Ana Johnson
and
Thérèse Stukel
Medical Practice Variations
Health Services Research
10.1007/978-1-4899-7573-7_162-1

Practice Guidelines and Practice Variation: Diagnostic Technology in Maternity Care

Jostein Grytten[1, 3], Lars Monkerud[1] and Rune Sørensen[2]
(1)Institute of Community Dentistry, University of Oslo, Box 1052, Blindern, 0316 Oslo, Norway
(2)Norwegian School of Management, Oslo, Norway
(3)Department of Gynecology and Obstetrics, Akershus University Hospital, Lørenskog, Norway

Jostein Grytten (Corresponding author)
Email: josteing@odont.uio.no
Email: j.i.grytten@odont.uio.no

Lars Monkerud
Email: lars.monkerud@bi.no

Rune Sørensen
Email: rune.sorensen@bi.no

Abstract

One of the most common measures for reducing practice variation is the use of practice guidelines. They are meant to limit the use of services that are of little or questionable value and to increase the use of services that are cost-effective, but underused. A general finding is that guidelines often have a limited effect on practice variation, mainly because adherence is weak. This is supported by the findings from extensive reviews of professional behavior change interventions. The conclusion from these reviews is that for most interventions, the effects are small. Guidelines can be effective in reducing practice variation if physicians experience that they improve clinical decision making and/or treatment outcomes. Adherence to guidelines may then be high. This is supported by studies from maternity care in Norway. During the 1970s and 1980s, ultrasound and cardiotocography were introduced as diagnostic tools for deciding on mode of delivery. The guidelines for these diagnostic tools provide clear indications for which conditions and test values indicate Cesarean section. Objective criteria then replaced more subjective criteria based on clinical judgment. Clinical uncertainty about diagnosis of risk factors of the mother and child during delivery was then reduced. Variation in type of delivery between hospitals was reduced accordingly.

Introduction

One of the most striking findings from health service research during the last decades is the enormous variation in the use of health services across geographical areas. Variations persist, even when health-care needs, age, gender, and socioeconomic variables have been controlled for (for a review, see Paul-Shaheen et al. *1987*; Phelps *1992*). These findings have raised the question of what is the appropriate level of care – when are services overused or underused? Within the literature the main focus has been on overuse (Wennberg *2010*). For example, it is claimed that about 25 % of patients get health care that might be harmful, and 30–40 % of patients do not get treatment of proven effectiveness (Shuster et al. *2005*; Fisher et al. *2003a*, *b*). Lately, the issue of underuse has also been raised as a serious problem, as it can have adverse effects on health and quality of life (Freeman et al. *1987*; McGlynn et al. *2003*; Grol *2001*). Further, underuse might lead to health-care expenditure in the future that would otherwise not have been necessary.

For several diseases and health conditions, the correct level of care is not always obvious. However, it is assumed that reducing practice variation will lead to an appropriate level of care (Wennberg *1984*; Fisher et al. *2003a*), resulting in lower health-care costs and efficient health service delivery in areas typical of high costs. Patients in areas with low costs are likely to receive appropriate care and to benefit from improvements in health. This assumption is based on the notion that outliers tend to practice outside regular practice.

The most common intervention for reducing practice variation is the use of practice guidelines (Grimshaw and Russell *1993*; Grilli et al. *2000*; Hutchinson et al. *2003*). Guidelines are mainly based on systematic reviews of the available evidence on the effectiveness of health-care interventions (Eden et al. *2008*). They are meant to limit the use of services that are of little or questionable value and to increase the use of services that are cost-effective, but underused (Grimshaw and Russell *1993*).

A brief review of the main findings from the extensive literature on systematic reviews and clinical guidelines is given below. The results from a study that shows that the use of diagnostic technology in maternity care has led to a marked reduction in practice variation for Cesarean sections in Norway are then presented (Grytten et al. *2012*). Several studies show that adherence to guidelines is weak – hence, guidelines often have a limited effect on reducing practice variation. The study from Norway is interesting, as it shows that practice guidelines can be effective.

Changing Doctors' Practice Profile: Evidence from Systematic Reviews

In systematic reviews, explicit scientific methods are used to "identify, select, access, and summarize similar but separate studies" (Eden et al. *2008*, p. 82; Haynes et al. *2006*). The first reviews came in the 1970s. Researchers began to identify representative studies within a given topic and to summarize the evidence on the effectiveness of different types of health-care interventions. As early as the end of the 1990s, Bero and co-workers (*1998*) and Grimshaw and co-workers (*2001*) pointed out that there were so many systematic reviews that it could be difficult for policy makers and professionals to identify the most relevant reviews. In particular, it was a challenge to know which interventions should be chosen in cases for which different reviews gave different recommendations. Therefore, Bero and co-workers (*1998*) and Grimshaw and co-workers (*2001*), as the first researchers, started to summarize existing systematic reviews in their own overviews.

In relation to practice variation, the work of Grimshaw and co-workers (*2001*) is relevant in that they found that in order to reduce variation, some doctors must provide more treatment and some must

provide less. This demands a change in doctors' professional behavior. There are different interventions that can be used to achieve such changes in behavior. The effects of these are summarized by Grimshaw and co-workers (*2004a*) in the following way: "Dissemination of printed educational materials and didactic educational sessions are generally ineffective, while reminders, educational outreach and multifaceted interventions are generally effective. There are mixed effects of audit and feedback and local opinion leaders' programmes." Continuing medical education, defined as "any attempt to persuade physicians to modify their practice performance by communicating clinical information," had only a minor effect (Grimshaw et al. *2001*). Grimshaw and co-workers (*2004b*) later pointed out some methodological weaknesses with their review from 2001. Therefore, they carried out a new review in 2004, based on the material from their article from 2001. They then found that printed educational materials and didactic educational sessions were also effective. There was no change in the effects of the other interventions. However, the main conclusion when all the interventions were seen as a whole was that the effects "are small to modest" (Grimshaw et al. *2004a*). This conclusion indicates that it is not easy to influence doctors' practice profile, which implies that it may also be difficult to reduce practice variation with the help of clinical guidelines. This has been shown in several studies, which have shown that for many diseases and types of treatment, the adherence to guidelines is weak.

Clinical Practice Guidelines: Adherence Is Often Weak

In a study by O'Malley et al. (*2007*), only 39 % of primary care physicians in the United States reported that guidelines played a significant role in their clinical decision making in 2005. For specialists the figure was 28 %. Grol (*2001*) found that national guidelines were not followed for more than 30 % of clinical decisions within the field of family medicine in the Netherlands. Hayward and co-workers (*1997*) in a study from Canada concluded that physicians "have not yet integrated the use of guidelines into their practices to a large extent." Bauer (*2002*) reviewed 26 studies on adherence to guidelines in the field of mental health. He concluded that adequate adherence was found in only 27 % of the studies.

Cabana and co-workers (*1999*) summarized the results from 76 primary studies in which the reasons why physicians did not follow clinical practice guidelines were examined. The most important reason was that the standard and the quality of the guidelines are often poor, and the physicians see through that. Accordingly, some researchers have argued that guidelines are "unreliable, invalid, and irrelevant" (Grilli et al. *2000*). Shaneyfelt and co-workers (*1999*) examined 279 guidelines that were in use from 1985 to 1997. Only 43 % of the guidelines satisfied the criteria for an acceptable standard. A major weakness is that they often lack a clear scientific base – in particular there is no documented effect on outcome. A related problem is that guidelines are not updated quickly enough. A review by Shekelle and co-workers (*2001*) showed that half of the guidelines had become outdated after 5 years. Further, guidelines have a tendency to be oversimplified. For example, they can often be used only for patients with a single health condition and not for patients with multiple comorbidities (O'Connor *2005*). Further, guidelines are usually developed for a particular group of patients and do not address the health-care needs of a larger group of patients (Boyd et al. *2005*). Hence, the applicability at the individual patient level is often poor. Another problem is that even though guidelines exist, physicians are often not aware of them. One study showed that lack of awareness as a reason for not using the guidelines was as high as 84 % (Christakis and Rivara *1998*).

Guidelines are more likely to be followed if the patient and/or the physician do not have incentives that encourage them to ignore the recommendations. Examples of such undesirable incentives are

generous insurance arrangements and remuneration systems that stimulate excessive treatment. Fee-for-item payment may even lead to supplier-induced demand (Grytten and Sørensen *2001*). Ideally, physicians should supply services based on medical evaluation and social and patient costs, without regard to their private economic interests. But, because the patient is poorly informed, the physician has the possibility to influence both the type of diagnosis and the amount of treatment provided. Greater competition provides an incentive to exploit the information advantage. In particular, this is the case if more competition leads to a lack of patients, which then can lead physicians to induce demand for their services (Grytten and Sørensen *2008*). They do this in order to maintain their income. Conversely, if physicians' income is independent of the number of patients seen, there is no inducement to overtreatment (Carlsen and Grytten *1998*). In that case it is more likely that physicians would adhere to the guidelines or, in order to reduce their workload, may undertreat or underdiagnose. This would be the case for physicians who are paid a fixed salary (Sørensen and Grytten *2003*). Grytten and co-workers (*2012*) have shown that obstetricians in Norway follow practice guidelines in maternity care. Their study was performed within a standardized institutional health-care setting with public funding. In Norway, all health services, maternity services included, are financed through taxes, and everyone has free health care and equal access given equal need (Ministry of Health *2002*). All hospitals are publically owned and financed, and all hospital doctors receive a fixed salary. Therefore, obstetricians have no private economic incentives to ignore the guidelines. Grytten and co-workers (*2012*) examined the effect of the introduction of advanced diagnostic technology on variation for Cesarean sections among hospitals in Norway.

Intervention by Use of Diagnostic Technology in Maternity Care

During the last decades there has been a rapid development in diagnostic technology in maternity care (Neilson *1998*; Placek et al. *1984*; Norwegian Society for Gynecology and Obstetrics *2008*). This has improved fetal monitoring both before and during delivery. In particular, there are two diagnostic tools that are important and that were introduced in maternity units during the 1970s and 1980s: ultrasound and cardiotocography (Neilson *1998*; Van Geijn *1998*). Grytten and co-workers (*2012*) collected detailed information about when maternity units in Norway started to use ultrasound and cardiotocography. In 1970 none of the maternity units had these two types of diagnostic technology; in 1985 nearly all the hospitals had them (Table 1). Obstetricians' adherence to the guidelines for use of these two diagnostic tools was high. In 1985, 96 % of all pregnant women were examined by ultrasound. The mean number of examinations with ultrasound per woman was 2.45 (Backe et al. *1987*).**Table 1**

The percentage of deliveries according to the type of diagnostic technology and year

Diagnostic technology	1970	1975	1980	1985	1990	1995–2005
Ultrasound	0	19	78	97	99	100
Cardiotocography	0	36	86	96	100	100

Ultrasound and/or cardiotocography	0	45	92	98	100	100

Ultrasound is used to check fetal circulation and anatomy, both before and during delivery, and to determine the date of delivery and the number of fetuses (Neilson *1998*). Cardiotocography is used to register fetal distress, a condition that can lead to lack of oxygen and for which Cesarean section may be indicated. Both for ultrasound and cardiotocography, there are fairly clear medical criteria for the conditions and test values that are regarded as normal. For example, for cardiotocography, the signs that something is wrong are based on deviations from the following criteria: "a baseline fetal heart rate frequency between 110 and 150 beats per minute, presence of periodic accelerations, a normal heart rate variability with a bandwidth between 5 and 25 beats per minute, and the absence of decelerations" (Rooth et al. *1987*; Van Geijn *1998*). There is an agreement that in the case of deviation from these test criteria, the child must be delivered without delay, preferably by Cesarean section. For both gestation length and fetal growth, which are determined by ultrasound, there are fairly clear criteria for normal values and indications for Cesarean section (Morrison et al. *1995*; Backe and Buhaug *1986*).

Grytten and co-workers (*2012*) merged data on the use of diagnostic technology with data from the Medical Birth Registry of Norway for the period 1970–2005. The Medical Birth Registry of Norway contains information about risk factors of the mother and child that predispose to having a Cesarean section (Kolås et al. *2003*). The following risk factors of the mother were included in the analysis: age, immigrant status, and the presence of predisposing factors such as preeclampsia, bleeding during pregnancy, asthma, diabetes, epilepsy, heart disease, chronic hypertension, chronic kidney failure, and rheumatoid arthritis. Risk factors of the baby were low birthweight, abnormal presentation, and multiple births. The data were analyzed using multilevel analysis. Thus, it was possible to calculate variations in the Cesarean section rate for different maternity units after controlling for relevant risk factors of the mother and the baby at the individual level. Previous studies of practice variation have usually been carried out with aggregated data at the level of an area or a district (for a review, see Paul-Shaheen et al. *1987*). Control variables for health-care needs have mainly been age and gender at the relevant level of analysis. The results from these studies can be biased for several reasons. Relevant control variables for health-care needs have been omitted. Analyses of aggregated data do not take into account the multilevel structure of the data. This weakens the statistical tests, and measurement errors are introduced with aggregated data (Diehr et al. *1990*).

The variation in Cesarean section between hospitals was large at the beginning of the period, until 1973 (Fig. 1) (Grytten et al. *2012*). This variation diminished markedly until the late 1970s. The fall in the inter-hospital variation in Cesarean section occurred at the time when there was a marked increase in use of ultrasound and cardiotocography. During this time period, the use of ultrasound increased about fourfold, and the use of cardiotocography increased about threefold (Table 1). In 1985, nearly all the hospitals had ultrasound and cardiotocography.

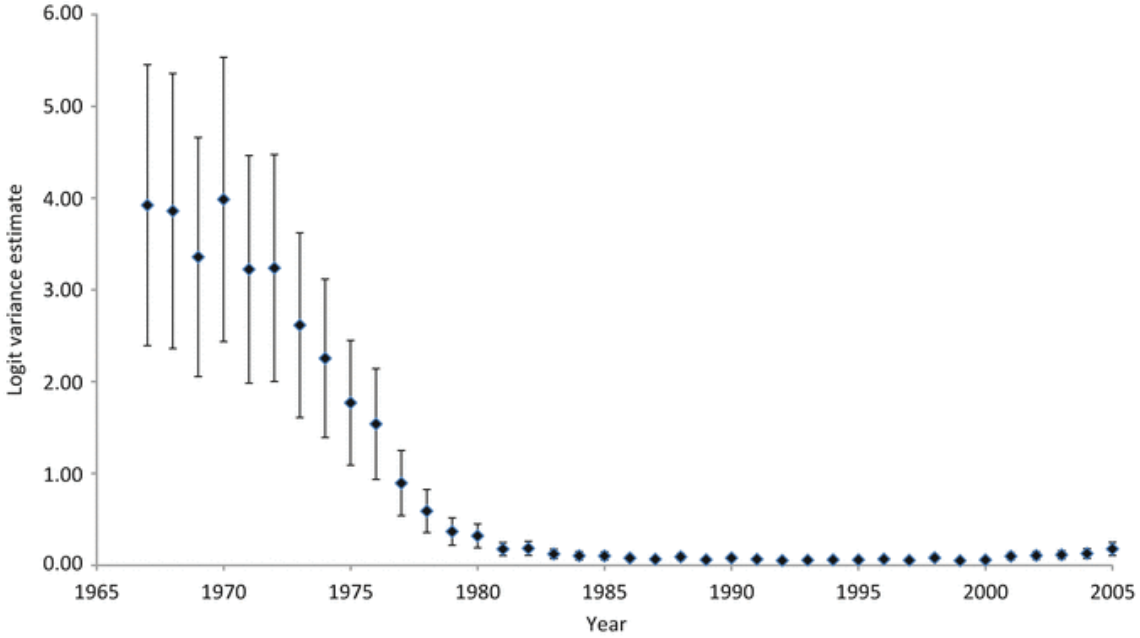

Fig. 1
Inter-hospital variation in Cesarean section (logit) according to year. Estimates of annual random effects and 90 % confidence intervals (*n* = 1,708,681)

Grytten and co-workers (*2012*) carried out several types of analyses to test whether there was an inverse causal relationship between the effects of the use of diagnostic technology and the Cesarean section rate among hospitals. The two most important analyses were a placebo test and a test of reversed causality.

The placebo test was performed by setting the time for receiving the new technology 5 years before the hospital actually got it. If the technology variable had an effect on practice variation for Cesarean sections 5 years before the hospital actually got the diagnostic technology, this would indicate that the technology effect was not causal. The placebo test showed that the effect was causal. In the main model of Grytten and co-workers (*2012*), it was assumed that the direction of the effect went from diagnostic technology to practice variation. The opposite direction was also a possibility. Reversed causality would occur if obstetricians were concerned about unwarranted variation in the Cesarean section rate between hospitals and as a response to that concern began to use ultrasound and cardiotocography. The authors found no support for reversed causality. Therefore, on the basis of both the placebo test and the test for reversed causality, there is reason to conclude that use of diagnostic technology had a causal effect on the reduction in the Cesarean section rate between hospitals.

For both ultrasound and cardiotocography, there are clear guidelines for the range of conditions and test values that represent deviations from normal values and for what range a Cesarean section is indicated. The results in Fig. 1 show that these guidelines have been followed by the obstetricians. By using diagnostic technology, clinical uncertainty about diagnosis of risk factors of the mother and child during delivery was reduced. Variation in type of delivery between hospitals was reduced accordingly. This follows directly from Wennberg and co-workers' (*1982*) theory on how practice variation arises.

Less Uncertainty in Diagnosis Leads to Less Practice Variation

According to Wennberg and co-workers (*1982*), variation in medical care exists, to a large extent, because of clinical uncertainty about diagnosis and treatment. This uncertainty has resulted in lack of unambiguous diagnostic criteria and appropriate standards for the care that is given. For example, diagnostic findings are not necessarily clear and can be interpreted differently by different doctors. Also, there is not necessarily a direct relationship between diagnostic findings and the treatment that should be chosen. Differences among physicians in their beliefs about how to interpret diagnostic signs and the benefits of alternative interventions lead to differences in clinical practice between physicians – they have developed their own practice style. These beliefs are formed by many factors, for example, where the physicians have been trained, their specialty and skills, the amount of scientific information they have on outcomes, the means by which such information is disseminated by physicians, and the practice culture of the area. The mix of physicians with different practice styles varies between areas. Therefore, level of utilization of health care per capita also varies between areas (Grytten and Sørensen *2003*).

Prior to the introduction of ultrasound and cardiotocography, diagnostic tools for deciding on the mode of delivery were not sufficiently precise. Obstetricians were dependent on their clinical judgment, which was mainly based on anamnestic information and clinical examination of the pregnant woman. It was particularly important to identify possible risk factors associated with complications and adverse health outcomes for the mother and the baby. This would indicate that a Cesarean section would be the preferred mode of delivery. Important risk factors that could easily be identified by the obstetrician are previous Cesarean section, mother's age, vaginal bleeding during pregnancy, and poor health status (Kolås et al. *2003*; Penn and Ghaem-Maghami *2001*; Cleary-Goldman et al. *2005*; Ananth and Savitz *1994*; O'Brien-Abel *2003*). In the United States, in 1981, only 3 % of all women who had previously had a Cesarean section had a vaginal delivery later (Harer *2002*). This figure had increased to 20.6 % in 2000 (Price and Simon *2009*).

Even if one or more of the risk factors mentioned above are present, most vaginal births appear to be successful. Only a few are likely to have a poor health outcome for the baby or the mother. Before diagnostic technology was available, obstetricians did not have sufficient information to predict which births would go well and which births would not go well. Therefore, in order to be safe, Cesarean section was often the first choice. Some Cesarean sections would have been unnecessary.

With the introduction of ultrasound and cardiotocography, obstetricians have been able to make informative decisions regarding the mode of delivery as the condition of the fetus and the mother can be observed directly. For example, the delivery could be vaginal, even if risk factors are present. Figures 2, 3, 4, and 5 show the reduction in the odds ratio for a Cesarean section delivery before and after diagnostic technology was introduced for the following risk factors of the mother: previous Cesarean section, age, vaginal bleeding during pregnancy, and poor health status. The latter is present if the mother has one or more of the following conditions: asthma, diabetes, epilepsy, heart disease, chronic hypertension, chronic kidney failure, or rheumatoid arthritis. The effect of diagnostic technology is particularly prominent for mothers who have had a Cesarean section previously. If the previous child had been delivered by Cesarean section, the odds ratio for a Cesarean section delivery was 36.2 before ultrasound and cardiotocography were introduced and 11.1 after. For women over 40 years of age, the odds ratio for a Cesarean section delivery was 4.3 before diagnostic technology was introduced and 2.9 after. Similar results were found for the other risk factors. Women with poor health status and vaginal bleeding during the pregnancy did not give birth so often with Cesarean section after ultrasound and cardiotocography were introduced.

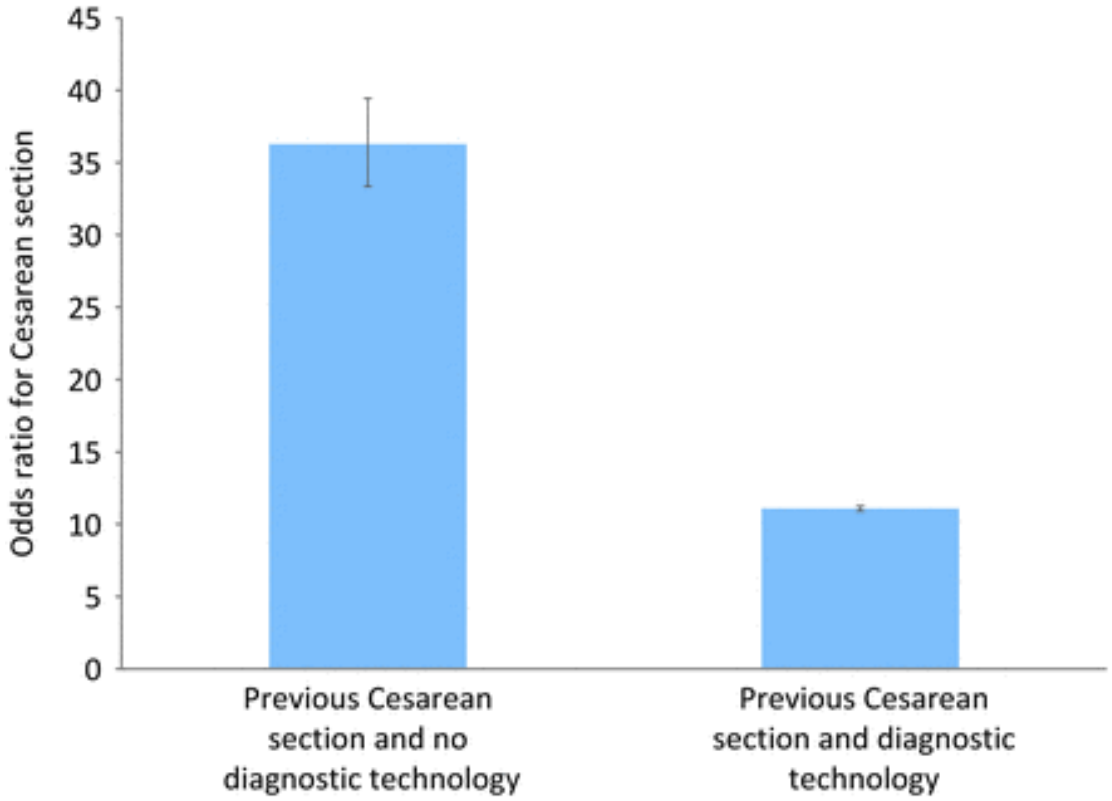

Fig. 2
Odds ratio with a 95 % confidence interval for Cesarean section according to previous Cesarean section and use of diagnostic technology. Reference category: no previous Cesarean section

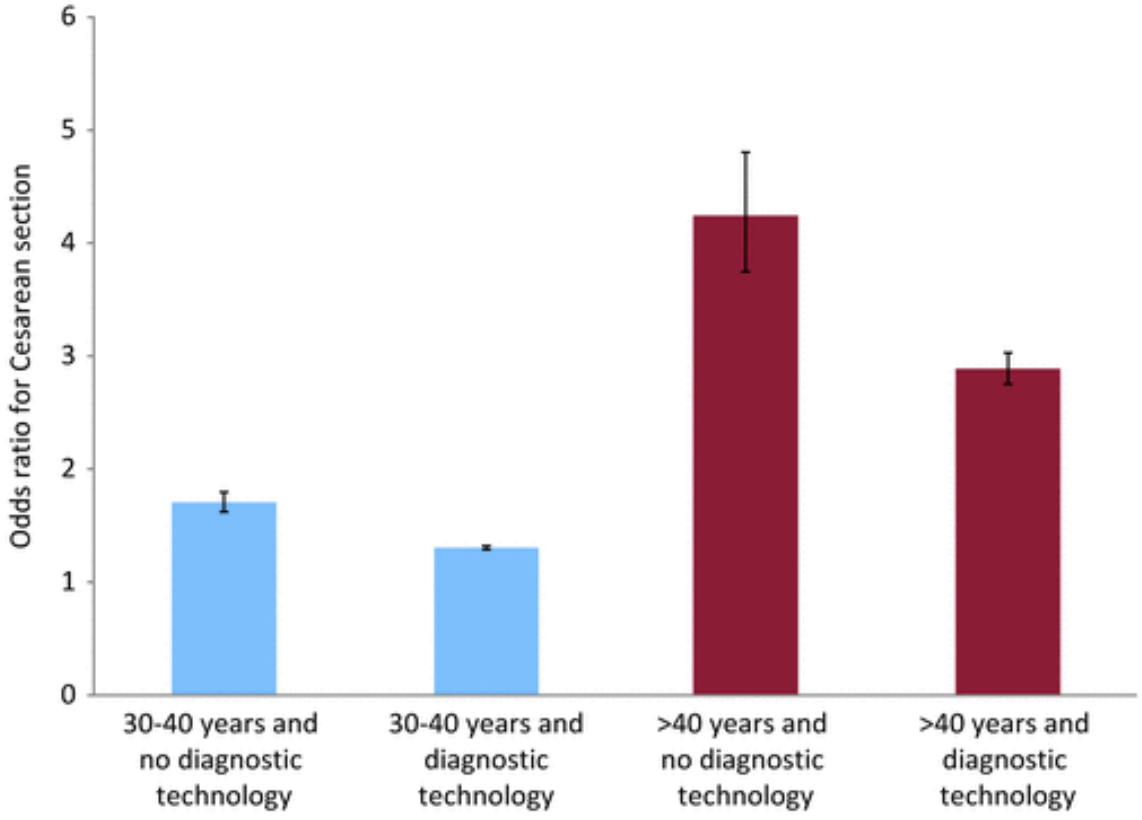

Fig. 3
Odds ratio with a 95 % confidence interval for Cesarean section according to mother's age and use of diagnostic technology. Reference category: <30 years

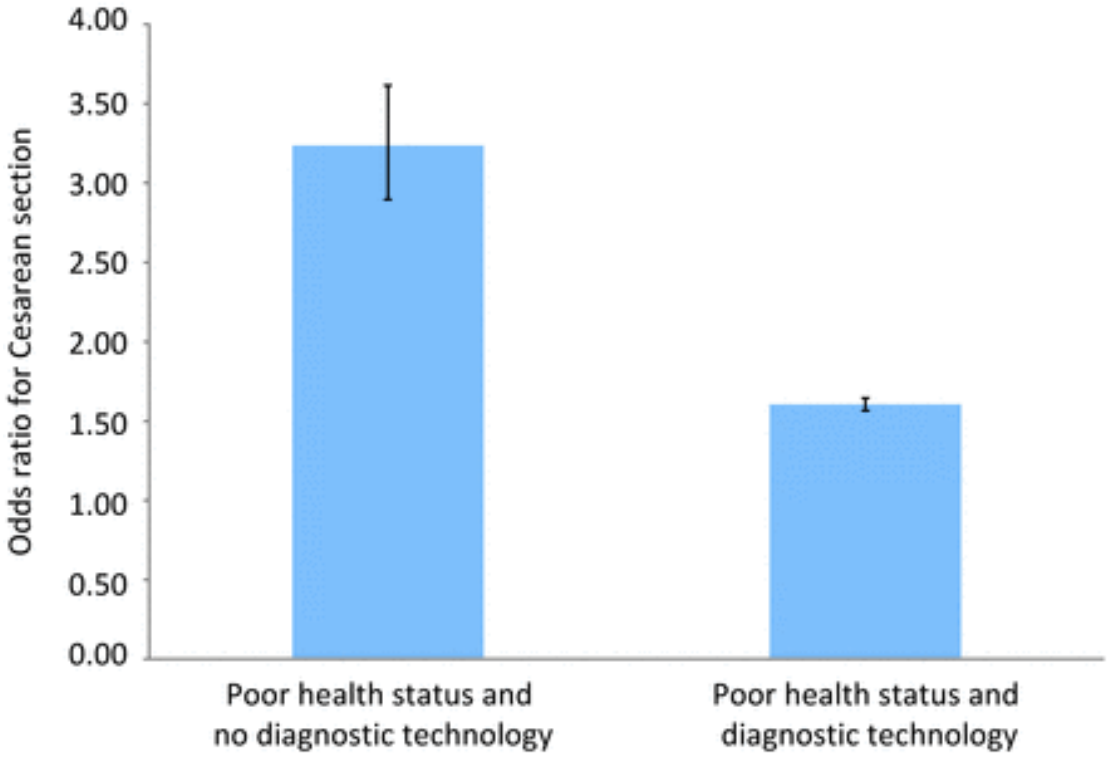

Fig. 4
Odds ratio with a 95 % confidence interval for Cesarean section according to mother's health status and use of diagnostic technology. Reference category: good health status

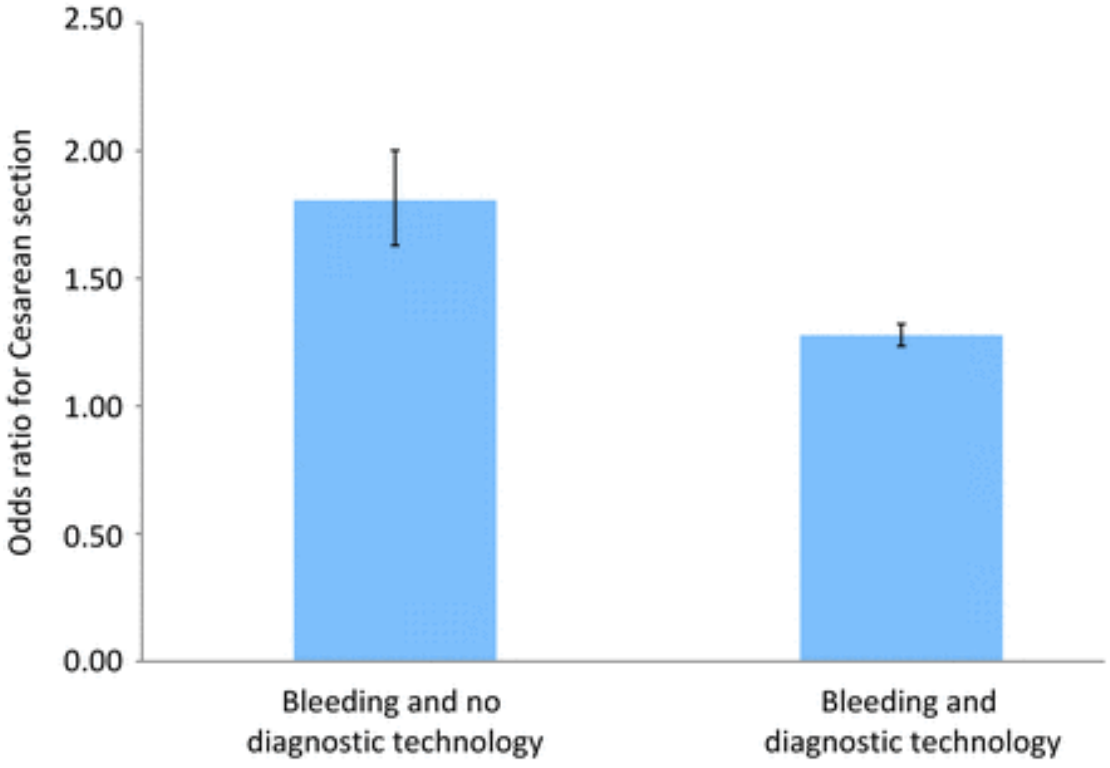

Fig. 5
Odds ratio with a 95 % confidence interval for Cesarean section according to bleeding during pregnancy and use of diagnostic technology. Reference category: no bleeding

The results in Figs. 2, 3, 4, and 5 show that fewer Cesarean sections are carried out when low-precision diagnostic criteria are replaced with high level of precision. Heterogeneity in choice of method of delivery between maternity units is reduced as a result of the introduction of new diagnostic technology. Therefore, practice variation between maternity units is also reduced (Fig. 1). It would have been desirable to have had some qualitative data demonstrating that obstetricians had changed their clinical practice because they trusted the new diagnostic technology. Unfortunately, such data are not available. There are numerous reports (written in Norwegian) that deal with different aspects of maternity care. None of these reports deal with whether the introduction of diagnostic technology made obstetricians change their clinical practice.

Less Uncertainty Weakens the Mothers' Position in Maternity Care

Another mechanism through which the use of diagnostic technology reduces practice variation in maternity care is through reduced influence on the part of the mother on the mode of delivery. When meeting the mother, the obstetrician must relate not only to diagnostic findings but also to the subjective symptoms she presents. Obstetricians must supplement clinical data with information about how the mother experiences her condition. They must listen to what she has to say. It may be that obstetricians listen more to mothers who are well informed and who are skilled at communicating. Grytten and co-workers (*2011b*) defined well-informed mothers as mothers who were either doctors, midwives, or obstetricians. Mothers with other types of university education and mothers with upper secondary school education were classified as partly informed. Mothers with only compulsory school education were classified as not well informed.

Well-informed mothers are able to efficiently register and assess the implications of symptoms for their health and their child's health and are probably also better able to communicate their assessment to the obstetrician compared with mothers who are less well informed. Further, the well-informed mother may be better able to respond to the information she receives from the obstetrician and to participate in the decision process about the choice of method of delivery than a mother who is not so well informed (Grytten et al. *2011b*). Therefore, well-informed mothers may more often engage in the mode of delivery they desire. If this is the case, then nonmedical characteristics of the mother, such as her preferences for mode of delivery, may have just as much impact on the outcome of treatment than purely medical indications (Balsa et al. *2003*; Grytten et al. *2013*).

There is a comprehensive literature, related to statistical discrimination, that describes how clinical uncertainty favors well-informed patients (Balsa and McGuire *2001*, *2003*; Maserejian et al. *2009*). This is one explanation for the great differences we see in health service consumption according to patients' level of education and ethnicity (Lê Cook et al. *2009*; Gaskin et al. *2012*; McKinlay et al. *2012*; Braveman et al. *1995*). For example, Balsa and co-workers (*2005*) and McGuire and co-workers (*2008*) found that ethnic minority patients are less likely than whites to be diagnosed with depression. This is partly due to concordance of language, culture, education, and class between patient and physician. The differences at the patient level are reflected in the variation at the district or area level, since the composition of the patients in relation to education and ethnicity varies according to geographical region.

Since the use of diagnostic technology reduces clinical uncertainty, the obstetrician is less dependent on judgment and interpretation of information from the mother for assessing whether the delivery is progressing without complications. That weakens the position of the mother in maternity care. Therefore, and according to statistical discrimination theory, we expect that improved diagnostic technology should lead to smaller differences in Cesarean section rates between well-informed and less well-informed mothers. This is supported by the results of Grytten and co-workers (*2011b*). Before ultrasound and cardiotocography were introduced, the odds ratio for having a Cesarean section delivery was 2.34 if the mother was a midwife (Fig. 6). This was reduced to 1.06 after the introduction of diagnostic technology. The corresponding odds ratios for doctors were 1.82 before and 1.01 after diagnostic technology was introduced. The difference in the Cesarean section rates between partly informed mothers also disappeared after diagnostic technology was introduced. Mothers' level of education is no longer decisive for the method of delivery that is chosen. This contributes to reducing variation in the rate of Cesarean section between maternity units.

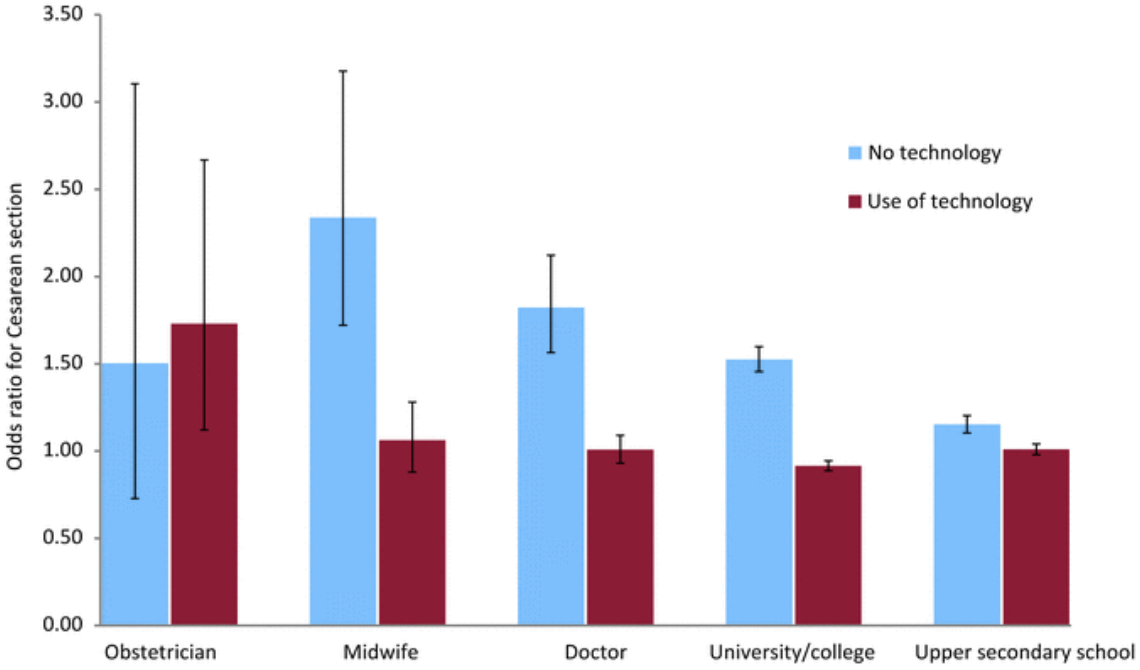

Fig. 6
Odds ratio with a 95 % confidence interval for Cesarean section for different education groups and use of diagnostic technology. Reference category: compulsory school

In another study, Grytten and co-workers (*2013*) found that preferences have a causal effect on the likelihood of Cesarean section. The effect diminished after diagnostic technology was introduced. Precise diagnostic tools improve the medical basis for taking decisions about the method of delivery. The level of clinical uncertainty has been reduced, which has weakened the position of the mothers in maternity care. This further contributes to less practice variation in maternity care.

Use of Diagnostic Technology: Effects on Infant Mortality and on the Organization of Maternity Services

Most likely, introduction of the use of diagnostic technology in maternity care in the 1970s contributed to a decrease in the infant mortality rate. In addition, the use of diagnostic technology has been important for maintaining decentralized maternity services in Norway.

From 1970 to 1980 the Cesarean section rate in Norway increased from 0.02 to 0.07 (Grytten et al. *2011a*). According to the World Health Organization, there is consensus that the Cesarean section rate should be in the range 10–15 % in any country in order to safeguard the medical needs of the mother and the fetus (WHO *1985*). Thus, it can be argued that the reduction in the variation in Cesarean section rate between hospitals during the 1970s reduced the level of underuse. There was an increase in the Cesarean section rate in hospitals with a low rate in the 1970s, while the rate remained unchanged in hospitals with a high rate. This may be one factor that has contributed to the marked reduction in maternal and infant mortality in Norway during the 1970s (Medical Birth Registry of Norway *2013*).

Availability and use of advanced diagnostic technology is important for maintaining a decentralized maternity service. In Norway hospitals are classified into two groups: those with their own neonatal department for dealing with high-risk deliveries (central and regional hospitals) and those that do not have their own neonatal department (local hospitals). As a general rule, mothers shall give birth at the maternity unit they belong to, according to the hospital's catchment area. High-risk mothers who live in the catchment area of a local hospital shall be referred to central or regional hospitals that have their own neonatal department for dealing with high-risk deliveries (Ministry of Health and Care Services *2009*). These mothers are identified at antenatal checkups in good time before the expected date of delivery. At these checkups all mothers have routine ultrasound screening. This screening is effective. For example, Grytten and co-workers (*2014*) have shown that the decentralized maternity service in Norway does not lead to increased neonatal and infant mortality. This is because very few high-risk deliveries occur at local hospitals. These mothers are identified by ultrasound screening and referred to a large hospital with sufficient perinatal resources to deal with these deliveries. For example, in 2011, only six babies with a very low birthweight (<1500 g) were born in local hospitals. The corresponding number for very preterm infants (gestation length < 32 weeks) was eight. These are small numbers considering that there were 10,618 deliveries in local hospitals in 2011.

Conclusion

One of the most common measures for reducing practice variation is the use of practice guidelines. There are numerous guidelines, and they have been developed for most diseases and types of treatment. A general finding is that guidelines often have a limited effect on practice variation, mainly because adherence is weak. The standard and quality of guidelines are often poor, and they often lack a clear scientific base. Another problem is that even though guidelines exist, physicians are not necessarily aware of them, and when they are aware of them, they do not necessarily follow them. This is supported by the findings from extensive reviews of professional behavior change interventions. The conclusion from these reviews is that for most interventions, the effects "are small to modest" (Grimshaw et al. *2004a*).

Guidelines can be effective in reducing practice variation if physicians experience that they improve clinical decision making and/or treatment outcomes. Adherence to guidelines may then be high. This is supported by studies from maternity care in Norway. During the 1970s and 1980s, ultrasound and cardiotocography were introduced as diagnostic tools for deciding on mode of delivery. The guidelines for these diagnostic tools provide clear indications for which conditions and test values indicate Cesarean section. Objective criteria then replaced more subjective criteria based on clinical

judgment. Clinical uncertainty about diagnosis of risk factors of the mother and child during delivery was then reduced. Variation in type of delivery between hospitals was reduced accordingly (Grytten et al. *2012*).

Acknowledgments
We wish to thank the Medical Birth Registry and Statistics, Norway, for providing data. This review had financial support from the South-Eastern Norway Health Authority, research grant number 2709002.

References

Ananth CV, Savitz DA. Vaginal bleeding and adverse reproductive outcomes: a meta-analysis. Paediatr Perinat Epidemiol. 1994;8:62–78.
PubMed CrossRef

Backe B, Buhaug H. Bruk av ultralyd i svangerskapet. NIS-rapport 8/86. Trondheim: Norsk institutt for sykehusforskning; 1986.

Backe B, Jacobsen G, Bakketeig LS, et al. Bruk av ultralyd ved norske fødeinstitusjoner. Tidsskr Nor Legeforen. 1987;107:471–3.

Balsa AI, McGuire TG. Statistical discrimination in health care. J Health Econ. 2001;20:881–907.
PubMed CrossRef

Balsa AI, McGuire TG. Prejudice, clinical uncertainty and stereotyping as sources of health disparities. J Health Econ. 2003;22:89–116.
PubMed CrossRef

Balsa AI, Seiler N, McGuire TG, et al. Clinical uncertainty and healthcare disparities. Am J Law Med. 2003;29:203–19.
PubMed

Balsa AI, McGuire TG, Meredith LS. Testing for statistical discrimination in health care. Health Serv Res. 2005;40:227–52.
PubMedCentral PubMed CrossRef

Bauer MS. A review of quantitative studies of adherence to mental health clinical practice guidelines. Harv Rev Psychiatry. 2002;10:138–53.
PubMed CrossRef

Bero LA, Grilli R, Grimshaw JM, et al. Closing the gap between research and practice: an overview of systematic reviews of interventions to promote the implementation of research findings. BMJ. 1998;317:465–8.
PubMedCentral PubMed CrossRef

Boyd CM, Darer J, Boult LP, et al. Clinical practice guidelines and quality of care for older patients with multiple comorbid diseases: implications for pay for performance. JAMA. 2005;294:716–24.
PubMed CrossRef

Braveman P, Egerter S, Edmonston F, et al. Racial/ethnic differences in the likelihood of cesarean delivery, California. Am J Public Health. 1995;85:625–30.
PubMedCentral PubMed CrossRef

Cabana MD, Rand CS, Powe NR, et al. Why don't physicians follow clinical practice guidelines? JAMA. 1999;282:1458–65.
PubMed CrossRef

Carlsen F, Grytten J. More physicians: improved availability or induced demand? Health Econ. 1998;7:495–508.
PubMed CrossRef

Christakis DA, Rivara FP. Pediatricians' awareness of and attitudes about four clinical practice guidelines. Pediatrics. 1998;101:825–30.
PubMed CrossRef

Cleary-Goldman J, Malone FD, Vidaver J, et al. Impact of maternal age on obstetric outcome. Obstet Gynecol. 2005;105:983–90.
PubMed CrossRef

Diehr P, Cain K, Connell F, et al. What is too much variation? The null hypothesis in small-area analysis. Health Serv Res. 1990;24:741–77.
PubMedCentral PubMed

Eden J, Wheatley B, McNeil B, et al., editors. Knowing what works in health care. A roadmap for the nation. Washington, DC: The National Academies Press; 2008. p. 81–120.

Fisher ES, Wennberg DE, Stukel TA, et al. The implications of regional variations in Medicare spending. Part 1: the content, quality, and accessibility of care. Ann Intern Med. 2003a;138:273–87.
PubMed CrossRef

Fisher ES, Wennberg DE, Stukel TA, et al. The implications of regional variations in Medicare spending. Part 2: health outcomes and satisfaction with care. Ann Intern Med. 2003b;138:288–98.
PubMed CrossRef

Freeman HE, Blendon RJ, Aiken LH, et al. Americans report on their access to health care. Health Aff. 1987;6:6–8.
CrossRef

Gaskin DJ, Dinwiddie GY, Chan KS, et al. Residential segregation and disparities in health care services utilization. Med Care Res Rev. 2012;69:158–75.
PubMedCentral PubMed CrossRef

Grilli R, Magrini N, Penna A, et al. Practice guidelines developed by specialty societies: the need for a critical appraisal. Lancet. 2000;355:103–6.
PubMed CrossRef

Grimshaw JM, Russell IT. Effect of clinical guidelines on medical practice: a systematic review of rigorous evaluations. Lancet. 1993;342:1317–22.
PubMed CrossRef

Grimshaw JM, Shirran L, Thomas R, et al. Changing provider behavior. An overview of systematic reviews of interventions. Med Care. 2001;39:II-2–45.
CrossRef

Grimshaw JM, Eccles M, Tetroe J. Implementing clinical guidelines: current evidence and future implications. J Contin Educ Health Prof. 2004a;24:S31–7.
PubMed CrossRef

Grimshaw JM, Thomas RE, MacLennan G, et al. Effectiveness and efficiency of guideline dissemination and implementation strategies. Health Technol Assess. 2004b;8:1–72.

CrossRef

Grol R. Successes and failures in the implementation of evidence-based guidelines for clinical practice. Med Care. 2001;39:II-46–54.
CrossRef

Grytten J, Sørensen R. Type of contract and supplier-induced demand for primary physicians in Norway. J Health Econ. 2001;20:379–93.
PubMed CrossRef

Grytten J, Sørensen R. Practice variation and physician-specific effects. J Health Econ. 2003;22:403–18.
PubMed CrossRef

Grytten J, Sørensen R. Busy physicians. J Health Econ. 2008;27:510–8.
PubMed CrossRef

Grytten J, Monkerud L, Hagen TP, et al. The impact of hospital revenue on the increase in Caesarean sections in Norway. A panel data analysis of hospitals 1976–2005. BMC Health Serv Res. 2011a;11:267.
PubMedCentral PubMed CrossRef

Grytten J, Skau I, Sørensen R. Do expert patients get better treatment than others? Agency discrimination and statistical discrimination in obstetrics. J Health Econ. 2011b;30:163–80.
PubMed CrossRef

Grytten J, Monkerud L, Sørensen R. Adoption of diagnostic technology and variation in Caesarean section rates: a test of the practice style hypothesis in Norway. Health Serv Res. 2012;47:2169–89.
PubMedCentral PubMed CrossRef

Grytten J, Skau I, Sørensen R. Do mothers decide? The impact of preferences in healthcare. J Hum Resour. 2013;48:142–68.
CrossRef

Grytten J, Monkerud L, Skau I, et al. Regionalization and local hospital closure in Norwegian maternity care – the effect on neonatal and infant mortality. Health Services Research. 2014;49:1184–1204.

Harer WB. Vaginal birth after cesarean delivery. JAMA. 2002;287:2627–30.
PubMed CrossRef

Haynes RB, Sackett DL, Guyatt GH, et al. Clinical epidemiology. How to do clinical practice research. 3rd ed. Philadelphia: Lippincott Williams & Wilkins; 2006.

Hayward RSA, Guyatt GH, Moore KA, et al. Canadian physicians' attitudes about and preferences regarding clinical practice guidelines. Can Med Assoc J. 1997;156:1715–23.

Hutchinson A, McIntosh A, Cox S, et al. Towards efficient guidelines: how to monitor guideline use in primary care. Executive summary. Health Technol Assess. 2003;7(18):1–97.
CrossRef

Kolås T, Hofoss D, Daltveit AK, et al. Indications for cesarean delivery in Norway. Am J Obstet Gynecol. 2003;188:864–70.
PubMed CrossRef

Lê Cook B, McGuire TG, Zuvekas SH. Measuring trends in racial/ethnic health care disparities. Med Care Res Rev. 2009;66:23–48.
PubMedCentral PubMed CrossRef

Maserejian NN, Lutfey KE, McKinlay JB. Do physicians attend to base rates? Prevalence data and statistical discrimination in the diagnosis of coronary heart disease. Health Serv Res. 2009;44:1933–49.
PubMedCentral PubMed CrossRef

McGlynn EA, Asch SM, Adams J, et al. The quality of health care delivered to adults in the United States. N Engl J Med. 2003;348:2635–45.
PubMed CrossRef

McGuire TG, Ayanian JZ, Ford DE, et al. Testing for statistical discrimination by race/ethnicity in panel data for depression treatment in primary care. Health Serv Res. 2008;43:531–51.
PubMedCentral PubMed CrossRef

McKinlay JB, Marceau LD, Piccolo RJ. Do doctors contribute to the social patterning of disease? The case of race/ethnic disparities in diabetes mellitus. Med Care Res Rev. 2012;69:176–93.
PubMedCentral PubMed CrossRef

Medical Birth Registry of Norway. MBRN's statistic bank. 2013. http://mfr-nesstar.uib.no/mfr/. Accessed 17 Jan 2013.

Ministry of Health. Behovsbasert finansiering av spesialisthelsetjenesten. Oslo: Ministry of Health; 2002.

Ministry of Health and Care Services. En gledelig begivenhet. St. meld. nr. 12 (2008–2009). Oslo: Ministry of Health and Care Services; 2009. p. 41–51.

Morrison JJ, Rennie JM, Milton PJ. Neonatal respiratory morbidity and mode of delivery at term: influence of timing of elective caesarean section. BJOG. 1995;102:101–6.
CrossRef

Neilson JP. Ultrasound for fetal assessment in early pregnancy (review). Cochrane Database Syst Rev. 1998:1–17. http://onlinelibrary.wiley.com/doi/10.1002/14651858.CD000182/pdf/standard. Accessed 17 Jan 2012.

Norwegian Society for Gynecology and Obstetrics. Veileder i fødselshjelp 2008. 2008. http://www.legeforeningen.no/id/131068.0. Accessed 17 Jan 2012.

O'Brien-Abel N. Uterine rupture during VBAC trial of labor: risk factors and fetal response. J Midwifery Womens Health. 2003;48:249–57.
PubMed CrossRef

O'Connor PJ. Adding value to evidence-based clinical guidelines. JAMA. 2005;294:741–3.
PubMed CrossRef

O'Malley AS, Pham HH, Reschovsky JD. Predictors of the growing influence of clinical practice guidelines. J Gen Intern Med. 2007;22:742–8.
PubMedCentral PubMed CrossRef

Paul-Shaheen P, Clark JD, Williams D. Small area analysis: a review and analysis of the North American literature. J Health Polit Policy Law. 1987;12:741–809.
PubMed CrossRef

Penn Z, Ghaem-Maghami S. Indications for caesarean section. Best Pract Res Clin Obstet Gynaecol. 2001;15:1–15.
PubMed CrossRef

Phelps CE. Diffusion of information in medical care. J Econ Perspect. 1992;6:23–42.
PubMed CrossRef

Placek PJ, Keppel KG, Taffel SM, et al. Electronic fetal monitoring in relation to cesarean section delivery, for live births and stillbirths in the U.S., 1980. Public Health Rep. 1984;99:173–83.
PubMedCentral PubMed

Price J, Simon K. Patient education and the impact of new medical research. J Health Econ. 2009;28:1166–74.
PubMed CrossRef

Rooth G, Huch A, Huch R. FIGO News. Guidelines for the use of fetal monitoring. Int J Gynaecol Obstet. 1987;25:159–67.
CrossRef

Shaneyfelt TM, Mayo-Smith M, Rothwangl J. Are guidelines following guidelines? The methodological quality of clinical practice guidelines in the peer-reviewed medical literature. JAMA. 1999;281:1900–5.
PubMed CrossRef

Shekelle PGE, Ortiz S, Rhodes SC, et al. Validity of the agency for healthcare research and quality clinical practice guidelines: how quickly do guidelines become outdated? JAMA. 2001;286:1461–7.
PubMed CrossRef

Shuster MA, McGlynn EA, Brook RH. How good is the quality of health care in the United States? Milbank Q. 2005;83:843–95.
CrossRef

Sørensen RJ, Grytten J. Service production and contract choice in primary physician services. Health Policy. 2003;66:73–93.
PubMed CrossRef

Van Geijn HP. Cardiotocography. In: Kurjak E, editor. Textbook of perinatal medicine, vol. 2. London: Parthenon Publishing; 1998. p. 1424–8.

Wennberg JE. Dealing with medical practice variations: a proposal for action. Health Aff. 1984;3:6–32.
CrossRef

Wennberg JE. Tracking medicine. New York: Oxford University Press; 2010. p. 209–24.

Wennberg JE, Barnes BA, Zubkoff M. Professional uncertainty and the problem of supplier-induced demand. Soc Sci Med. 1982;16:811–24.
PubMed CrossRef

WHO. Appropriate technology for birth. Lancet. 1985;326:436–7.
CrossRef