

# Decision-making in general practice

- the effect of  
financial incentives  
on the use of  
laboratory analyses

**Siri Fauli Munkerud**

*The Norwegian Medical Association,  
Norwegian Quality Improvement of  
Laboratory Services in Primary  
Care (NOKLUS) & HERO*

**UNIVERSITY  
OF OSLO**

HEALTH ECONOMICS  
RESEARCH PROGRAMME

Working paper 2008: 13

**HERO**

# Decision-making in general practice

## The effect of financial incentives on the use of laboratory analyses

Siri Fauli Munkerud

*2. December 2008*

**Health Economics Research Programme at the University of Oslo  
HERO 2008**

**Keyword:** Financial incentives, laboratory analyses, diagnostic uncertainty, medical practice.

### **Acknowledgements:**

This work was funded by the Norwegian Medical Association's Quality Improvement Fund III, which was established by the Norwegian Government, the Norwegian Association of Local and Regional Authorities, and the Norwegian Medical Association (NMA). The author is grateful to Tor Iversen, the Health Economic Research Programme at the University of Oslo (HERO), and John Dagsvik, Statistics Norway for valuable guidance. Thanks also to Burkhard Hehenkamp at Universitaet Dortmund, Knut Wangen, Geir Godager and Sverre Grepperud at HERO and, finally, to Geir Thue at NOKLUS for helpful comments and suggestions regarding a previous version of this paper. The author has no conflicts of interest.

### **Author:**

Siri Fauli Munkerud, Norwegian Quality Improvement of Laboratory Services in Primary Care (NOKLUS), P.O.Box 1152 Sentrum, N-0107 Oslo, Norway, 004723109115  
fax 004723109100, [siri.fauli.munkerud@legeforeningen.no](mailto:siri.fauli.munkerud@legeforeningen.no)

## **Abstract**

This paper examines the reaction of general practitioners (GPs) to a reform in 2004 in the remuneration system for using laboratory services in general practice. Data from Norway make it possible to distinguish between income motivation and service motivation. The purpose of this paper is to study whether income motivation exists, and if so, the degree of income motivation among general practitioners (GPs) in Norway regarding the use of laboratory services in general practices. We argue that the degree of income motivation among GPs is stronger when the physicians are uncertain about the utility of the laboratory service in question.

We have panel data from actual physician-patient encounters in general practices in the years 2001-2004, and use discrete choice analysis and random effects models.

Our results indicate that there may be an income motivation among GPs regarding using laboratory services as, after the reform, the GPs chose to use laboratory services less frequently where the fees had been most reduced. In addition, estimation results show that an increase in the fees will lead to a small but significant increase in use. The reform led to minor changes in the use of laboratory analyses in GPs' offices, and we argue that financial incentives were diluted because they were in conflict with medical recommendations and existing medical practice. The patient's age has the most influence on GPs' choice regarding use of laboratory services. The results support the hypothesis that the impact of income increases with increasing uncertainty about diagnosis and treatment.

The policy implication of our results is that financial incentives alone are not an effective tool for influencing the use of laboratory services in GPs' offices.

## 1. Introduction

Economic theory is largely based on the hypothesis of self-interest. In health care, the expected behaviour of physicians is different. Arrow (1963, 1986) and others argue that physicians have different motivation than other business people, and that they are concerned for their patient's health and this influences their choices. Do physicians give advice that is completely divorced from self-interest? Hillmann (1990) suggests that most physicians will act in the patient's best interest when the medical decision is clear-cut, and that the effect of financial incentives may be more important when the correct decision is not evident. Thus, income motivation is more likely to exist when the correct decision is not obvious or where there is a lack of medical recommendations.

There is a lack of studies on the effects of remuneration on GP behaviour (Scott and Shiell, 1999) and in particular on GPs' behaviour regarding their use of diagnostic tests. The purpose of this paper is to study whether income motivation exists regarding the use of laboratory services in general practice, and if so, the degree of income motivation among general practitioners (GPs) in Norway. If a change in laboratory-related fees changes the use of laboratory services, we call it income-motivated behaviour. In this paper, we argue that the degree of income motivation is stronger when physicians are uncertain of the benefit of the medical action in question in regard to the health status of the patient.

Usually, financial incentives are used to reduce health care resources, change clinical practice or to improve the quality of health care. In these circumstances it can be difficult to distinguish the effects of financial incentives from a wanted change without financial incentives. We study an example from Norway where a reform took place in July 2004 in the remuneration system for the use of laboratory services in general practice. The reform was carried out for political reasons, and intended to move the out-of-pocket expenses from a service utilised by the patient out of office (telephone contacts) to in-office services (Norwegian Medical Association, 2004). This reform led to financial incentives that were *in conflict* with medical recommendations and existing medical practice. As the remuneration system was the only change and the intention was not to change the GP's practice profile this gives us a unique opportunity to study the effect of the financial incentives.

Laboratory analyses are widely used in diagnostic work-up and monitoring of patients in primary care. Interest in using laboratory analyses in GPs' offices has been increasing in Europe

(DeNeef, 1986; Fischer, 1991; Hjortdahl, 1990). In Norway (Source: NOKLUS<sup>1</sup>), Switzerland and the Netherlands almost all practices have their own laboratory facilities (Leurquin et al. 1995). In Norway (National Insurance Administration, 2007) about 24.3% of all fees in general practices in 2006 were from the use of laboratory services, and the fee for taking a blood test was the third most used fee after consultation fees.

Factors that contribute to the extensive use of laboratory analyses in primary care are: patients' increasing intolerance of late diagnosis, sicker patients in the community due to earlier discharge from hospital and the trend that the GPs should manage the surveillance of chronic diseases (Hobbs, 1996).

In spite of the extensive use of laboratory analyses, there is a lack of studies on the impact of laboratory results and the economic consequences of near patient laboratory results in primary care (Delaney et al. 1999). Better knowledge of the GPs' routines for use of laboratory services when there is diagnostic uncertainty may lead to interventions directed at reducing unnecessary testing (Van der Weijden et al. 2002). It is important to avoid unnecessary testing because false-positive outcomes may increase the risk of somatisation by labelling healthy persons as unhealthy and also by raising the cost of health care.

The availability of laboratory analyses in the office varies across general practices. The advantage of having the test available in the office is that the GP can get the results of the test immediately and use the results during the consultation. An immediate result from a near patient test for C reactive protein (CRP) might enable a doctor to avoid unnecessary prescription of an antibiotic by distinguishing between viral and bacterial infection (Hobbs et al. 1995). In contrast, if the GP sends a blood sample for analysing outside the office, it takes 1-7 days to get the results, and this may delay the treatment and usually demands more follow-up by the GP.

The reform in the remuneration system for laboratory services consisted of an increase in the out-of-pocket payment and a decrease in the fees for the most used laboratory analyses in general practice in order to retain the total revenue from laboratory services inside general practices. Before the reform, patients paid a fee only when a laboratory test was analysed outside the office. After the reform, an out-of-pocket fee is paid every time a blood test is taken. There are recommendations (Hunskår, 2003) listing which laboratory analyses should be performed in the GP's office (here called the basic repertoire) and which laboratory analyses should be used, depending on the patient's history and symptoms. The fees for laboratory analyses in the basic

---

<sup>1</sup> NOKLUS (The Norwegian Quality Improvement of Laboratory Services in Primary Care) is an organization to ensure that laboratory tests performed outside hospitals are ordered, carried out, and interpreted correctly and in accordance with patients' needs for investigation, treatment, and follow-up.

repertoire were subject to the highest reduction. Before the reform, the fees exceeded the marginal costs, and there was a clear economic incentive to do more laboratory analyses and upgrade the laboratories in general practices. After the reform, the fee is lower than the marginal costs for the test equipment for the most-used basic analyses. The main incentive after the reform is to stop performing basic analyses in-office and to perform more laboratory analyses out of the office.

We examine the effect of the remuneration, the age of the patient and the characteristics of the GP regarding use of laboratory services in general practice. Before a GP decides whether or not to take a blood test, he must first consider which laboratory test to take and secondly where to perform it. Thus, the decisions we study are (ref. Fig. 1).

A. Whether or not to take a blood test.

B. Given that a blood sample is taken, whether to perform one or several of the laboratory analyses in the GP's office or whether to perform all the laboratory analyses out of the office.

C. Given that the laboratory analyses are performed in the office, whether these are only basic analyses or only non-basic analyses. The alternative "both basic and non-basic" was excluded because we particularly wanted to study how the reform influenced the use of laboratory analyses with the most reduction in fees (basic analyses) versus laboratory analyses where there was the least reduction in fees (non-basic analyses).

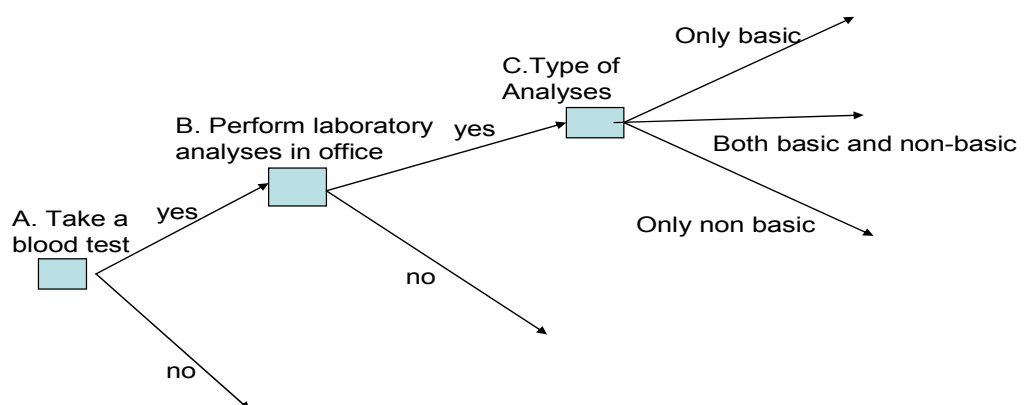


Fig. 1. Overview of the decisions studied

According to Wennberg (1985), uncertainty has a major impact on physician behaviour. Uncertainty is involved in making a diagnosis, in predicting the effects of treatment of a given condition, and in perceiving patient preferences.

We assume that the impact of income depends on the GP's uncertainty about the diagnosis and appropriate treatment of the patient. The degree of uncertainty depends on the patient's age and on the knowledge of the GP regarding the use of laboratory analyses. We investigate this hypothesis to some extent in this paper. We are not aware of any other study that addresses this issue.

We have panel data from actual physician-patient encounters in general practices in the years 2001-2004, and use discrete choice analysis and random effects models.

Our results indicate that income motivation among GPs regarding the use of laboratory services exists, because, since the reform, GPs have slightly reduced the use of laboratory services where the fees have been reduced. The reform led to minor changes in the use of laboratory analyses in the GPs' offices and we suggest that financial incentives were diluted because they were in conflict with medical recommendations and existing medical practice. Patient's age has the greatest influence on the GPs' choice regarding the use of laboratory services. Fees have a relatively small influence on the GPs' choice, and the impact of fees is reduced when we reduce the uncertainty about diagnosis and treatment. Thus, the degree of income motivation among the GPs is stronger when the physicians are uncertain about the utility of the medical action in question. The policy implication of our results is that financial incentives alone are not an effective tool for influencing the use of laboratory services in GPs' offices.

We are not aware of any study that has focused on the consequences of financial incentives when they are in conflict with medical recommendations.

This paper gives insight into how GPs respond to financial incentives when the financial incentives are in conflict with medical practice and recommendations. It also discusses what other factors influence the GPs' use of laboratory services in the office and in particular the importance of medical uncertainty. Hence, our results have interest beyond the setting used in this article.

## **2. Literature related to our study approach**

Physicians order laboratory analyses primarily to establish or exclude a diagnosis, to indicate a prognosis, to select the most appropriate therapy, and to monitor therapy. An alternative to taking a blood test is to "wait and see" or to use other diagnostic utilities such as, for example, upper endoscopy or a breath test for the *Helicobacter pylori* test. Some GPs use laboratory analyses in their offices to save time and to improve patient satisfaction (Higgins, 2000). Guthrie (2001) found

that the most important reasons for taking blood tests were monitoring, looking for a specific diagnosis, general screening/excluding anything serious, reassurance of the patient or doctor, and uncertainty. Van der Weijden et al. (2002) studied which factors influenced GPs' decisions for ordering laboratory analyses when they are confronted with patients with unexplained complaints. They found numerous motives such as personnel routines, tolerance of diagnostic uncertainty, time pressure and tactical motives for test ordering.

There is a lack of studies on the effects of remuneration on GP behaviour (Scott and Shiell, 1999) and in particular on GPs' behaviour regarding their use of diagnostic tests. It has been argued by Mooney and Ryan (1993) that ethical values may dilute the influence of economic incentives and in health care there might be relatively less need for strong external incentives. Frey (1997) states that use of extrinsic (or financial) incentives may crowd out intrinsic motivation.

Previous studies have examined the effects of different remuneration systems, and it is well-known that the fee-for-service system results in a higher quantity of service than the salary payment system, and also a larger number of laboratory analyses (Gosden et al. 1999). Chaix-Couturier et al. (2000) concluded that financial incentives are effective with practice guidelines. Van Walraven et al. (1998) found that financial incentives were effective with practice guidelines and with modification of the requisition form.

Carlsen et al. (2003) studied how the number of laboratory analyses taken by a sample of 44 GPs in 1991-1994 in Norway was influenced by the GPs' private economy outside the practice and by changes in remuneration. They found that changes in fees had little or no effect on four types of laboratory analyses (including one basic analysis: CRP). Grytten et al. (2008) examined how the service production (the relative change in the number of laboratory tests per consultation, consultations per physician per month and the proportion of consultations lasting more than 20 min per physician) of 2 650 primary physicians in Norway was influenced by changes in the fees in the years 1995-2000. In this period the mean increase in fees for laboratory analyses was 24.3%. They found that the fee changes had no effect on the service production.

There is a huge literature on risk and uncertainty (see web link to "new school" and to Wikipedia). In this paper we deal with uncertainty as according to Knight's (1921) definition: The GP does not assign mathematical probabilities to the randomness he faces because he does not know the probabilities. A possible explanation is that each meeting with a patient in the GP's office is usually unique and complex, consisting of many factors influencing the medical decision. It is a natural variation in the way people respond to a medical procedure and many of the diseases are rare. The cases are few in number so there is no scientific basis for any calculable probability.



### 3. Institutional set-up: primary physicians in Norway

In Norway, the municipalities are responsible for planning, organising and running primary health care services. The payment system in Norway for general practitioners (GPs) is a mixture of capitation fee and fee-for-service. General practice financing is split between the State, the municipalities and the patient. Each GP has a list of patients to take care of, and the municipalities pay a per capita fee. The State pays a fee-for-service component and the patient pays an out-of-pocket expense according to a fixed schedule (“normal tariff”<sup>2</sup>) negotiated between the Ministry of Government Administration and the Norwegian Medical Association (NMA). Fees in the normal tariff are valid from July 1st to June 30th of the following year. Changes in fees are therefore exogenous to individual physicians. The patient’s out-of-pocket expenses have an annual ceiling; if the ceiling is reached, the fee is reimbursed by the National Insurance Administration (NIA).

Before the reform, patients paid a fee (2001/2002 and 2002/2003: €4.4, 2003/2004: €5 (NMA, 2001, 2002, 2003) only when a laboratory test was analysed out of office. Since the reform, a fee (€8.9) is paid every time a blood test is taken in-office (NMA, 2004). Thus, the cost for the patient after the reform is independent of where the blood test is analysed. This fee is paid only once, even if the GP chooses to perform laboratory analyses both in and out of office. In addition, the fees for laboratory analyses performed in-office have been reduced.

Table 1 shows the fees and costs of the laboratory analyses where the fees had been most reduced by the reform in 2004. The table includes all the analyses in the basic repertoire. The costs of the laboratory equipment and the fees were nearly constant between 2001 and 2004 and we therefore include only the year before the reform in Table 1. The fee for taking a blood test is not included in the table. The sum of income from laboratory services per encounter is €8.9 after the reform if all laboratory analyses are performed out of office. If any laboratory analyses are performed only in-office, the total income is €8.9 plus the fee for the laboratory analyses chosen.

The cost of using laboratory services in-office includes costs for laboratory equipment, personnel, and storage. The cost columns include only the cost of laboratory equipment (average cost of strip and laboratory instrument). Most basic analyses are simple tests for single use, and the only equipment cost is the test itself. For CRP and Glucose, a laboratory instrument is needed in addition to a strip that is used once. Table 1 shows that for CRP the strip costs €3.2 and the instrument costs €900. Personnel take blood tests and perform laboratory analyses. The GPs’ cost for laboratory analyses performed outside the office includes personnel packing the blood samples

---

<sup>2</sup>A list of fees for consultations, house calls, procedures etc. (NMA, 2001, 2002, 2003, 2004)

and following up the laboratory results (telephone calls, return visits). The cost of transport to and from the external laboratory is zero because external laboratories covered these costs during this period.

*Table 1. Overview of the laboratory analyses where the fees had been most reduced by the reform*

Type of laboratory analyses	2003/04	2004/05	Change	Cost of lab. equipment	
	Fee - €	Fee - €	%	Strip - €	Instrument- €
<b>Fecal occult blood (FOB)</b>	<b>2.4</b>	<b>1.2</b>	<b>-50</b>	<b>2.5</b>	<b>0</b>
<b>FOB – Home</b>	<b>4.4</b>	<b>1.9</b>	<b>-57</b>	<b>3.8</b>	<b>0</b>
<b>CRP</b>	<b>6.0</b>	<b>2.5</b>	<b>-58</b>	<b>3.2</b>	<b>900</b>
<b>Pregnancy test</b>	<b>6.0</b>	<b>3.3</b>	<b>-46</b>	<b>3.2</b>	<b>0</b>
Microalbumin uri	6.0	2.5	-58	6	723
<b>Group A streptococcus (Strep A)</b>	<b>6.0</b>	<b>2.6</b>	<b>-57</b>	<b>4.0</b>	<b>0</b>
<b>Mononucleosis</b>	<b>6.0</b>	<b>3.5</b>	<b>-42</b>	<b>5.8</b>	<b>0</b>
<b>Glucose</b>	<b>3.8</b>	<b>0.8</b>	<b>-79</b>	<b>1.1</b>	<b>585</b>
Prothrombin Time and International Normalized Ratio (PT-INR)	8.8	5.9	-33	3.0	1737
Helicobacter Pylori	11.2	7.5	-33	10.0	0

Analyses in bold type are analyses in the basic repertoire.

After the reform, the marginal net revenue for performing laboratory analyses in-office for most basic analyses was negative. Including the fee for taking the blood test (€8.9), the total revenue from laboratory services could be higher than before the reform, particularly if no or only one basic analysis was performed.

#### **4. Hypotheses**

In this paper, we study the effect of the patient's age, the characteristics of the GP, and income on the three medical decisions A, B and C.

As a result of the existence of medical recommendations about the use of laboratory analyses (Hunskår, 2003), we expect that an increase in laboratory fees will not affect or only slightly increase the probability of taking a blood test, performing laboratory analyses in-office, and performing only basic analyses.

Regarding *patient's age*, the benefit of a laboratory analysis is higher with older patients because they have a higher risk of disease (Van der Weijden et al. 2002). We thus expect the use of blood tests to be higher among the older than the younger age groups. Infections and the use of basic analyses (CRP, Strep A) dominate in the encounters with the youngest. Among the oldest patients, more complex issues (e.g. atrial fibrillation) dominate (non-basic analysis: PT-INR). We therefore expect the estimated signs of the effects to be negative for older patients compared to younger patients regarding decisions B and C.

The repertoire of available laboratory analyses in-office varies across general practices. *Group practices* have more patients than solo practices, and the average cost of having laboratory equipment will decrease with the number of patients. Thue and Sandberg (1994) have shown that group practices have a wider spectrum of laboratory analyses, and Fauli and Thue (2005) found that having a test available in a practice increases its use. Other studies show inconsistent results: Vinker et al. (2007) found that GPs in a larger practice ordered more laboratory analyses per patients than physicians in smaller practice. Verstappen et al. (2004) concluded with the opposite result regarding group practices with more than two GPs, and the authors assume that this is probably a result of general discussions and reflections on practice behaviour in these group practices. We expect that GPs in group practices will take laboratory analyses more often than GPs in solo practices and, as their repertoire is wider, they will perform analyses in-office more often and have fewer encounters with only basic analyses versus only non-basic analyses. Thus we expect the estimated signs of the effect to be positive for decisions A and B, and negative for decision C.

General practices in urban *areas* face competition for patients. One way of competing is to give quicker service (Iversen, 2005). The benefit of analysing in-office is patient convenience: patients do not need to endure delays until the results are available. On the other hand, for GPs in less central areas, it may take longer to send blood tests by mail. Thus, GPs in both central and less central areas will benefit from analysing in-office, and we therefore do not have any hypothesis regarding the effect of this variable.

Several studies (Hartley et al. 1984; Franks et al. 2000; Salloum and Franssen, 1993; Verstappen et al. 2004) have found that older GPs use fewer laboratory analyses than younger GPs. This may be because more experienced (*older*) GPs or GPs with additional education (a *speciality*) know which laboratory analyses to use, and therefore use fewer laboratory analyses than younger GPs or GPs without a speciality. In addition, younger GPs were trained in an environment more oriented to laboratory investigations than their more senior colleagues, and more laboratory analyses were available to them during their training (Salloum and Franssen,

1993). Thus, we expect that the probability of taking a blood test is lower for more experienced GPs or for GPs with additional education than for other GPs.

The GP's uncertainty about the utility of the laboratory service depends on the GP's knowledge and experience with the use of laboratory services and the type of patient. According to Hillman (1990) income is expected to have less influence when the GP has more knowledge about laboratory services than when the GP has less knowledge. To check this, we introduced *interaction terms* between the income variable and variables related to knowledge and experience (GP's age, speciality, and type of practice). Like Verstappen (2004), we expect that GPs in group practices have collegial discussions and are less uncertain of the expected health outcome of the use of the laboratory service in question than less experienced GPs or GPs in solo practices. Thus, we expect that the estimated signs for the interaction terms between income and the two oldest GP age groups, speciality and group practices, to be negative.

Uncertainty about diagnosis and treatment depends on patient age and on the laboratory service in question. As the benefit of a blood test is higher in older than in younger patients, we expect the uncertainty about the utility of using laboratory services to be lower for older age groups regarding decision A. Therefore we expect the estimated signs on the interaction term between the income variable and older vs. younger patient to be negative. In the youngest patients, basic analyses are the most often used laboratory analyses, in the oldest patients the GPs are often more uncertain about diagnoses when basic analyses are used compared to when they use non-basic analyses. Non-basic analyses are often used to monitor older patients and in these situations the GP is certain of the diagnosis. Thus, we expect the estimated sign of the effects of the interaction terms for decisions B and C to be positive for older patients compared with the youngest patients.

## **5. Data and descriptions**

We used actual physician-patient data from encounters in general practices obtained from the Norwegian Social Science Data Services (NSD). 10% of the GPs are municipal employees on a fixed salary. Our data include only GPs with a private practice and a contract with the municipality. Our data is from an unbalanced panel of 1668 GPs from the month of November for three years (from 2002 to 2004) with 1 557 498 encounters, and from 3849 GPs from the month of October for four years (from 2001 to 2004) with 1 399 626 encounters. Table 2 gives an overview of the number of GPs and number of encounters per year in our datasets per decision.

Data from November include physician-patient encounters with and without blood tests. We used these data to study decision A, whether or not a blood test was taken. Data from October only

include encounters where a blood test had been taken. To study the decisions B and C, which are both based on a blood test having been taken, we used data from October.

As data from November included all type of encounters, there were more encounters per GP during November.

Our data is from all electronic bills to NIA from GPs during November and October. These bills are sent to different local departments of NIA in Norway. Since a growing number of GPs are sending bills electronically and there is an increased storage capacity at NIA, the number of encounters are increasing every year, NSD data showed 3 times more encounters in the last year (2004). Because we wanted to study the change in practice after the reform, NSD only included encounters from the last year from GPs who had also been included in the two previous years. We then had 1 721 927 encounters. Due to computer capacity, the maximum limit was 1.6 million encounters, and NSD therefore excluded encounters from the three first days of the month, thus reducing the number of observations in the dataset from 1 721 927 to 1 557 498.

*Table 2. Overview of number of GPs and number of encounters (N) per tariff year for decisions A, B and C*

Decisions\ Dataset	2001/2002		2002/2003		2003/2004		2004/2005		GPs in data
	GPs	N	GPs	N	GPs	N	GPs	N	
A. Blood test\ November			1019	415246	1214	481068	1381	647233	1668
B. In-office\ October	2698	303750	3076	348340	3114	366941	3031	371684	3849
C. Basic only\ October	2681	153946	3056	176272	3109	186524	3016	180119	3841
Population of GPs in Norway	3635		3704		3711		3747		

Because we have data from 2 to 3 years before the reform, we can detect trends in the use of laboratory services before the reform (ref. Fig. 2). We excluded groups with encounters with deviant characteristics—GPs over 67 years of age and patients younger than 0 years and older than 100 years—thus reducing our November dataset to 1 543 547 and our October dataset to 1 390 715 encounters. At decision C, the population is reduced from 1 390 715 to 696 861 encounters. Here we only studied patient-physician encounters where laboratory analyses were analysed in-office, and excluded 506 760 encounters where laboratory analyses were only performed outside the office. We grouped the encounters where laboratory analyses were performed in-office into three groups: only basic analyses used, only non-basic analyses used, and both types of laboratory

analyses used. After we excluded the group of encounters with both basic and non-basic analyses (n= 187 094) we had a dataset with 696 861 encounters.

### Dependent variables

Our datasets only provide information about laboratory analyses that are reimbursed and performed in-office.

Fig. 2 shows descriptive statistics of the dependent variables. The dependent variables are coded as follows:

- 1 if the GP used a blood test and 0 otherwise
- 1 if the GP performed one or several laboratory analyses in the office and 0 if the analyses were only performed out of office
- 1 if the GP only performed basic analyses in-office, and 0 if only non-basic analyses were performed in the office

After the reform it was not possible in our data to differentiate between encounters where the GP chose to analyse laboratory tests both in and out of office from those where he/she chose only to perform the analyses in office. Therefore the first alternative at decision B includes all encounters where the laboratory analyses are performed in office regardless of whether or not laboratory analyses were also performed out of office.

A blood test was taken in 27.2% of the encounters from November. In 63.6% of the encounters from October, one or several laboratory analyses were done in GPs' offices and in 64% of these encounters only basic analyses were performed.

Fig. 2 shows the development in the use of laboratory services in general practice. In particular, we are interested in changes in medical practice as a result of last year's reform. In regard to decision A, there has been an increase in the share of encounters where a blood test is taken (from about 25 % in 2002/2003 to about 28 % in 2004/2005). With regard to decision B, there has been an increase in the percentage of encounters with blood tests where laboratory analyses are only done outside the office (from around 35 % to almost 40% after the reform). The proportion in which only basic analyses (decision C) were performed in office has decreased from about 33% to about 30%. The figure shows that the decrease in the percentage of encounters with blood tests where the GP performs laboratory analyses in-office (from 65% to 60%) is due to a decrease in the share of encounters where only basic analyses were performed, and a decrease in the share of encounters with both types of analysis. The figure also shows that, after the reform, the GPs took more blood tests. However, since there was also an increase in the share from

2002/2003 to 2003/2004, the increase in the share for the last year may be because of a trend towards more frequent use of laboratory services (Hobbs, 1996).

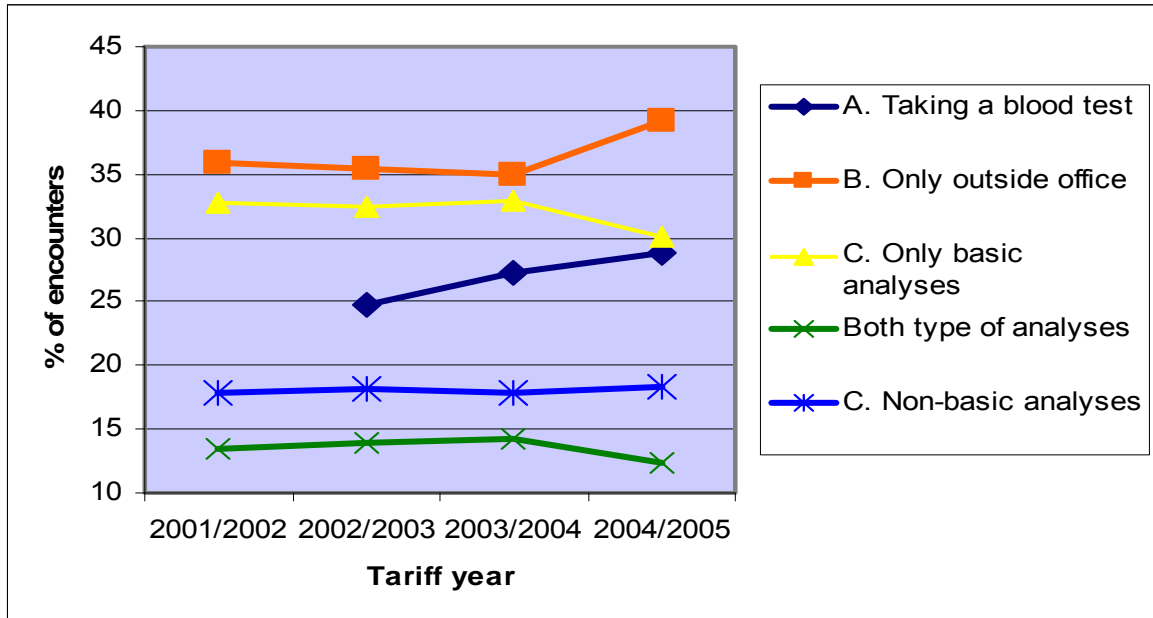


Fig. 2. Development in the use of laboratory services in general practice

### Explanatory variables

The explanatory variables are listed in Tables 4 and 5.

We face a simultaneity problem because income from fees depends directly on the laboratory services chosen and is endogenous. In addition we do not have information about laboratory analyses that are performed out of the office. To deal with these problems, we use price indices and construct exogenous proxies for the income variables. Based on the fees in the normal tariff (NMA, 2001, 2002, 2003, 2004) and information from the NIA about quantities per fee per tariff year, we created Fisher ideal indices, Fisher (1922)<sup>3</sup>, for the relevant fees for all alternatives regarding decisions A through C. As the index is based on aggregate quantities of real encounters, we assume that the Fisher index is representative of the average GP and that the GP is aware of his own practice and how the fees influence his income.

Price indices take care of the fact that the impact of a change in fee depends on the change in the fee and how often the laboratory analysis is used. In our study the most often used laboratory analyses (basic analysis) were subject to the highest reduction. Usually a price index measures the

<sup>3</sup> The Fisher ideal index is calculated as the geometric mean of the Paasche and the Laspeyres indices

average price of a set of goods and services in one period compared with the average price of the same goods and services in another period (Grabouen, 1994). In our study we use price indices to measure the average income of laboratory services in general practices from one period to the next. We assume that the GP are more sensitive to changes in fees when the volume is high. The GPs are also influenced by trends regarding use of laboratory services in primary care (Hobbs, 1996), changing the volume in spite of unchanged fees.

Before the GP decides whether or not to take a blood test (decision A), he first consider which laboratory test to take and, secondly, where to analyse it. Because the net income from taking the blood test depends on both which laboratory test is used and where it is analysed, we assume that the GPs take the general development in fees into consideration and include all the laboratory fees in the index when choosing to take a blood test. At decision A the most plausible alternative to taking a blood test is “wait and see”. Therefore, as a proxy for the alternative not to take a blood test, we created an index for fees related only to the encounter. In deciding whether or not to perform laboratory analyses in-office (decision B) or which type of laboratory analyses to use (decision C), we assume that the GPs sum the income per encounter associated with the different alternatives, which consist of both the out-of-pocket fee and the related fees, depending on the choice. The details of how this was done and the development in the Fisher indices appear in Appendix A. We wanted to study the effect of a change in the differences between the Fisher indices for the alternative choices. Table 3 shows the development of these differences.

*Table 3. The development of indices regarding decisions A, B and C*

Difference in Fisher indices	2001/2002	2002/2003	2003/2004	2004/2005
A. Lab-Encounters		0	0.04	-0.18
B. In-office-Out of office	0	0.08	- 0.04	-0.27
C. Only basic-Only non-basic	0	0.12	0.16	-0.16

The differences are 0 in the basis year, as all indices start at 1. The table shows that the last year the difference for decision A is - 0.18, meaning that the Fisher index for encounters has increased by 0.18 compared to the Fisher index for using the lab.

The reform included an increase in the out-of-pocket fee for laboratory services in-office. Based on data from NMA (2003, 2004), we have calculated that the patient share of the payment to the GP for laboratory services increased from 34% before the reform to 57% after the reform. Table 4 gives an overview of the variables included.



Table 4. Overview of the mean of the variables during all periods with different decisions

		<b>November 2002-2004</b>	<b>October 2001-2004</b>	
<b>Variable</b>	<b>Definition</b>	<b>A. Blood test N= 1543547</b>	<b>B. In-office N=1390715</b>	<b>C. Only basic N=696861</b>
Fisher indices	The difference in the Fisher indices	-0.063	-0.063	0.030
Patient age	< 16 yrs old	0.093	0.080	0.108
	16-39 yrs old	0.269	0.254	0.239
	40-59 yrs old	0.317	0.281	0.249
	60-74 yrs old	0.162	0.201	0.201
	>74 yrs old	0.159	0.184	0.203
GP sex	Male (1) Female (0)	0.800	0.755	0.769
GP age	<35 yrs old	0.051	0.082	0.079
	35-54 yrs old	0.720	0.718	0.714
	>54 yrs old	0.229	0.200	0.207
	Continuous in years	48.5 (7.9)	47.4 (8.3)	47.5 (8.3)
Postgraduate qualifications	Specialist in family medicine / community medicines (1), None (0)	0.709	0.669	0.675
Group practice	Type of practice: Binary variable: 1 if group. 0 If solo	0.771	0.801	0.804
Centrality	Least central	0.077	0.112	0.112
	Less central	0.057	0.072	0.071
	Central	0.268	0.273	0.275
	Most central	0.598	0.543	0.542

\* The standard deviation is in brackets.

Centrality refers to the geographical location of the general practice in a municipality in relation to densely populated areas according to the 1994 standard for municipal classification<sup>4</sup> (Statistics Norway, 1994).

Disease frequency varies systematically with age and gender with many diseases. In our data we only included patient age because patient gender was erroneously coded.

We compared the GP characteristics of age, sex, and share of specialists in our samples of GPs with the population of Norwegian GPs in the period (see web link NIA). We found that our samples of GPs from October and November had the same mean value regarding age as the

<sup>4</sup> There are four main categories of centrality ranging from least to most central. A municipality is in the category “most central” when its main population centre lies within 75 minutes' travel time (90 minutes for Oslo) from a densely populated area with at least 50 000 inhabitants (with the fastest mode of transport except for plane). The category “central” means that there is a maximum of 60 minutes' travel time to a densely populated area with a minimum of 15 000 inhabitants. The category “less central” means that there is a maximum of 45 minutes' travel time to a densely populated area with a minimum of 5 000 inhabitants. Municipalities that do not meet any of these criteria are assigned “least central”. If a municipality fulfils two or more criteria at the same time, the category with the highest number is selected.

population of GPs, but had a slightly higher percentage of males (76 % Oct.- 80% Nov. versus 71%), and a higher proportion of GPs with a speciality (70 % Oct.-67 % Nov. versus 56%) than the population of GPs in the period. NIA informs that some data may have been deleted due to storage capacity and this, in addition to the fact that we only have electronic bills, may be the reason why our data versus the population of GPs have a higher share of male GPs and a higher proportion of GPs with a speciality.

## **6. Empirical models**

The modelling framework is based on the theory of discrete choice analysis (see, for example, Ben-Akiva and Lerman (1985) or Greene (2000)).

We want to establish models regarding the use of laboratory services in the GP's office for predicting the probability of:

- a GP taking a blood test during the encounter (decision A),
- performing (one or several) laboratory analyses in the office versus performing all the laboratory analyses out of the office (decision B),
- performing only basic analyses versus performing only non-basic analyses in the GP's office (decision C).

As the models are estimated based on different data sets (see Table 3), it was not possible to study the three decisions in a joint model.

We model the probabilities as a function of the GP's income (represented by the difference in the Fisher indices), the GP in question (age, sex, etc.) and patient's age. At all decisions the GP has two possible choices that are mutually exclusive, so we use binary models.

We assume that the GPs have preferences regarding the different alternatives that can be represented by a random utility function. The utility of an alternative depends on the GP's income or cost associated with the alternative, ethical considerations, reputation and status. Increases in the patient's health are assumed to increase the GP's utility. The trade-off between the welfare of the patient and the GP's income will differ depending on the situation, the GP's knowledge of laboratory services and the GP's loyalty to medical recommendations.

As regards decision A, there is a trade-off between the utility of a blood test compared to the utility of other services in the GP's office. Important factors regarding the utility of taking a blood test are the time spent analysing the test and following up the test results versus the "wait and see" strategy. The last strategy sometimes includes a follow-up consultation to check whether the

medical condition has improved. As regards decision B, the trade-off is between the utility of receiving the test result during the consultation compared to the utility of the quality of the laboratory test. The utility of analysing in-office versus out of the office depends on how important it is to get the answer during the same consultation. Analysing in the GP's office is time-saving with less following up than analysing outside, but the quality of the laboratory analyses in an external lab may be better (higher sensitivity and specificity). As regards decision C, the utility of using only basic analyses versus non-basic analyses depends on the type of patient.

Recall that we assumed that income motivation is more likely to exist for GPs with less knowledge of laboratory services or if the GP is uncertain about the utility of diagnosis and treatment. We test these hypotheses by including interaction terms between the income variable and variables related to knowledge and experience (GP's age, education (speciality), and type of practice), and between the income variable and the patient's age, as explained in the previous section.

Encounters are nested; each GP has several observations in the dataset. GPs are in turn nested in local geographical areas and grouped by centrality, thus giving our data a hierarchical structure. We have three levels: the observation, the GP, and the local area. The GPs are grouped into four categories of centrality, and we choose to include these as dummies. Because the Fisher indices change yearly (ref. Table 3), we could not estimate a model with time dummies.

Previous studies (Yager et al. 1986; Van der Weijden et al. 2002) have found that test-ordering behaviour is partly based on the GP's personal habits and characteristics, and not only on the patient's signs and symptoms. Since we have many observations by the same decision-makers, we take account of unobservable GP-specific heterogeneity in our specification. We represent permanent unobserved heterogeneity in preferences at the GP level by using random effect specification.

Let  $U_{iejt}$  be the utility of  $GP_i$  as evaluated by  $GP_i$  in encounter number  $e$ , given alternative  $j=1,2$  at time  $t$ . For decision A,  $j=1$  if a blood test is taken and  $j=2$  if no blood test is taken. For decision B,  $j=1$  if laboratory analyses are performed in the office and  $j=2$  if all laboratory analyses are performed outside the office. For decision C,  $j=1$  if only basic analyses are performed in-office and  $j=2$  if only non-basic analyses are performed in-office.

We assume that  $U_{iejt}$  is stochastic. The motivation for this is that the GP's preferences may depend on variables known to him but unobserved by the researcher. For example, there will be non-systematic variation in choices that cannot be explained by the variables available to the researcher, especially those stemming from patient-physician interaction. In addition, the GP may have unstable preferences, meaning that he or she may have problems in evaluating the utility of

the different alternatives. This means that the GP may make different choices in seemingly identical choice settings. The variation in responses across seemingly identical choice settings may change because colleagues, medical case records and experience gained from treating other patients have a continuous influence on the GP<sup>5</sup>.

We assume that the utility function has the structure

$$(1) \quad U_{iejt} = a_j + u_{ij} + \beta_j X_{ie} + \gamma_j I_{jt} + \theta_j Z_{ie} I_{jt} + \varepsilon_{iejt},$$

where  $X_{ie}$  is a vector of characteristics of the GP  $i$  (gender, age, education, type of practice and location) and the patient's age in encounter  $e$ .  $a_j$ ,  $\beta_j$ ,  $\gamma_j$  and  $\theta_j$  are vectors of parameters and  $u_{ij}$  is a time constant random term that is independent of  $\varepsilon_{iejt}$ . The terms  $u_{ij}$  and  $\varepsilon_{iejt}$  are random variables that are assumed to be normally distributed with zero mean. The term  $\varepsilon_{iejt}$  is the stochastic term that is supposed to account for unobserved variables of the GP that affect his preferences. Moreover,  $\varepsilon_{iejt}$ ,  $t = 1, 2, \dots$ , are i.i.d. over time. The term  $u_{ij}$  represents the individual specific level of the utility difference across seemingly identical choice settings accounts for the GP's practice style. The variable  $I_{jt}$  is a proxy variable for the incomes associated with alternative  $j$  at time  $t$ , and it is constructed as the difference between the Fisher indices, cf. Table 3.  $Z_{ie}$  is a vector of variables (patient's age, GP's age, group practice and postgraduate qualifications) we assume influence the impact of the income variable. From (1) it follows that the difference in utility equals

$$(2) \quad \begin{aligned} U_{ie1t} - U_{ie2t} &= (a_1 - a_2) + (u_{i1} - u_{i2}) + (\beta_1 - \beta_2)X_{ie} + (\gamma_1 - \gamma_2)(I_{1t} - I_{2t}) + (\theta_1 - \theta_2)Z_{ie}(I_{1t} - I_{2t}) + (\varepsilon_{ie1t} - \varepsilon_{ie2t}) \\ &= a + u_i + \beta X_{ie} + \gamma I_t + \theta Z_{ie} I_t + \varepsilon_{iet}, \end{aligned}$$

As mentioned above, the two random components are assumed to be independent of each other, with  $E(u_i) = E(\varepsilon_{iet}) = 0$ ,  $E(u_i^2) = \sigma_u^2$ ,  $E(\varepsilon_{iet}^2) = \sigma_\varepsilon^2$ . Hence, we have

$$\text{Corr}(\varepsilon_{iet} + u_i, \varepsilon_{ifs} + u_i) = \rho = \sigma_u^2 / (\sigma_u^2 + \sigma_\varepsilon^2),$$

---

<sup>5</sup> In addition, the GP's choice setting can be viewed as a case of choice between lotteries because of the uncertainty of both the initial health status of the patient and the laboratory analyses. An example of the latter is the use of the Helicobacter Pylori Rapid Test (HPRT). The uncertainty of the laboratory analysis occurs because the HPRT measures antibodies to the HP bacteria and not the disease as such, and because healthy carriers of the HP bacteria exist. Because some patients are carriers of HP without being ill, the use of the test result will depend on the GP's knowledge of HP. In this sense, the utility function defined above can be interpreted as a (subjective) expected utility of the GP. However, we shall continue to refer to it simply as the utility function.

representing the correlation coefficient between the utility differences of two encounters from the same GP.

We assume that the GP will choose the alternative with the highest utility. Let  $P_1(V_{iet})$  denote the probability of GP<sub>i</sub> choosing alternative 1 for a patient in encounter  $e$  in year  $t$ , where we for convenience use  $V_{iet}$  to represent the vector of explanatory variables. Similarly, let  $P_1(V_{iet}|u_i)$  be the corresponding choice probability of choosing alternative 1 given the random effect  $u_i$ . Hence the conditional probability of choosing alternative 1 given the random effect is given by

$$(3) \quad P_1(V_{iet}|u_i) = \Phi(u_i + V_{iet}b) = \Phi(a + u_i + \beta X_{ie} + \gamma I_t + \theta Z_{ie}I_t),$$

where  $\Phi$  is the standard cumulative normal distribution function and  $b$  represents the vector of unknown parameters associated with the variable vector  $V_{iet}$ . One can show (see McFadden and Reid (1975) and appendix C) that the corresponding unconditional choice probability is given by

$$(4) \quad P_1(V_{iet}) = E_u P_1(V_{iet}|u_i) = \Phi(ka + k\beta X_i + k\gamma I_t + k\theta Z_{ie}I_t),$$

where  $E_u$  denotes the expectation with respect to  $u$  and  $k = \sqrt{1 - \rho}$ . The expression in (4) is used to compute the effect of changes in explanatory variables.

The model is estimated by the maximum likelihood method. The corresponding likelihood is given in the appendix C.

If the variance of the random effect is different from zero, the random effects model fits the data better than the model without random effects. We use standard econometric software provided by NLOGIT (Greene, 2007) to estimate a random effects probit model. We tested for unobserved heterogeneity by estimating a model with random effects and a  $H_0$  saying that all estimated individual effects are equal. The LR-test rejected  $H_0$  with a level of statistical significance of less than 1% at all decisions.

## 7. Results

### Estimation results

Table 5 shows the estimated parameters of the random effects probit models, given by (3), and provides information about the sign and relative magnitude of the effects. At all three medical

choices,  $\rho$  is significant. At decision A,  $\rho$  is small, so the results are very similar to the standard model.

Table 5. Results of the estimation of the random effects probit models

Independent variables		A. Blood test		B. In-office		C. Basic only	
		Para- meters	Std. dev	Para- meters	Std. dev	Para- meters	Std. dev
Constant		<b>-0.618</b>	0.027	<b>0.266</b>	0.008	<b>1.278</b>	0.012
I = Difference in the Fisher indices		<b>0.353</b>	0.121	<b>0.170</b>	0.036	<b>0.439</b>	0.055
<b>Patient age</b>	0-14	<b>Reference</b>					
	15-39	<b>-0.102</b>	0.009	<b>-0.488</b>	0.004	<b>-0.729</b>	0.005
	40-60	<b>-0.168</b>	0.010	<b>-0.477</b>	0.004	<b>-0.863</b>	0.005
	61-74	<b>0.089</b>	0.012	<b>-0.253</b>	0.004	<b>-1.356</b>	0.005
	>75	<b>0.045</b>	0.011	<b>-0.118</b>	0.004	<b>-1.632</b>	0.005
<b>GP age</b>	<35	<b>Reference</b>					
	35-54 yrs	<b>-0.054</b>	0.020	<b>0.097</b>	0.005	<b>0.043</b>	0.008
	>54 yrs	<b>-0.076</b>	0.021	<b>0.218</b>	0.006	<b>0.038</b>	0.009
Specialist vs. not specialist		0.012	0.010	<b>-0.036</b>	0.003	<b>0.073</b>	0.004
Group vs. solo practice		<b>-0.046</b>	0.015	<b>0.057</b>	0.004	<b>-0.079</b>	0.004
GP male vs. female		<b>-0.103</b>	0.017	<b>0.225</b>	0.004	<b>-0.068</b>	0.005
Centrality	Least central	<b>Reference</b>					
	Less central	-0.026	0.030	<b>0.054</b>	0.007	<b>0.081</b>	0.010
	Central	0.025	0.024	<b>0.138</b>	0.006	<b>0.230</b>	0.007
	Most Central	-0.021	0.021	<b>0.121</b>	0.005	<b>0.214</b>	0.007
<b>Interactions with Fisher indices</b>							
Patient age	15-39	<b>-0.354</b>	0.095	<b>0.172</b>	0.033	-0.087	0.051
	40-60	<b>-0.299</b>	0.097	0.054	0.033	<b>0.130</b>	0.050
	61-74	<b>-0.466</b>	0.107	0.027	0.035	<b>0.149</b>	0.050
	>75	<b>-0.532</b>	0.101	0.051	0.034	<b>0.182</b>	0.049
GP age	35-54 yrs	<b>-0.489</b>	0.109	<b>0.107</b>	0.024	<b>-0.252</b>	0.036
	>54 yrs	<b>-0.393</b>	0.128	<b>0.169</b>	0.026	<b>-0.185</b>	0.038
Specialist vs. not specialist		0.093	0.063	-0.017	0.012	<b>0.075</b>	0.020
Group vs. solo practice		-0.072	0.068	-0.015	0.013	<b>-0.184</b>	0.019
Rho		<b>0.065</b>	0.002	<b>0.139</b>	0.001	<b>0.209</b>	0.001
McFadden's R <sup>2</sup>		0.88		0.41		0.24	
Log L full model random		-110 509		-536 498		-347 093	
Log L intercept only		-903 619		-912 172		-455 538	
Log L full model std		896 858		-892 144		-412 732	
% y=1 predicted correctly		0%		94%		84%	
% y=0 predicted correctly		100%		10%		39%	
% Correctly predicted		72.8%		63.8%		68.1%	
No. of obs.		1 543 547		1 390 715		696 861	

Bold figures are significant at the 5% level.

The goodness of fit for the models is high as the models correctly predict in the range of 64%-73% for all the observations and the McFadden's  $R^2$  range from 0.24 to 0.88.

The results in Table 5 support most of our hypotheses. For all decisions the effect of the income variable ( $I_{jt}$ ) is positive. The total effect of income for decision A and C will depend on the situation we study as the interaction terms are negative. For decision B, the negative interaction terms are small in magnitude so an increase in income (the differences between the Fisher indices) will increase the probability of using the laboratory service. The use of blood tests is more likely among older than younger patient groups, and this is also consistent with a study by Kristiansen and Hjortdahl (1992). Using only basic analyses versus non-basic analyses is less likely in older patients. As expected, older GPs take fewer blood tests. They are also more likely to carry these out in-office, and to analyse only basic rather than only non-basic analyses compared to younger GPs. GPs in group practices are more likely to perform laboratory analyses in-office and less likely to have encounters with only basic analyses versus only non-basic analyses. This is as expected due to a larger repertoire of laboratory analyses in-office. GPs located in the central areas (central and most central) are more likely than GPs in less central areas to analyse in-office, or to use only basic versus non-basic analyses. We suggest that the reason for this is that they want to give quicker service due to more competition in central areas. Female GPs use more blood tests than male GPs, and this also agrees with previous studies (Franks et al. 2000; Kristiansen and Hjortdahl, 1992; Salloum and Franssen, 1993; Vinker et al. 2007). The results (shown by the sign of the parameter of the interaction terms) support our hypothesis that the degree of income motivation is stronger when GPs are uncertain about the utility of the laboratory service, and based on the situation or type of patient. As regards decision A, the impact of income is less for older patients versus younger patients. This is as expected because older patients have a higher risk of disease; thus, the expected benefit of a blood test is higher. As regards decision C, the impact of income (and the uncertainty) is less for younger patients than for older patients. This is also as expected because basic analyses are the most used laboratory analyses among the youngest because of a high level of infections.

We expected that the impact of income would be less for GPs with more knowledge of laboratory services. The results support most of our hypotheses. The exceptions are that the impact of the index is higher for older versus younger GPs at decision B, and for GPs who are specialists at decision C.

### The effect of changes in significant variables

We wanted to study the magnitude of the effect of significant variables on the probability of decisions A through C. Estimations of models including only significant variables appear in Table B1 in Appendix B. We use these results in calculating the effect of changes in significant variables. The marginal effects on the probability are listed in Table 6, and some of the calculations are shown in Appendix C. As an example, we use a GP with the following characteristics (after this named “our case”):

- Male, 34 years old
- Works in a group practice located in the most rural area
- Has a specialist licence in general practice/family medicine
- Has a patient who is 10 years old.

In Table 6 we study the effect of changing one variable at a time while keeping everything else constant. In particular, we want to study how the marginal effect of the independent variables varies with the income variable (differences between the Fisher indices,  $I_{jt}$  from equation (4)). To find the effect of the income variable and the interaction terms, we study the marginal effects when  $I_{jt} = 0$  (as for the basis year) and when  $I_{jt} = 0.1$ . When  $I_{jt} = 0.1$  the Fisher index for the first alternative has increased by 0.1 versus the second alternative.

The table shows that patient’s age has the most influence and the income variable has the least influence on the probability for all decisions. The probability for taking a blood test, decision A, increases from 22.9% to 23.7% when  $I_{jt}$  increase from 0 to 0.1. Similarly, for decision B the probability increases from 67.7% to 67.9% and from 85.8% to 86.3% for decision C. In interpreting the results, we focus on when  $I_{jt} = 0$ . If the Fisher index for the first alternative increases by 0.08 compared to the second alternative choice, the probability for choosing alternative 1 will increase by 0.7 percentage point for decision A, by 0.5 percentage point for decision B and by 0.4 percentage point for decision C. The results show that the income variable has most relative impact on decision A versus decision B and C. The reason for this may be that the uncertainty about the utility of whether or not to take a blood test is higher, versus the decisions B and C. When a blood test is taken, the types of laboratory analyses depend on the patients – and the laboratory used depends on the repertoire of available laboratory analyses in the GP’s practice.

The patient’s age has the most effect, particularly at decision C. If the patient’s age is 76 years versus 10 years, the probability of performing only basic versus non-basic analyses decreases by 50.6 percentage point (from 85.8% to 35.2%).



Table 6. The marginal effects of significant variables on the probability

Independent variables		A. Probability of taking a blood test		B. Probability of performing analyses in office vs. all analyses out of office		C. Probability of only performing basic in office vs. only non-basic	
		22.9%	23.7%	67.2%	67.9%	85.8%	86.3%
Difference between the Fisher indices		I = 0	I = 0.1	I = 0	I = 0.1	I = 0	I = 0.1
Increase in the difference in the Fisher index: 0.08		+ 0.7	+ 0.7	+ 0.5	+0.6	+ 0.4	+ 0.4
Patient age	10→ 34	- 3.3	- 4.2	- 18.3	- 17.9	- 19.5	- 19.1
	10→ 55	- 5.2	- 6.1	- 17.2	- 17.3	- 24.0	- 22.9
	10→ 62	+ 2.7	+ 1.3	- 10.7	- 9.2	- 41.2	- 40.0
	10→ 76	+ 1.3	- 0.2	- 3.8	- 3.8	- 50.6	- 49.4
GP age	34→ 40	0	- 1.1	+ 3.7	+ 4.1	+ 0.9	+ 0.4
	34→ 55	0	- 0.8	+ 5.9	+ 6.3	+ 0.7	+ 0.4
Male→female GP		+ 3.7	+ 3.7	- 7.6	- 7.5	+ 1.3	+ 1.4
Group – solo practice		ns	ns	- 1.4	- 1.3	+ 1.5	+ 1.9
Specialist – not specialist		ns	ns	+ 1.2	+ 1.1	- 1.7	- 1.6
Least → less central		ns	ns	+ 1.6	+ 1.6	+ 1.5	+ 1.6
Least →central		ns	ns	+ 4.6	+ 4.6	+ 4.1	+ 4.0
Least → most central		ns	ns	+ 3.9	+ 3.8	+ 3.9	+ 3.7

ns = not significant

Female GPs are more likely to take blood tests than male GPs (26.6% vs. 22.9%). If a blood test is taken, female GPs are less likely to perform laboratory analyses in-office (59.6% vs. 67.2%) and more likely to perform only basic analyses versus only non-basic (87.1% vs. 85.8%). As expected, GPs in solo practices are less likely than GPs in group practices to analyse in office (65.8% vs. 67.2%) and more likely to perform basic analyses, probably because of a smaller repertoire of laboratory analyses in office (87.3% vs. 85.8%).

Our results regarding use of blood tests agree with results by Leurquin et al. (1995). They studied use of blood tests in general practice in eight European countries and found that characteristics of general practitioners and practice were either only slightly or not associated with the use of blood tests. The request rate for blood tests varied considerably between countries.

We are particularly interested in the effect of the interaction terms. We find the effect of the interaction terms in table 6 by comparing the effect of the variables: patient's age, GP's age, group practice and postgraduate qualifications, when  $I_{jt} = 0$  versus when  $I_{jt} = 0.1$ . When  $I_{jt} = 0$  the interaction terms have no effect. Table 5 showed that the impact of income is less for older patients versus younger patients for decision A. If the patient's age is 35 years versus 10 years, table 6

shows that the probability of taking a blood test will decrease by 3.3 percentage point at  $I_{jt} = 0$  and by 4.2 percentage point at  $I_{jt} = 0.1$ . The difference by minus 0.9 percentage point ( $-4.2 - 3.3$ ) is the effect of the interaction term. If the patient's age is 76 years versus 10 years, there is a difference in the probability by minus 1.5 percentage point ( $-0.2 - 1.3$ ) when  $I_{jt}$  increases from 0 to 0.1. Thus the probability for taking a blood test is decreasing for older versus younger patients due to the income.

Table 6 shows that the impact of income is less for older versus younger GPs, because when  $I_{jt}$  increases from 0 to 0.1, this leads to a decrease in the probability for older versus younger GPs at decision A by 1.1 for GPs in the category 35-54 yrs and by 0.8 percentage point for GPs in the category >54 years. Similarly at decision C, the effect of the interaction terms is a decrease in the probability by 0.5 for GPs in the category 35-54 yrs and 0.3 percentage point for GPs in the category >54 years.

## 8. Discussion

We will now discuss our results and some objections to the arguments and methods we have used.

Recall that time dummies can be used to study whether there are trend effects in the use of the laboratory service in question. We could not include time dummies in our model because this would lead to identification problems because the Fisher indices do not vary across individuals. In our case, any existing trend in using laboratory services is partly included in the effect of the Fisher index because the yearly quantities of laboratory service are included in the index. Thus, it is not possible to distinguish between the income and the trend effect.

Secondly, analysing only out of office often demands more follow-up from the GPs, partly because more analyses are easily ordered when filling out a requisition form. The cost of the follow-up will vary between GPs and depends on the type of laboratory analyses and local routines for follow-up. These costs reduce the utility of analysing only outside the office and reduce the effect of the financial incentives to analyse outside the GP's office.

Thirdly, we only have information about the analyses performed in-office and reimbursed, and not about the laboratory analyses that are analysed outside the office or about laboratory analyses e.g. Erythrocyte Sedimentation Rate (SR) and Haemoglobin (HB) which are not reimbursed. As GPs do not carry any costs of the laboratory analyses done outside the office (except the cost of more follow-ups), it is possible that, in order to compensate for the loss after the reform, the GPs may add laboratory analyses that are not reimbursed on the requisition form. In our further

research we would like to study the development in the use of laboratory analyses analysed outside the office.

Fourthly, data from the last year was only collected 4-5 months after the reform, and there may be a time lag before the reform has any effect on a GP's practice profile. When it was introduced, the reform was widely debated in Norway among GPs on their discussion page on the Internet, so it seems likely that the GPs were aware of the changes in the remuneration system. The long-term effects may come later for laboratory analyses when it is necessary to invest in new laboratory equipment. In Fauli and Thue (2005) we found that the remuneration system did not have any significant effect on the decision to have or to use the test in a specific situation. However, the HPRT is a rapid test for single use only and does not demand any investment except buying the test itself. We assume that physicians will weigh benefits and costs before investing in new instruments. For analyses in the basic repertoire the necessary investments are not major (ref. Table 1), but after the reform GPs may have felt insecure about the development in fees and delayed investment in new equipment. The long-term effect of the reform may be a decline in analysing in-office, and lower test quality due to old laboratory equipment.

Fifthly, we have only studied the use of laboratory services in-office and do not know whether GPs may have compensated for the reform by using more of other services with positive net marginal revenue.

Finally, our results show that there has been an increase in the use of taking blood tests after the reform. We have only studied the use of type of analyses (basic or non-basic) in office. Recall that the total revenue from using lab and basic analyses can be higher after the reform, particularly when only one laboratory analysis is performed due to the out-of pocket fee. There are also differences between laboratory analyses regarding how much the fee was reduced (see Table 1), and the GP may for example compensate by using more basic analyses that are less reduced than those that are most reduced (e.g. Glucose), or using fewer laboratory analyses per encounter. In our future research we would like to study the development in composition of laboratory analyses per encounter by studying the development in number of laboratory analyses per encounter and the use of typical sets of laboratory analyses.

## 9. Conclusion

Physicians are influenced by many different motives such as ethical conduct, standards of care and service motives (Arrow, 1986). Our results indicate that income motivation among GPs regarding the use of laboratory services exists, because since the reform GPs have slightly reduced the use of laboratory services in office in line with the reduction in fees that have occurred. Since the reform led to minor changes in the use of laboratory analyses in the GPs' offices, we suggest that the financial incentives were diluted because they were in conflict with medical recommendations and existing medical practice. Our result illustrates that the presence of ethical considerations in the GP's utility function counteracts incentives that may arise from the remuneration system.

We suggest that our results are valid before the GP needs to invest in new laboratory instruments. In the long run, the reform may lead to a decline in analysing in-office because the occurrence of fees below costs prevents investment in new instruments.

A policy implication of our results is that economic incentives alone are not an effective tool for influencing the use of laboratory services in GPs' offices. The results support the suggestions by several (Chaix-Couturier et al. 2000; Hillman, 1990; Mooney and Ryan, 1993; Wennberg et al. 1982): Income motivation is more likely to exist when the correct decision is not obvious. When using financial incentives, it is important to consider the degree of professional uncertainty, as financial incentives will have more influence when the degree of professional uncertainty is high or in areas where there is a lack of professional consensus (medical recommendations). Correspondingly, financial incentives are less likely to influence medical decisions when the incentives conflict with medical guidelines that are internalized by physicians.

## References

- Arrow K.J. Uncertainty and the Welfare Economics of Medical Care. *The American Economic Review* 1963;53;941-973.
- Arrow K.J. Agency and the market, in K. J Arrow and M.D. Intriligator, eds. *Handbook of mathematical economics*, vol III.Elsevier. Amsterdam. 1986.
- Ben-Akiva M, Lerman, S *Discrete Choice Analysis: Theory and Application to Travel Demand*, MIT Press, Cambridge, Massachusetts, 1985.
- Carlsen F, Grytten J, Skau I. Financial incentives and the supply of laboratory analyses. *European Journal of Health Economics* 2003;4;279-85.
- Chaix-Couturier C, Durand-Zaleski, I, Jolly D., and Durieux P. Effects of financial incentives on medical practice: results from a systematic review of the literature and methodological issues. *International Journal of Quality in Health Care* 2000;12;133-42.
- Delaney BC, Hyde CJ, McManus RJ, Wilson S, Fitzmaurke DA, Jowest S, Ros T, Thorpe GH, Hobbs R. Systematic review of near patient test evaluations in primary care. *British Medical Journal* 1999, 319; 824–827.
- DeNeef P. The expanding role of the office laboratory. *Journal of Family Practice* 1986; 22; 215-216.
- Fauli S, Thue G. Decision making in general practice: the availability and use of a specific laboratory analysis. *International Journal of Technology Assessment in Health Care* 2005;21;386-92.
- Fauli S. Thue G. Economic consequences of near-patient test results: the case of tests for the *Helicobacter Pylori* bacterium in dyspepsia. *European Journal of Health Economics* 2008;9;221-8.
- Fischer PM. Laboratory testing in the 1990s. *Journal of Family Practice* 1991; 33;453-454.

Fisher J. *The Making of Index Numbers*. Houghton Mifflin, Boston, 1922.

Franks P, Williams GC, Zwanziger J. Why do physicians vary so widely in their referral rates? *Journal of General Internal Medicine* 2000;15;163–168.

Frey B. On the relationship between intrinsic and extrinsic work motivation. *International Journal of Industrial Organisation* 1997;15;427- 439.

Gosden T, Pedersen, L, Torgerson, D. How should we pay doctors? A systematic review of salary payments and their effect on doctor behaviour. *International Journal of medicine* 1999; 92: 47-55.

Graboyes RF. *Medical Care Price Indexes*. Federal Reserve Bank of Richmond Economic Quartely 1994;Volume 80/4.

Greene WH. *Econometric Analysis*, Prentice-Hall International, Englewood Cliffs, New Jersey, 2000.

Greene WH. *NLOGIT version 4.0, Users manual*, Econometric Software, Australia, 2007.

Grytten J, Carlsen F, Skau I. Primary physicians' response to changes in fees. *European Journal of Health Economics* 2008; 9;117-25.

Guthrie B. Why do general practitioners take blood? A cross-sectional study of use of blood tests in UK general practice. *European Journal of General Practice* 2001; 7;138-42, 160.

Hartley RM, Charlton JR, Harris CM, Jarman B. Influence of patient characteristics on test ordering in general practice. *British Medical Journal (Clin Res Ed)* 1984;289;735–738.

Higgins JC. The status of physician office labs since CLIA '88. *The Journal of Medical Practice Management* 2000;16;99-102.

Hillmann A. Health maintenance organizations, financial incentives and physician judgement (editorial). *Annals of Internal Medicine* 1990;112; 891-893.

- Hjortdahl P. The silent revolution. *Scandinavian Journal of Primary Health Care* 1990;8;188-190.
- Hobbs FDR, Kenkre JE, Carter YH, Thorpe GH. Cost and clinical effectiveness evaluation of using a C-reactive protein near patient test (NycoCard) in primary care. *Journal of Family Practice* 1995;12;132.
- Hobbs R. Near patient testing in primary care, *British Medical Journal* 1996; 312; 263-264.
- Hunink M, Glasziou P, Siegel J, Weeks J, Pliskin J, Elstein A, Weinstein M. Decision making in health and medicine. Integrating evidence and values. University Press, Cambridge. 2001.
- Hunskår, S (ed). *Allmennmedisin* (Norwegian textbook of general practice). Ad Notam Gyldendal. 2003.
- Iversen T. A study of income-motivated behavior among general practitioners in the Norwegian list patient system. University of Oslo. Health research programme. Working paper 2005; 8.
- Knight F.H. Risk, Uncertainty and Profit. Boston MA. Hart, Schaffner & Marx: Houghton Mifflin Company, 1921.
- Kristiansen IS, Hjortdahl P. The general practitioner and laboratory utilization: why does it vary? *Journal of Family Practice* 1992;9;22-27.
- Leurquin P, Van Casteren V, De Maeseneer J. Use of blood tests in general practice: a collaborative study in eight European countries. Eurosentinel Study Group. *British Journal of General Practice* 1995;45; 21-25.
- McFadden D, Reid F. Aggregate Travel Demand Forecasting from Disaggregate Models. *Transportation Research Record* 1975;534;24-37.
- Mooney G, Ryan M, Agency in health care: getting beyond first principles, *Journal of health Economics* 1993;12;125-138.

National Insurance Administration (NAV). What's going on in general practice? Report nr. 4, 2007.

Norwegian Medical Association. 2001. Fixed Fee schedules for 2001/2002.

Norwegian Medical Association. 2002. Fixed Fee schedules for 2002/2003.

Norwegian Medical Association 2003. Fixed Fee schedules for 2003/2004.

Norwegian Medical Association 2004. Fixed Fee schedules for 2004/2005.

Salloum S, Franssen E. Laboratory investigations in general practice. *Canadian Family Physician* 1993;39:1055–1061.

Scott A, Shiell A. Do fee descriptors influence treatment choices in general practice? A multilevel discrete choice model. *Journal of Health Economics* 1997;16:323-42.

Statistics Norway, Standard for municipal classification. 1994. NOS C192, Statistics Norway, Oslo

Thue G. and Sandberg S. Survey of office laboratory analyses in general practice. *Scandinavian Journal of Primary Health Care* 1994;12:77-83.

Verstappen W HJM; Riet G ter; Dubois WI, Winkens R, Grol R-van der Weijden T. Variation in test ordering behaviour of GPs: professional or context-related factors? *Journal of Family Practice* 2004;21;4:387-395.

Vinker S, Kvint I, Erez R, Elhayany A, Lezion R, Kahan E. Effect of the characteristics of family physicians on their utilisation of laboratory tests. *British Journal of General Practice* 2007; 57; 377–382.

Van der Weijden T, van Bokhoven MA, Dinant GJ, van Hasselt CM, Grol RP. Understanding laboratory testing in diagnostic uncertainty: a qualitative study in general practice. *British Journal of General Practice* 2002; 52;974-80.



Van Walraven C; Goel V; Chan B. Effect of Population-based Interventions on Laboratory Utilization. *Journal of the American Medical Association* 1998;280;2028-2033.

Wennberg J.E. On patient need, equity, supplier-induced demand, and the need to assess the outcome of common medical practice. *Medical Care Research and Review* 1985;23;512-520.

Wennberg J.E, Benjamin A, Zubkoff B and M. Professional uncertainty and the problem of supplier-induced demand. *Social Science and Medicine* 1982;16;811-824.

Yager J, Linn LS, Leake B, et al. Initial clinical judgments by internists, family physicians, and psychiatrists in response to patient vignettes: II. Ordering of laboratory tests, consultations, and treatments. *General Hospital Psychiatry* 1986; 8;152–158.

### **Web references**

<http://cepa.newschool.edu/het/essays/uncert/intrisk.htm>

[http://en.wikipedia.org/wiki/Risk#Risk\\_vs.\\_uncertainty](http://en.wikipedia.org/wiki/Risk#Risk_vs._uncertainty)

National Insurance administration: <http://77www.nav.no/page?id=1073743257>.

## Appendix A. Calculating the Fisher indices

The Paasche and Laspeyres indices are two main methods for calculating price indices.

Laspeyres' index is based on the old quantities of laboratory services. These are the weights; they are fixed and equal to the quantities in the first period or the basic period. The index gives information about the change in income of using the same basket of laboratory services in the current period as in the first period.

$$(1A) P_{LA} = \sum p_i q_0 / \sum p_0 q_0$$

The Laspeyres index is calculated for the second year like this:

$$(2A) P_{LA}^{01} = [\sum_{j=1}^n p_1^j q_0^j] / [\sum_{j=1}^n p_0^j q_0^j] = \sum p_1 q_0 / \sum p_0 q_0$$

where 0 is the first year and 1 represents the current year, p is the fee per laboratory service and q is the frequency of use of the laboratory service. J represents laboratory service nr. j. There are n number of fees in the index.

Paasche's index is based on the new quantities and compares the current basket of laboratory services with the same basket with fees from the first period. The fees are weighted by the quantities of the current period. This means that each time the index is calculated, the weights are different:

$$(3A) \Delta P_P = \sum p_i q_i / \sum p_0 q_i$$

$$(4A) P_{PA}^{01} = [\sum_{j=1}^n p_1^j q_1^j] / [\sum_{j=1}^n p_0^j q_1^j] = \sum p_1 q_1 / \sum p_0 q_1$$

We chose to use the Fisher ideal index by Fisher (1922). The Laspeyres index systematically overstates the effect of a change in the fee as the index does not take into account that the GPs react to changes in the fees by changing the use of laboratory analyses. The Paasche index understates the effect of a fee change. The Fisher index minimizes the biases associated with the Laspeyres and the Paasche indices by calculating the index as the geometric average

[ $F = (P_{LA} * P_{PA})^{0.5}$ ] of the Paasche and the Laspeyres indices.

### Principles used in creating the price index

For each medical choice we created a price index. For new fees, the basis year is the first year. The fees included in the indices depend on the medical choice. Table 1A shows an overview of the fees.

Table 1A. Overview of the fees included in the Fisher indices

<b>Alternative choices</b>	<b>Fees included in the Fisher indices</b>
A1. Take a blood test	All laboratory fees
A2. Do not take a blood test	Fees for all types of encounters
B1. Analyse in-office	All fees for laboratory analyses in-office
B2. Analyse out of office	All fees for analysing only out of office
C1. Only basic analyses	All fees for basic analyses
C2. Only non-basic analyses	All fees for non-basic analyses

Recall that before the GP has decided whether to take a blood test (decision A) he must consider which laboratory test to take and where to perform it. Because the net income from taking the blood test depends both on which laboratory test is used and where it is analysed, we assume that the GPs take the general development in fees into consideration, and include all the laboratory fees in the index when choosing to take a blood test (A1). The alternative to using the lab is to “wait and see”; therefore, we also created an index for fees related to the encounter (A2).

In general, we assume that GPs sum the income of the different alternatives. Since the reform (the last year), the out-of-pocket fee is paid every time a blood test is taken, so this fee must be included in the Fisher indices for all alternatives for decisions B and C in the last year. The frequency of the out-of-pocket fee is based on the share of encounters for the different alternatives. We have information about the shares from our datasets and the frequency of out-of-pocket fee during the last year is based on data on the aggregate level from NIA.

### Development in Fisher indices

Figures 1A and 2A show the development in the indices used in our empirical estimations. Fig. 1A shows that fees for encounters have increased while fees for using the lab have decreased in the last year.

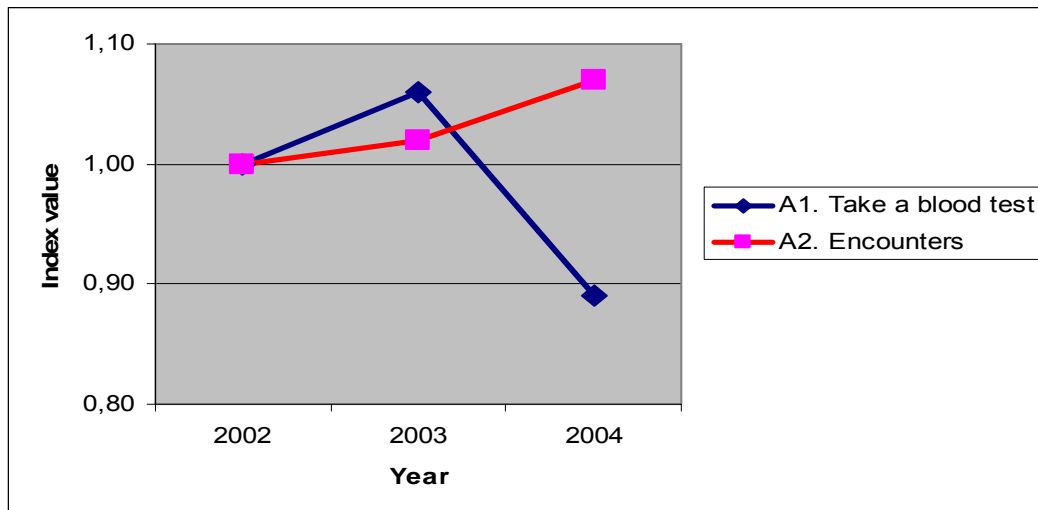


Fig. 1A. The development in the Fisher indices regarding decision A

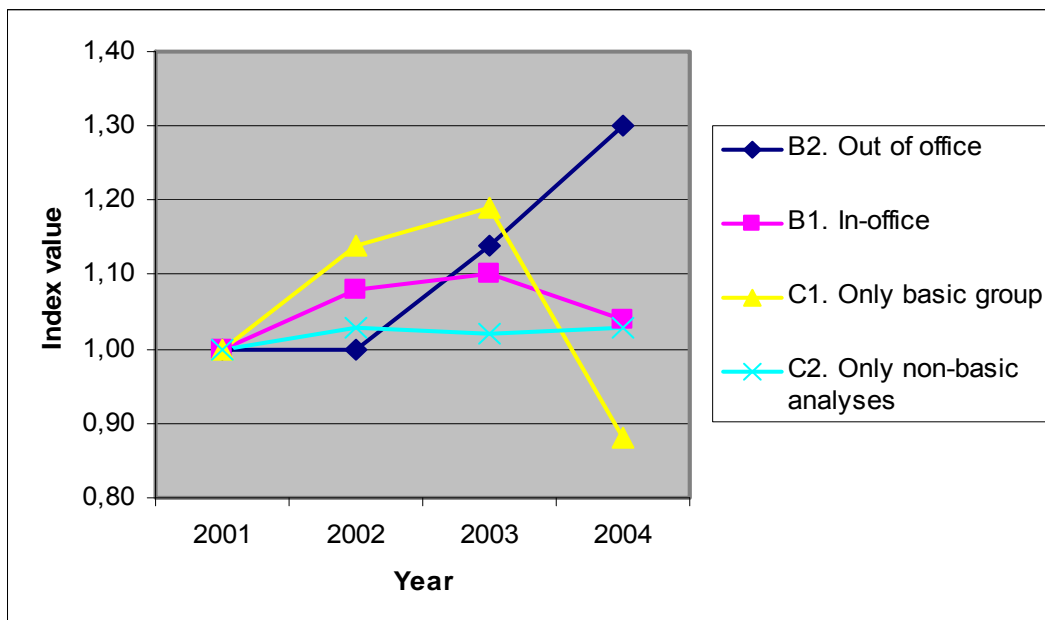


Fig. 2A. The development in the Fisher indices regarding decisions B and C

Fig. 2A shows that the Fisher index for analysing out of office increased while analysing inside decreased after the reform. The reduction in the Fisher index for analysing inside is mostly due to the decrease in the Fisher index for the basic repertoire.

## Appendix B. Estimation results of significant models

Table B1. Results of the estimation of the random effects probit models

Independent variables		A. Blood test		B. In-office		C. Only basic	
		Para- meters	Standard dev.	Para- meters	Standard dev.	Para- meters	Standard dev.
Constant		<b>-0.644</b>	0.016	<b>0.254</b>	0.008	<b>1.280</b>	0.011
I = Difference in the Fisher indices		<b>0.282</b>	0.115	<b>0.210</b>	0.023	<b>0.369</b>	0.038
<b>Patient age</b>	0-14	<b>Reference</b>					
	15-39	<b>-0.115</b>	0.009	<b>-0.510</b>	0.004	<b>-0.731</b>	0.005
	40-60	<b>-0.189</b>	0.010	<b>-0.486</b>	0.003	<b>-0.865</b>	0.004
	61-74	<b>0.089</b>	0.012	<b>-0.264</b>	0.004	<b>-1.358</b>	0.004
	>75	<b>0.046</b>	0.011	<b>-0.112</b>	0.003	<b>-1.633</b>	0.004
<b>GP age</b>	<35	<b>Reference</b>					
	35-54 yrs			<b>0.112</b>	<b>0.005</b>	<b>0.042</b>	0.007
	>54 yrs			<b>0.181</b>	0.006	<b>0.038</b>	0.009
Specialist vs. not specialist				<b>-0.033</b>	0.003	<b>0.073</b>	0.004
Group vs. solo practice				<b>0.042</b>	0.004	<b>-0.079</b>	0.005
GP male vs. female		<b>-0.121</b>	0.016	<b>0.217</b>	0.004	<b>-0.069</b>	0.005
Centrality	Least central	<b>Reference</b>					
	Less central			<b>0.049</b>	0.008	<b>0.081</b>	0.010
	Central			<b>0.140</b>	0.006	<b>0.230</b>	0.007
	Most Central			<b>0.115</b>	0.005	<b>0.214</b>	0.007
<b>Interactions with Fisher indices</b>							
Patient age	15-39	<b>-0.355</b>	0.095	<b>0.137</b>	0.017		
	40-60	<b>-0.298</b>	0.095			<b>0.200</b>	0.030
	61-74	<b>-0.463</b>	0.107			<b>0.219</b>	0.030
	>75	<b>-0.531</b>	0.100			<b>0.252</b>	0.028
GP age	35-54 yrs	<b>-0.386</b>	0.108	<b>0.096</b>	0.023	<b>-0.252</b>	0.036
	>54 yrs	<b>-0.269</b>	0.112	<b>0.161</b>	0.025	<b>-0.187</b>	0.038
Specialist vs. not specialist						<b>0.076</b>	0.020
Group vs. solo practice						<b>-0.183</b>	0.019
Rho		<b>0.065</b>	0.002	<b>0.142</b>	0.001	<b>0.209</b>	0.001
McFadden's R <sup>2</sup>		0.88		0.41		0.24	
Log L full model random		-110 542		-536 435		-347 095	
Log L intercept only		-903 619		-912 172		- 455 538	
Log L full model std		896 996		-892 147		-412 733	
% y=1 predicted correctly		0%		92%		84%	
% y=0 predicted correctly		100%		14%		39%	
% Correctly predicted		72.8%		63.7%		68.1%	
No. of obs.		1 543 547		1 390 715		696 861	

Bold figures are significant at the 5% level

## Appendix C. The effect of changes in significant variables

To study the effect of changes in significant variables, we have several observations per GP and used random effects probit models regression because we have a large dataset with few periods. We have a non-linear model, and cannot interpret the coefficients as we do in ordinary linear regression.

Recall that  $P_1(V_{iet})$  denotes the probability of GP<sub>i</sub> choosing alternative 1 in encounter e in year t, where  $V_{iet}$  is the vector of explanatory variables. Similarly, we recall that  $P_1(V_{iet}|u_i)$  is the corresponding choice probability of choosing alternative 1 given the random effect  $u_i$ . Based on our assumptions of a probit model, we established in (3) that

$$(1.C) \quad P_1(V_{iet}|u_i) = \Phi(u_i + V_{iet}b) = \Phi(a + u_i + \beta X_{ie} + \gamma I_t + \theta Z_{ie}I_t).$$

We shall now write down the likelihood function that is the basis for estimation. Let  $Y_{iet} = 1$  if GP i chooses alternative 1 in encounter e in year t, and zero otherwise. Then the likelihood function for GP i can be expressed as

$$(2.C) \quad L = \prod_i E_u \left( \prod_t \prod_{e=1}^{N_{it}} P_1(V_{iet}|u_i)^{Y_{iet}} (1 - P_1(V_{iet}|u_i))^{1-Y_{iet}} \right),$$

where  $E_u$  is the expectation with respect to  $u$  and  $N_{it}$  is the number of encounters of GP i in year t. The output of NLOGIT gives the estimates of the vector  $b$  in the conditional model above. The corresponding unconditional choice model is given by

$$(3.C) \quad P_1(V_{iet}) = E_u P_1(V_{iet}|u) = \Phi(ka + k\beta X_{ie} + k\gamma I_t + k\theta Z_{ie}I_t),$$

where  $k = \sqrt{1 - \rho}$ .

*Proof of (3.C):*

Recall that if  $W$  and  $Y$  are two random variables we have that  $EW = EE(W|Y)$ . This implies that

$$(4.C) \quad P(u_i + V_{iet}b + \varepsilon_{iet} > 0) = E_u P(u_i + V_{iet}b + \varepsilon_{iet} > 0 | u_i) = E \Phi(u_i + V_{iet}b) = E_u P_1(V_{iet}|u_i).$$

Moreover, since  $u_i + \varepsilon_{iet}$  is normal with zero mean and variance  $1 + \sigma^2$ , it follows that

$$(5.C) \ P(u_i + V_{iet}b + \varepsilon_{iet} > 0) = P(V_{iet}b > -(u_i + \varepsilon_{iet})) = P\left(\frac{V_{iet}b}{\sqrt{1 + \sigma^2}} > -\frac{u_i + \varepsilon_{iet}}{\sqrt{1 + \sigma^2}}\right) = \Phi\left(\frac{V_{iet}b}{\sqrt{1 + \sigma^2}}\right).$$

By combining (4.C) and (5.C) we obtain that

$$P_1(V_{iet}) = E_u P_1(V_{iet}|u_i) = E_u \Phi(u_i + V_{iet}b) = \Phi\left(\frac{V_{iet}b}{\sqrt{1 + \sigma^2}}\right).$$

Since

$$k = \frac{1}{\sqrt{1 + \sigma^2}},$$

we have proved (3.C).

Our explanatory variables are binary, and here they have the value 1 or 0, depending on the characteristics of the GP or the patient. The exception is the income variable which in our case has the value 0 or 0.1. If we want to study the effect of a change in a binary variable, we do this by calculating the difference;  $P_1(V_i^*) - P_1(V_i)$ , where  $V_i^*$  is the vector of variables after the change and  $V_i$  is the vector of variables before the change.

We want to study the magnitude of the effect of significant variables on the probability of decisions A - C. As an example, we have used a GP with the following characteristics (after this named “our case”):

- male, 34 years old
- works in a group practice which is located in the most rural area
- has a specialist licence in general practice/family medicine

The patient is a child of 10 years and the level of income variable (the difference between the Fisher indices) is 0 or 0.1.

We study the effect of significant variables and calculate the probability for different laboratory services in general practice, by inserting the values of the coefficients from the results of the estimation of the random effects probit models and the characteristics in the formula

$$\Phi(k\alpha + k\beta X_{ie} + k\gamma I_t + k\theta Z_{ie} I_t).$$

When we consider the signs for the variables with interaction terms (patient’s age, GP’s age, specialist and group practice) we also need to include the parameter of the interaction term.

Below we show examples from decision B of how the calculations are done.

Decision B: The probability of analysing in the GP's office

$$\rho = 0.142 \text{ (ref appendix B), } k = \sqrt{1 - \rho} = (1 - 0.142)^{0.5} = 0.93.$$

When  $I_t = 0$ :

$$ka + k\beta X_{ie} + k\gamma I_t + k\theta Z_{ie} I_t = [(0.254 - 0.033 + 0.042 + 0.217 + 0.210(0))]0.93 = 0.446.$$

Tables of areas under the Standard Normal CDF show that the corresponding probability is 0.672, thus the probability for our case of analysing in-office is 67.2%.

$$P_1(V_i) = 67.2\%$$

Marginal effects

Effect of an increase in  $I_t$  by 0.08:  $0 \rightarrow 0.08$

$$ka + k\beta X_{ie} + k\gamma I_t + k\theta Z_{ie} I_t = [0.254 - 0.033 + 0.042 + 0.217 + 0.210(0.08)]0.93 = 0.461.$$

Tables of areas under the Standard Normal CDF show that the corresponding probability is 0.677, thus the probability for our case of analysing in-office is 67.7%.

The effect of the patient age in this case is an increase in the probability by

$$P_1(V_i^*) - P_1(V_i) = 67.7\% - 67.2\% = 0.5\%.$$

The effect of an increase in the patient age:  $10 \rightarrow 34$

$$ka + k\beta X_{ie} + k\gamma I_t + k\theta Z_{ie} I_t = [0.254 - 0.033 + 0.042 + 0.217 + 0.210(0)] - 0.51 + 0.137(0) 0.93 = -0.028$$

Tables of the Standard Normal CDF show that the corresponding probability is 0.489, thus the probability for our case of analysing in-office is  $P_1(V_i^*) = 48.9\%$ .

The effect of the patient age in this case is a decrease in the probability by

$$P_1(V_i^*) - P_1(V_i) = 48.9\% - 67.2\% = -18.3\%.$$