

ORIGINAL ARTICLE

## The effects of gatekeeping: A systematic review of the literature

MARCIAL VELASCO GARRIDO, ANNETTE ZENTNER & REINHARD BUSSE

*Department of Health Care Management, Berlin University of Technology, Berlin, Germany*

### Abstract

**Objective.** To assess the effects of physician-centred gatekeeping on health, health care utilization, and costs by conducting a systematic review of the literature. **Methods.** Systematic search in PubMed (MEDLINE and Pre-MEDLINE), EMBASE, and the Cochrane Library, from the databases' respective inception dates up to January 2010, using the search words "gatekeeping", "gatekeeper", "first contact", and "self-referral". We included RCTs, CCTs, cohort studies, CBAs, and interrupted time-series. We included only studies in which the gatekeeper function was exercised by a physician and that reported health and patient-related outcomes including quality of life and satisfaction, quality of care, health care utilization, and/or economic outcomes (e.g. expenditures or efficiency). Selection was made independently by two reviewers and discrepancies were solved by consensus after discussion. Data on target population, intervention, additional interventions, study results, and methodological quality were extracted. Methodological quality was assessed independently by two reviewers following the previously defined criteria. Discrepancies were solved by consensus after discussion. **Results.** This review includes 26 studies in 32 publications. The majority of studies (62%) reported data from the United States and in most gatekeeping was associated with lower utilization of health services (up to -78%) and lower expenditures (up to -80%). However, there was great variability in the magnitude and direction of the differences. **Conclusion.** Overall, the evidence regarding the effects of gatekeeping is of limited quality. Many studies are available regarding the effects on health care utilisation and expenditures, whereas effects on health and patient-related outcomes have been studied only exceptionally and are inconclusive.

**Key Words:** *Gatekeeping, health care costs, health care utilization, primary health care, systematic review*

Orienting a health system towards primary care can enhance the continuity and coordination of care, thus reducing the inappropriate use of specialty services and improving a population's health [1]. One of the features of primary care-based health systems is the requirement to visit a generalist – acting as gatekeeper and coordinator of care – prior to accessing further specialty care. In Europe, gatekeeping is encountered both in tax-funded health systems, such as those in the United Kingdom and Spain, and in social health insurance systems, such as those in Switzerland and the Netherlands. Ecological evidence from multi-country comparisons suggests that lack of direct access to speciality care is associated with lower levels of expenditure on ambulatory care [2].

Gatekeeping is common in the United States, especially in the context of managed care [3]. In past

years, the cons of managed care and particularly of gatekeeping arrangements have been debated (e.g. restrictive access to care, potential delays). Since the late 1990s a managed-care backlash has been made evident by a decline in health maintenance organizations' enrolment, especially when enrolees are given a choice [4]. Despite increasing criticism, gatekeeping has continued to be one of the major tools of health maintenance organizations [5].

The question remains whether gatekeeping can contribute to improving health and quality of care and at the same time containing health expenditures. The purpose of this systematic review is thus to assess the effects of physician-centred gatekeeping on health outcomes, including health-related quality of life, quality of care, use of health services, and expenditures.

Gatekeeping is an international policy issue. There are a considerable number of publications on gatekeeping requiring a systematic approach to summarize the evidence.

- Overall, the evidence regarding the effects of gatekeeping is of limited quality.
- The majority of studies are from the US and have focused on health care utilization and expenditures.
- The effects on health and patient-related outcomes have been studied only exceptionally and are inconclusive.

## Material and methods

### Search strategy

Using combinations of “gatekeeping”, “gatekeeper\*”, “first contact”, and “self-referral”, the databases PubMed (MEDLINE and Pre-MEDLINE), EMBASE, and the Cochrane Library were searched from their respective inception dates through to April 2008. The search was updated in January 2010. No restrictions were made for publication type, study design, or language. The reference lists of included papers were scanned to identify additional relevant studies.

### Study selection

Studies were selected independently by two reviewers (MVG and AZ). In accordance with recommendations made by the Cochrane Effective Practice and Organization of Care Group (EPOC) [6] and the US Task Force on Community Preventive Services (USTFCPS) [7], randomized controlled trials (RCTs), cluster RCTs, non-randomized controlled trials (CCTs), cluster non-randomized controlled trials, controlled before–after studies (CBAs), cohort studies, case-control studies, and interrupted time-series (ITS) were considered acceptable for this review. Studies were included if they analysed the effects of gatekeeping on at least one of the following outcomes: health- and patient-related outcomes (mortality, morbidity, health-related quality of life, and satisfaction); quality of care; health care utilisation; or economic outcomes.

In this review, gatekeeping is defined as the requirement to visit a general practitioner, family practitioner, general internal medicine physician, or general paediatrician in an ambulatory setting and to obtain a referral from him/her prior to accessing specialist care. Studies in which the gatekeeper function was not exercised by a physician were excluded.

### Extraction and assessment of studies

Using a standardized abstraction form, the following data were extracted from the publications: study setting, population, gatekeeping arrangements, accompanying interventions, control-group intervention, study design and quality, length of follow-up, and outcomes.

The quality of the studies was assessed independently by the reviewers using criteria specific to each of the study designs and discrepancies were resolved through consensus. Quality was assessed following the recommendations of the EPOC [6], the USTFCPS [7], and the US Task Force on Preventive Services (USTFPS) [8]. The suitability of the study design was classified according to definitions of the USTFCPS [7] and the EPOC [6]. The quality of each study was rated as *good*, *fair*, or *poor* according to the USTFCPS [7] and the USTFPS [8] based on allocation, outcome assessment, data sources, and risk of contamination and of attrition bias [7,8]. Statistical adjustment for characteristics that could explain differences in results played an important role in our quality assessment.

### Synthesis

The results are summarized in a narrative fashion and presented graphically, following the approach of a review of managed care performance [9]. For each study, one observation per outcome parameter was created by calculating the relative difference between the comparison groups. Studies comparing two or more types of gatekeeping with no gatekeeping contributed more than one observation (i.e. more than one bar in the figures e.g. Hurley 1991-1, Hurley 1991-2). Studies reporting results for subgroups for one comparison produced a single observation, i.e. a single bar with several segments which – depending on the subgroup results – can simultaneously show both positive and negative results (i.e. bar left and right of zero) as well as statistically significant and non-significant results (i.e. several colours within the same bar). For example a study with three subgroups showing statistically significant differences of 90% and 300% and a statistically non-significant difference of 120% would be represented as a bar with three segments, the first ranging from 0% to 90% coloured in dark grey, the second segment ranging from 90% to 120% in light grey representing the subgroup with statistically non-significant difference, and the third segment from 120% to 300% again in dark grey. Confidence intervals are not represented in the figures.

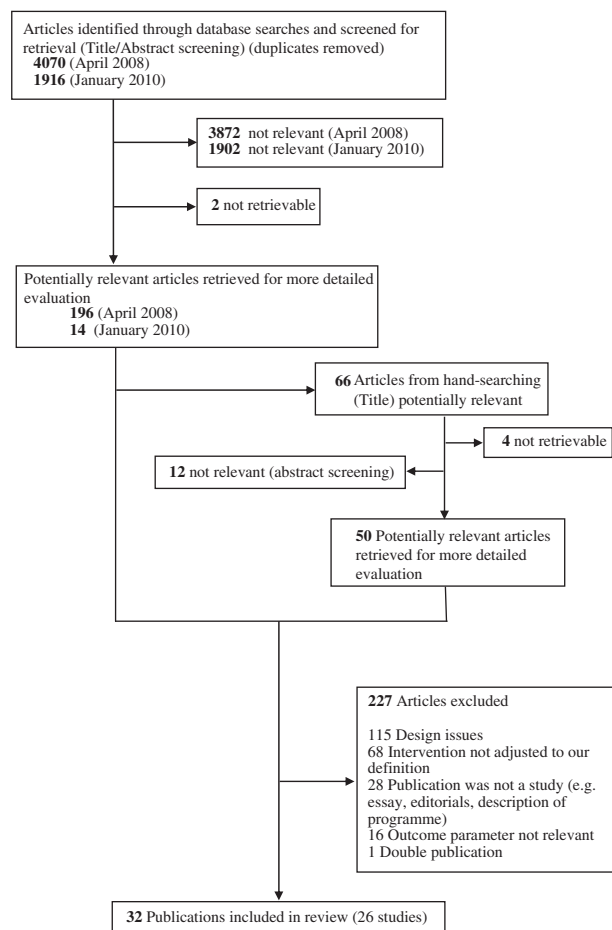


Figure 1. Study selection flow diagram.

## Results

### Study pool

The initial search yielded 4070 publications and the update search 1916. The selection process (Figure 1) left a final pool of 26 studies in 32 publications.

### Study characteristics

Table I summarizes the characteristics of the 26 studies included. All but 10 studies were conducted in the United States (62%), five in Switzerland [10–14] and one each in Denmark [15], Germany [16], Scotland [17–20], and The Netherlands [21]. Another study included data from two countries (Germany and The Netherlands) [22].

The number of individuals whose data had been analysed ranged from 234 to 4 210 000. A specified sample size was lacking in two studies [14,23]. The level of detail regarding the reporting of selection criteria and the quality of reporting the characteristics of the selected study population were heterogeneous. All but nine studies [14,16,21–23,36–39] reported mean

age of included persons and all but five studies [14,23,36–39] provided information on sex distribution. The majority of studies covered the wide spectrum of medical conditions encountered in general practice.

The accuracy of the descriptions of the gatekeeper arrangements varied. In 17 studies, gatekeeping was one of many reported components within managed-care plans [10–14,23–30,32,33,35,39–41]. Nine studies reported visits to specific specialists to be excluded from gatekeeping arrangements (obstetricians/gynaecologists [11,13,16,27,28,40], psychiatrists/psychotherapists [16,27,28,34], paediatricians [10,11], ophthalmologists [13,15,16], optometrists [40], dermatologists [27,28], dentists [40], ENT [15], and family planning specialists [34]).

Fifteen studies explicitly reported the requirement for patients to choose and register with a primary care physician (PCP) as the gatekeeper for a specific period of time [10–13,15,16,22–26,29–31,36–39].

### Study design and quality

The 26 studies under review included one RCT [30], two quasi-randomized controlled clinical trials (CCTs) [31,34], four prospective cohort studies [16–22,41], five CBAs [10–12,29,37,38], 12 retrospective cohort studies [13–15,21,23–26,32,33,35,36,39,40], and two ITs [27,28]. According to the hierarchy of study designs 12 studies were considered to have a *suitable* design and 14 a *moderate suitability* for evaluating the effects of gatekeeping.

Only two of the studies [13,24–26] were rated as *good* quality, three as *fair* quality [16,29,33], and the vast majority as *poor* quality. The only RCT was judged as *poor* quality because a description of allocation concealment was lacking, the dropouts were poorly reported, the risk of detection bias for some outcomes was high, and the study had also a high risk of contamination bias.

Few studies analysed data collected specifically for the purpose of assessing at least one of the studied outcomes [11,12,17–19,30,31,34,41]. The majority analysed routinely collected administrative data (e.g. data collected for reimbursement) [10,13–16,21,23–30,32–34,37–40]. Four studies analysed routinely collected clinical data [31–33,35] and two used survey data [22,36].

The most common shortcoming was the lack of control for relevant clinical and sociodemographic characteristics. Only five studies reported that results had been controlled for morbidity [13,16,24–26,29,33] and self-reported health status was controlled for in two studies [13,36].

Other frequent shortcomings were potential exposure misclassification (e.g. lack of accurate ascertainment of exposure to gatekeeping), contamination

Table I. Overview of included studies.

Study	Population			Gatekeeping <sup>d</sup>		Study design <sup>e</sup>	Observation length <sup>f</sup>	Data sources <sup>g</sup>	Adjustment	Parameters studied <sup>h</sup>	Assessment of study quality
	Country <sup>a</sup>	Age Group	Conditions <sup>b</sup>	Enrol.	Co-pay.						
Anderson 1996 [23]	USA	16-64 yrs	acute	+	+	RC	18 mo	Ad	nr	U, E	III-1-poor
Escarce 2001 [24-26]	USA	working age	all	+	+	RC	2 yrs	Ad	good	U, E	III-1-good
Etter 1997 [11]	CH	18-44 yrs	all	+	-	CBA	12 mo	Qu	poor	U	II-3-poor
Etter 1998 [10]	CH	adults	all	-	-	CBA	12 mo	Ad	poor	E	II-3-poor
Ferris 2001 [29]	USA	< 18 yrs	all	+	+	CBA	2 yrs	Ad	good	U, E	II-3-fair
Ferris 2001a [27]	USA	> 18 yrs	all	nr	nr	ITS	18 mo	Ad	poor	U	III-3-poor
Ferris 2002 [28]	USA	< 18 yrs	all	nr	nr	ITS	18 mo	Ad	poor	U	III-3-poor
Forrest 1999 [41]	USA	children	all	nr	nr	PC	10 d	Ad, Cl	fair	U	II-2-poor
Holdsworth 2007 [17-20]	UK	> 16 yrs	physiotherapy	nr	nr	PC	12 mo#	Qu	nr	H, U, E	II-2-poor
Hurley 1989 [36]	USA	adults	all	+	+	RC	3 mo	Su	fair	U	III-1-poor
Hurley 1991 [37,38]	USA	all ages	all	+	+	CBA	12 mo	Ad	poor	U	II-3-poor
Laditka 2001 [39]	USA	nr	all	+	+	RC	12 mo	Ad	poor	U, E	III-1-poor
Leemrijse 2008 [21]	NL	all ages	physiotherapy	+	nr	RC	LT	Ad, Cl	fair	H	III-1-poor
Linden 2003 [22]	D, NL	adults	all	+	nr	PC	12 mo	Su	nr	U	II-2-poor
Martin 1989 [30]	USA	< 65 yrs	all	+	+	RCT	12 mo	Ad, Qu	nr	H, U	I-poor
Meyer 1996 [31]	USA	adults	all	+	-	CCT	12 mo	Ad, Qu	poor	H, U, E	II-1-poor
Olivarius 1994 [15]	DK	all ages	all	+	-	RC	12 mo	Ad	nr	U, E	III-1-poor
Ose 2008 [16]	D	adults	all	+	-	PC	24 mo	Ad	good	U	II-2-fair
Paone 1995 [32]	USA	adults	CHD-surg	nr	nr	RC	HS	Ad, Cl	nr	H, Q, U	III-1-poor
Perneger 1996 [12]	CH	adults	all	+	-	CBA	12 mo	Qu	poor	H	II-3-poor
Rask 1999 [33]	USA	> 30 yrs	chest pain	nr	nr	RC	12 mo	Ad, Cl	fair	H, Q, U, E	III-1-fair
Schillinger 2000 [34]	USA	adults	all	nr	nr	CCT	12 mo	Ad, Qu	poor	H, Q, U	II-1-poor
Schwenkglekns 2006 [13]	CH	> 18 yrs	all	+	nr	RC	12 mo	Ad, Cl	good	U, E	III-1-good
Shatin 1998 [40]	USA	< 18 yrs	chronic	nr	-	RC	2 yr	Ad	poor	U	III-1-poor
Swetter 2007 [35]	USA	adults	CM	nr	nr	RC	nr	Cl	nr	H, Q	III-1-poor
Werblow 2005 [14]	CH	all ages	all	nr	nr	RC	3 yrs	Ad	poor	E	III-1-poor

Notes: <sup>a</sup>CH: Switzerland, D: Germany, DK: Denmark, NL: The Netherlands, UK: United Kingdom, USA: United States of America; <sup>b</sup>acute: acute conditions in primary care, all: any conditions presenting to primary care, CHD-Surg: Patients with coronary heart disease with surgery indication; chronic: asthma, attention deficit disorder/attention deficit hyperactivity disorder, epilepsy, diabetes, sickle-cell anaemia, CM: cutaneous melanoma, physiotherapy: conditions requiring physiotherapy treatment; <sup>c</sup>n: number of study participants, <sup>d</sup>inconsistency across publications: 55 954/55 011; <sup>e</sup>n = 2177 for outcome satisfaction; <sup>f</sup>Enrol.: Enrolment with primary care physician +; required, -: not required, nr: not reported, Co-pay.: co-payments +; yes, -: no, nr: not reported; <sup>g</sup>CBA: controlled before-after study; CCT: controlled clinical trial; ITS: interrupted time-series; PC: prospective cohort; RC: retrospective cohort; RCT: randomized controlled trial; fHS: hospital stay, LT: length of therapy, nr: not reported; <sup>h</sup>Ad: routine administrative data, Cl: routine clinical data, Pr-Ad/Cl: primary administrative/clinical data; Qu: satisfaction questionnaire, Su: survey, #4 wks for outcome satisfaction; <sup>i</sup>H: health and patient-related; Q: quality of care; U: utilization, E: economic.

(e.g. same caregivers for patients in both groups), the risk of observational bias (e.g. a lack of blinding for subjective outcomes), the high risk of selection bias, and limited reporting of items included in expenditures calculations.

*Health- and patient-related outcomes*

*Symptoms, morbidity.* The results suggested no relevant differences between gatekeeping and free access (Figure 2). No differences were reported regarding the need for percutaneous transluminal coronary angiography (PTCA) in patients with chest pain [33], in the tumour stage at diagnosis in patients with cutaneous melanoma [35], or in the symptom severity in patients with musculoskeletal conditions [17–19]. Additionally, one study reported a non-significant lower myocardial infarction rate in gatekept patients with chest pain compared with patients with free access (1% vs. 2%,  $p = 0.17$ ) [33] and another study reported a non-significant lower mortality post-CABG (1.9% vs. 2.2%,  $p = 0.794$ ) [32]. In the studies with patients undergoing physiotherapy, fewer patients in the gatekeeping group were reported to have achieved their therapeutic goal, with the difference statistically significant in one study (63% vs. 73%,  $p < 0.001$ ) [21] and not significant in the other one (67% vs. 68%,  $p = 0.82$ )

[17–20]. All studies used appropriate data sources (e.g. clinical files). Only one study reported adjusting for relevant patient characteristics, although it did not detail which ones [33].

*Quality of life.* Quality of life (QoL) was an outcome in two studies [12,31] and both used the SF-36 instrument for assessment. Both showed results favouring gatekeeping in single items (bodily pain [31] and role limitations [12]), but no statistically significant differences in overall QoL (see Figure 2). It is, however, questionable whether the studies had enough power to detect differences.

*Satisfaction with care.* Patient satisfaction was assessed in five studies (see Figure 2) [12,17–20, 30,31,34]. It was the primary outcome in only one study, which showed decreased satisfaction under gatekeeping (change between baseline and follow-up  $-2.6$  vs.  $+2.0/+3.3$ ,  $p < 0.05$ ) [12]. In the RCT more participants were reported to be very satisfied with their care in the group with free access (64% vs. 75%,  $p < 0.05$ ) [30]. Similar results were reported for physiotherapy in Britain [20].

*Quality of care*

Two studies reported the percentage of patients treated according to evidence-based guidelines.

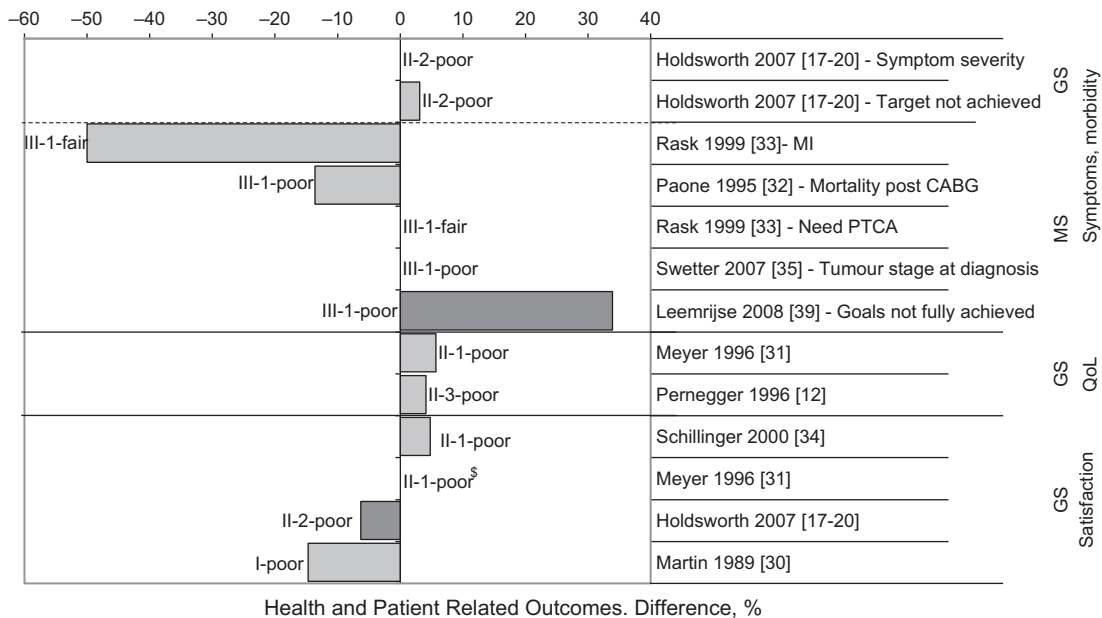


Figure 2. Percentage difference in gatekeeping arrangements (vs. free access) for health- and patient-related outcomes (symptoms, fatalities, quality of life, satisfaction).

Notes: Dark grey bars indicate statistically significant results, grey bars indicate results that are not statistically significant, and white bars indicate that the study did not report significance. A lack of bars indicates that there was no difference. Studies are grouped according to level of evidence (GS: greatest suitability; MS: moderate suitability), and labelled with their quality assessment. §: Study reported “no significant differences” but no data were provided to calculate percentage difference. PTCA: percutaneous transluminal coronary angiography; CABG: coronary artery bypass graft; MI: myocardial infarction.



One found no differences in the provision of an annual ophthalmologic control for patients with diabetes [34]. The other found no differences in the medication prescribed prior to the hospitalization of patients scheduled for cardiovascular surgery, with the exception of the prescription of heparin, which was significantly higher in the gatekeeping group (35.7% vs. 26.7%,  $p = 0.015$ ) (Figure 3) [32].

The diagnosis of melanoma was made without delay significantly more frequently in patients with free access, which, however, did not lead to any differences in the tumour stage at diagnosis [35]. Another study suggested better coordination of care within the gatekeeping system for patients presenting with chest pain [33]. Both studies used appropriate data sources but neither adjusted for relevant clinical characteristics.

Utilisation of health care

**Length of stay** Five studies reported length of hospital stay in days [11,30–32,39]. Overall, the results suggest shorter length of stay under gatekeeping, although the results were not statistically significant (Figure 4). Two of the three studies with greater suitability reported the greatest differences in this direction [30,31]. All studies used appropriate data sources (e.g. hospital claim data), but lacked adjustment for relevant clinical characteristics.

**Hospitalizations** This parameter was operationalized in several ways (percentage of patients with at least one inpatient episode in the observation period [11,15,22,33], number of hospitalizations per patient or per 1000 patients [13,23,29,31,34,40], number of hospital days per 1000 persons [30], observed/expected ratio of hospitalizations [39]). The studies with greater suitability suggest fewer hospitalizations under gatekeeping (see Figure 4). Only one study adjusted for morbidity and it showed a significantly lower hospitalization rate for children in gatekeeping systems (4 vs. 18 per 1000 enrollees,  $p = 0.02$ ) [29]. None of the studies adjusted for disease severity.

The results from the subgroup of studies with moderate suitability are more inconclusive. Only one study showed significant results (higher hospitalization rates among gatekeeping), although this may be the result of confounding [40]. The study of the highest quality, however, showed statistically non-significant lower hospitalizations under gatekeeping [13].

**Ambulatory care** Twelve studies reported visits to both PCP and specialists [11,13,15,22,24–31,34,37,38], four exclusively reported visits to specialists [16–19,33,41], and two reported ambulatory care visits without differentiating between PCP and specialist care [23,40]. Overall the results suggest lower use of specialist care under gatekeeping (Figure 5). Few studies controlled for morbidity, at

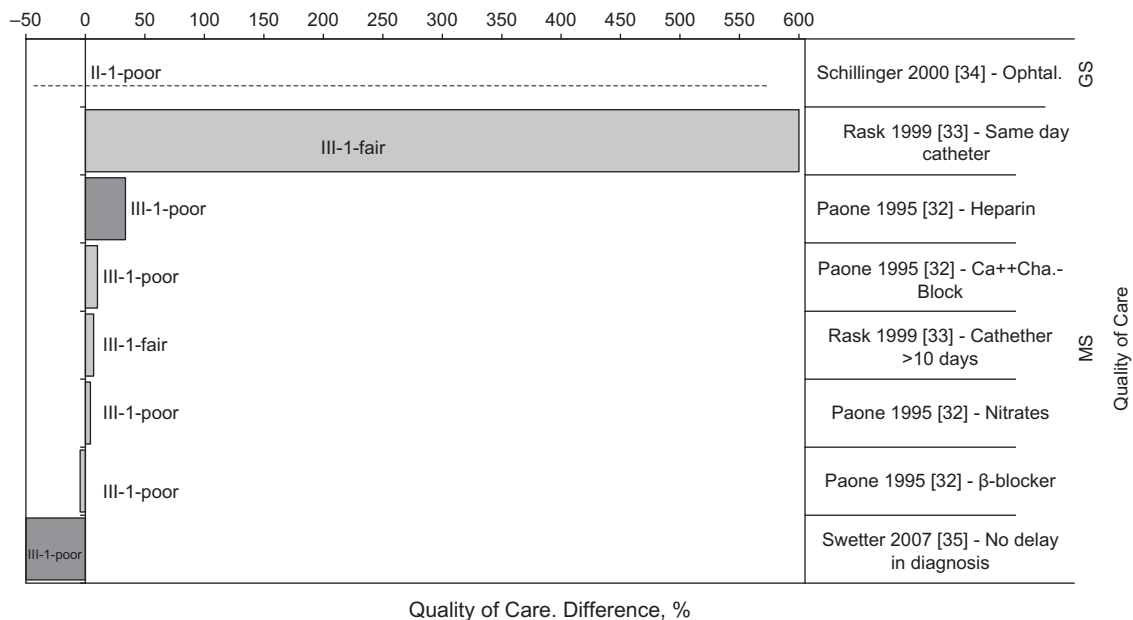


Figure 3. Percentage difference in gatekeeping arrangements (vs. free access) for quality of care parameters.

Notes: Dark grey bars indicate statistically significant results, grey bars not statistically significant results, and white bars indicate the study did not report significance. Lack of bars indicates no difference at all. Studies are grouped according to level of evidence (GS: greatest suitability; MS: moderate suitability) and labelled with their quality assessment.

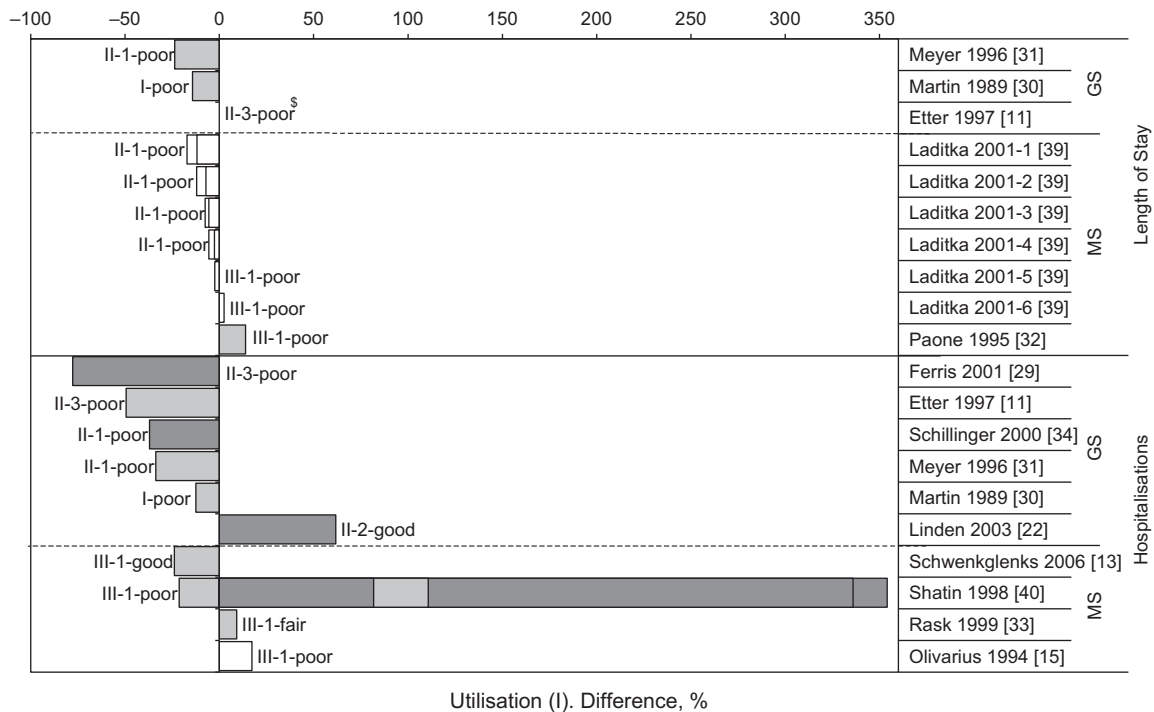


Figure 4. Percentage difference in gatekeeping arrangements (vs. free access) for length of stay and hospitalizations.

Notes: Dark grey bars indicate statistically significant results, grey bars indicate not statistically significant results, and white bars indicate the study did not report significance. Studies are grouped according to level of evidence (GS: greatest suitability; MS: moderate suitability) and labelled with their quality assessment. \$: Study reported “length of stay remained stable in both cohorts” but no data were provided to calculate percentage difference. Multiple segments within a bar indicate results for subgroups, which may differ in their direction, magnitude and statistical significance.

least to some extent, all of them showing less utilization under gatekeeping (two being significant [27,33] and two non-significant [13,24–26]).

Lower utilization of specialty care was not necessarily associated with higher utilization of PCP. Four studies reported slightly higher utilization of PCP paralleling the lower utilization of specialists [15,30,31,34] and four showed an overall lower utilization of ambulatory care [11,22,29,37,38]. Two studies reported referrals to specialized care following PCP visits. Both showed statistically significantly more referrals in the gatekeeping group (3.1% vs. 1.4%,  $p < 0.05$  [17–19] and 3.41% vs. 2.19% [41]).

**Emergency department visits** Emergency department (ED) visits were reported as the percentage of patients with at least one [11,36–38], or the number of ED visits per (1000) patients [31,34,36,40]. The results were inconclusive (see Figure 5). The majority of observations suggested a reduction in the utilization of ED [11,37,38]. However, only two were statistically significant.

*Expenditures*

Health care expenditures were reported in 12 studies (Figure 6). Seven studies reported overall health

care expenditures [10,13,17–19,23–26,29,30], five studies reported hospitalization costs [10,13,24–26,30,39], five studies reported expenditures for ambulatory specialist care [15,24–26,29,30,33], and three studies reported drug expenditures [10,24–26,31]. In the majority of studies, cost information was collected using administrative sources (e.g. reimbursement claim data), thus presumably only including the perspective of the reimbursor. In two studies, however, primary cost information was collected for the study purpose [17–19,31]. One of these reported higher overall expenditures under gatekeeping (89.99 vs. 79.99 pounds sterling/episode,  $p$  not reported) [17–19] whereas the other reported considerably, although statistically not significant, lower drug treatment costs under gatekeeping (491.7 vs. 754.8 US dollars per person per year,  $p = 0.09$ ) [31]. The only RCT reported lower overall costs under gatekeeping (239 vs. 254 US dollars per person per year,  $p = 0.09$ ), which was the net result of a significant reduction of ambulatory specialist care expenditures and a non-significant slight increase in hospitalization costs under gatekeeping [30]. Overall, the majority of observations suggest 6% to 80% lower expenditures under gatekeeping.

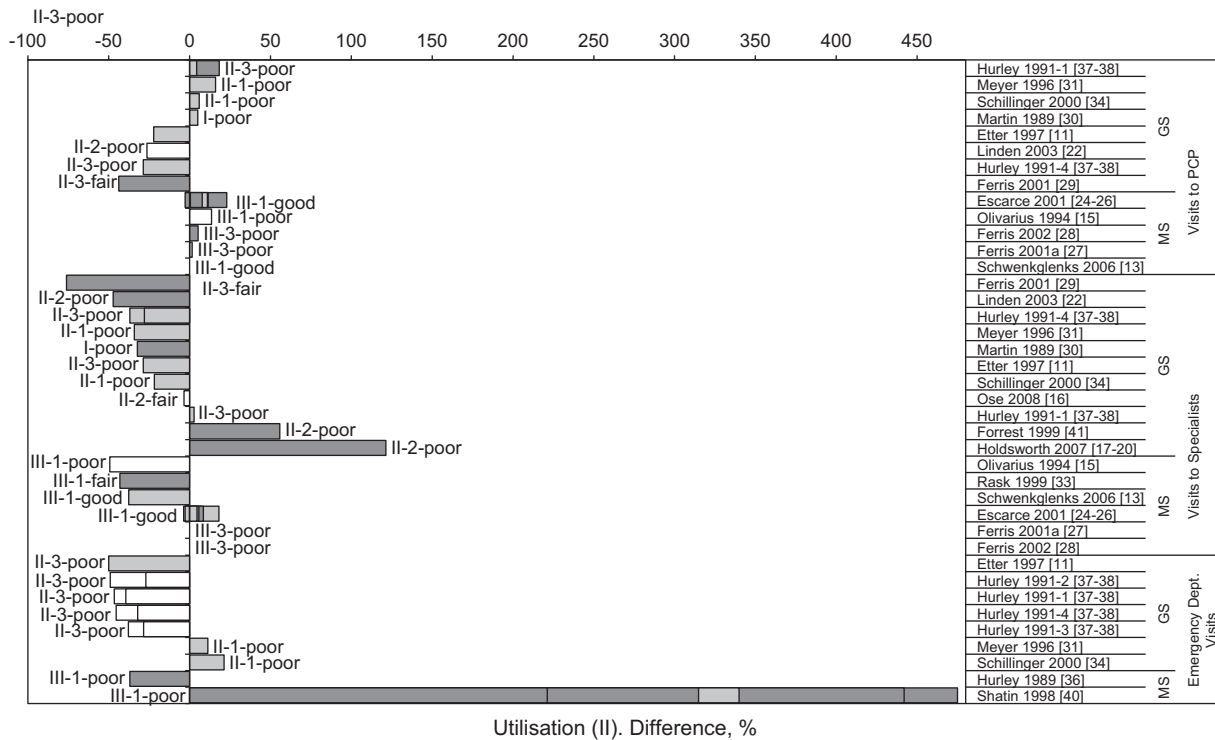


Figure 5. Percentage difference in gatekeeping arrangements (vs. free access) for visits to primary care physicians, to specialist physicians, and to the emergency department.

Notes: Dark grey bars indicate statistically significant results, grey bars indicate not statistically significant results, and white bars indicate the study did not report significance. Lack of bars indicates no difference at all. Studies are grouped according to level of evidence (GS: greatest suitability; MS: moderate suitability) and labelled with their quality assessment. PCP: Primary care physician. Multiple segments within a bar indicate results for subgroups, which may differ in their direction, magnitude, and statistical significance.

### Discussion

The purpose of this review was to systematically examine the evidence on the effects of gatekeeping by primary care physicians on health- and patient-related outcomes, on quality of care, on utilisation of health care, and on expenditures. The evidence regarding the effects of gatekeeping is of limited quality and the effects on health and patient-related outcomes have been studied only exceptionally. The available evidence indicates that gatekeeping is associated with lower utilization of health services (up to -78%) and lower expenditures (up to -80%). However, there was great variability in the magnitude and direction of the differences.

To our knowledge, this is the only extensive systematic review specifically assessing physician gatekeeping. A previous review on gatekeeping in Dutch included only 14 studies, although it also accepted studies in which gatekeeping was exercised by non-physician personnel [42]. In a systematic review on the effectiveness of general practice, the issue of gatekeeping was addressed only secondarily summarizing four studies [43].

Our review may have failed to include some potentially relevant studies. Publication bias is a

potential issue in any systematic review. However, in light of the broad range of results identified, its impact here is probably nearly negligible.

Despite the number of studies included in this review, the evidence on the effects of gatekeeping on health outcomes is still limited. Health outcomes were reported only exceptionally and the validity of results was considered poor. The analysis of frequently reported utilization or economic parameters also revealed methodological shortcomings. The majority of the 26 studies included in this review can be considered analyses of natural experiments because only three reported on interventions with an experimental character [30,31,34].

The lack of adjusting for relevant clinical and socio-demographic characteristics is particularly problematic, since these are important determinants of the utilization of health services and the expenditures incurred by a patient. In many studies, these characteristics were not evenly distributed among the gatekeeping and free-access groups; thus one must consider that the differences observed in the utilization of health services or in the costs reflect underlying differences in relevant patient characteristics and are therefore attributable to such differences, rather than to gatekeeping.



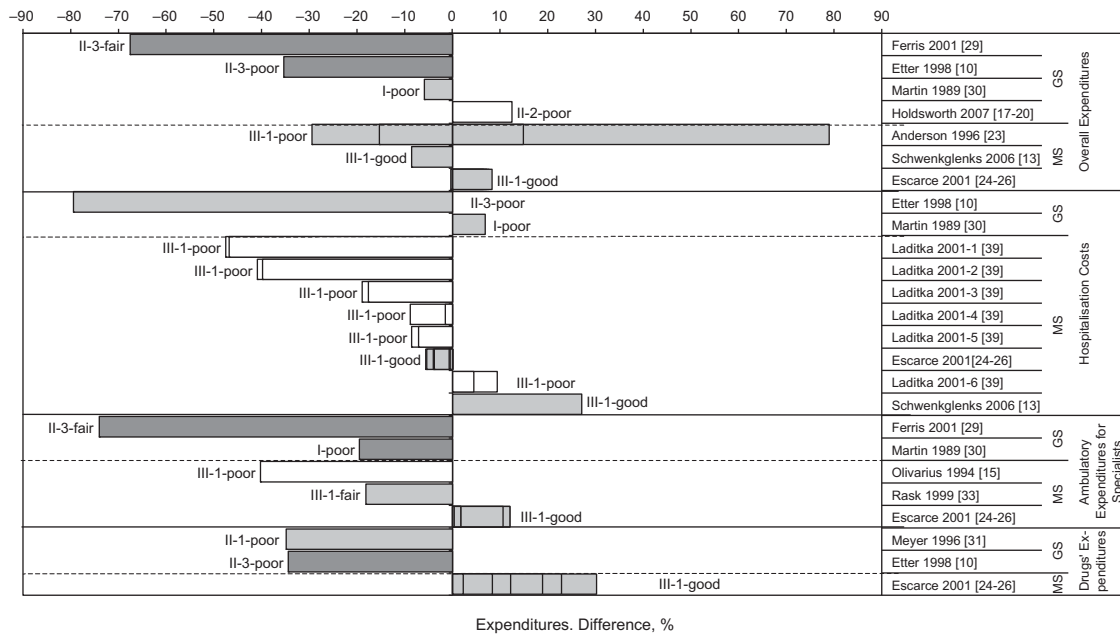


Figure 6. Percentage difference in gatekeeping arrangements (vs. free access) for expenditures (overall expenditures, hospitalization costs, expenditures for ambulatory specialist care, and drug expenditures).

Notes: Dark grey bars indicate statistically significant results, grey bars indicate not statistically significant results, and white bars indicate the study did not report significance. Studies are grouped according to level of evidence (GS: greatest suitability; MS: moderate suitability) and labelled with their quality assessment. Multiple segments within a bar indicate results for subgroups, which may differ in their direction, magnitude, and statistical significance.

Another issue was the interpretation of utilization and expenditure. In general, the studies' authors considered lower utilization of health services (or expenditures) a desirable and positive effect of gatekeeping programmes. Such an interpretation assumes that there is an overutilization of health services. However, none of the studies included in this review provided evidence on the delivery of unnecessary care or an estimation of the problem of overuse in the specific setting. The studies also lacked a definition of appropriate utilization and did not address the undesirable effect of limiting the utilization of care that is actually needed (i.e. of inappropriate reductions). Only one study addressed the potential for inappropriate reductions in access to specialty care by assessing whether changes in overall specialty utilization were related to changes in the number of diabetics receiving an annual ophthalmological examination [34].

Especially relevant to this review was the fact that the majority of studies compared groups whose health care service access and delivery management arrangements differed not only in gatekeeping arrangements but also in many other elements, which could have affected the outcomes. In 10 studies [10–12,15,22,23, 30,33,39,40], there were at least reported differences in the way providers were paid, in the amount of patient co-payments, in the services portfolios, in the availability of providers, and in the implementation of

tools for controlling service delivery, such as utilization reviews. Thus, it is not possible to attribute the observed differences solely to the gatekeeping arrangement.

In conclusion, gatekeeping has mainly been studied as a feature of managed care and not in the context of system-wide implementation. The bulk of research is of limited quality and has focused on health care utilisation and expenditures, whereas effects on health- and patient-related outcomes have been studied only exceptionally and are inconclusive. When considering gatekeeping, policy-makers need to be aware of the limitations and uncertainties uncovered by this review. Future research should focus on studying effects on health outcomes and on patients' satisfaction in health system contexts other than managed care in order to warrant strong recommendations.

### Acknowledgements

The authors would like to thank Matthew D. Gaskins and Mary K. Renaud for copy-editing the manuscript. They would also like to thank Karen Hofmann for her support in retrieving the publications.

### Conflict of interest

The authors do not have any conflicts of interest.

## Sources of support

This review originated in the authors' work for the German Advisory Council on the Assessment of Developments in the Health Care System for its report "Coordination and Integration: Health Care in a Society of Longer Living". The Council did not have any influence on the design and conduct of the study; collection, management, analysis, and interpretation of the data; or preparation, review, or approval of the manuscript.

## Prior presentation

Results of this review were presented as a poster at the 7th Annual Meeting of HTAi in June 2010. An abridged and previous German-language version of this systematic review covered publications only up to 2007 [44].

## References

- [1] Starfield B, Shi L, Macinko J. Contribution of primary care to health systems and health. *Milbank Q* 2005;83:457–502.
- [2] Delnoij D, van Merode G, Paulus A, et al. Does general practitioner gatekeeping curb health care expenditure? *J Health Serv Res Policy* 2000;5:22–6.
- [3] Franks P, Clancy CM, Nutting PA. Gatekeeping revisited, protecting patients from overtreatment. *N Engl J Med* 1992; 327:424–9.
- [4] Cooper PF, Simon KI, Vistnes J. A closer look at the managed care backlash. *Med Care* 2006;44:I–4–I–11.
- [5] Fang H, Liu H, Rizzo JA. Has the use of physician gatekeepers declined among HMOs? Evidence from the United States. *Int J Health Care Finance Econ* 2009;9:183–95.
- [6] Cochrane Effective Practice and Organization of Care Group (EPOC). EPOC resources for review authors. <http://www.epoc.cochrane.org/en/handsearchers.html> (last accessed May 25, 2009).
- [7] Briss PA, Zaza S, Pappaioanou M, et al. Developing an evidence-based guide to community preventive services – methods. *Am J Prev Med* 2000;18:35–43.
- [8] Harris RP, Helfand M, Woolf SH, et al. Current methods of the US Preventive Services Task Force: A review of the process. *Am J Prev Med* 2001;20:21–35.
- [9] Miller RH, Luft HS. Managed care plan performance since 1980: A literature analysis. *JAMA* 1994;271:1512–9.
- [10] Etter JF, Perneger TV. Health care expenditures after introduction of a gatekeeper and a global budget in a Swiss health insurance plan. *J Epidemiol Community Health* 1998; 52:370–6.
- [11] Etter JF, Perneger TV. Introducing managed care in Switzerland: Impact on use of health services. *Public Health* 1997; 111:417–22.
- [12] Perneger TV, Etter JF, Rougemont A. Switching Swiss enrollees from indemnity health insurance to managed care: The effect on health status and satisfaction with care. *Am J Public Health* 1996;86:388–93.
- [13] Schwenkglens M, Preiswerk G, Lehner R, et al. Economic efficiency of gate-keeping compared with fee for service plans: A Swiss example. *J Epidemiol Community Health* 2006;60:24–30.
- [14] Werblow A. [Efficiency of general practice in Switzerland]. *Gesundheitswesen* 2005;67:107–11.
- [15] Olivarius ND, Jensen FI, Gannik D; et al. Self-referral and self-payment in Danish primary care. *Health Policy* 1994;28:15–22.
- [16] Ose D, Broge B, Riens B, Szecsenyi. Mit Überweisung vom Hausarzt zum Spezialisten – Haben Verträge zur Hausarztzentrierten Versorgung (HZV) einen Einfluss? [Contacts with specialists with referrals by GP – Do GP Centred Health Care (HZV) contracts have an impact?] (English summary). *Z Allg Med* 2008;84:321–6.
- [17] Holdsworth LK, Webster VS, McFadyen AK, and the Scottish Physiotherapy Self-Referral Study Group. Are patients who refer themselves to physiotherapy different from those referred by GPs? Results of a national trial. *Physiotherapy* 2006;92:26–33.
- [18] Holdsworth LK, Webster VS, McFadyen AK and the Scottish Physiotherapy Self-Referral Study Group. Self referral to physiotherapy: Deprivation and geographical setting: is there a relationship? Results of a national trial. *Physiotherapy* 2006;92:16–25.
- [19] Holdsworth LK, Webster VS, McFadyen AK, on behalf of the Scottish Physiotherapy Self-Referral Study Group. What are the costs to NHS Scotland of selfreferral to physiotherapy? Results of a national trial. *Physiotherapy* 2007;93:3–11.
- [20] Webster VS, Holdsworth LK, McFayden AK, Little H. Self-referral, access and physiotherapy: Patients' knowledge and attitudes – results of a national trial. *Physiotherapy* 2008;94: 141–9.
- [21] Leemrijse CJ, Swinkels ICS, Veenhof C. Direct access to physical therapy in the Netherlands: Results from the first year in community-based physical therapy. *Phys Ther* 2008;88:936–46.
- [22] Linden M, Gothe H, Ormel J. Pathways to care and psychological problems of general practice patients in a "gate keeper" and an "open access" health care system: a comparison of Germany and the Netherlands. *Soc Psychiatry Psychiatr Epidemiol* 2003;38:690–7.
- [23] Anderson S, Cave D. The gatekeeper effect on managing acute medical conditions. *Med Interface* 1996;9:122–9.
- [24] Escarce JJ, Kapur K, Joyce GF, et al. Medical care expenditures under gatekeeper and point-of-service arrangements. *Health Serv Res* 2001;36:1037–57.
- [25] Joyce GF, Kapur K, Van Vorst KA, et al. Visits to primary care physicians and to specialists under gatekeeper and point-of-service arrangements. *Am J Manag Care* 2000;6:1189–96.
- [26] Kapur K, Joyce GF, Van Vorst KA, et al. Expenditures for physician services under alternative models of managed care. *Med Care Res Rev* 2000;57:161–81.
- [27] Ferris TG, Chang Y, Blumenthal D, et al. Leaving gatekeeping behind: Effects of opening access to specialists for adults in a health maintenance organization. *N Engl J Med* 2001;345:1312–7.
- [28] Ferris TG, Chang Y, Perrin JM, et al. Effects of removing gatekeeping on specialist utilisation by children in a health maintenance organization. *Arch Pediatr Adolesc Med* 2002;156:574–9.
- [29] Ferris TG, Perrin JM, Manganello JA, et al. Switching to gatekeeping: Changes in expenditures and utilisation for children. *Pediatrics* 2001;108:283–90.
- [30] Martin DP, Diehr P, Price KF, et al. Effect of a gatekeeper plan on health services use and charges: A randomized trial. *Am J Public Health* 1989;79:1628–32.
- [31] Meyer TJ, Prochazka AV, Hannaford M, et al. Randomized controlled trial of residents as gatekeepers. *Arch Intern Med* 1996;156:2483–7.
- [32] Paone GH, Higgins RS, Spencer T, et al. Enrollment in the health alliance plan: HMO is not an independent risk factor for coronary artery bypass graft surgery. *Circulation* 1995;92:II69–72.

- [33] Rask KJ, Deaton C, Culler SD, et al. The effect of primary care gatekeepers on the management of patients with chest pain. *Am J Manag Care* 1999;5:1274–82.
- [34] Schillinger D, Bibbins-Domingo K, Vranizan K, et al. Effects of primary care coordination on public hospital patients. *J Gen Intern Med* 2000;15:329–36.
- [35] Swetter SM, Soon S, Harrington CR, et al. Effect of health care delivery models on melanoma thickness and stage in a university-based referral center: An observational pilot study. *Arch Dermatol* 2007;143:30–6.
- [36] Hurley RE, Freund DA, Taylor DE. Gatekeeping the emergency department: Impact of a Medicaid primary care case management program. *Health Care Manage Rev* 1989;14:63–71.
- [37] Hurley RE, Freund DA, Taylor DE. Emergency room use and primary care case management: Evidence from four Medicaid demonstration programs. *Am J Public Health* 1989;79:843–6.
- [38] Hurley RE, Freund DA, Gage BJ. Gatekeeper effects on patterns of physician use. *J Fam Pract* 1991;32:167–74.
- [39] Laditka SB, Laditka JN. Utilization, costs, and access to primary care in fee-for-service and managed care plans. *J Health Soc Policy* 2001;13:21–39.
- [40] Shatin D, Levin R, Ireys HT, et al. Health care utilisation by children with chronic illnesses: A comparison of Medicaid and employer insured managed care. *Pediatrics* 1998;102:E44.
- [41] Forrest CB, Glade GB, Starfield B, et al. Gatekeeping and referral of children and adolescents to specialty care. *Pediatrics* 1999;104:28–34.
- [42] Willems D, Veldhuis M. Vormen van selectie aab de poort en hun effecten. Eenliteratuuronderzoek [Does the gatekeeper enhance appropriateness of care? A literature study] (English summary). *Huisarts Wet* 2004;47:625–30.
- [43] Engström S, Foldevi M, Borgquist L. Is general practice effective? A systematic literature review. *Scand J Prim Health Care* 2001;19:131–44.
- [44] Zentner A, Velasco Garrido M, Busse R. Macht der Hausarzt als Lotse die Gesundheitsversorgung wirklich besser und billiger? Ein systematischer Review zum Konzept Gatekeeping [Do primary care physicians acting as gatekeepers improve health outcomes and decrease costs? A systematic review of the concept of gatekeeping] (English summary). *Das Gesundheitswesen* 2010;72:e38–e44 (DOI: 10.1055/s-0029-1234126).