The importance of study design strategies in gender bias research: the case of respiratory disease management in primary care

Maria Teresa Ruiz-Cantero, Elena Ronda and Carlos Álvarez-Dardet

J. Epidemiol. Community Health 2007;61;11-16
doi:10.1136/jech.2007.060301

Updated information and services can be found at:
http://jech.bmj.com/cgi/content/full/61/Suppl_2/ii11

These include:

References
This article cites 28 articles, 8 of which can be accessed free at:
http://jech.bmj.com/cgi/content/full/61/Suppl_2/ii11#BIBL

1 online articles that cite this article can be accessed at:
http://jech.bmj.com/cgi/content/full/61/Suppl_2/ii11#otherarticles

Rapid responses
You can respond to this article at:
http://jech.bmj.com/cgi/eletter-submit/61/Suppl_2/ii11

Email alerting service
Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article

Notes

To order reprints of this article go to:
http://journals.bmj.com/cgi/reprintform

To subscribe to Journal of Epidemiology and Community Health go to:
http://journals.bmj.com/subscriptions/
RESEARCH REPORT

The importance of study design strategies in gender bias research: the case of respiratory disease management in primary care

Maria Teresa Ruiz-Cantero, Elena Ronda, Carlos Álvarez-Dardet

Background: Most research on gender bias has been carried out in hospitals and focuses mainly on severe diseases. Consequently, little is known about gender bias in relation to other illnesses and healthcare settings.

Aim: To explore the existence of gender bias in the management of patients seeking primary care for respiratory complaints.

Method: An observational, prospective blind follow-up study was performed in a primary care centre in Alicante, Spain. 830 patients were monitored from first visit to their general practitioner with a respiratory complaint until final diagnosis. Information was obtained about the diagnostic process (anamnesis, clinical examination and diagnostic tests) and therapeutic procedures (concession of unfit to work status and the patient’s destination following the visit). Logistic regression was used to compare the diagnostic/therapeutic procedures in men and women.

Results: Although men (318) and women (512) had similar respiratory complaints, after adjustment by age, marital status, employment, education, comorbidity and severity, men were more likely to be asked about smoking habits: RRa: 2.41 (95% CI: 1.57 to 3.70), auscultated: RRa: 1.30 (0.90 to 1.75), provided with a defined diagnosis: RRa: 1.77 (0.98 to 3.32) and considered unfit to work: RRa: 5.43 (1.64 to 9.96). Women were more likely to receive a pharyngotonsillar exploration: RRa: 0.63 (0.41 to 0.97).

Conclusions: Despite having the same respiratory symptoms, women were less likely to undergo diagnostic procedures in men and women.

Gender bias refers to a differential management of men and women, the impact of which may be positive, negative or neutral. Nevertheless, for research purposes, it could be redefined as an erroneous approach to the similarities or differences between men and women which may lead to incorrect or partial knowledge arising from past and current research, causing inequitable responses from health services and discrimination against either of the two sexes.

It was during the last decade of the 20th century that empirical evidence began to be gathered on the possible mistreatment of women by the medical profession as a result of gender bias. The emerging body of research concerning the appropriate management of women’s health problems by health services has provided information that is related mainly to the area of specialised care and to the application of hi-tech procedures—mostly related to cardiovascular problems and to the area of specialised care and to the application of hi-tech procedures—mostly related to cardiovascular problems and to primary care, its impact will be greater than in specialised care because of the extremely large number of patients attending at this type of healthcare centre.

In countries with publicly funded health services, the existence of gender bias poses some additional problems, as the national health services were created to reduce inequalities and to provide free services to all citizens regardless of their ethnicity, social level or gender. In these equity based systems, the discrimination of more than half of the population—women—has more explicit political implications.

Studies on gender bias in diagnostic and therapeutic procedures have been criticised because of their failure to take into account how the level of severity leads physicians to evaluate and manage the illness less aggressively in women than in men. The retrospective design of studies based on previously diagnosed patients as cases may also cause another potential methodological problem related to selection bias. If more women than men are underdiagnosed, as has been mentioned above, such women would not be included in the study. Another important issue is that gender bias research has frequently focused on those diseases that may be diagnosed through one single test—cardiac catheterisation. However, we believe that gender bias research should also take into account the previous anamnesis and physical examination, as well as the diagnostic test.

In an attempt to overcome these problems, a prospective study was designed whereby patients were classified by their disease diagnosis but rather by their initial complaints. The study was restricted to respiratory problems, as such conditions are important causes of death and reasons for seeking care at the primary level, and also the main diagnoses are gender dependent. The aim was to explore the existence of gender bias in the management of patients seeking primary care for respiratory complaints in Spain through its national health service.

Abbreviations: COPD, chronic obstructive pulmonary disease; GP, general practitioner; RR, relative risk
SUBJECTS AND METHODS
Design and data collection
A cohort of 830 patients at their first visit (318 men and 512 women) who attended a primary care centre in the city of Alicante, Spain, with respiratory complaints was prospectively monitored from the first time they visited their general practitioner (GP) with these complaints until the end of this consultation. The whole diagnostic and therapeutic process was followed up and registered.

According to literature published on the subject,24–27 and with the collaboration of a panel of experts, consisting of two GPs and one pneumologist, the most frequent reasons for seeking care with respiratory complaints were identified. These included eight symptoms: sore throat, coughing, expectoration, nasal congestion, dysphonia, breathlessness, sneezing and fever. Patients newly seeking care for at least one of these reasons were included in the study. Meanwhile, patients with other complaints or those collecting routine prescriptions or reasons were included in the study. Meanwhile, patients with other complaints or those collecting routine prescriptions or enrolled on specific preventive or therapeutic respiratory programmes were excluded.

The five GPs in the studied primary healthcare centre gave their informed consent to participate in a study about the natural history of respiratory diseases, while only the director of the centre was aware of the main gender bias hypothesis. This blind procedure was chosen after the research team had discussed the ethical aspects and in order to avoid changes in the outcome of the physician-patient encounter induced by the Hawthorne effect. Two specialised interviewers were present in the GP’s office during the patients’ visits, to collect the relevant data. These interviewers reviewed the list of patients daily and selected those who fulfilled the case criteria. These patients were then asked if they would allow the interviewers to be present during the consultation and if they were willing to participate in the study. Data were collected between 1 February 1994 and 31 June 1994 and 15 November 1995 and 15 February 1995.

A pilot study had been carried out previously in order to test the questionnaire, establish the location of the interviewers in the doctor’s surgery and to allow the interviewers to familiarise themselves with the running of the healthcare centre. Consensus was tested between the interviewers and an external observer and a 90% consistency was achieved. Finally, in order to obtain the necessary data to calculate sample size, the frequency of lung auscultation was used as an indicator of diagnostic procedure, as this is the minimum exploration process required according to the Joint Commission on Accreditation of Health Care Organizations.25 The estimated prevalence in women was 40% according to the information obtained in the pilot study, with a significance level of 5% for type I errors and 8% for type 2. In order to detect a risk increase of 20% (relative risk (RR) = 1.2) for women compared to men, a sample size of 826 patients (44% men; 56% women) was needed.

The research team and panel of experts designed a specific questionnaire based on conditioning factors inherent in diagnostic and therapeutic gender bias. The questions were related to the measurement of the diagnostic and therapeutic effort and were selected after guidelines and recommendations regarding good clinical practice in Primary healthcare settings had been reviewed.24–27

Variables
The patient’s sex was considered to be the main explicative variable, and the diagnostic and therapeutic procedures were the outcome variables.

| Table 1 Distribution of personal characteristics, respiratory complaints and diagnoses according to sex of the patients attended at a primary healthcare centre |
|------------------|------------------|------------------|------------------|------------------|------------------|
|                  | Men (n = 318)    | Women (n = 512)  | p Value          |
| No (%)           | No (%)           |                  |
| Patient symptoms* |
| Group 1          | 110 (34.6)       | 145 (28.3)       | 0.14             |
| Group 2          | 36 (11.3)        | 74 (14.5)        |                  |
| Group 3          | 172 (54.1)       | 293 (57.2)       |                  |
| Marital status   |                  |                  |                  |
| Single           | 138 (43.4)       | 160 (31.3)       | -0.01            |
| Married          | 171 (53.8)       | 288 (56.3)       |                  |
| Divorced/separated| 2 (0.6)         | 13 (2.5)         |                  |
| Widow/er         | 7 (2.2)          | 51 (10.0)        |                  |
| Occupational status|               |                  |                  |
| Employed         | 137 (43.1)       | 141 (27.5)       | -0.01            |
| Unemployed       | 181 (56.9)       | 371 (72.5)       |                  |
| Level of education|                |                  | 0.54             |
| None or few qualifications | 118 (37.1) | 233 (45.9) |                  |
| Primary studies  | 141 (45.6)       | 216 (42.1)       |                  |
| Secondary studies| 32 (10.1)        | 45 (8.7)         |                  |
| University degree| 7 (2.2)          | 18 (3.7)         |                  |
| Age: mean (SD)   | 38 (19.9)        | 40.9 (19.0)      | 0.04             |
| Patient’s diagnosis at the end of the visit† |
| High tract respiratory processes | 249 (78.3) | 387 (75.6) | 0.02             |
| Low tract respiratory processes | 34 (10.7) | 53 (10.4) |                  |
| Respiratory chronic diseases | 11 (3.5) | 10 (2.0) |                  |
| Undefined diagnosis | 15 (4.7) | 46 (9.0) |                  |
| Non respiratory diagnosis | 9 (2.8) | 16 (3.1) |                  |

*Group 1: patients with a temperature and any of the other symptoms: cough, sore throat, dysphonia, hypersecretion, breathlessness.
Group 2: patients with breathlessness and any of the other symptoms: cough, dysphonia, sore throat and/or hypersecretion.
Group 3: the remaining patients, with no temperature or breathlessness, but with any of the other symptoms.
†High tract respiratory processes: including codes 460, 461, 463, 464, 477, 487, 3820 Wonca classification; low tract respiratory processes: including codes 466, 486 Wonca classification; respiratory chronic diseases: including codes 462, 491-493, Wonca classification.
Diagnostic effort was measured through: anamnesis: (1) if the doctor asked the patients during the visit about the duration of their current symptoms (yes/no) and (2) about smoking (yes/no); clinical examination, including (3) pharyngotonsillar exploration (yes/no) and (4) pulmonary auscultation (yes/no); (5) diagnostic tests (yes/no) prescribed, including blood tests, radiology and complementary analyses.

The therapeutic effort was measured by: (6) the concession of unfit to work status (yes/no); (7) the patient destination following the visit, with four options: (i) sending the patient to hospital, (ii) assessment by a specialist, (iii) monitoring in primary care, or (iv) discharge.

As each participating patient could present any combination of the eight symptoms chosen, a method of classification and grouping was required for the analysis in order to avoid misclassification bias. The purpose was to classify respiratory symptoms into groups of a similar prognosis and expected healthcare procedures. After discussion with the panel of experts, three groups were defined:

- **Group 1**: patients who visit their GP with a temperature and any of the other symptoms—cough and/or sore throat and/or hypersecretion and/or sore throat and/or breathlessness.
- **Group 2**: patients who visit their doctor with breathlessness and any of the other symptoms—cough and/or sore throat and/or hypersecretion and/or sore throat.
- **Group 3**: the remaining patients, with no temperature or breathlessness, but with any of the other symptoms. All the patients enrolled in this study were then allocated to one of the three categories in a mutually exclusive way.

With the collaboration of a panel of experts, a similar procedure was used to classify the clinical diagnosis into groups with a similar clinical prognosis and thus the same expected diagnostic and therapeutic response. The panel divided all the factors related to the respiratory system in the WONCA Classification in Primary Care into four groups:

- **Group 1**: high tract respiratory processes
- **Group 2**: low tract respiratory processes
- **Group 3**: chronic respiratory diseases
- **Group 4**: undefined diagnosis.

An additional group 5 with non-respiratory diseases was added. The diagnoses were classified into one of the above categories.

The three co-variables used as indicators of severity were:

- (a) The severity of the disease perceived by the physician and by the patient him/herself, which was evaluated by asking them both about the patient’s self perceived severity on a Likert scale as low, medium and high.
- (b) co-morbidity—that is, if the patient was diabetic or had a history of previous respiratory disease, as these two diseases were considered by the panel of experts to be the main processes that could influence diagnostic or therapeutic efforts.
- (c) The duration of the symptoms as acute (less than 14 days) or chronic.

Finally, personal characteristics were also recorded: level of education, age, marital status and occupational status.

### Data analysis

First of all a descriptive analysis was carried out on the distribution by sex according to the previously mentioned variables. The $\chi^2$ and Student $t$ tests were used to discover the significance of differences in the proportions and means, after the parametric distribution had been checked.

To examine how sex affects the diagnostic and therapeutic effort, a simple analysis was first performed in order to calculate the crude relative risks with 95% confidence intervals.

The second step involved using non-conditional logistic regression models to calculate the magnitude of the association. In each case, all the explicative variables (age, education, marital status, occupational status, group of symptoms, diabetes, previous history of respiratory disease, severity evaluated by the doctor and by the patient him/herself), were taken into account in the regression models and were represented by dummy variables. All the necessary calculations were made using the SPSS statistical program.

### RESULTS

Table 1 shows the main personal characteristics of the 318 men and 512 women who visited their GP with respiratory symptoms. The non-response rate was 1.8%. Women were slightly older and more often unemployed than men. Table 1 also shows the distribution of men and women according to the symptoms which led to the visit to the GP. No differences between the sexes were detected. However, there were statistically significant differences in the distribution of men and women patients according to the final diagnosis (p = 0.02), mainly because of the higher frequency of undefined diagnosis in women (9.0%) than in men (4.7%).
Table 2 shows the distribution of the variables used as severity indicators. The presence of diabetes was statistically more frequent in women (9.2%) than in men (7.8%) (p = 0.03). Conversely, a previous history of respiratory disease was more frequent in men than in women (27.7% and 18.2%) (p = 0.01). There were differences in the doctor’s evaluation of the severity of the disease as GPs tend to evaluate complaints by men as more severe than those presented by women: 7.9% of male patients were classified as having a very severe process while only 2.9% of women were given such a diagnosis. Furthermore, 59.8% of women were included in the low severity category compared with 51.6% of men (p = 0.01). However, in patients’ self-evaluations, there were no differences according to sex.

Table 3 shows the results of the analysis regarding sex and diagnostic and therapeutic effort. The probability of whether a GP will ask about smoking habits is 63% higher for men than for women (<p<0.01). Pulmonary auscultation was slightly more frequently performed in men (50.6%) than in women (45.1%), and the likelihood of a pharyngotonsillar exploration was higher for women (88.5%) than for men (84.0%). Also in table 3, it can be observed that although both sexes visited their GP with the same complaints, doctors tended to classify women more often than men in the category of undefined diagnosis: RRA = 1.77 (95% CI 0.98 to 3.32).

As regards therapeutic effort, men were sent to a specialist or a hospital or to be monitored in primary care more frequently than women (RRA = 1.48 (95% CI 0.88 to 2.45) in table 3). Meanwhile, the concession of unfit to work status for those in paid employment (137 men and 141 women of those studied) was more frequent for men (16.7%) than for women (6.4%).

DISCUSSION

By observing the medical management of patients with respiratory complaints, we have confirmed the need to develop prospective designs in gender bias studies, as those of a retrospective nature based on previously diagnosed patients may have resulted in the exclusion of underdiagnosed women, solely the result of gender bias in the diagnostic procedures. In fact, the most interesting finding of this study is that despite there being no differences between men and women as regards symptoms, there were indeed differences as regards the final diagnosis, as doctors tended to classify women more often than men in the undefined diagnosis category at the end of the diagnostic process. Also, these findings provide new knowledge about possible gender bias in the anamnesis and physical examination, which support results from other studies on diagnostic tests—particularly those based on cardiovascular diseases dealt with in specialised care centres, which affirm that women are less likely than men to undergo cardiac catheterisation.4–10 It was also observed that women are less likely than men to receive “unfit to work” status when in paid employment.

Despite the fact that healthcare practice is known to be strongly influenced by factors related to the physicians themselves and to the organisation of the healthcare centre, there are clear guidelines and a strong general consensus between practice groups as regards respiratory disease management. However, similar research involving other and more centres should be considered as an implication of the study.

Even after efforts to maintain the blinded aspect of the study’s main objective in order to avoid the Hawthorne effect and its influence in differential misclassification bias, a certain degree of bias may have occurred because of the presence of the interviewer in the surgery. Moreover, our results could be considered as an underestimation of the true scale of sex differences caused by a gender bias, as they were obtained in a publicly financed primary care centre, where a lower degree of discrimination could be expected compared to the private sector.

The type of symptom(s) presented and the various combinations may affect diagnostic and prognostic values, as well as diagnosis and treatment within primary care settings.19 Therefore, to avoid this effect, and because of the lack of specific information about the diagnostic process obtained in retrospective studies, the present study developed a prospective design in order to search for sex differences in the diagnostic process and therapeutic interventions in patients with the same respiratory symptoms. By the end of the process, more men than women had obtained a respiratory diagnosis, which raises the question as to whether gender stereotypes may influence clinical decision making, and whether this is justifiable as regards clinical outcome.

It is possible that physicians may carry out more diagnostic procedures on men than women as they tend to believe that men are more likely to have a respiratory disease, as there is a higher incidence of previous history of such conditions in

Table 3: Gender bias in diagnostic and therapeutic effort in respiratory complaints: crude and adjusted * relative risks (RR) and 95% confidence intervals

<table>
<thead>
<tr>
<th></th>
<th>Men (n=318)</th>
<th>Women (n=512)</th>
<th>Crude RR(95% CI)</th>
<th>Adjusted RR(95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Anamnesis</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Smoking</td>
<td>64 (20.1)</td>
<td>47 (9.2)</td>
<td>1.63 (1.35 to 1.96)</td>
<td>2.41 (1.57 to 3.70)</td>
</tr>
<tr>
<td>Duration of symptoms</td>
<td>129 (40.6)</td>
<td>215 (42.0)</td>
<td>0.96 (0.80 to 1.15)</td>
<td>0.90 (0.60 to 1.20)</td>
</tr>
<tr>
<td><strong>Clinical examination</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pulmonary auscultation</td>
<td>167 (50.6)</td>
<td>231 (45.1)</td>
<td>1.16 (0.96 to 1.36)</td>
<td>1.30 (0.90 to 1.75)</td>
</tr>
<tr>
<td>Pharyngotonsillar exploration</td>
<td>267 (84.0)</td>
<td>453 (88.5)</td>
<td>0.79 (0.64 to 0.99)</td>
<td>0.63 (0.41 to 0.97)</td>
</tr>
<tr>
<td><strong>Diagnostic test</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Complementary tests ordered</td>
<td>6 (1.9)</td>
<td>12 (2.3)</td>
<td>0.78 (0.42 to 1.47)</td>
<td>0.80 (0.32 to 2.03)</td>
</tr>
<tr>
<td>Defined diagnosis given</td>
<td>303 (95)</td>
<td>476 (92)</td>
<td>1.14 (1.06 to 1.45)</td>
<td>1.77 (0.98 to 3.32)</td>
</tr>
<tr>
<td><strong>Therapeutic effort</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>If the patient is discharged</td>
<td>280 (88.1)</td>
<td>473 (92.4)</td>
<td>1.25 (0.98 to 1.55)</td>
<td>1.48 (0.88 to 2.45)</td>
</tr>
<tr>
<td>If the patient is unfit to work</td>
<td>23 (16.7)</td>
<td>9 (6.4)</td>
<td>2.31 (1.19 to 4.47)</td>
<td>5.43 (1.64 to 9.96)</td>
</tr>
</tbody>
</table>

*Adjusted by age, education, marital status, occupational status, group of symptoms, diabetes, previous history of respiratory disease, severity evaluated by the doctor and by the patient himself/herself.

†Including only patients in paid employment: 137 men and 141 women.
men, or because of the higher smoking prevalence in men than in women. Furthermore, although no differences were detected in the judgment of severity by patients themselves, physicians evaluated men’s symptoms as more severe than women’s.

It is easy to understand this logic, as men and women seemed to present different risk profiles. However, until the year 2000, the sex differences in pulmonary disease, such as chronic obstructive pulmonary disease (COPD) were an understudied subject. Furthermore, there is increasing evidence to support the theory that although more women are at risk of COPD, primary care physicians underdiagnose COPD in women. In addition, it should be considered that epidemiological data comparing the prevalence and severity of this type of disease by sex depend on accurate diagnoses being given by physicians, which are put in doubt in this study, alongside other evidence. Moreover, until the use of spirometry there was no single, clearly defined test for certain respiratory diseases and current diagnostic criteria are confusing and vague, with overlapping features between syndromes. Most diseases are not diagnosed through complementary tests, especially in individuals attending primary healthcare centres. Therefore, in order to achieve the aims set by gender bias research on sex differences in diagnostic and therapeutic processes, information must be obtained about the anamnesis and physical examination processes carried out by physicians.

In this study, as in others, it has been observed that despite the fact that more women than men with respiratory complaints seek care, the regional practitioner diagnoses register shows similar patterns in the prevalence of respiratory disease among sexes. This controversy could be explained by demonstrating that a higher prevalence of undefined diagnoses is observed in women than in men, possibly because physicians may be influenced by historical beliefs about sex differences in respiratory diseases, which may cause an underestimation of the risk of a particular respiratory disease in an individual woman patient. Consequently, physicians may tend to develop more diagnostic strategies in men than in women. Once again, in order to obtain this information, a prospective study design is necessary to observe the diagnostic process (anamnesis, physical examination and complementary tests).

Although sex differences in smoking rates have decreased in Spain during the last two decades, this study shows that men are more frequently asked about their smoking habits than women. A similar result was found in another study about variability in preventive activities in primary care in Spain. This phenomenon may be the result of inertia related to the historically higher rates of cigarette smoking among men that, in turn, produce this difference in gender biased anamnesis regarding smoking habits. It has also been observed that women were less likely to have received a spirometry test in the COPD diagnosis, which is highly relevant as this test reduces the risk of underdiagnosing COPD. Thus, a gender bias that may compromise the early diagnosis of COPD in women has been described. Once physicians were shown the abnormal results of the spirometry, a defined diagnosis was given and the gender bias was no longer apparent. Because of this biased approach, physicians believe that the risk of respiratory disease is truly higher in men than in women.

As regards sex differences in therapeutic strategies, in the case of patients in paid employment, GPs more frequently prescribe men the concession of unfit to work status than women. This could be related to the sex differences in the type of occupation or because women in Spain have more unstable occupational positions than men. However, gender bias also plays an important part in the diagnostic process, as does the GP’s differing perception of the severity of the disease according to sex.

In summary, in some cases, results from this study may not discriminate between gender bias and other explanations for some of the differences observed in the diagnosis and treatment of respiratory problems in men and women. However, in order to be able to deal with an appropriate management of certain diseases in women, such as respiratory diseases, future research should concentrate on the lack of evidence regarding sex/gender differences in the natural history of diseases. Furthermore, a possible gender biased diagnostic suspicion could contribute to an incorrect estimation of prevalence, which could, in turn, lead to an unequal management of one sex related to the other.

**ACKNOWLEDGEMENTS**

This study was partially funded by a grant from the Fondo de Investigaciones Sanitarias (95/0688) (Health Research Fund) and the Instituto Valenciano de Estudios e Investigación (Valencian Institute for Study and Research).

**Authors’ affiliations**

Maria Teresa Ruiz-Cantero, Elena Ronda, Carlos Alvarez-Dardet, Preventive Medicine and Public Health Area, University of Alicante, Spain

**Conflict of interest:** none.

**REFERENCES**