Do maximum waiting times guarantees change clinical priorities for elective treatment? Evidence from Scotland

Silviya Nikolova [a,*], Arthur Sinko [b], Matt Sutton [c]

[a] Academic Unit of Health Economics, Leeds Institute of Health Sciences, Charles Thackrah Building, 101 Clarendon Road, University of Leeds, Leeds LS2 9LJ, UK
[b] Economics, School of Social Sciences, Arthur Lewis Building, Oxford Road, University of Manchester, Manchester M13 9PL, UK
[c] Manchester Centre for Health Economics, Institute of Population Health, Jean McFarlane Building, Oxford Road, University of Manchester, Manchester M13 9PL, UK

ARTICLE INFO

ABSTRACT

The level and distribution of patient waiting times for elective treatment are a major concern in publicly funded health care systems. Strict targets, which have specified maximum waiting times, have been introduced in the NHS over the last decade and have been criticised for distorting existing clinical priorities in scheduling hospital treatment. We demonstrate the usefulness of conditional density estimation (CDE) in the evaluation of the reform using data for Scotland for 2002 and 2007. We develop a modified goodness of fit test to discriminate between models with different numbers of bins. We document a change in prioritisation between different patient groups with longer waiting patients benefiting at the expense of those who previously waited less. Our results contribute to understanding the response of publicly funded health systems to enforced targets for maximum waiting times.

© 2015 Elsevier B.V. All rights reserved.

1. Introduction

Waiting times are of public concern in state healthcare systems because they are a key determinant of satisfaction with public services (Sanmartin et al., 2007; Cutler, 2002), a perceived indicator of public sector inefficiency (Cullis and Jones, 1983, 1985; Oliver, 2005; Smith, 2002), and a source of discomfort and anxiety (Lindsay and Feigenbaum, 1984; Propper, 1995; Siciliani and Hurst, 2005). It is possible also, though the evidence is very limited, that delays in treatment may have negative health consequences (Siciliani and Gravelle, 2008; Appleby et al., 2003; Noseworthy et al., 2005; Garbuz et al., 2006; Escobar et al., 2009; Oudhoff et al., 2007; Nikolova et al., 2014). Also of concern are variations in waiting times across geographical areas and personal characteristics since such variations may represent a source of inequity (Dimakou et al., 2009; Askildsen et al., 2011). In a number of OECD countries, individuals with higher socioeconomic status (as measured by income or educational attainment) tend to wait less for publicly funded hospital care than those with lower socioeconomic status (see (Siciliani and Verzulli, 2009; Cooper et al., 2009; Laudicella et al., 2012) for England, (Johar et al., 2010; Sharma et al., 2013) for Australia, (Monstad et al., 2014; Carlsen and Kaarboe, 2010) for Norway.)

The setting for this paper is Scotland where high-profile, political guarantees on waiting times for elective hospital admissions were introduced in 2003. We compare the structure of waiting lists for elective surgery before and after the reform. We find that prioritisation between different patient groups changed with longer waiting patients benefiting at the expense of those who previously waited less. Our results contribute to understanding the response of publicly funded health systems to enforced targets for maximum waiting times.

1.1. Institutional background

Long waiting times for NHS treatment were a significant source of public and policy concern across the United Kingdom throughout the 1990s. Prior to devolution in 1999, targets for elective wait were set at 12 months for Scotland by Patients' Charter (1995). However, this Charter was not rigorously enforced.

* Corresponding author. Tel.: +44 01133430578.
E-mail addresses: s.k.nikolova@leeds.ac.uk (S. Nikolova),
arthur.sinko@manchester.ac.uk (A. Sinko), matt.sutton@manchester.ac.uk
(M. Sutton).

The first attempt in earnest by Scottish Executive (the new devolved administration responsible for health policy) to reduce waiting times was announced in 2000 (Scottish Executive, 2000). The maximum waiting time for elective patients was to be reduced from 12 months to nine months in December 2003. In addition, patients were to wait no more than 12 weeks for angiography or 24 weeks for revascularisation by end of 2002. Finally, the target of two months from urgent referral to treatment for all cancers (and one month for breast cancer) was set, to be achieved by the end of 2005.

A more ambitious target of six months for 2005 was announced in a 2002 press release (Audit Scotland, 2006). This press release also announced targets of eight weeks for angiography or 18 weeks for revascularisation by end of 2004. However, the strongest signal that health policy in Scotland was changing was when the White Paper issued in February 2003 offered patients guarantee that "waiting times targets will be met...[and] monitored". It also emphasised the way waiting lists were managed expressing concern that clinical activity response to waiting time pressure has "not resulted in sustained service improvement and... sometimes distorted clinical priorities". A further White Paper in 2004 pledged to reduce waiting times to 18 weeks for inpatients by 2007 Scottish Executive (2004a,b)Scottish Executive, 2004. The targets for cardiac surgery was shortened to 16 weeks, and new targets were set for cataract surgery of 18 weeks from referral to treatment and nine weeks for eight key diagnostic tests, to be achieved by the end of 2007.

Hospitals in Scotland did not incur any economic penalty if waiting times were violated. However, the regional health boards were monitored on a monthly basis on their complete achievement of the maximum waiting times targets. Individual "breaches" of the waiting times targets had to be reported to the Scottish Executive and were rigorously investigated. This monitoring regime was similar in approach to the "targets and terror" regime that had been adopted in England some years earlier (Propper et al., 2008), and the dissolution of one regional health board in 2006 was credited to its poor performances on waiting times and finances.

Elective patients in Scotland are those that are pre-booked for treatment. Thus elective waiting times reflect the time that elapses between the hospital specialist's decision that a patient needs treatment to the date at which this treatment episode begins. This is only a part, and sometimes less than a majority, of the total delay between when a patient initially seeks and receives treatment. The Scottish NHS operates a gatekeeping system under which, for elective treatments, patients must first seek the advice of a general practitioner (GP) (with a trivial wait), second receive a referral from the GP to a hospital specialist (often a more substantial wait, that may involve waiting for diagnostic test results), and third a decision by the specialist that hospital treatment is necessary (which may also involve waiting for diagnostic test results). Direct access by patients to hospital specialists is only possible for emergency care through hospital Accident and Emergency departments.

Patients at risk of breaching the targets were diverted to a national waiting times centre, a dedicated hospital that the NHS had bought from the private sector. It has been estimated that about GBP 116 million was spent on tackling waiting times in 2004/2005. Approximately 40% (GBP 45.7 million) of this was spent on the national waiting time centre (Auditor General of Scotland, 2006). This additional expenditure on reducing waiting times was made at a time of substantial growth in the general resources spent on the hospital sector in Scotland. Annual growth rates in expenditure in the hospital sector in Scotland were 6.0% in 2000/01, 7.4% in 2001/02, 9.6% in 2002/03, 7.0% in 2003/04, 11.1% in 2004/05, 7.2% in 2005/06 and 4.8% in 2006/07.

1.2. Patient prioritisation

There has been widespread concern that the policy of waiting time guarantees would result in fraudulent statistics and distortion of clinical priorities. The National Audit Office (2001) reported that 20% of consultants surveyed in three specialties claimed that they changed the ordering of patients for treatment in order to meet the 18-month target in England. Given the similarity of reforms in England and Scotland, the behaviour of consultants is likely to be similar. Siciliani and Hurst (2005) suggested that maximum waiting time guarantees in theory may be effective in reducing long waiting times, but might distort the incentives for hospitals: "they are not very effective in reducing mean or median waiting times, if the provider simply gives higher priority to less severe patients (who have waited longest), as they approach the maximum" (p.212). Appleby et al. (2003) conducted "before-and-after" comparison of waiting times distributions for English trauma and orthopaedic patients to evaluate the implications of the reform on patient prioritisation. They calculated that the number of admissions around the 15-month target at the time increased by 2.2% of all orthopaedic admissions in the post-reform period. While they could not unambiguously establish whether additional admissions had lead to delayed treatment for other patients, there was no evidence that very short wait patients suffered. Askildsen et al. (2011) compared actual waiting times to the recommended maximum waiting times in Norway. They found that the reduction in waiting times favoured patients who had longer waits.

Evaluating the impact of government targets for waiting times on patterns of average waits for different patient groups is related to the broader literature of patient ordering for treatment in health care. One principle for decision-making is the "rule of rescue" (Hadorn, 1991). This implies that patients with most serious conditions are treated first. Thus, severity of a patient's illness establishes priority for health care treatment. Cullis et al. (2000) argue that one of the criteria for determining waiting times should be the severity of the condition. This point of view concurs with NICE Citizens Council view on clinical need (NICE Citizens Council, 2002). Using a Cox proportional hazard model, Arnesen et al. (2001) showed that perceived or verified severity of patient health condition is the strongest predictor of a physician's decision regarding wait for inpatient treatment. They also find that age is not a significant predictor although a tendency to longer waiting times for patients aged 70 or older was present. This is consistent with the public's preference to assign higher utility of health to younger, rather than older, patients (Busschback et al., 1994; Cropper et al., 1994).

Despite the fact that the reform towards reduction in waiting times was carried out over the course of the previous decade, it is unclear whether patient prioritisation changed. This paper examines whether the ordering for elective hospital treatment changed after the introduction of maximum waiting time guarantees. We compare waiting times distribution before the implementation of the political guarantees (2002) with distribution after their implementation (2007). Characteristics of patient's health (age, disease severity, disease type) influence doctors' perception of urgency for treatment (Cullis et al., 2000). For example, an increase in disease severity might not reduce the waiting times for elderly patients older than 85 years who already have several health problems, but speed up treatment for younger adults. The conditional density estimation (CDE) allows for such variations in covariate effects. In addition, its flexible specification of conditional probability functions and, hence, conditional expectations of the outcome of interest avoids restrictive assumptions about error distribution and functional form.