1. Introduction: What is a systematic literature review and why doing it?

1.1. Systematic literature reviews in contrast to narrative reviews

Systematic literature reviews are common practice in medical research but are known less well as a methodology in the social sciences although they are becoming increasingly popular in studies on health policy or education.

Probably every political scientist has done a literature review at some point in their career as this is the usual starting point of most pieces of research in the social sciences and often a standard requirement of a PhD dissertation. Such ‘conventional’ narrative literature reviews typically describe the theories, methodologies and findings of previous key or even classic studies within the topic area. The aims of such exercise include locating their own research in the wider subject area, deriving testable hypotheses, summarising the state-of-the-art research and identifying gaps in the existing research in order to provide a rationale for their own study. There is no explicit methodology for identifying the literature which will be reviewed and it is up to the discretion of the researcher how many and which studies are included in such a review although the omission of an influential piece of work would present a serious shortcoming and evoke serious criticism by any examiner.

In contrast to such traditional forms of reviewing literature, a systematic literature review tries to identify and then summarise all relevant research evidence on a particular question which fulfils certain criteria that are clearly described at the outset, following a transparent research protocol. A protocol is the plan or set of steps to be followed in preparing a systematic review. It clearly describes the review question and how the reviewer authors will go about developing the review. It details how they will seek, select as relevant, critically appraise studies, and collect and analyse data from the included studies. The last
step of the process is the data synthesis; such combination of studies which were identified as valid usually applies quantitative statistical techniques, such as meta-analysis.

“High-quality systematic reviews take great care to find all relevant studies published and unpublished, assess each study, synthesise the findings from individual studies in an unbiased way and present a balanced an impartial summary of the findings with due consideration of any flaws in the evidence” (Davies and Crombie, 2001). In other words, to identify the valid and applicable evidence, explicit methods for performing a thorough literature search and selecting relevant studies which will be critically appraised and summarised are the basis of any systematic literature.

1.2. The potentials of systematic literature reviews
The advantages of systematic literature reviews are widely acknowledged in the field of medicine where the abundance of published (and unpublished) studies with often contradictory findings makes keeping track with primary research impossible. As the website of the Cochrane Collaboration, the central organisation of health care specialists which systematically review the effectiveness of interventions in clinical research, states “[...] healthcare providers face a serious challenge to keep up-to-date with the latest health care. Research is being published every day, all over the world. Information appears in thousands of medical, scientific and health-related journals worldwide. Furthermore, the results of one healthcare study may be different from, or even contradict, the results of another study making it difficult to draw conclusions” (http://www.cochrane.org/consumers/sysrev.htm, accessed on 14 March 2008).

The Cochrane Handbook for Systematic Reviews of Interventions - the ‘bible’ of systematic reviews in clinical studies - praises systematic literature reviews as efficiently integrating valid information and providing a basis for rational decision making (Higgins and Green, 2006). Furthermore, systematic reviews can help those wanting to make better decisions about healthcare, particularly where randomised controlled trials, the gold standard’ of study design in medical research, have not been done and may not be possible or appropriate. As Curran et al. (2007: 290) note, “[s]ystematic reviews can avoid the need for costly

---

1 Meta-analysis is the most common form of synthesising quantitative evidence. For a good overview of other techniques, also for summarising qualitative evidence, see Dixon-Woods et al. (2005).
primary research and provide a transparent, hopefully robust method for managing information using an agreed protocol”.

Indeed, systematic reviews are leading the hierarchy of evidence in the medical sciences reflecting the fact that, when well conducted, they should give the best possible estimate of any true effect of an intervention (Davies and Crombie, 2001).

In the context of evidence-based policy-making, systematic literature reviews seem to be a very useful methodological tool not only in clinical research but also in policy analysis. Although some authors in the social sciences are rather critical of the concept of ‘evidence-based policy’ (e.g. Glasby and Beresford, 2006), the importance of research evidence for ‘what works’ in health services policy is emphasised in a number of health policy documents under New Labour. Evidence-based policy making is widely advocated at national and international levels, for example, one relatively recent Department of Health’s research strategy papers emphasises that the NHS Service Delivery and Organisation (SDO) programme was established “to consolidate and develop the evidence base on the organisation, management and delivery of healthcare services to increase the quality of patient care, ensure better patient outcomes and contribute to improved population health” (Department of Health, 2006a: 23). In like manner, the WHO's World Health Report 2000 sought to ‘to push forward national and global development of the skills and information required to build a solid body of evidence on the level and determinants of performance, as a basis for improving how [health] systems work.’” (WHO, 2000: 12). It is thus not surprising that systematic attempts to summarise existing research findings are also being made in non-medical academic disciplines with a high policy-relevance.

This paper will present the methodology of systematic literature reviews with the example of a study on organisational performance of not-for-profit providers of health and social care. This study was commissioned by the Service Delivery and Organisation Programme of the National Institute for Health Research of the British NHS. For the purpose of exemplifying the use of systematic literature reviews for a non-medical audience, the focus will not be on the findings of the individual studies which were included in our review, but rather on the methodology of identifying and selecting relevant studies for a systematic literature review in a policy-related field. The second major aim of this paper is to
highlight both the potentials and pitfalls for using systematic literature reviews in subject areas like social policy or public policy.

In the following, I will thus first give a short overview of the background of not-for-profit health care organisations in the UK. I will then briefly outline the theoretical assumptions about the performance about different providers in health and social care before I describe the methodology we applied in our systematic literature review. Finally, I will discuss the methodological problems we encountered in this study when it came to drawing any conclusions for policy-makers.

2. Theories of not-for-profit sector performance and the UK policy background

2.1. Theories of ownership effects on efficiency and the particularities of the health and social care services

In recent years, cost containment through the substitution of competing providers for public monopolies has been a priority of health system reform in Britain and elsewhere in Europe (Mossialos and Le Grand, 1999). The opening-up of universal public health systems to private providers which we are currently witnessing in the case of the British NHS is premised on economic models about the benefits of private ownership and competition, such as the property rights model (Alchian and Demsetz, 1973; Demsetz, 1967), which predict that private provision will lead to greater consumer choice, more innovative services, more efficient management, a leaner public workforce and thus lower costs.

This might provide an argument for using private providers; however, it does not provide an adequate rationale for using the not-for-profit sector in health care, where the market and the ‘product’ differ from other markets. There is asymmetric information in health, childcare and nursing services leaving the potential purchaser/user vulnerable to supplier exploitation. Furthermore, the complexity and uncertainty of service requirements mean that it is difficult to specify the precise requirements and thus contracts between the purchaser and provider are deemed to be ‘incomplete’, in the sense that not all possible eventualities can be written into a contract. Managers of for-profit organizations have an incentive, which experience shows is not always resisted, to exploit this situation and maximize profits by reducing service quality. Hansmann (1980) argues that the non-distribution constraint of not-for-profit organizations in public
services, where quality is at a premium, prevents those in control of such an organization from financially benefiting from its activities. This, he argues, means that they can be trusted not to exploit the purchaser. This provides the theoretical justification for the turn to not-for-profit provision in health care and social services.

2.2. New Labour’s emphasis on strengthening the not-for-profit sector

In August 2007, the Minister for Care Services for England announced the launch of the Department of Health’s £73m Social Enterprise Investment Fund to support and encourage the development of a social enterprise sector in the delivery of health and social care services. The government defines a social enterprise basically as a not-for-profit organisation, namely as “a business with primarily social objectives whose surpluses are principally reinvested for that purpose in the business or in the community, rather than being driven by the need to maximise profit for shareholders and owners” (Department of Health, 2007).

The promotion of not-for-profit providers of welfare marks an important policy change since the not-for-profit sector has played only a marginal role in the provision of health services in the UK since the inception of the NHS in 1948 (see, e.g., Mohan, 1984; 1985). It is part of a general shift from state to market provision of public services.

The ‘NHS Plan’ in 2000 saw a key policy shift in relation to the non-state provision of healthcare, with the New Labour government declaring its intention to develop partnerships with the independent sector (both private for-profit and the voluntary sector) (NHS, 2000: 5). The government subsequently introduced legislative and other measures designed to lower the barriers to entry for voluntary agencies (also called “Third Sector”) and to abolish the public sector’s monopoly (see Department of Health, 2006b; HM Treasury, 2002, 2006; NHS Primary Care Contracting, 2006). These measures included access to start-up funds and long term investment, which were intended to enable the not-for-profit sector to compete on a level playing field (Department of Health, 2006b: 175-176). A Third Sector Commissioning Task Force was established in 2005 to forward these changes (Third Sector Commissioning Task Force 2006). The setting up of the Social Enterprise Investment Fund was the last step in this development so far.
Her Majesty’s Treasury (2004: 3) mentions that not-for-profits have “the capacity to build users’ trust” while the Department of Health has presented the policy of favouring social enterprises over for-profit providers as a means of enhancing community based care. It furthermore describes not-for-profit organizations as having better relations with particular patient groups and expertise in specific areas and praises them as innovative providers of primary care which are value-driven (Department of Health, 2006b). However, there is little evidence to support these a priori beliefs about the advantages of not-for-profits (Beresford, 2005; Farnsworth, 2006).

3. Examining the performance of the not-for-profit sector in health care

Whether the performance of the not-for-profit sector justifies the recent UK policy preference for not-for-profit provision of health care is a legitimate question to ask of a government committed to evidence-based policy. We thus examine whether the evidence supports the government’s policy of expanding the not-for-profit sector, drawing on a systematic literature review of the performance of the not-for-profit sector in health and social care.

3.1. The research protocol: Searching for and identifying relevant literature

The literature reviewed in the present study was obtained from 26 sources (14 literature databases from medicine, social science and economics, and 12 websites), using 9 different search terms which were various synonyms and spellings of ‘not-for-profit’ in combination with the key words health care and social care. This multitude of search terms reflects the absence of a universally agreed terminology of organisations which do not distribute any profits to the owners or stakeholders, including unincorporated community and voluntary groups, co-operatives, registered charities, friendly societies, and the increasing number of social enterprises (for comprehensive definitions see Kendall, 2003 and Salamon and Anheier, 1997). The huge number of searched literature sources, in turn, results from the fact that not only published research articles from a number of academic disciplines were likely to deal with the topic of organisational performance, but that government papers, unpublished research reports and other forms of grey literature had to be considered as well (see Bravata et al., 2005, for a discussion of challenges particular to systematic reviews on topics related to health care organisation, finance and delivery). As a
consequence, our literature search extended beyond the traditional medical and social scientific literature databases to include websites of health departments and voluntary sector research institutions and the like.

Our search period was between 2001 and 2006 (i.e. ending in the year in which we conducted our search). This timeframe was chosen in order to review the most up-to-date evidence that had not been covered in previous systematic reviews. The initial search yielded over 14,000 hits including duplicates. These were reduced in three steps: Firstly, the titles and subtitles of studies were scanned in order to identify the most obvious irrelevant hits and to sift out duplicates of articles. Such duplication occurred due to the fact that different databases overlapped in the inclusion of the same journals. Second, we read all abstracts or first paragraphs of the remaining studies to decide about their suitability for our systematic review. For example, all non-empirical studies and those which did not compare not-for-profit organisations with other providers of health or social care were excluded. If necessary, a second opinion about the relevance of a study was being sought. In case of doubt, studies were included in our intermediate search result and the final decision about its inclusion was left to the reviewer. Finally, we double-checked for cross-references to identify further relevant studies which were not captured by our electronic search.

In the final analysis we included 163 studies which met the criterion of having compared the performance by different ownership types of health and social care providers. The studies were divided between six reviewers. The reviewers first briefly described the setting, methodology, variables and main findings of each study and noted any limitations or further comments. Then each study was reviewed according to as many of our six analytical themes as appropriate. These findings were finally summarized by two members of the research team and then peer-reviewed by all others. In addition, some overview tables were compiled about different aspects of the 163 studies, e.g. the employed methodology or the country setting.
Table 1: Research protocol and search strategy

<table>
<thead>
<tr>
<th>Task</th>
<th>Step 1</th>
<th>Step 2</th>
<th>Step 3</th>
<th>Step 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Define research question</td>
<td>Review of theoretical literature on nonprofit performance</td>
<td>Arrival at six key themes for later analysis</td>
<td></td>
<td></td>
</tr>
<tr>
<td>and key themes</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Identify search terms</td>
<td>Synonyms of “nonprofit” in different spellings (10 terms)</td>
<td>Trial search using five databases</td>
<td>Refinement of search terms: 9 search terms*</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(combined with either health care or social care)</td>
<td></td>
</tr>
<tr>
<td>Search</td>
<td>Initial search of 26 sources with specified search terms, confined to publications 2001-2006 in English language</td>
<td>first sifting by scanning of titles, subtitles; checking for duplicate articles which were yielded by different sources</td>
<td>Second sifting by scanning of abstracts/first paragraphs)</td>
<td>checking for cross-references to identify further relevant studies which were not captured by electronic search</td>
</tr>
<tr>
<td>Review</td>
<td>Distribution of studies among research team</td>
<td>Brief description of each study according to common template</td>
<td>Analysis according to 6 themes</td>
<td></td>
</tr>
</tbody>
</table>

* The following 9 terms were used: “not for profit”, “non profit” (first two terms allowing for different spellings), “social enterprise”, “co-operative”, “foundation trusts”, “community owned”, “independent sector”, “voluntary sector”, and “third sector”.

We also looked at the findings of previous literature reviews but did not include these in our systematic review. The focus of the few systematic reviews which have been conducted previously on not-for-profit sector performance was mainly on comparisons with the commercial for-profit sector, not the public sector, since their theoretical assumption was that not-for-profits deliver higher quality than for-profits. But since UK-policy entails a shift from the public sector to not-for-
profits, we had to ascertain whether there was any literature comparing the not-for-profit sector with the public sector. In addition, since the UK has a universal health system, performance needed to be considered in relation to equity, access and the regulatory requirements that governed healthcare providers. In contrast to previous literature reviews we therefore examined studies to assess the degree to which universal coverage, geographic population, and local and central government regulation were considered in the design of the studies and the interpretation of findings. Third, we also included studies that went beyond a strictly quantitative nature and encompassed service user and staff perspectives.

Based on the central theoretical arguments about the performance of not-for-profit provision of services discussed in this literature we derived six key criteria which we then used in summarising the existent literature on not-for-profits: costs, efficiency, quality of care, uncompensated care², staffing ratios and organisational isomorphism³.

We decided that a meta-analysis or some other form of quantitative synthesis was neither feasible nor appropriate for our review. Already previous literature reviews highlighted a number of issues that rendered a sophisticated meta-analysis of studies very difficult, if not impossible (e.g. Currie, Donaldson and Lu, 2003), including: the heterogeneity of the studies, the inability to standardise them with respect to regulation and various methodological problems. For example, some studies were qualitative or did not use comparable data sets. There is some debate whether it is acceptable to conduct syntheses of qualitative evidence (Dixon-Woods et al., 2005). Previous systematic literature reviews on this or similar topics often simply undertook a crude numerical synthesis of study findings, e.g. by counting how many studies found not-for-profits performing better than for-profit organisations. We decided against such a method, not least because many study findings were indeterminable in their outcome. Instead it became clear to us that it was much more important to examine arising policy issues when attempting to transfer policy findings within or across countries and we decided for a thematic analysis which much more resembles a traditional social science literature review.

² In order to enjoy tax-exempt status, US not-for-profits are required to provide uncompensated care, that is, care provided to a patient for which the hospital is not reimbursed, and so-called community benefits, for example, health education, screening tests to specific vulnerable populations, and activities benefiting the greater public good, such as education for medical professionals and medical research (GAO, 2005: 2)

³ Many studies argue that there is a convergence of behaviour among different providers of medical care depending on the context they are embedded in.
3.2. Main findings in a nutshell

Studies of not-for-profits mainly focused on the analysis of health care affordability and cost containment strategies (around one in five studies). These included measuring the impact of cost control and changes in reimbursement methods on providers and the impact of ownership on costs and efficiency. Another major theme was quality of care. Nine US studies also focused on uncompensated care, reflecting the fact that the US has high numbers of uninsured people (46.1 million in 2005, see U.S. Census Bureau, 2006).

In line with the theoretical predictions, the literature found a weak tendency for not-for-profit providers of health and social care to provide better quality than for-profit institutions whereas the for-profits performed better in relation to efficiency measures than not-for-profit providers. As one would expect, the for-profit hospitals were more likely to offer the more profitable medical services and were more responsive to changes in service profitability, while government hospitals were more likely to offer less profitable services. Not-for-profits often fell in the middle. One study concluded that there appears to be an inverse relationship between hospital efficiency and social responsibility.

There was, however, no consistent support for contentions that the growth of the for-profit sector led to declining access and quality. Neither was there unanimous support for concerns that not-for-profits were less efficient than for-profit organizations. For instance, one study found not-for-profit hospitals were less inefficient than for-profit hospitals, with public hospitals falling in the middle. Comparisons including public hospitals revealed mixed findings. For example, a study on German hospitals claimed that both public and not-for-profit hospitals are more efficient than private for-profit-hospitals while a Swiss study argued that not-for-profit nursing homes were more cost-efficient than public sector nursing homes.

Many findings were rather indeterminate depending on the performance measure used. For example, one study found that on one quality of care measure not-for-profit hospitals performed better whereas on another quality of care measure for-profit hospitals performed better. Significant findings depended on the control variables employed in the statistical modelling. A majority of studies concluded that ownership is not important in determining performance. Crucially market structure and the regulatory framework under which hospitals and health
insurance providers operate proved much more important than ownership form in explaining the performance of different medical care providers. When they were in a hostile market environment, there was little to distinguish not-for-profits from for-profit organizations. Findings seemed to be consistent with a growing suspicion about increasing commercialisation in the health care sector and the view that many not-for-profit institutions were for-profits ‘in disguise’. It was therefore crucial to consider the market conditions under which not-for-profit organizations operated and the requirements of the regulatory framework.

This issue raises considerable doubts regarding the role of trust in not-for-profit organisations. Although the “value-driven” character of not-for-profits is repeatedly emphasized by UK-policy guidance and, according to theory (Hansmann, 1980), the mission statement of such organizations is crucial for their trustworthiness, the role of the not-for-profits’ mission is increasingly questioned in the US. For example, vol. 25(1) of Health Affairs in 2006 devoted an entire issue on this debate and concluded that when it came to ‘mission versus market’, the market ruled. Rathgeb Smith and Lipsky (1993) argued that by competing for contracts the not-for-profit sector had been forced to adopt government norms and standards and thus had lost its niche character and innovativeness, another of the alleged advantages of not-for-profits. In the US, contracting and the incorporation as mainstream providers made not-for-profits more business-like, focusing on their financial position at the expense of flexibility and responsiveness. Finally, this need for financial viability of not-for-profits in a system aiming first and foremost at cost containment tended to reduce their value-driven character.

In sum, there was great ambiguity in the literature regarding the issue of whether one ownership form is superior to another or even that different from another. The only exception to this general ambiguity was the theme of staffing ratios, finding consistently lower nursing staff levels at for-profit than at not-for-profit organizations, especially in nursing homes. Crucially for the UK health care context, two studies found that government-run facilities had the best staffing ratios. A number of other studies concluded that the public sector seems to be the provider of ‘last resort’ in a mixed economy of health care providing for those who cannot afford other forms of health care (e.g. Green et al., 2005; Horwitz, 2003).
3.3. Methodological Problems of studies comparing performance of different health care provider forms

Our systematic review of the performance literature of the not-for-profit sector highlighted a number of methodological problems relating to comparisons of not-for-profit health care with other sectors. First, two thirds of the studies derived from the US where a completely different – namely an unequal, inequitable and non-universal – health care system is in place, so that the findings have limited applicability to the UK and most European countries. Second, most of the studies suffered from weaknesses in their study design when comparing different types of providers and particularly failed to define key variables appropriately. Third, most studies assessed the importance of contextual factors such as the market structure and, where it was considered, the regulatory framework in which providers operated as much more important than the ownership form of providers. Finally, provider comparisons could not take account of the problem of patient and treatment selection. Without a geographic population perspective, it was impossible to draw any conclusions at the provider level about access or service outcomes. Each of these is considered in turn.

Limited applicability to the UK health system context

One key question of our review was the transferability of findings across systems and countries, in particular to the UK with its universal and not market-based health care system. It is important to note that even universal systems vary in the range of services covered, and the extent of coverage. For example, most long-term care for older people in the UK is not included in the NHS. Another difference between the countries is the payment systems which can be multi-payer or single payer. In multi-payer systems, health care is funded by a variety of public and private contributions and there is no one central organization which administers the collection of fees and payment of health care costs. These three dimensions, universality, coverage and payment system, need to be taken into account when comparing health systems.

In the US, where the public sector plays only a minor role, the comparison was between the not-for-profit sector, which was historically dominant in providing health and social care, and the increasingly important for-profit sector, whereas the move in the UK is from a public health system to a mixed economy, including private for profit and not-for-profit providers. This limits the value of this literature for UK policy making because there was almost no attempt to compare performance outside the market paradigm. Indeed US studies explicitly excluded
the few public sector hospitals in the databases from their analyses. For example, the Veterans Healthcare System is a publicly funded, integrated healthcare service, but its hospitals were usually deliberately excluded from the studies we have reviewed. This means that the findings about quality of care, efficiency of services, etc. usually referred only to those with insurance coverage, that is, the wealthier parts of the population. This factor introduces additional complexities for any comparison and makes the transfer of such findings to the UK very difficult due to the non-comparability of the health systems. Moreover, the literature did not deal with the issue of universal access except in the very narrow sense of uncompensated emergency care. Interestingly, an official US government study which (exceptionally) included public hospitals in a comparison of uncompensated care benefits found that government hospitals generally devoted substantially larger shares of their patient operation expenses to uncompensated care than did not-for-profit and for-profit hospitals, thus casting doubt on the current policy of tax exemptions for not-for-profit entities (GAO, 2005).

**Study design weaknesses**

Already a literature review by Sheaff et al. (2003: 7) on organisational performance criticised “the weak methodologies underpinning many of the studies”, “uncertainty about the generalisability of the results”, and “the lack of clear conceptual and theoretical frameworks to guide the analysis and interpretation of the evidence”. The New York Academy of Medicine (1999) complained that there has so far been no good study design comparing not-for-profits and for-profits. Two further systematic reviews comparing not-for-profit and for-profit hospitals identified the absence of any randomised controlled trials and the potential for confounding factors. “Ideally, studies would have adjusted for, or considered as explanatory factors, other variables for which data were not available” (Devereux et al., 2002: 1404). Eggleston et al. (2006) pointed out that conventional methods of meta-analytic synthesis should be applied with great caution, since the same hospitals were frequently included in several different studies.

We confirm these scepticisms. Furthermore, the large scale studies could not establish causality due to the use of cross-sectional data. In addition, study sub-samples, e.g. the number of hospitals of a particular ownership type, were often small, limiting the statistical power of the models. The qualitative studies, in turn,
were often case studies with very limited generalisability. Most studies did not follow the Cochrane Effective Practice and Organisation of Care Group (EPOC) recommendations for case studies. For example, there are only a few controlled before and after studies following a change of policy or a change of provider in a certain area (see http://www.epoc.uottawa.ca/FAQIncludedStudies.doc). A study by Vita and Sacher (2001) was one rare exception. They took an ex post event study approach and compared pre- and post-merger prices charged by the merged entity and one remaining rival, controlling for case-mix, input prices, and other cost- and demand-side characteristics. They found significant price increases after the merger that were not due to higher post-merger quality thus raising considerable antitrust concerns and, in accordance with our findings, highlighting the effect of market structure on hospital behaviour.

It was especially problematic that a number of studies which compared cost efficiency between the different ownership types did not attempt to control for quality of the services provided nor for case mix, that is, patient characteristics which are likely to influence the medical outcome. In other words, we do not know whether efficiency was achieved by sacrifices relating to quality of care or by choosing the healthiest patients or cherry-picking the most profitable services. Studies did little to adjust for patient characteristics of the various health institutions included in the analyses, including disease severity.

A particularly serious shortcoming was the failure to define key terms and establish valid measures of key variables (see also Sheaff et al., 2003). While many of the studies purported to measure efficiency, that is the relationship of inputs to outputs, most studies were in fact measuring economy as reflected in operating costs. Furthermore, the data sets were usually aggregate hospital statistics, created and generated for other purposes, resulting in some of the measures having little sensitivity or specificity. Some of the quality of care measures relied on only imperfect measures or did not include any objective assessment of quality. Cost comparisons were difficult to undertake on the basis of available data.

*Disregard of regulatory context*

Although many studies highlighted the importance of the market conditions under which not-for-profit organizations operate and what the regulatory framework stipulates, there was little concern for differences in regulation between states. Studies which sought to compare provider performance across a number of states...
did not address state differences in regulation, with regard to reimbursement mechanisms, eligibility for health care, Medicaid benefits or community benefits. However, quality standards and eligibility for health care were often the product of state regulations which vary between states.

The literature on uncompensated care itself constituted one of the few attempts to come to grips with the variation in regulatory standards, although attempts to describe regulatory mechanisms in this context were limited and difficult to summarise in the form of a conventional systematic review. We had to bear in mind that the specific regulations defining uncompensated care varied between each of the 50 US-states. Yet, the different definitions of uncompensated care and community benefits which had to be provided by not-for-profits in order to enjoy tax-exemption have not been systematically captured in the literature. A GAO report showed that US state definitions of community benefits varied considerably and most strikingly, some states did not define such a benefit at all (GAO 2005: 24). This meant that findings of studies which had as their focus just one region or state only, thereby controlling for regulation differences, were not generalizable to other states. Moreover, the problem of uncompensated care is only an issue in a non-universal insurance-based health system.

Already Schlesinger and Gray (2006) in a summary of the current public policy debate on the question of not-for-profit versus for-profit organization in the USA pointed out that the regulatory environment is important. Some studies compared hospitals operating under relatively benign market conditions with others operating in rather harsher contexts. Comparisons of hospitals before and after conversion from not-for-profit to for-profit status showed few differences because the hospitals that converted were typically struggling, and changed status in order to achieve certain strategic goals, with the underlying characteristics of the hospitals remaining largely the same.

Currie, Donaldson and Lu (2003) highlighted two important factors to be considered when evaluating cost differences between different providers: first, reimbursement mechanisms influence the results. For example, in a cost-based reimbursement system, it would be rational to seek higher profits by charging more, rather than reducing costs, since the fact that all reasonable expenses were reimbursed meant that inter-hospital competition was based on quality, amenities, and availability of technology (what the authors called ‘a medical arms race’). Second, information on cost differences could not be meaningfully
interpreted in the absence of information on service quality. For instance, the apparently low efficiency of not-for-profits may indicate that they provide a high-quality service; conversely low administrative costs may indicate that less attention is given to monitoring service quality. Comparisons may also be difficult because of the problem of self-selection – not-for-profits may appear less efficient than for-profits, but this could be because not-for-profits were operating in less favourable locations, pursuing their goal of providing access to care, whereas for-profits chose to operate in areas most likely to generate a good return on investment. “Overall, the literature contains a few studies which favour for-profits hospitals, some which favour not-for-profits and a majority of studies which suggest that there is no significant difference” (Currie, Donaldson and Lu, 2003: 13). This reflected exactly the findings from our own systematic literature review. When not-for-profit organisations faced sharp competition or financial stress, the empirical evidence pointed to isomorphism or convergence of behaviour among different providers of medical care.

Population - provider focus
A major limitation of these studies, not addressed in previous literature reviews, was their focus on providers within a market for healthcare. Their main concern was to identify behavioural changes on the part of hospitals, hospices, nursing homes or insurance plans when reimbursement systems, ownership form or the regulatory framework changed. That is, their starting point was not the population covered, with the result that the effects on patients or the implications for access to care were unclear. Because of the non-universal and market character of the US health care system, the literature (with few exceptions) does not, and in most cases cannot, account for variations in access to the different provider schemes and in the entitlements of whole populations. Universal and comprehensive coverage is, however, a key goal of the British NHS. Because of the absence of a universal base in the US, the measurement of performance is difficult as it is confounded by issues of adjustment for case-mix, selection bias, etc. This, plus the difficulty of assessing efficiency due to the lack of outcome measures, means that results may not be interpretable and explains the conflicting findings in the literature.

4. Conclusion

Systematic reviews in the medical sciences have traditionally focussed on estimating the effectiveness of specific therapeutic interventions. The literature
searches, definition of inclusion and exclusion criteria, and choice of methods for combining evidence follow readily from such research questions, as Bravata et al. (2005) argue. Although using systematic literature reviews as a methodology for evidence-based policy research offers some promising potential if done in a thorough way, following a detailed research protocol and a meticulous search for relevant literature, in practice there are some problems with transferring this methodology from the medical sciences, where it is well established and highly regarded, to the social sciences. For once, the literature search as well as the development of inclusion and exclusion criteria are less straightforward for systematic reviews of topics like organisational performance of health care providers. For example, the relevant literature may not only constitute of published articles which are easily searchable by databases, but of government papers, research reports and various forms of grey literature which will be found on websites of health care providers or voluntary sector associations.

Secondly, this transfer of methodology from the medical to the social sciences is problematic due to the fact that research designs vary enormously in the social sciences and it is difficult to appraise the quality of many studies in comparison to others. For example, it seems difficult to assess the reliability of qualitative studies. If the studies analysed in a systematic literature review are questionable, then any conclusions drawn from them are equally inconclusive. This does not only refer to qualitative studies which sometimes rely only on a handful of cases, but as this literature review has shown, also to quantitative studies which constituted the main research corpus included in our review. Obviously this is not only a problem inherent to the social sciences but also to the medical sciences; however, in the latter clear quality criteria such as the use of randomised control trials make a critical appraisal of underlying studies easier.

Another question is, how do we weigh the importance of a qualitative study with a small N relative to a large N quantitative study? Whereas large-N quantitative studies, if done well, offer a great potential for generalisability and theory-testing, small-N qualitative studies have other benefits such as in-depth understanding of social phenomena. It is also debatable to what extent different studies can be synthesised in a meaningful way if they are based on completely opposite epistemological assumptions.

A technically refined meta-analysis is therefore often not feasible if the included studies are as heterogeneous as in the example presented here. We decided t
review the literature in a less sophisticated way, that is, according to six analytical themes, in order to draw any conclusions about the evidence base.

These limitations not withstanding, our example has shown that a systematic literature review can be useful in examining the evidence base of policy priorities. By comparing all recent studies on the topic of not-for-profit sector performance we could establish in a systematic manner that many theoretical arguments brought forward in favour of these organisations do not stand the empirical test. As a result of our analysis we could make explicit policy recommendations, e.g. that contextual and regulatory factors have to be taken into account. It is for the future to see what the reaction of policy-makers is to this systematically gained evidence. Is evidence-based policy just the latest fashionable trend or is there a real determination of those in charge to transform such evidence into real politics?
References


