sections explore these subjects further in the context of policy strategies used by other research funding agencies, or reported in the literature, and our own experiences with the development and evaluation of research policies. The research funding agencies have been chosen for their innovative approaches and their influence on international clinical research policies. Included are the Framework Programme for biomedical research of the European Commission (FP-EU), the National Institutes of Health in the US (NIH), the UK Medical Research Council (MRC), and the Norwegian Research Council (NRC). Information was obtained from their websites and from that of ZonMw.

6.3 Developing new strategies to focus on diversity in clinical research

6.3.1 Conceptualisation of diversity

Differences in health outcomes can be based on a large variety of biomedical and socio-cultural factors that may also interact with each other. In the development of medical science, the strong focus on scientific rationality and the biomedical perspective have limited the capacity of medical professionals to recognise and deal with those issues that cannot be understood in these terms, but which may be important for the ways in which patients experience and respond to health problems (Grin 2004). Review 2 revealed a number of these 'blind spots' in the daily practice of diabetes care. Also, this review showed that the differences that are relevant to clinical practice reach beyond the categories of sex/gender, age and ethnicity as differential categories. These widely used categories are in fact important barriers to focus on diversity in research. In this section we consider strategies to reach beyond this concept. These strategies link the patient categories age, sex/gender and ethnicity to the mechanisms that underlie differences in favourable outcomes, and identify other dimensions of diversity that matter to patients and practitioners. Ultimately, these strategies may offer new ways to develop research hypotheses concerning diversity in clinical research.

Linking categories to mechanisms of diversity

Review 1 found many biomedical and socio-cultural factors and mechanisms underlying differences in diseases and outcomes of interventions, in relation to the patient characteristics age, sex/gender and ethnicity. In addition, several reviews showed that using these patient characteristics has led to the exclusion of populations from research. A strategy to overcome this exclusionary barrier is to develop research questions and hypotheses by addressing the underlying biomedical and/or socio-cultural differential mechanisms. In this way, the patient categories sex/gender, ethnicity and age can be linked to the biomedical and socio-cultural meaning of diversity issues. Using this strategy for the development of research hypotheses may stimulate research that is highly relevant to medical practice. Also, this strategy offers the possibility to move away from the social implications of the wording of studying minorities. Clinical research should address diversity because differences are a social reality for every individual; categorisations that bring in these biomedical and/or socio-cultural differences should lead away from the association with under-privileged populations, because that may further emphasise inequalities.

The following examples from review 1 illustrate how this strategy can be applied. In the treatment of hypertension, relevant questions concern the comparability of treatment effects between younger and older patients, because elderly patients have been less studied. The components of

hypertension and their relevance to developing cardiovascular diseases were found to change with increasing age. Studying age effects on antihypertensive treatment therefore needs to be operationalised by using different variables, i.e. diastolic blood pressure in younger persons and systolic blood pressure in elderly. As a result of these age-related mechanisms in hypertension, it can be hypothesised that the patterns of side effects also differ between age groups.

In a second example cultural values or norms for (ab)normal child behaviour were found to influence the diagnosis of attention-deficit hyperactivity disorder (ADHD). This can lead to differences in the incidence and prevalence of the disorder between boys and girls, among various ethnic (or cultural) groups in a country and between countries. To address this variation in diagnostic criteria, it is necessary to understand how cultural differences affect the valuation or interpretation of child behaviour. This knowledge can be used to differentiate between ADHD diagnostic criteria or symptom thresholds for both sexes and for different cultural groups.

The last example concerns effective dietary management to prevent osteoporosis. In Western countries, preventive strategies generally focus on increased intake of dairy products rich in calcium. However, this strategy may not apply to people originating from Asia and Africa many of whom have lactose intolerance as a result of biological differences. In addition, cultural variation in dietary habits is likely to affect calcium consumption and strategies to increase this. Thus, to address ethnic differences in dietary strategies for osteoporosis prevention, biological and cultural variation must be taken into account.
In these examples, research questions are operationalised using biomedical and/or socio-cultural variables to give meaning to age, sex/gender and ethnic differences. This approach provides a new perspective on studying diversity, because the issues are made relevant in terms of medical practice and the people involved.

Searching for other dimensions of diversity

It is increasingly recognised that input from patients or users of health services, and clinical practitioners, is essential to broaden the perspective of research, especially in chronic diseases. The differing perspectives of patients and professionals may help to identify research questions about diversity issues that matter in patients' lives and in medical practice. The process of mutual learning between patients, practitioners, and specialists/researchers will produce knowledge that combines experiences in daily life and practice with academic knowledge. This may increase our understanding of differences between populations and their origins. In this section, we discuss strategies for involving patients/consumers and practitioners in the research process. Although consumer involvement is not specifically aimed at increasing the focus on diversity in research, it can be developed and used for this purpose. ZonMw also recognises

patient and consumer involvement as a key strategy to improve health care, because it may help to identify health needs and differences between patient groups (ZonMw 2004).

Diversity issues relevant to patients' lives and health care practices become apparent when introducing differing perspectives into the development of research questions and hypotheses. It is increasingly recognised that the way patients talk about their experiences is to be considered as knowledge in its own right of their diseases (Caron-Flinterman 2005). The experiential knowledge and skills of patients complement those of researchers, and they are likely to have good ideas about how research questions might be asked differently (Abma 2004; Rabeharisoa 2003; Trivedi 2002). Strategies for collaborating with patients/consumers and practitioners in the conceptualisation of diversity are important for the aims of this project. In general, such strategies involve the conduct of workshops with various stakeholders and a range of participants, qualitative research methods (Philipsen 2004), and methods of technology assessment. Addressing diversity from the patient/consumer perspective implies that new topics, such as the patient's every day life and her use of technologies and medication are introduced into the processes of research development. Dependent on the nature of the research, for example whether it concerns cure or care, appropriate strategies and methodologies can be applied.

This approach raises the question who, or which stakeholders, should be involved in the conceptualization of diversity. Who can be considered to represent a specific patient population, community, or profession? How does representation relate to the needs, perspectives and values of the patients concerned? The UK-group 'Consumers in NHS Research', set up by the Department of Health's Director of Research and Development in 1996, has developed the position that involvement of patients or consumers should not aim to represent users, but to seek different perspectives. Involving a range of people introduces a range of perspectives (Hanley 2004; Williamson 1999).

Experiences with interactive technology assessment (ITA) demonstrated the value of this analytic tool to systematically study different perspectives between various stakeholders (Reuzel 2002). As a method, ITA can be used for example to better understand the nature of controversies in health care, specifically when side effects (in the broad sense of the word) are partly rooted in system features that are normally taken for granted (Grin 2004). An example is the development of cochlear implants for deaf people (Reuzel 2002). The debate on cochlear implants started when, to the surprise of many involved in the development of this technology, deaf peoples' organisations were not overly enthusiastic. Many felt that the technology would deny and hamper the existence of a deaf culture, and they felt threatened and offended by the presumed assumption that deafness was a handicap to be eradicated. Thus, cochlear implants may be characterised as a promising technology that is perceived to have side

effects not anticipated by health professionals: prevailing standards of merit disregarded the experiences and identities of deaf people (Grin 2004).

In general not enough is know about methods and strategies for involving patient/consumer perspectives into the conceptions of diversity, which aspects matter to their well being, and how these might differ from those of researchers and health practitioners. Review 2 showed that in clinical practice a broad range of knowledge from patients and practitioners is available and applied on a daily basis. Ethnographic and participative research in such practice may help patients and practitioners to articulate the diversities that matter and facilitate their implementation in other domains of research. In this way, patients' perspectives and their experiential knowledge can be incorporated in clinical studies. Moreover, patients' perspectives are not viewed in isolation but in the interaction with practitioners. This approach may also sensitize researchers to differences in medical practices, since the diversities that matter in an out-patient clinic may be different from the ones that matter at the general practitioner. Hence, increasing our knowledge and experience with such methods is needed when aiming at diversity in clinical research.

Strategies

In sum, the following strategies to further develop the conceptualisation of diversity were discussed:

- Designing research questions and methodology that address the underlying biomedical and/or socio-cultural mechanisms of difference between men and women, patients of different age groups, and various ethnicities may increase and deepen the meaning of 'diversity' in clinical research.
- Using patient/consumer and practitioner involvement to broaden the perspective on diversity in clinical research, in particular for the conceptualisation of diversity (what differences matter to whom and when) and how to incorporate these perspectives into research.
- Involvement of patients/consumers in clinical research should not aim to represent users, but to seek different perspectives from a range of people, and their interaction in practices.
- Methodological approaches to seeking different perspectives in clinical research include diverse methodologies: qualitative methods and techniques of health technology assessment, workshops.

6.3.2 Incorporating diversity into the research agenda

Biomedical and socio-cultural differences between people are likely to influence the development and prognosis of many diseases, and may lead to effect modification of interventions for diagnosis, prevention or treatment. The rapid pace of scientific understanding has led ZonMw to defining 'diversity' as one of the organisation's main focus points. ZonMw

has developed and conducted a specific research programme on diversity, but the question now is how the agency can expand its policy to reach other research programmes as well (ZonMw 2004). In general, two approaches can be identified: challenging the dominant research paradigms, and defining under-researched areas. The following sections discuss these strategies.

Defining alternative research paradigms

From an epistemological point of view, diversity issues do not reflect the dominant paradigm in biomedical scientific research. Focusing on diversity in clinical research is therefore reaching beyond the dominant hypotheses. Which biomedical paradigms are dominant in specific research areas, and which alternative paradigms can be found? How well tested is the extrapolation of clinical data from one population to another, and what arguments can be found against extrapolation, thus providing grounds for research questions about diversity? Such questions asked systematically will challenge the mainstream translational research of basic knowledge to clinical studies, and consequently clinical practice. This approach may also identify areas relevant to specific populations that are relatively neglected. Strategies for identifying alternative paradigms to introduce more diversity in clinical research were developed by various research institutes, in particular by the US National Institutes of Health (NIH) and the European Commission's Framework Programmes (FP-EU). These strategies involve collaborative efforts to define research questions concerning diversity issues, and their inclusion in research programmes. Our research revealed that the NIH and the FP-EU differ in their organisational structures and in their approaches to redirect the research agenda. We consider a number of practical experiences and implications, discuss the effects of these strategies, and their feasibility for ZonMw.

The NIH research programmes are directed and co-ordinated by institutes and centres. In addition to this organisational structure, offices can be established to increase research on specific subjects. This was done to stimulate research on women's health issues when the NIH Office for Research on Women's Health (ORWH) was established in 1990. Research offices have a special position, because they cannot fund research projects directly, unlike NIH institutes and centres. Instead, they have to develop research projects through partnerships with institutes and centres (Pinn 2004). Research on women's health issues can thus be implemented in NIH funding, but only in areas of shared interests.
An example of shared research interests in which mainstream paradigms were challenged, is the Women's Health Initiative (WHI) that the ORWH developed in co-operation with the National Heart, Lung and Blood Institute. This study critically questioned the existing evidence on the preventive effects of hormonal replacement therapy on cardiovascular diseases. A 15-year research programme was developed, involving 16,600 postmenopausal women. At the time of its conception, the WHI was called by some the "mother of all clinical trials", because it was the biggest trial

world-wide at that time. Many saw the study as being too big, too expensive, too ambitious, too interdisciplinary, and as testing questions that were already answered. For example, one objection was that hormonal replacement therapy was already known to be good for the heart and questioned whether a placebo-controlled trial was possible, necessary, or even ethical (Healy 2003). The clinical study actually learned that oestrogen-alone hormone therapy has no effect on coronary heart disease, but increases the risk of stroke (NIH News 2004). In a reflection on the research process, it was acknowledged that the "scientific self-assurance looks a little silly now" (Healy 2003). The unexpected results of this trial affected medical practice because guidelines were changed, the pharmaceutical promotion of drugs deceased, and women are being prescribed fewer hormones (Brass 2004; Majumdar 2004).

The same policy was also applied to address health issues of minorities. In 1990, the Office of Research on Minority Health (ORMH) was established. It was less successful than the ORWH in developing research projects. At the 10-year anniversary of the ORMH it was noted that NIH had not directed as much of its power to studying and alleviating health disparities as it should (Helmuth 2000).

Implementation of the same strategy, i.e. establishment of research offices, to focus on diversity issues concerning women's and minorities' health showed different results. Why was the ORWH more successful than the ORMH in challenging the NIH research agenda? One possibility is that the underlying paradigms driving the implementation of research on women's issues differ less from the mainstream research paradigms, than do paradigms concerning health issues of minority groups. A wide range of scientific, social and organisational issues contribute to health disparities, therefore requiring different research approaches to be developed and implemented. The policy evaluation of ORMH at its 10-year anniversary led to an organisational change. In 2000, the US Congress approved to establish the National Centre on Minority Health and Health Disparities (NCMHD) to give it both greater visibility and the power to fund its own projects. From it aims and mission, the NCMHD seeks collaboration with other institutes and centres, thus continuing the path of the ORMH, but it is hoped that the new centre will get real power, because "the office has faced obstacles and ghettoization" (Helmuth 2000).

In contrast to women's and minorities' health issues, which came onto the NIH research agenda in the 1990s, the situation was different for clinical research in children. Here, a special NIH Institute was established already in 1962, the Institute of Child Health and Human Development (NICHD). However, the Institute's research agenda does not appear to cover age-related diversity issues. In 1996, a Congressional Committee "strongly encourages the NIH to strengthen its portfolio of basic, behavioural and clinical research conducted and supported by all of its relevant institutes, to establish priorities for paediatric research and to ensure the adequacy of translational research from the laboratory to the clinical setting (...) The

Committee expects the NIH to develop performance indicators to measure specific progress on the above, demonstrated by the development of new, or strengthening of existing programs" (NIH Guide 1998). Therefore, age-related health issues of children can also be understood as alternative research hypotheses, which were not adequately addressed within the existing structures. Establishment of the Paediatric Pharmacology Research Unit Network within the NICHD, that facilitates and promotes paediatric labelling of new or already marketed drugs, can be identified as a strategy to help fill this gap.

These experiences show that diversity issues concerning women, ethnic minorities and children were strategically addressed in different ways in the US. The organisation of the research processes varied, but additionally, it seems that research paradigms and methods to address health problems of ethnic minorities and children require more complex changes in research policies, as compared to studying women's health issues. The data indicate that US research policy changes concerning women's health have been more successful, than those addressing ethnic minorities and children.

European research strategies are organised differently from those in the US. The EU develops central research programmes in which new research paradigms can be implemented. In particular, women's health issues were placed on the research agenda. This policy was developed in 1996, after the United Nations World Conference on Women in Beijing in 1995. The launch of gender mainstreaming, or integrating gender into all major European policy areas, has formed the strategic approach to the question of equal opportunities between men and women. According to the WHO, gender mainstreaming in health is a strategy that promotes the integration of gender concerns into the formulation, monitoring and analysis of policies, programmes and projects, with the objective of ensuring that women and men achieve the highest health status (Bekker 2003). This policy is to be implemented in all institutions, policies, programmes and policies of the EU.

A gender assessment of the Fifth Framework Programme (FP5) was conducted to evaluate its policy. It was noted that projects addressing sex/gender differences did so in a very limited way. Projects did not address differentiated human populations. Also, many project designs did not discuss the composition of the research population and ended by under-representing women and making males the norm. Similarly, data collection methods were not explicitly explained in terms of their suitability for both sexes. The report concludes that in research projects "the evaporation of gender was evident" (Klinge 2001).

Current developments in FP6 address several of these shortcomings. The EU recognised a threefold relationship between women and research, and has articulated its action around the following themes:

- Women's participation in research must be encouraged both as scientists/technologists and in the evaluation, consultation and implementation process.
- Research must address women's needs, as much as men's needs.
- Research must be done to contribute to better understanding of gender issues (Science & Society in Europe 2004).

Thus, at a general level, the need to study sex/gender issues is acknowledged, and is considered to reflect scientific excellence. In contrast, the programme description of FP6 Priority Theme 1, *Life sciences, genomics and biotechnology for health*, does not refer to any diversity or gender specific issues. A number of research areas are mentioned, for example diabetes and cardiovascular diseases in which sex differences are known to be relevant, and research into relationships between functional genomics and fundamental biological processes is envisioned, but no reference is made to gender sensitive research questions. Another research area of FP6 is the programme on human development and ageing. Again, the programme description does not mention studying gender-related differences in diseases. Thus, in the information for researchers, the gender policy is inadequately addressed. Guidelines for project proposals require the development of a gender action plan and mention the importance to integrate sex/gender, age (child health) and genetic variation (inter-individual variations) into research. The Guide for proposers states that "the possibility of gender/sex must therefore be considered in all areas of health research, unless it can be demonstrated that gender/sex is inappropriate with respect to the health of the subjects or the objectives of the research." Sex/gender issues should be considered in:

- The formulation of research hypotheses, in the development of research protocols, choice of research methods and in the analysis of results.
- Biological, pre-clinical and epidemiological, behavioural research/studies on both human and animal subjects.
- The use of cells, tissues and other specimen, where appropriate.
- The choice of a particular study population which should be thoroughly justified and the sex of the participants fully described (Guide for proposers 2004).

We conclude that the EU has implemented a range of recommendations of the FP5 gender assessment, in particular increasing awareness for sex/gender balances in research. The recommendation to build a socio-cultural dimension into research would have made an innovative research strategy, but this was not implemented. Additionally, information about the main biomedical research programmes did not mention the focus on sex/gender issues, and more strength could be added to implement the new research strategy into the programming level. At present, FP6 is still in action and the focus on gender issues in research is monitored regularly.

The various approaches of the NIH and the EU for programming the research agenda on diversity issues may have different social impacts, i.e. building partnerships on mutual scientific interests that challenge mainstream paradigms, in contrast to more or less imposing a new scientific paradigm on researchers. Concerning the latter strategy, experience and knowledge on how to deal with the new scientific approaches may be considered a relevant condition. A lack of such experiences has been acknowledged within the EU internal reviewing processes of research protocols, and the need to increase the level of knowledge on studying sex/gender issues was recognised (Klinge 2001). Subsequently, manuals were developed for scientific and project officers for guidance how to implement concretely the gender mainstreaming throughout the whole process of the call to the follow-up of contracts (DG RTD 2003). Therefore, a key element is to build up experience inside and outside research agencies on diversity in clinical research and to gradually expand this strategic approach to the research agenda. An alternative approach for the general implementation of diversity in the research agenda is therefore to define a minimum number of projects that specifically deal with diversity issues in order to build up scientific experiences, and from there on gradually expand the range of projects. The conduct of the US study on hormonal replacement therapy is a valuable example, because it highlights the need to critically review the existing evidence and consider alternative research questions. Also, stimulation of multidisciplinary research projects may increase the focus on diversity. These approaches can be used in top-down as well as in bottom-up procedures. Careful programming of the nature of the projects is essential.

Information from the other research funding agencies included in this review, i.e. MRC in the UK and NRC in Norway, revealed no general research policies concerning diversity issues at the same level as those developed by the US and EU, so these agencies are not discussed further.

Identifying under-researched diversity issues

The second strategy to change the research agenda is to identify under-researched areas. A well-recognised method is collaboration with stakeholders that bring in different perspectives on health issues. Interactive and participative mobilisation of different knowledge stocks, including social sciences and various stakeholders, may lead to new ways of framing health problems, identifying which diversity issues matter to whom and when, and different scientific paradigms.

Health professionals working with under-researched populations form a group of relevant stakeholders. Feedback from their experiences can offer a better understanding of the knowledge gaps and the socio-economic and cultural aspects of diversity. Also, they may help to discover the plurality of diversity issues; which aspects matter when and for whom in medical practice? Experiences from health professionals may help to identify hypotheses in which different disease perceptions or observed effect

modification may result in differences in health outcomes in practice, and meaningful translation of these issues into clinical research questions. Also, implementation of research findings into diverse medical practices and settings may require diversity research. Developing strategies for systematic input from practitioners into the process of agenda setting, recognises that a particularly important part of innovation takes place in the context of medical practice (Gelijns 1998; Vos 1991).

Another relevant source of practice information is the website of the Dutch College of General Practitioners (NHG). It provides information on knowledge gaps in practice guidelines, as experienced by general practitioners. Many of these gaps refer to diversity issues, in particular concerning age and sex/gender differences in diagnosis, treatment and prevention (NHG website). With respect to ethnicity, a recent study revealed that more relevant clinical information exists, than has been considered in the development of practice guidelines (Manna 2003; Manna 2003 b). This study also illustrated the lack of information in medical practice on clinical treatment of various patient groups. To identify under-researched areas of health issues in children and the elderly, various professionals can be identified, also to address gaps in methodological aspects of studying diversity.

Clearly, representatives of the specific patient population can help in identifying and prioritising under-researched areas. In the Netherlands, interaction between clinical researchers and patients with muscular diseases has generated new research topics and research collaboration. These processes also reveal mismatches between the interests of researchers and those of patients in defining research topics. Perspectives on research differ genuinely, but the research community does not sufficiently acknowledge this (Abma 2004).

An example of collaboration with stakeholders is the *Gender and HIV viral load workshop*, organised by the NIH (NIAID 2000). This workshop brought together a small group of basic and clinical researchers, as well as community representatives, to review the research to date, define potential research gaps and future directions and facilitate a more comprehensive and co-ordinated research agenda on the issue of gender and viral load. The concise topic allowed an in-depth approach to the complex interactions between biomedical and socio-cultural parameters affecting clinical outcomes (effect modification) in diagnosis, treatment and prevention.

Internationally, patient/consumer involvement in clinical research is recognised as an important strategy. This can also be applied to developing research agendas in the perspective of diversity issues in priorities and outcomes. Two aims for the policy of public involvement in research are often mentioned, i.e. to increase democratic input into decision making, and to increase responsiveness of patients to services (Florin 2004). It is argued that representation of consumers in the decision-making and

research process is a political priority; stakeholders have a 'right to be involved' (Hanley 2004). Within the research community, notions of objectivity in research are increasingly being challenged. Debates about (consumer) involvement in research involve challenges to the norms of research structures, where knowledge development is increasingly seen not just as the domain of academics, but as a more inclusive activity (Reed 2004). The public will become more knowledge-empowered, and therefore better able to challenge beliefs and practices of researchers (Hanley 2001). Participatory research strategies were developed in particular in the UK. A recent evaluation study looked at the processes and outcomes of involving consumers in research and development agenda setting for the NHS (Oliver 2004). Good leadership, purposeful outreach to consumers, investing time and effort in good communications, training and support and thereby building good relationships were important to overcome barriers. In the Netherlands, a recent analysis of the patients' and consumers' movements revealed that their main efforts are concerned with patients' interests in health care, and patients' self help groups (Nederland 2004). Although consumer/patient involvement in research has evolved less strongly in the Netherlands than in the UK, experiences in collaborative processes have increased the relevance of the outcomes for both patients and professionals (Abma 2004).

Strategies

In sum, the following strategies for programming the research agenda were identified and discussed:

- Challenging the dominant paradigms in research and identifying alternative research questions and hypotheses that focus on diversity issues.
- Identifying under-researched areas in specific patient populations or communities.
- Systematic consultation of various stakeholders, in particular users/patients and health professionals working with under-researched patient groups, to identify knowledge gaps and various perspectives on diversity issues.
- Careful planning and incorporation of new paradigms into research programmes, including consistent implementation of general research policies into research programmes and guides for researchers.
- Developing expertise among researchers as well as in the reviewing processes.
- The need for monitoring, regular evaluations, and feedback on the outcomes of research policies and subsequent strategic developments.

6.3.3 Policy aspects of methodological strategies

The present practice of RCTs as main method to study diversity issues was identified as a third barrier to focus on diversity. Methodological implications to increase the focus of RCTs on studying effect modification

were discussed in review 5. Here we consider policy strategies that aim to stimulate the use of a broader range of methods to study diversity issues in clinical research. This may broaden the biomedical focus on diversity to include other perspectives.

Effect modification, or differences in health outcomes of diagnostic, preventive or treatment interventions, must be considered because it affects individual health care. Therapeutic effect modification is uncommon without obvious causes. Additionally, biomedical and socio-cultural causes may interact in the nature and size of effect modification. As a result, detecting effect modification is highly relevant to medical practice, but complicated to study in methodologically sound ways. In principle, the RCT is a good instrument to study effect modification, provided that sufficient precautions are taken in study design and conduct. In review 5 a broad range of methodological implications were discussed. The preceding reviews also argued that it is relevant to consider 'diversity' in a broader concept than studying differences between men and women, elderly and younger patients, and people from various ethnic origins. Strategies that aim to broaden the biomedical perspective on clinical research therefore focus on combining qualitative and quantitative research methods, involving consumers, patients and practitioners in the design and conduct of clinical research, and rethinking the procedures for research proposals. These strategies will be further elaborated.

Stimulating diversity of methods

As discussed in section 6.3.1, a methodological strategy is to define research questions and hypotheses that address the underlying biomedical and/or socio-cultural mechanisms of diversity. In addition, research is needed to increase our understanding of concepts of diversity in the context of patients' lives. Therefore, other methods than RCTs are needed to explore the ways in which patients, consumers, or trial participants experience and operationalise diversity in disease parameters, outcomes, and in the use of health care facilities. Especially in chronic diseases, this approach is likely to augment understanding of how diseases affect patient's lives and well being differently, and how their valuation of treatments and side effects can be understood in the contexts of the life-worlds in which they arise. Addressing different perspectives, notions and values of patients prospectively will generate new research hypotheses on diversity, for example concerning the implementation of medical practices in daily life, or studying effect modification. Thus, the results from qualitative studies of diversity issues may be used subsequently in RCTs to quantify effect modification using parameters that are particularly relevant to patients. There is increasing understanding that combining qualitative and quantitative research methods, therefore encouraging multidisciplinary research, is essential to optimise health care (Philipsen 2004).

Qualitative research methods to be applied differ according to the nature of the clinical problem, for example in the domain of care or cure. Also, the

stage of the clinical research process in which qualitative methods will be applied, may direct the choice of method (workshop, interviews, panels, focus groups, survey, ethnographic study) and stakeholders to involve (patients, clinical practitioners). A novel approach is integrating qualitative research methods into quantitative studies, in particular RCTs (Dixon-Woods 2004). This strategy is likely to be an effective way to broaden the biomedical focus on diversity. New approaches to using the internet to define study populations or communities can be expected to expand research methods for studying patient experiences, perspectives and values (Eysenbach 2004).

The above strategies aim to carefully consider how to take into account diversity and broaden the biomedical research perspective on diversity. If diversity is relevant, because of underlying biomedical and/or socio-cultural mechanisms, it needs to be made the core perspective in study design and conduct (Doyal 2001). This strategy can be found to have taken a different direction than international methodological debates on the lack of focus on diversity in clinical studies. Many studies in the US and Europe have revealed the general under-representation of women, minorities, and children in clinical research and RCTs in particular. These findings have played an important role in the present debates on how to address the shortcomings and inequalities. Increasing the inclusion of women, minorities and children was the main counteracting methodological strategy of the NIH (see also Review 3). This policy involves the development and implementation of inclusion guidelines for women, minorities, and children, in order to analyse the results to detect possible subgroup differences. The development and implementation of these guidelines have gained much attention internationally. For example, the Norwegian Research Council issued comparable guidelines in 2001, making sex/gender a variable in all medical research (Forskningsetiske komiteer 2001).

Evaluation studies of the effects of the NIH Guidelines on the inclusion of women found results that varied between achievements in the desired direction (Pinn 2003), and no changes at all (Ramasubbu 2001; Vidaver 2000). However, these studies distract attention from the actual bottleneck, which continues to be a lack of subgroup analyses (Greenberger 2000).

An important argument underlying the NIH guidelines is representation. However, it is not clear that this is the way to proceed. Participation in itself does not lead to new knowledge on diversity issues. A pitfall of the representation strategy in trial methodology is that the sample is too small; the trial result may not just be a null, but an uninformative null (Buring 2000). Thus, although research showing under-representation of various patient groups has usefully boosted the discussions on underserved populations, increasing representation can be regarded as a reductionist strategy for studying diversity issues. The present analysis has therefore taken a different direction.

Participatory strategies

A strategy to broaden the perspective on trial design and conduct is patient and consumer involvement at all stages of design and conduct of clinical research (Hanley 2004). This strategy may also involve clinical practitioners with experiences in relevant patient populations. The general aim of participatory strategies is to ensure the development of patient centred research questions and outcome measures in order to increase the quality of care (Hanley 2001, Hanley 2004). Also, clinical research can be more transparent and publicly accountable.

The literature contains many examples of patient/consumer involvement at different stages of research. Involvement of stakeholder umbrella organisations and consumers/consumer groups can help facilitate access to marginalised communities (e.g. minorities) and facilitate recruitment. Consumers can thus be involved in managing processes of explanation, information and support (Thornton 2003; Hanley 2004). Collaboration with consumers/patients may be beneficial for recruitment of trial participants through advising on, or revising trial protocols and information leaflets, ensuring that the patient will be able to make a balanced decision, and is willing to accept trial conditions (risk assessment) (Thornton 1997; Hanley 2001; Maslin-Prothero 2003). Consumers can help solve ethical problems associated with research (Koops 2002). In the dissemination stage of research, consumers can help and work to ensure that changes are implemented by identifying existing research that is not being disseminated or implemented and could improve treatments or services (Hanley 2004). Also, consumers can improve the quality of information provided to patients. In many cases consumers have links to consumer networks to help publicise results of clinical research (Hanley 2001). Involvement allows more direct access to research findings, which might enhance consumer's abilities to influence policy (O'Donnell 2004).

Experiences with public involvement in research are numerous, but vary between the domains of care and cure. For instance, collaboration with users is particularly important in the social care field (Hanley 2004). Experiences with consumer involvement in the design of RCTs for curative interventions can also be found, for example concerning the use of thrombolysis for acute ischaemic strokes (Koops 2002). In the following, we will further elaborate on two examples of consumer involvement.
In The Netherlands, the perspective to place the development of clinical research in the context of patients' lives has been developed in the collaboration of patients with muscular diseases and researchers that started fifteen years ago. An example of this collaboration is a study of people with spinal cord lesions which found that they considered rehabilitation treatment to be limited to mobility and the person's ability to live or do things independently. In the patients' perspective, rehabilitation should have a longer time window and focus more on the perception of the disorder and the changes that come with it. These experiences have been

translated into research topics and projects in rehabilitation in a collaborative process of patients and researchers. The researchers acknowledged the outcomes were more relevant both for the patients and the professionals (Abma 2004).

In the second example, concerning the dialogue between AIDS activists' organisations and researchers in the 1990s, the researchers' choices in the conduct of clinical research on antiviral drugs for HIV/AIDS were challenged (Epstein 1997). In the early years of the AIDS epidemic, clinical trials of anti-viral drugs were designed to evaluate the effectiveness of the drugs, i.e. reduction of mortality. These rigorous trials, however, were time consuming and this was exactly what those infected with HIV were running short of; they needed quick access to drugs. When an uncontrolled trial of a new combination revealed an unexpected increase of CD4 cells of the participants, AIDS activists challenged the scientific community to use this surrogate parameter to study whether a drug 'worked' against HIV/AIDS. The underlying scientific question at that time was whether the suggested parameter, CD4 cell count, was related to mortality, that is, if it was a true predictor of effectiveness. Despite a lack of proof, the debates with the AIDS activists persuaded the research community and the drug regulators to accept this switch in parameters. A social argument for their decision was the belief that patients would no longer take part in the long lasting effectiveness trials, which would therefore have to be abandoned. In the end, the switch in parameters was proven wrong, because the increase in CD4 cell counts was unrelated to reduced mortality, but some argued that the stage of the disease was a relevant variable in this conclusion. Reflecting on the course of the events, AIDS activists stressed that they needed "access and answers", referring to accelerated approval of new drugs, but also good science providing the relevant answers (Epstein 1997).

These examples highlight the relevance of public involvement and lay-expertise in clinical research, not only concerning research questions, where new perspectives are introduced into the design process, but also in the scientific debate on the implications of methodological aspects of research (Rabeharisoa 2003). In both examples, the research aims and consequently research parameters were contested: are the choices made by the scientists the most relevant to the patients or participants in clinical research? These questions are equally important in the study of diversity issues.

A number of drawbacks and difficulties can be identified in the process of consumer involvement in clinical research. Collaboration with consumers/patients can be time consuming, and as a result trials may take longer to complete and cost more (Hanley 2004). Consumer preferences often reflect individual experiences, rather than a general view or representativeness of participating consumers. It should be recognised that people may want to get involved for various political or personal reasons (O'Donnell 2004 b). Therefore, much depends on the individual's ability to

capture the essence of experience and generalise from that (Williamson 1999; Hanley 2001; Lockwood 2004; Wensing 2003). Researchers may perceive the process of sharing control and the loss of power as disadvantageous or as a threat (Hanley 2004; Reed 2004). In some cases, researchers have accused consumer representatives of politicising science (International Cancer News 1997). Other drawbacks concern possible conflicts of interest between the needs of a trial and those of a patient group. For example, the need for reliable assessment of the cost effectiveness of expensive new drugs may clash with patients' requests for immediate availability for all patients. It was also mentioned that response rates may be reduced as a result of consumer input, for instance, consumer-participants may object to sending multiple requests for response when selected trial subjects do not react after the first request (Hanley 2001).

Experiences with consumer/patient involvement in clinical research have shown that building relationships with patients/consumers in whom they feel that their contribution is respected and valued is crucial. This can be expected to be equally important in collaborations with practitioners. New roles between patients, practitioners and researchers need to be defined and developed, and differences in perspectives acknowledged and respected. Frequent communication and collaboration between those involved is essential, not only on research topics in a narrow sense, but also related to ownership, trust, expectations and who will be able to use and benefit from the results (Abma 2004; Trivedi 2002; Reed 2004; Adams 2004). Giving people accurate, high quality, up to date information is an important starting point (Richards 1999). It is relevant to make clear what kind of range of experience, perspectives and expertise of patients/consumers is needed. Ideally, the 'end user' of the research should be involved, which can be a carer, patient or both, depending on the situation. Also, it may be necessary to offer resources and support, and provide training for members of the public and researchers to facilitate processes (Hanley 2004).

Research procedures

Methodological strategies that focus on studying diversity in RCTs, stimulating the combination of qualitative and quantitative research, and involvement of patients, consumers and clinical practitioners in research, need translation and implementation into the processes of ZonMw. Rethinking various procedures, guidelines and criteria for researchers and assessors is needed after decision-making about the general aim and direction of strategies to focus on diversity in clinical research. A number of implications emerge from the above that can be addressed here.
To guide the development of project proposals in which diversity is the core perspective of research, a strategy for ZonMw is to always request a concise literature search in which biomedical and socio-cultural mechanisms of diversity and effect modification are considered. This strategy corresponds to the EC-FP6 guidelines for project proposers, as

discussed in section 6.3.2. According to the outcomes and interpretation of the evidence, methodological consequences for the design and conduct of the research will follow. If effect modification in various subgroups is plausible, the choices for study population, enrolment criteria and sites, research measures, and analyses can consequently be argued for. A study by Tallon indicates that given the diversity of individual experiences, it may be more appropriate to use patient-centred measures or individualised measures, rather than a single standard research instrument for all patients (Tallon 2000). Also, it is suggested that research might have to become less rigid about methodology and begin to see diversity as strength of the study.

Moerman and Van Mens critically evaluated many steps in study protocols on their unintended effects as barriers to include specific patient groups (Moerman 2004). Screening these topics for specific criteria to guide the development of research proposals and subsequent selection may decrease selection biases. In review 4 examples were given to adjust visit schedules for specific patient groups, such as elderly or adolescents, to making participation in research 'do-able'.

The shift in ethical assessment of study protocols from protection to consider the benefits of participation implies critical assessment of patient selection criteria. Research into the processes of patient selection for research is relevant to gain a better understanding of exclusionary practices and mechanisms. According to the findings of such research, specific strategies can be applied to avoid barriers to participation. A strategy is thus to develop new approaches for selection criteria that create opportunities for participation in research, more than to function as barriers. Feedback and training of researchers is needed to increase awareness of these issues.

In addition to increasing research hypotheses concerning effect modification, ZonMw can also facilitate research that explores different perspectives, notions and values of patients on disease parameters, outcomes, and in the use of health care facilities. Therefore, ZonMw can stimulate the development of clinical research in multidisciplinary teams, including biomedical and social scientists, and application of a diversity of methodologies. Involvement of relevant populations affected by the research project in the assessment of study protocols and consultation by Research Ethics Committees was mentioned in review 4. Also, ZonMw can actively stimulate the involvement of patients, consumers and clinical practitioners in research design and conduct. As mentioned before, a strategy is to carefully develop these collaborations and provide feedback on the processes and outcomes to the research communities and other stakeholders.

The costs of clinical research in which diversity is a core issue of design and conduct can be expected to increase. Patient groups may need to be larger to allow for comparison of outcome measures. Also, consumer

involvement often increases costs because more people are involved in consultations. Because of the potential benefits of these novel approaches, researchers should be encouraged to submit proposals that include a diversity of methods and budget for people's involvement (Hanley 2004). A last consideration of introducing novel methodological strategies in ZonMw's procedures is that expertise may need to be broadened. Expertise on biomedical and socio-cultural aspects of diversity is necessary to guide research programming, development of project criteria, and assessment of project proposals, both on the clinical relevance to study diversity, as well as methodological implications.

Strategies

In sum, the following policy strategies were discussed:

- Stimulating and facilitating the combination of qualitative and quantitative research methods to broaden the biomedical perspective on diversity in clinical research. As a result, studying effect modification and exploring diversity from the patients' perspective can be integrated to generate new research hypotheses and increase our understanding of diversity.
- Stimulating and facilitating clinical research development in multidisciplinary teams.
- Assessing diversity in clinical research can be achieved by making it the core perspective in the design and conduct of research.
- Involving the perspective of users/patients and clinical practitioners at all stages of research planning, decision-making and conduct in order to increase relevance of the outcomes for patients and health professionals.
- Requesting concise literature overviews on diversity issues of the clinical topic and research parameters for all project proposals.
- Performing critical reflection on ZonMw's project criteria and procedures to identify exclusionary barriers in research practices, for example on patient selection criteria. Study protocols and patient selection criteria need to create opportunities for diverse patient groups to participate in clinical research.
- Further development of expertise on diversity issues, clinical topics, methodology, selection processes, and participatory research is essential.

References

Abma TA. Proefpersoon wordt participant. Onderzoek met en voor patiënten is interessant en uitdagend. Medisch Contact 2004: 38: 1484-1486.

Adams A, Miller-Korth N, Brown D. Learning to work together: developing academic and community research partnerships. WMJ 2004: 103(2): 15-19.

Bekker MHJ. Integrating gender within health research is more than sex disaggregation of data: a multi-facet gender and health model. Psychology, Health and Medicine 2003; 8: 231-243.

Brass LM. Hormone replacement therapy and stroke: clinical trials review. Stroke 2004: 35(11 suppl.1): 2644-7.

Buring JE. Women in clinical trials: a portfolio for success. N Engl J Med 2000: 343: 505-506.

Caron-Flinterman JF, Broerse JEW, Bunders JFG. The experiential knowledge of patients: A new resource for biomedical research? Soc Sci Med, 2005 (in press).

DG RTD, Unit C05: Women & Science. Gender mainstreaming in the 6th Framework Programme – Reference guide for scientific officers/project officers. European Commission Community Research, March 2003.

Dixon-Woods M, Agarwal S, Young B, Jones D, Sutton A. Integrative approaches to qualitative and quantitative evidence. London: Health Development Agency, 2004. www.had.nhs.uk

Doyal L. Sex, gender, and health: the need for a new approach. BMJ 2001: 323: 1061-1063.

Epstein S. Activism, drug regulation, and the politics of therapeutic evaluation in the AIDS era: a case study of ddC and the 'surrogate markers' debate. Social Studies of Science 1997: 27(5): 691-726.

Eysenbach G, Powell J, Englesakis M, Rizo C, Stern A. Health related virtual communities and electronic support groups: systematic review of the effects of online peer to peer interactions. BMJ 2004: 328: 1166-1171.

Florin D, Dixon J. Public involvement in health care. BMJ 2004; 328: 159-161.

Forskningsetiske komiteer. Guidelines for the inclusion of women in medical research – gender as a variable in all medical research. Guidelines set by the National Committee for Medical Research Ethics in Norway, 2001-09-01. www.etikkom.no/Engelsk/NEM/publications/women

Gelijns AC, Rosenberg N, Moskowitz AJ. Capturing the unexpected benefits of medical research. N Engl J Med 1998; 339(10): 693-698.

Greenberger P, Gevers NN. Improving clinical research for women. J Womens Health Gend Based Med 2000; 8: 809-811.

Grin J. Health technology assessment between our health care system and out health. Exploring the potential of reflexive HTA. Poiesis Prax 2004; 2: 157-174.

Guide for proposers. Priority 1, Life sciences, genomics and biotechnology for health. Call 3, integrated projects, 15th June 2004.

Hanley B, Truesdale A, King A, Elbourne D, Chalmers I. Involving consumers in designing, conducting, and interpreting randomised controlled trials: questionanaire survey. BMJ 2001; 322: 519-523.

Hanley B. et al. Involving the public in NHS, public health and social care research: briefing notes for researchers, 2nd Ed. INVOLVE, February 2004. www.invo.org.uk

Healy B. Keynote address 2 to A Report of the Task Force sponsored by the NIH Office of Research on Women's Health Recruitment and Retention of Women in Clinical Studies, and the Critical Role of Relevance. Challenging sameness: Women in clinical trials. January 6, 2003. www4.od.nih.gov/orwh/SMR_Final.pdf

Helmuth L. Biomedical Policy: NIH, under pressure, boosts minority health research. Science 2000: 288: 596-597.

International Cancer News. Eur J Cancer 1997; 33: 802.

Klinge I, Bosch M. Synthesis Report 2001 of the European Commision. Gender in research. Gender impact assessment of the specific programmes of the fifth framework programme. An overview.
http://europa.eu.int/comm/research/rtdinfo.html

Koops L, Lindley RI. Thrombolysis for acute ischaemic strike: consumer involvement in design of new randomised controlled trial. BMJ 2002; 325: 415-418.

Lockwood S. "Evidence of me" in evidence based medicine? BMJ 2004; 329:1033-1035.

Majumdar SR, Almasi EA, Stafford RS. Promotion and prescribing of hormone therapy after report of harm by the Women's Health Initiative. JAMA 2004; 292(16): 1983-1988.

Manna DR, Bruijnzeels MA, Mokkink HGA, Berg M. Minder etnische kennis in de standaarden van het Nederlands Huisartsen Genoorschap over diabetes mellitus type 2, hypertensie en over astma bij volwassenen dan in de onderbouwende literatuur. Ned Tijdsch Geneeskd 2003: 147(35): 1691-1696.

Manna DR, Bruijnzeels MA, Mokkink HGA, Berg M (b). Ethnic specific recommendations in clinical practice guidelines: a first exploratory comparison between guidelines from the USA, Canada, the UK, and the Netherlands. Qual Saf Health Care 2003: 12: 353-358.

Maslin-Prothero S. Developing user involvement in research. J Clin Nurs 2003; 12: 412-421.

Moerman CJ, Mens-Verhulst J van. Gender-sensitive epidemiological research: suggestions for a gender-sensitive approach towards problem definition, data collection and analysis in epidemiological research. Psychology, Health and Medicine 2004: 9(1): 41-52.

Nederland T, Duyvendak JW. De kunst van effectieve belangenbehartiging door de patiënten- en cliëntenbeweging. Utrecht: Verwey-Jonker Instituut, 2004.

NIAID. Gender and HIV viral load workshop, January 29, 2000. www.niaid.nih.gov/diads/adulttrials/gender_and_hiv.htm

NIH News. Why study finds no heart disease benefit, increased stroke risk with estrogen alone. NIH News 13 April 2004. Website: www.nhlbi.nih.gov/new/press/04-04-13.htm

NIH Policy and guidelines on the inclusion of children as participants in research involving human subjects. NIH Guide for grants and contracts, 1998.

O'Donnell M, Entwistle V. Consumer involvement in research projects: the activities of research funders. Health Pol 2004; 69: 229-238.

O'Donnell M, Entwistle V (b). Consumer involvement in decisions about what health-related research is funded. Health Pol 2004; 70: 281-290.

Oliver S, Clarke-Jones L, Rees R, Milne R, Buchanan P, Gabbay J et al. Involving consumers in research and development agenda setting for the NHS: developing an evidence-based approach. Health Technol Assess 2004; 8(15).

Philipsen H, Vernooy-Dasssen M. Kwalitatief onderzoek: nuttig, onmisbaar en uitdagend. Huisarts & Wetenschap 2004: 47(10): 454457.

Pinn VW. Preface and introduction to A Report of the Task Force sponsored by the NIH Office of Research on Women's Health Recruitment and Retention of Women in Clinical Studies, and the Critical Role of Relevance. Science meets reality. January 6, 2003. www4.od.nih.gov/orwh/SMR_Final.pdf

Pinn VW, Begg L, Rudick JE. Office of Research on Women's Health: 2004 Priorities for women's health research. www4.od.nih.gov/orwh

Rabeharisoa V. The struggle against neuromuscular diseases in France and the emergence of the 'partnership model' of patient organization. Soc Sci Med 2003: 57: 2127-2136.

Ramasubby K, Gurm H, Litaker D. Gender bias in clinical trials: do double standards still apply? J. Womens Health Gend Based Mewd 2001; 10: 757-764.

Reed J, Weiner R, Cook G. Partnership research with older people – moving towards making the rhetoric a reality. International Journal of Older People Nursing, in association with Journal of Clinical Nursing 2004: 13(3a), 3-10.

Reuzel RPB. Health technology assessment and interactive evaluation: different perspectives. PhD dissertation, Faculty of Medicine, University of Nijmegen, 2000: 146-188.

Richards T. Patients' priorities need to be assessed properly and taken into account. BMJ 1999; 318: 277.

Science and Society in Europe 2004. Integrating the gender dimension in FP6 projects. http://europa.eu.int/comm/research/science-society/gen_mainstr_en.htm

Tallon D, Chard J, Dieppe P. Relation between agendas of the research community and the research consumer. Lancet 2000; 355: 2037-2040.

Thornton H. The voice of breast cancer patients – a lonely cry in the wilderness. Eur J Cancer 1997; 33(6): 825-828.

Thornton H, Edwards A, Elwyn G. Evolving the multiple roles of 'patients' in health-care research: reflections after involvement in a trial of shared decision-making. Health Expectations 2003; 6: 189-197.

Trivedi P, Wykes T. From passive subjects to equal partners. British Journal of Psychiatry 2002: 181: 468-472.

Vidaver RM, Lafleur B, Tong C, Bradshaw R, Marts SA. Women subjects in NIH-funded clinical research literature: lack of progress in both representation and analysis by sex. Journal of Women's Health & Gender-Based Medicine 2000: 9(5): 495-503.

Vos R. Drugs looking for diseases. Innovative drug research and the development of the beta-blockers and the calcium channel blockers. 1991. Dordrecht: Kluwer Academic Publishers.

Wensing N, Baker R. Patient involvement in general practice: a pragmatic framework. Eur J Gen Pract 2003; 9: 62-65.

Williamson C. The challenge of lay partnership. BMJ 1999; 319: 721-722.

ZonMw 2004. Thema Opmaat, diversiteit en patiëntenperspectief bij ZonMw. ZonMw-werkgroep Vraaggericht Werken, februari 2004: 1-6.

APPENDIX

Invitational expert meeting

Diversity in Clinical Research

November 12-13 2004, Bergen, The Netherlands

Programme

Friday November 12 2004

9.00 hour	Arrivals and registration
	Discussions
	Chair: prof. dr. Martin Offringa
10.00 – 11.00 hour	Review *"Why is it important to focus on diversity in clinical research?"* Authors: Dr. Nicolien Wieringa, Prof.dr. Menno Reijneveld, Dr. Karien Stronks. Referee: Dr. Richard Koopmans, Department of Medicine and Pharmacology, Amsterdam Medical Center - Universiteit van Amsterdam
11.00 - 12.00 hour	Review *"Which diversity matters?"* Authors: Dr. Amâde M'charek, Drs. Mirjam Kohinor Referee: Dr. Jeanette Pols, Trimbos Insitute, Utrecht
12.00 – 14.00 hour	Lunch
	Chair: Prof.dr. Patrick Bossuyt
14.00 – 15.00 hour	Review *"Medical and socio-cultural assumptions in research"* Authors: Prof.dr. Anita Hardon, Drs. Eline van Haastrecht Referee: Prof.dr. Trudy Dehue, Heymans Institute, Department of Psychology, Education and Social Sciences, University of Groningen

15.00 – 16.00 hour	Review *"Ethical considerations when taking into account diversity in clinical research"* Author: Prof.dr. Dick Willems Referee: Prof. dr. Niek Klazinga, Department of Social Medicine, Amsterdam Medical Center - Universiteit van Amsterdam
16.00 – 17.00 hour	Tea break Chair: Prof.dr. Menno Reijneveld
17.00 – 18.00 hour	Review *"Methodological implications of focusing on diversity in clinical research"* Authors: Dr. Madelon van Wely, Prof.dr. Patrick Bossuyt, Prof.dr. Rick Grobbee, Prof.dr. Martin Offringa, Dr. Karien Stronks Referee: Prof.dr. Klim PcPherson, Department of Public Health Epidemiology, Oxford University
18.00 – 19.00 hour	Review *"Strategies for change"* Authors: Dr. Nicolien Wieringa, Drs. Eline van Haastrecht, Dr. Andrew Herxheimer, Prof.dr. Anita Hardon, Prof.dr. Niek Klazinga, Dr. Amâde M'charek Referee: Mrs. Hazel Thornton, Dsc, Honorary Visiting Fellow, Department of Health Sciences, University of Leicester
19.30 hour	Dinner

Saturday November 13 2004

	Chairs: Prof.dr. Anita Hardon, Prof.dr. Trudy Dehue
9.00 – 11.00 hour	Two parallel sessions to discuss conclusions and implications for the steering, design and conduct of clinical research Chair: Prof.dr. Niek Klazinga
11.00 – 12.30 hour	General discussion on conclusions and recommendations for ZonMw and other actors
12.30 hour	Closure

Participants

Invitational Conference on Diversity in Clinical Research,

November 12-13 2004, Bergen, The Netherlands

Drs. Teus van Barneveld	Dutch Institute for Health Care, CBO
Prof.dr. Stuart Blume	Universiteit van Amsterdam, Department of Sociology and Anthropology, Dynamics of Science and Technology
Drs. Tjitske Boland	University of Maastricht, Department of Health Care Studies
Drs. Hannie Bonink	Dutch Organisation for Health Research and Development (ZonMw)
Mrs. Froukje Bos	The Pandora Foundation
Prof.dr. Patrick Bossuyt	Academic Medical Centre, Department of Clinical Epidemiology and Biostatistics
Dr. Marc Bruijnzeels	Institute of Health Policy and Management, Erasmus University Rotterdam
Prof.dr. Trudy Dehue	Rijksuniversiteit Groningen, Heymans Institute, Department of Psychology, Education and Social Sciences
Dr. Annette de Deugd	Dutch Organisation for Health Research and Development (ZonMw)
Dr. Rob Dillmann	Orde van Medisch Specialisten
Prof.mr. Sjef Gevers	Academic Medical Centre, Department of Social Medicine
Prof.dr. John Grin	Universiteit van Amsterdam, Department of Political Science
Drs. Eline van Haastrecht	Universiteit van Amsterdam, Department of Political Science
Prof.dr. Anita Hardon	Amsterdam School for Social Sciences Research (ASSR), Universiteit van Amsterdam
Dr. Andrew Herxheimer	UK Cochrane Centre
Mrs. Christine Hogg	Consumers for Ethics in Research CERES UK
Dr. Karin Jansen	Dutch Organisation for Health Research and Development (ZonMw)
Dr.ir. Jessika van Kammen	Dutch Organisation for Health Research and Development (ZonMw)
Drs. Debby Keuken	Academic Medical Centre, Department of General Practice
Prof.dr. Niek Klazinga	Academic Medical Center, Department of Social Medicine
Dr. Ineke Klinge	University Maastricht, Department of Health Care Studies
Drs. Mirjam Kohinor	Universiteit van Amsterdam, Faculty of Science
Dr. Richard Koopmans	Academic Medical Center, Department of Medicine and Pharmacology
Drs. Dineke Korfker	TNO Prevention and Health

Participants

Invitational Conference on Diversity in Clinical Research,

November 12-13 2004, Bergen, The Netherlands

Mrs. Yolan Koster	Dutch Organisation for Health Research and Development (ZonMw)
Dr. Joke Lanphen	Dutch Organisation for Health Research and Development (ZonMw)
Dr. Jos van Loenhout	Health Care Insurance Board, CVZ
Dr. Amade M'charek	Universiteit van Amsterdam, Department of Political Science
Dr. Clara Moerman	Academic Medical Center, Department of General Practice, Women's Studies
Prof.dr. Klim McPherson	Oxford University, Department of Public Health and Epidemiology
Dr. Nico de Neeling	Health Council of the Netherlands
Prof.dr. Martin Offringa	Academic Medical Center, Department of Pediatrics; UK Cochrane Center
Drs. Martin van Oostrom	Aids Fonds Stichting AIDS, Policy Department
Prof.dr. Naomi Pfeffer	Consumers for Ethics in Research CERES UK
Dr. Jeanette Pols	Dutch Institute of Mental Health and Addiction
Prof.dr. Menno Reijneveld	University Groningen, Department of Health Sciences
Dr. Floor Rikken	Health Care Insurance Board, CVZ
Dr. Ronals Stolk	University Medical Center Utrecht, Julius Center for Health Sciences and Primary Care
Dr. Karien Stronks	Academic Medical Centre, Department of Social Medicine
Mrs. Hazel Thornton, DSc	Honorary Visiting Fellow, Department of Health Sciences, University of Leicester
Drs. Evert van der Veen	Dutch Organisation for Health Research and Development (ZonMw)
Dr. Cees Vos	Ministry of Health, Welfare, and Sports
Dr. Madelon van Wely	Academic Medical Centre, Department of Social Medicine
Dr. Nicolien Wieringa	Universiteit van Amsterdam, Department of Political Science
Prof.dr. Dick Willems	Academic Medical Centre, Department of General Practice

List of authors

P.M.M. Bossuyt, PhD, MD

Professor of Clinical Epidemiology. Academic Medical Center – Universiteit van Amsterdam, Department of Clinical Epidemiology and Biostatistics, P.O. Box 22660, 1100 DD Amsterdam, the Netherlands, Email: p.m.bossuyt@amc.uva.nl

D.E. Grobbee, PhD, MD

Professor of Clinical Epidemiology, Julius Center for Health Sciences and Primary Care, Clinical Epidemiology, University Medical Center Utrecht (UMCU), P.O. Box 85500, 3508 GA Utrecht, the Netherlands, Email: d.e.grobbee@umcutrecht.nl

E.K. van Haastrecht

Universiteit van Amsterdam, Department of Political Sciences / Department of Biology OZ Achterburgwal 237, 1012 DL Amsterdam, the Netherlands, Email: elinevh@gmail.com

A.P. Hardon, PhD

Professor of Care and Health, Anthropology, Universiteit van Amsterdam, Amsterdam School for Social Science, Department of Cultural Anthropology, Kloveniersburgwal 48, 1012 CX Amsterdam, the Netherlands, Email: a.p.hardon@uva.nl

A. Herxheimer, PhD, MD

UK Cochrane Centre, Middle Way, Oxford OX2 7LG, England, Email: andrew_herxheimer@compuserve.com

N.S. Klazinga, PhD, MD

Professor of Social Medicine, Academic Medical Center – Universiteit van Amsterdam Department of Social Medicine, P.O. Box 22660, 1100 DD Amsterdam, The Netherlands, Email: n.s.klazinga@amc.uva.nl

M. Kohinor

Universiteit van Amsterdam, Department of Political Sciences / Department of Biology, OZ Achterburgwal 237, 1012 DL Amsterdam, The Netherlands, Email: m.j.e.kohinor@uva.nl

A.A. M'charek, PhD

Associate Professor, Universiteit van Amsterdam, Department of Political Sciences / Department of Biology, OZ Achterburgwal 237, 1012 DL Amsterdam, The Netherlands, Email: a.a.mcharek@uva.nl

M. Offringa, PhD, MD

Professor of Paediatric Clinical Epidemiology, Academic Medical Center – Universiteit van Amsterdam, Department of Paediatrics, P.O. Box 22660, 1100 DD Amsterdam, The Netherlands, Email: m.offringa@amc.uva.nl

S.A. Reijneveld, PhD, MD

Professor of Social Medicine, University of Groningen, Department of Health Sciences, P.O. Box 196, 9700 AD Groningen, The Netherlands, Email: s.a.reijneveld@med.umcg.nl

R.P. Stolk, PhD, MD

Julius Center for Health Sciences and Primary Care, Clinical Epidemiology, University Medical Center Utrecht (UMCU), Present address: Professor of Clinical Epidemiology, University of Groningen / University Medical Center Groningen, Department of Epidemiology and Bioinformatics, P.O. Box 30001, 9700 RB Groningen, The Netherlands, Email: r.p.stolk@med.umcg.nl

K. Stronks, PhD

Associate Professor, Academic Medical Center – Universiteit van Amsterdam, Department of Social Medicine, P.O. Box 22660, 1100 DD Amsterdam, The Netherlands, Email: k.stronks@amc.uva.nl

T.E. Swierstra, PhD

Associate Professor, Faculty of Behavioral Sciences, Department of Philosophy, University of Twente, P.O. Box 217, 7500 AE Enschede, The Netherlands, Email: t.e.swierstra@wmw.utwente.nl

M. van Wely, PhD

Post-doc, Academic Medical Center – Universiteit van Amsterdam, Department of Social Medicine, P.O. Box 22660, 1100 DD Amsterdam, The Netherlands, Email: m.vanwely@amc.uva.nl

D.L. Willems, PhD, MD

Professor of Medical Ethics, Academic Medical Center – Universiteit van Amsterdam. Department of General Practice, P.O. Box 22660, 1100 DD Amsterdam, The Netherlands, Email: d.l.willems@amc.uva.nl

N.F. Wieringa, PhD, MPharm

Post-doc, Department of Political Sciences, Universiteit van Amsterdam, Present address: Science & Society, Biology Centre, University of Groningen, P.O. Box 14, 9750 AA Haren, The Netherlands, Email: n.f.wieringa@rug.nl

ISBN 90-901-1965-1