Gender differences in general practice consultations: methodological challenges in epidemiological research

Kirsti Malterud and Inge Okkes


**Background.** Women consult their GP more often than men do. The distribution of complaints and diagnoses are different for women and men patients. Although several findings on gender differentials on mortality and health care consumption are rather consistent across studies, detailed findings and subsequent conclusions diverge in several important fields.

**Objectives.** Our aim was to explore methodological aspects of research on gender differences in general practice.

**Methods.** We reviewed empirical studies within this field, aiming to identify methodological and interpretative intricacies which deserve special attention in epidemiological research on GP consultations.

**Results and Conclusions.** We found that descriptive and explanatory levels of research are frequently confused. Simple questions, answers and explanations are commonly raised for complex issues within a poorly defined theoretical explanatory framework. There is a need to assess relevant approaches for various purposes, and to develop more uniform conceptual terms. Findings from one level are often transferred to another, incompatible level. Epidemiological issues must be considered, especially matters related to denominator level and standardization/confounders—not in order to decide which level represents ‘reality’, but to clarify the consequences of different measures for different research questions. The contents of the core variables and the potentials for bias should be discussed in order to provide a sound basis for future explanatory studies.

**Keywords.** Gender differences, general practice, health care consumption, morbidity.

---

**Background**

Gender has a substantial impact on outcome as well as process levels in health care. Empirical research from hospital settings demonstrates that the gender of the patient makes a significant difference regarding strategies for investigation and treatment of disease.\(^1\) \(^4\) Primary care research has shown that medical communication is gendered, meaning that doctors’ symptom interpretation and conversation styles are different for women and men patients.\(^5\) \(^7\) Adequate health care for women as well as men requires knowledge about gender differences in health and health care for women and men.\(^8\) Knowledge about such matters can also contribute to further understanding of the mutual relationships between sociocultural mechanisms and health, with regard to gender.\(^9\)

General practice is the health care context most closely connected to the borderland between everyday illness and medical care. It is also the interface between health care and population which is most comprehensively utilized. GPs’ comprehension of the impact of gender in their practice will therefore bear important consequences for quality of care, for organization of services and for strategies which aim to provide evidence-based practice. Epidemiological differences between men and women patients as observed in general practice represent important sources of knowledge about health, disease and the adequacy of the health care system.\(^10\)
Previous and more recent consultation studies from developed countries confirm that women consult their GP more often than men do. More diagnoses are registered in women than in men, even when reproductive and preventive matters are excluded.11-17 The distribution of complaints and diagnoses are different for men and women patients,14,18 but seem to vary considerably across studies. The patterns of disease and use of health care seen by the GP can be explained in numerous ways. Macintyre and co-workers conclude that gender differences in health care vary by age, morbidity measure and social context.19 Before a causal hypothesis can be raised and tested, descriptive and updated studies are needed in order to provide an adequate map of this terrain. Such a map is also a worthwhile tool for illustrating and documenting aspects of everyday life which might otherwise be taken for granted. Maps allowing for careful analysis of the various aspects of gendered health differentials are still lacking.

Although several findings on mortality and health care consumption are rather consistent across studies, the conclusions are diverging in several important fields. Conceptual and methodological matters as well as the content of differences are not always clearly specified, and are often presented and explained in various ways by different authors. Comparisons across studies based on different denominators such as listed population, per visiting patient, per episode or per encounter may be misleading and perhaps explain the diversity of diagnostic profiles on gender. As repeatedly pointed out by previous and more recent authors, theoretical, conceptual and interpretative matters related to gender differences in health are only occasionally discussed.12,20,21 The unstable conclusions might therefore be due to methodological inconsistency.

Mechanic stated several decades ago that the literature within this field suffered from oversimplification and neglect of conceptual and methodological problems, and from ignorance of the rich available sources of data on the topic.20 He pointed to a widespread confusion among concepts belonging to different levels, such as experienced illness, reported illness, illness behaviour, health care utilization, symptom presentation, disease and disability. More recently, Kandrack and co-workers recapitulated these viewpoints by characterizing the literature in this field as atheoretical, and insensitive to social structural dimensions of gender relations.21

Measurement and interpretation of gender differences in health and health care are complex issues, which require considerations about underlying models and definitions as well as the variables to be selected and compared. Interpretation and explanations of differences are influenced by theoretical frames of reference as well as positions for perception, observation and interpretation. Such perspectives are often implicit and terms and explanations may be confused and biased. Several authors therefore suggest the term ‘gender difference’ (emphasizing social, psychological and cultural explanations) to be preferred rather than the term ‘sex differences’, which might sustain a narrow scope of biological perspectives.21-23 In the text below, the confused terminology will be reflected by us as authors preferring the term ‘gender’ (according to prevailing standard concepts), while authors to whom we refer will be quoted by the terminology used in their own reports.

There is a need to explore the issue of gender differences at the specific level of general practice encounters within a descriptive approach, scrutinizing the capability to transfer findings from other health care levels to this context. An important issue is to identify methodological and interpretative intricacies which deserve special reflection.

Below, we shall pose the question of gender to the epidemiology of general practice consultations, considered as part of a larger social and demographic context. Our approach is a critical review of epidemiological studies on gender and consultation patterns, supplemented by studies on potentially explanatory structural issues as gender, mortality and population morbidity.

Method

We shall address the following questions:

• What are the consistent and diverging conclusions concerning gender difference in health and health care across various studies?

• What is the relevance of these findings for general practice?

• Which methodological challenges can be identified, and what are the implications of these for future research which aims to provide adequate descriptions of and explanations for gender differences in general practice?

Our review includes studies on gender differences in health and health care from developed countries over the last 20 years, with the emphasis on consultation epidemiology from general practice, supplemented by studies intended to contribute to our understanding of the former. Several of the most extensive studies within this field have been conducted during the early part of this period.

Findings and conclusions

Men’s disadvantage in mortality

A consistent finding on gender and health across studies from developed countries is the longevity of women. For all ages and for all leading causes of death, the death
rate and risks of death are lower for women than for men.\textsuperscript{12,13,22,24-26} The mortality numbers are drawn from official vital statistics and can probably be assumed to be valid facts. The divergence of men’s versus women’s death rates had been steadily increasing, although this divergence seems to have slowed or even perhaps reversed during the last decade.

Reasons for men’s higher mortality rates have been extensively discussed according to models concerning biological and acquired risks. It is commonly concluded that the higher mortality of men of younger ages is largely due to higher accident rates, and of men of older ages, to heart disease.\textsuperscript{27} Verbrugge summarizes existing studies from Western countries, in which controlling for certain social risks has very little effect on the difference in mortality, which still remains unexplained.\textsuperscript{28} Koskinen and Martelin discuss why socioeconomic differences expected to explain differences in mortality rates are smaller in women than in men, considering the possibilities that (i) conventional socio-economic measures might be inadequate for women; (ii) confounding sociodemographic factors might mask women’s inequalities or accentuate them in men; or (iii) there could be a gender-different structure of causes of death.\textsuperscript{29} Mortality data regarding the cause of death are usually based on information from ICD-codes from death certificates. Nathanson\textsuperscript{12} examines potential sources of errors related to recording and classification procedures for assignment of causes of death that may affect the analysis of gender differences in mortality. Although this was written several years ago, the procedures and problems are still the same.

For the GP, statistics on mortality and gender difference in mortality are helpful to their understanding the demographic distribution of their practice population. Because women live longer, older women will outnumber men of the same age among the GP’s patients. Age may be considered as a confounder in morbidity statistics and must be controlled for. However, the factual consequence of the gender mortality differential is that the morbidity of old age constitutes a notable proportion and characteristic of women’s total morbidity and subsequently results in the large number of older women patients seen by the GP.

**Women’s excess reported morbidity**

Statistics on morbidity originate mainly from population-based surveys. The findings are consistent: women report more morbidity than their male counterparts, even when reproduction-related events are excluded.\textsuperscript{12,13,16,30,31} Women have excess morbidity and limited activity due to acute conditions (infective, respiratory, digestive and ‘other’ diseases), especially in reproductive age groups, and irrespective of whether there are children in the household or not.\textsuperscript{16} The gap between morbidity rates for men and women are especially high for genitourinary disorders and headache, and sizeable for unspecified viruses, influenza, bronchitis, upper gastrointestinal disorders and musculoskeletal problems.\textsuperscript{12,32} According to Cleary, the ranking of conditions in which women report a higher prevalence than men, varies somewhat, although gall bladder conditions, diseases of the urinary system, thyroid conditions and varicose veins seem to appear across studies.\textsuperscript{14} Macintyre and co-workers find an excess in total ‘malaise’ symptoms at all ages in women, but for ‘physical symptoms’ only in certain age groups.\textsuperscript{19} Only for injuries are rates lower in women than in men.\textsuperscript{28}

Adult women more often have a chronic condition than adult males, but males are more severely disabled by chronic conditions.\textsuperscript{14,16,30} Until about the age of 50, more men suffer from cardiovascular disease, after that age, the rates for men exceed those of men. In all adult age groups, women have higher prevalences of anaemia, diabetes, osteoarthritis and rheumatoid arthritis.\textsuperscript{16} Women more often report affective disorders and minor physical morbidity than men do.\textsuperscript{33} Macintyre and co-workers found that the direction and magnitude of gender differences in health vary according to the particular symptom or condition in question and according to the stage of life.\textsuperscript{39} Women’s excess of psychological distress is nevertheless consistently found across the life span.

Verbrugge suggests that gender health differentials originate from differential risks acquired from roles, stress, lifestyles and preventive health practices, while psychosocial factors (perception, evaluation and therapeutic actions resulting from symptoms), prior health care, biological risks, health reporting and carer effects have little importance.\textsuperscript{22} Popay et al. could not confirm hypothesized explanations regarding social positions and psychosomatic origins of minor illness in men and women, and recommend for future research the use of more gender-sensitive measures.\textsuperscript{33} However, conclusions have been conflicting,\textsuperscript{34} probably because of differences in theoretical approaches and interpretations.

Nathanson discussed 20 years ago potential bias when gender morbidity differentials are observed in health surveys during the process of data collection: (i) the definition of illness; (ii) who asks the questions; and (iii) who answers the questions.\textsuperscript{35} Verbrugge suggests that survey data on gender differentials in morbidity may be confounded during data collection because health behaviour is included in measures interpreted as physical status difference.\textsuperscript{32} Several authors warn against a commonly observed confusion among illness experience and health behaviour, which of course may be related measures although they represent different conceptual levels.\textsuperscript{12,32,21} Nathanson calls attention to the fact that international comparisons of morbidity statistics may be biased by a diversity of data sources, including not only household interviews but even
hospital in-patient data, representing a confounding with utilization of services. However, when morbidity data from interviews are compared with medical examination data, the agreement is nevertheless rather high.

When an individual perceives a symptom, the subsequent course of health behaviour is influenced by several matters. Pain, worry and disability may be experienced very differently by different individuals, and to a varying extent be reflected as consumption of health care services. The decision about seeking professional care represents an important marker for the limitations of self-care. The pattern of morbidity among men and women patients in general practice is therefore not necessarily identical to the pattern of morbidity perceived and reported in population surveys. Health care measures recorded on different levels of the system should not be interpreted as anything beyond what is actually measured, as for instance consultation numbers (describing people who call on health care) or diagnoses (describing the doctor’s medical conclusion). Differences observed between these two levels of morbidity may however indicate promising paths to a fuller understanding of symptom experience and behaviour in men and women, which may be pursued in analytical studies.

**Women utilize medical care more frequently**

Women of all ages use out-patient medical service more frequently than men do, expressed in survey studies as reported visits per person per year. The difference clearly remains when reproductive matters are excluded. Women have a shorter average interval between visits than men do. The largest gender differences in utilization of care are occurring in diagnostic categories of mild morbidity and conditions with symptoms of undiagnosed disease.

Women appear to consult more often than men for preventive purposes. However, findings are diverging regarding gender differences in preventive medical examinations, probably reflecting the fact that various measures have been studied.

Hospitalization data show a slightly different picture than those for out-patient care: between the ages of 15 and 64 years, women have a higher annual rate of discharge from hospital, while for ages 65+, men have higher rates. Findings are diverging regarding gender difference when hospitalization for childbirth has been excluded. Men tend to have longer hospital stays than women do.

Suggested explanations for the gender differential in utilization of medical care include real differences in health and morbidity, attitudes and help-seeking behaviour, accessibility of care, psychosocial distress, social role and related status characteristics, and marital status. So far, most studies confirm the correspondence between self-reported morbidity and utilization of care, although this still does not account for explanation of the difference.

These findings are assumed to be regained in patient statistics from general practice, although they do not provide unambiguous explanations about gender differentials. It should, however, be noted that numbers and denominators are not identical (and even not directly comparable) when the position of study is population based (annual visits per person per gender) as opposed to practice based (annual consultations per gender). An essential methodological matter is thus to state the denominator level. Another matter making comparison even more complex is the incongruence between numbers from list-patient practices and practices with no defined patient population.

**Men and women as patients in general practice**

As expected from the findings presented above, there is a consistent excess of consultations with women patients compared with men in general practice. Gender differences may be expressed by means of population-based measures (female: male ratio of annual visit rates per person), or from practice recordings as patient gender per visit (female: male ratio of consultations). Several studies have found an approximate female: male ratio of 2:1 for consultations. Bertakis found evidence that the variance can be explained by differences in morbidity and severity of disease in men and women. In a Dutch study of general practice, women were found to have an excess number of diagnoses compared with men (9.57 and 7.04, respectively per 5 years), and prior studies of the same data show a strong correlation between number of diagnoses and number of visits.

More complicated is the question of which health problems in women are most often seen by the GP. An answer to this question presupposes that gender-specific patterns of diagnoses in general practice can be summarized consistently across studies. However, different studies present results based on various measures, most often as prevalence tables where male and female morbidity are provided separately and in parallel, without comparative analysis or discussion. Such studies present gender differences by comparing the most prevalent complaints or diagnoses for female patients with those of male patients.

On her list of the 25 leading diagnoses after gender, Verbrugg found that the most prevalent diagnoses (medical exam, medical/surgical aftercare, hypertension, neurosis) are the same for men and women, but further down the lists the differences are emerging. For women, non-sick conditions account for a larger share of all visits (including prenatal care), two conditions rank much higher for women (obesity, and observation of a condition without need for further care), and two diagnoses appear on the women’s list only (cystitis and bronchitis). For men the following conditions were
found to rank higher and account for more visits than for women: heart disease, respiratory conditions and musculoskeletal problems. The following diagnoses appeared only on the men’s list: symptomatic heart disease, influenza, emphysema, sprains/strains of back and sprains/strains of sacroiliac region. Various diagnostic classifications and categories have been used for the purpose of research in general practice, and Verbrugge’s findings do not match those of similar studies. Sayer and Britt found that respiratory, skin and musculoskeletal problems were managed most often for men, while for women respiratory cardiovascular and musculoskeletal problems were managed more often.10

Van Wijk et al. studied the diagnostic groups of (i) sex specific conditions; (ii) symptoms without disease; (iii) prevention and diagnostics; and (iv) trauma.17 Gender difference was measured as the ratio of the mean number of female diagnoses to the mean number of male diagnoses. Trauma was the only diagnostic category with male excess. Reproduction and urogenital conditions showed extreme female excess, especially at fertile ages. A large female excess was also demonstrated for the prevention and diagnosis category, including contraceptive diagnoses. Female rates for symptoms without disease (physical symptoms, nervous functional complaints and general symptoms) strongly surpassed the male rates. However, for some subcategories the ratios were based on very small mean numbers of diagnoses. Van Wijk et al. conclude that the excess morbidity of women would be halved if gynaecological and obstetrical health problems were excluded. However, they remark “The aim for controlling of sex specific conditions in research on sex differences in health should not be to disregard them in order to obtain a ‘true’ sex difference in morbidity. Such a strategy would mainly ignore those categories of illness, researchers should be concerned to explore.”11 The excess morbidity in women for “symptoms of undiagnosed disease” (as well as “acute conditions involving micro-organisms”) has previously been demonstrated by Hibbard and Pope.36

**Crude rates or standardization?**

From an epidemiological point of view, standardization of consultation data on different levels may at first sight seem relevant. Age represents a major example. If the purpose of the study is explanatory, statistical strategies for controlling for the effect of age on diagnostic distribution in men and women by means of logistic regression may be most essential. We might find, for instance, that the excess rates of degenerative musculoskeletal disorders in women were due to age rather than to gender. However, the everyday reality for the GP is that these disorders most often are attributed to the individual women, simply because they survive and get old. Standardization procedures may thus disguise the descriptive and clinical realities, and do not eliminate the fact that general practice deals with a large quantity of conditions in women patients which may certainly be explained by age, occupation or other variables beyond gender.

Several studies present findings in which consultations for reproductive and other gender-specific conditions are excluded. Such an approach is adequate when the question to be answered is to identify the contribution of these matters regarding the excess health care utilization of women or certain biological differences between men and women patients. However, such approaches can also disguise the importance of these conditions as a part of the GP’s medical activity, and may also lead to neglect of the indirect influence of such consultations on other health problems.

Standardization on the denominator level is another example of potential conflicts between explanatory and descriptive levels. Health care consumption and morbidity patterns may be represented in terms of numbers of consultations, episodes of care, diagnoses or reasons for encounter. Gender differences may vary on different levels, for example if women typically have more consultations within one episode of care than men do. Differentiation between levels of denominator is necessary for the specification of the context of analysis. We assume that standardization according to denominator in some cases will eliminate gender differences initially recorded at other denominator levels, and probably in some cases reveal gender differences which are hidden by first sight.

We do not deny the advantage of standardization procedures. However, our point is that statistical manipulations simply produce different versions of reality—none of them necessarily representing ‘the truth’. If the objective of the study is to describe the activity in the GP’s office, crude rates may for many purposes be more adequate than standardized numbers. Our concern here is to demonstrate the methodological point that the relevance of standardization procedures depends on the purpose of the study, and that such matters need to be clearly specified. Different findings on different levels should be regarded as interesting explanatory clues.

**Comparing diagnostic patterns**

The most inconclusive findings regarding gender differences in general practice are observed when questions of diagnostic distribution are being considered. Definitions of gender difference are often omitted in presentation and discussion. Is, for example, a gender difference observed when the number of a specific diagnosis or diagnostic group (without regard to the matter of denominator) in women exceeds that of men—as for example in fibromyalgia? Or are women’s diseases the most commonly occurring diagnoses in women irrespective of their prevalence in men—as, for example, hypertension? Some authors discuss diagnostic
patterns recorded on the level of the individual diagnosis (as cystitis), while others refer to the major diagnostic groups (such as genitourinary disorders). We propose conceptual discussion and clarification of what is meant by a gender difference, and how it should be expressed and measured most adequately. An answer to such questions should precede analysis of the meaning and consequences of observed differences.

Reported differences in diagnostic patterns often seem to imply an understanding of the diagnostic process as an unambiguous matter of combining biological facts. However, the diagnostic process is a complex interaction between doctor and patient, where several sociocultural matters may be interfering as a natural part of the game. Verbrugge calls attention to the intriguing questions of gender bias in the diagnostic process, asking how the doctor perceives the patient’s complaint and then applies it to diagnosis.18 Doctors attribute different meanings to identical symptoms for presenting male and female patients.9,40 The communicative interaction in the consultation differs according to the gender of the patient.6,7,39 Such matters, which clearly influence the resulting diagnosis, are seldom discussed, even in studies where diagnosis is the basic variable.

Recommendations for future research

This review has demonstrated serious potential shortcomings in the literature on gender differences in health regarding theoretical, conceptual and methodological matters, which need to be clarified in future studies.

Descriptive and explanatory levels of research are commonly confused, when simple questions, answers and explanations are raised for complex issues within a poorly defined theoretical explanatory framework. There is a need to assess relevant approaches for various purposes, and to develop more uniform conceptual terms.

Findings from one level do not automatically allow for conclusions transferred to another level—although it often seems to be done in order to construct explanations about potential interrelationships.

Epidemiological issues must be considered, especially the roles of denominator level and standardization/confounders—not in order to declare which level represents reality, but to clarify the consequences of different measures for different research questions. The contents of the core variables and the potentials for bias should not be taken for granted, but be discussed and further explored.

In the Dutch Transition project, data on all encounters between 38 GPs in 22 practices, with an average of 40.796 listed patients have been coded with ICPC over a period of at least one full year. These data have been structured into a database with 108.704 episodes of care.41,42 In this database, all data concerning incidences and prevalences, and the related reasons for encounter and interventions, are retrievable per 1000 listed patients, per 1000 visiting patients, and separately, for all 14 standard sex/age groups. The availability of data sets like these may help us to gain more insight into the influence of gender, age and utilization on data, regarding incidences and prevalences.43 This will be the domain of further empirical research, where the presented methodological intricacies can be challenged.

References


