

Do Targets Matter? A comparison of English and Welsh national health priorities

Katharina Hauck^a and Andrew Street^{b*}

^aCentre for Health Economics
Monash University
Clayton, Victoria 3800
Australia

^bCentre for Health Economics
University of York
YORK YO10 5DD
England

Katharina Hauck: katharina.hauck@buseco.monash.edu.au

Andrew Street: ads6@york.ac.uk

*** Corresponding author**

Acknowledgements

Funding for this work was received from Shropshire and Staffordshire Strategic Health Authority. KH was sponsored by a Health Services Research Program Grant (number 334114) from the National Health and Medical Research Council, Australia. Thanks are due to John Appleby, Michael Kuhn and Marisa Miraldo for comments on early drafts. The views expressed are those of the authors and not necessarily those of the funding bodies.

Do Targets Matter? A comparison of English and Welsh national health priorities

Abstract

National priorities and performance management regimes in the National Health Services of England and Wales diverged following devolution, most notably with respect to the use of waiting time targets, which have been progressively strengthened in England but were abandoned in Wales in the immediate post-devolution period. We analyse routine data collected over a six-year period from three English and one Welsh hospital trust close to the English-Welsh border to ascertain whether: (a) there is evidence of differential performance over time that relates to the country where the hospital is located; (b) within each hospital, there is evidence that English and Welsh patients faced different waiting times. Over the period the English hospitals recorded increased levels of activity, undertook proportionately more day case activity, and mortality rates fell. Activity levels remained constant in Wales, the proportion of day case activity fell, proportionately more non-elective patients were admitted, and mortality rates rose. There is partial evidence that English patients faced lower waiting times than their Welsh counterparts and were more likely to be admitted with a target waiting period. The stronger performance management regime operating in England appears to have contributed to higher levels of performance in the English hospitals over the period.

Key words: hospital performance, waiting times, targets

1. Introduction

The “internal market” that existed in United Kingdom’s National Health Service (NHS) during 1990-1997 coupled directive policy demands made by central government with decentralisation of operational matters to subordinate organisations (1). Central directives were most evident in the form of target waiting times for admission to hospital, this being the key performance indicator against which many managers considered themselves held to account (1).

In 1998 political power was devolved to the constituent countries of the UK. While each country had exercised a degree of autonomy over health policy prior to devolution, afterwards policy divergence became more pronounced. This divergence is exemplified by the approach to managing waiting times adopted in England and Wales. Immediately after devolution, the Welsh Office abandoned waiting time targets and reduced its reliance on directive policy, concentrating instead on promoting co-operative working between health, local government and the voluntary sector (2). The English Department of Health, in contrast, relied more heavily on central direction, in particular by setting and monitoring increasingly demanding waiting time targets and by strengthening the performance management regime more generally (3).

Targets are appealing because they send strong signals about what type and level of achievement is desired. Governments use targets to focus the attention of public sector agencies; private companies use them to spur the workforce toward corporate objectives. But their use is controversial: targets can produce perverse as well as positive effects.

Perverse effects can take many forms (4-6). Those subject to targets may claim to have met the required standard by making inaccurate statements about the amount, type and quality of their achievements. Targets may concentrate attention exclusively on targeted outcomes at the expense of other desirable, but untargeted, objectives. Targets may fail to reward efficient behaviour if inadequate consideration is given to the different circumstances in which agencies are required to function.

Isolating the effect of the target is difficult, for two main reasons. First, as has long been recognised, it is insufficient to judge the effect of the target solely in relation to the objective to which it applies. Targets may induce various forms of gaming and may divert attention from other important, but untargeted, objectives (7). Evaluation of targets, therefore, requires examination of a more rounded picture of performance that better reflects the overall objectives that the principal wants its agents to pursue.

Second, evaluating the impact of targets requires disentangling what proportion of achievement is due to imposition of a target from other factors that might have contributed to performance, such as institutional characteristics or the contextual features of the agent's operating environment.

Much of the literature examines the effects of targets in the context of a vertical hierarchy, with a principal setting targets that it wants its agents to achieve. In contrast, we analyse a situation in which agents are subject to two distinct types of principal that differ with respect to the emphasis they place on targets in their dealings with the agent. This facilitates disentangling the effect of the target-setting arrangements from the institutional and contextual features of the agent. This set-up is detailed in section 2. We then provide a description of policy with respect to waiting times as it applied in England and Wales from 1997/98 to 2003/04. In section 4 we describe our data and analytical approach. Section 5 provides results and section 6 concludes.

2. Institutional set-up

We analyse a set-up in which each provider (a hospital Trust) contracts with two purchasers, E and W , who purchase care on behalf of English and Welsh patients according to national policy guidelines. The purchasers aim to maximise social welfare, and judge benefits to derive from (i) the number of patients treated x , with patients comprising non-electives u and electives v , with $x=u+v$; (ii) the waiting times, r , faced by elective patients prior to their admission; and (iii) other dimensions of

quality q . Benefits increase in the number of patients treated and quality $\partial b / \partial x > 0, \partial b / \partial q > 0$ and decrease by waiting time $\partial b / \partial r < 0$. u , v and r are observable and contractible, whereas q is much less precisely verifiable. The social welfare function for both purchasers consists of these benefits, the payoff g to the provider, and the payment P each purchaser makes to the provider. Hence, total social welfare is:

$$\sum_{j=1}^2 b^j(u^j, v^j, r^j, q^j) + g^j - (1 - \sigma)P^j$$

$$= b^E(u^E, v^E, r^E, q^E) + g^E - (1 - \sigma)P^E + [b^W(u^W, v^W, r^W, q^W) + g^W - (1 - \sigma)P^W]$$

where j indexes the purchasers and σ reflects the shadow cost of public funds, assumed common to the two purchasers.

Like the national governments they represent, the purchasers attach different weights, ϕ^j , to each type of benefit, with:

$$b^E = \phi_1^E u + \phi_2^E v + \phi_3^E r + \phi_4^E q, \quad \phi_1^E + \phi_2^E + \phi_3^E + \phi_4^E = 1$$

$$b^W = \phi_1^W u + \phi_2^W v + \phi_3^W r + \phi_4^W q, \quad \phi_1^W + \phi_2^W + \phi_3^W + \phi_4^W = 1$$

Compared to the Welsh purchaser, the English purchaser attaches greater importance to waiting times, hence $\phi_3^E > \phi_3^W$. Consequently the contract that the provider has with each purchaser differs. The Welsh purchaser makes payments according to the number of non-elective and elective patients treated, with $P^W = P_u^W(u^W) + P_v^W(v^W)$. The contract is not specified explicitly in terms of quality, although the purchaser may monitor quality *ex post*, with the provider being aware that there may be some non-financial cost associated with this argument. Waiting times are not specified as part of the contract.

As well as making a volume related payment, the English purchaser's contract includes a payment related to the waiting time achieved (A) for elective patients, with achievement being inversely related to the waiting times, $A = f(r(v^E))$, $\partial A / \partial r < 0$. If

setting a target waiting time for elective patients, $\bar{r}(v^E)$ the payment function can be expressed as $P^E = P_u^E(u^E) + P_v^E(v^E) + P_r^E(\bar{r}(v^E))$. The setting of a target implies a ‘kinked’ reward schedule, with rewards applying for achievement above the target waiting time, $\bar{r}(v^E)$, and sanctions being imposed otherwise. Rewards may not necessarily be financial: they could, for example, be expressed as levels of opprobrium or praise.

Rewards or penalties may be tapered according to distance from the target, depending on how important it is to the principal that the agent hits the target precisely. The general form of the reward schedule associated with the target can be written as $P_r^E = \alpha + \gamma p[\bar{r}(v^E) - r(v^E)]$, where α represents a lump sum payable on achieving the target, p represents a unit payment related to deviation from the target and γ represents the proportion of this payment made (or withheld) if the target is exceeded (or not met).

Various forms of the reward function are illustrated in figure 1. If $\gamma = 1$, the agent is subject to the linear reward function, ORA. If the agent hits its target exactly, it will receive reward amounting to $\bar{P}_r^E = \alpha$. But the linear function provides no specific incentive to hit this target.

Figure 1 around here

Precision might be incentivised by withholding the lump sum (or, equivalently, dismissing the agent) if the target is not met, and setting $\alpha > 0$ and $\gamma = 0$ if it is, there being no reward for achievement above the target. The reward function would take the form $OMRB$.

Moving from these two special cases, the value of γ can be adjusted to produce a kink in the reward function. One possibility would be to provide weak incentives for the agent to exceed the target. If $0 < \gamma < 1$ applies when patients have shorter waiting times that the target, i.e. $r(v^E) < \bar{r}(v^E)$, marginal rewards are offered at a lower rate than rewards up to the target amount. This would imply that the marginal value to the

principal of an additional unit of achievement above the target is less than the value of a unit of achievement at the level to which the target pertains. The function ORC is one example. The arrangements introduced in the early 1990s to tackle waiting in the Victorian State of Australia were of this form (8).

To encourage achievement up to the target without resort to the extreme sanction of withholding the entire lump sum, the principal might gradate penalties according to distance below the target. One approach would involve reducing the size of the lump sum at a lower rate than the average implicit upon meeting the target level of achievement. This would require that $\gamma > 1$ when $r(v^E) > \bar{r}(v^E)$. For example, an agent reaching a level of achievement $r_1(v^E)$ might receive revenue $P_{r_1}^E$, funds having being withheld at a rate of $\gamma > 1$ over the range $r_1(v^E), \bar{r}(v^E)$. During the operation of the NHS market, some health authorities stipulated gradated penalty clauses in their contract with providers so that Patients Charter standards were met (1).

But generally in the NHS the incentive to meet the target has been non-financial. The main lever has been the threat that the senior management of the provider would be removed from post, with targets, particularly those for waiting times, viewed as the key performance indicator against which managers were judged (1, 9), an incentive regime described as akin to “management by terror” (10). This threat has force in the NHS because, until recently, providers have not been autonomous organisations but subordinate units within the centralised bureaucracy of the NHS.

The provider wishes to maximise a utility function that depends on the payment received from the two purchasers and the level of effort e exerted, such that $\partial U / \partial P^E > 0, \partial U / \partial P^W > 0$, and $\partial U / \partial e_k < 0$, where $k=u,v,q$. The utility function can be expressed as:

$$\max E(U(P^E, P^W, e_k))$$

subject to

$$P^E = P_u^E [u^E(e_u)] + P_v^E [v^E(e_v)] + P_r^E [\bar{r}(v^E(e_v))] + \tilde{q}^E(e_q)$$

$$P^W = P_u^W [u^W(e_u)] + P_v^W [v^W(e_v)] + \tilde{q}^W(e_q)$$

$$q^E = \tilde{q}^E(e_q) + \delta$$

$$q^W = \tilde{q}^W(e_q) + \delta$$

Quality is measured imperfectly as \tilde{q} , which is influenced by the level of effort exerted. δ captures the uncertainty inherent to the imperfectly understood relationship between true and measured quality.

The provider has a number of operational options available to meet the contractual obligations it has to its two purchasers:

- First, it could exert more effort e_u, e_v perhaps by managing patients more efficiently, in order to free up capacity. This would allow the hospital to undertake more activity and admit patients more rapidly from the waiting list. Evidence of such behaviour might include changes in levels of activity, day case rates and average length of stay.
- Second, the provider could reduce attention to aspects of care that are not well specified in the contract. This might entail reduced quality of care (q).
- Third, it could reorganise the queue of those waiting for treatment. This would involve reducing the length of wait for those at the upper tail of the distribution (the “long waiters”) and delaying admission for those who would otherwise have been admitted well before the maximum cut-off. Such behaviour would be evidenced by a change in the distribution of waiting times, with an upper truncation at the target maximum and an increased mean wait (11). Appleby et al found no evidence that waiting time targets in England led to a re-prioritisation of orthopaedic patients (despite surgeons voicing concerns that they would) (12).
- Finally, the provider may prioritise patients who are the subject of the more completely specified contracts (ie those including waiting time targets) ahead of other patients. This behaviour would be evidenced by English patients experiencing shorter waits or/and being more likely to be treated within the maximum target time than their Welsh counterparts.

We explore the extent to which these behaviours occurred in the empirical sections, but first describe the evolution of waiting time policy in the two countries.

3. Waiting list policy in England and Wales

The demand for hospital care is subject to random fluctuation on a daily basis, making it impossible to operate at full capacity without a queue developing (13). The policy challenge is to ensure that patients do not have to endure unduly long waits before they are admitted. The political approach to the meeting this challenge has been the most striking area in which English and Welsh health policy has diverged since devolution. Prior to 1998, policy with respect to waiting times was common across the two countries, there being standard targets specified in the Patients Charter (14). The incoming Labour government dropped the Patients Charter and set about making good its election pledge to reduce the size of the NHS waiting list it had inherited by 100,000 (15). The reduction was achieved by March 2000, after which the English Department of Health shifted attention from how many were waiting to how long patients had to wait, by setting maximum waiting time targets. In March 2000 it was promised that no-one should wait more than 18 months for admission to hospital, with the maximum wait being steadily reduced over time (table 1).

In contrast, in the immediate post-devolution period the Welsh Office abandoned targets relating to waiting lists/times in Wales (2, 16). However, criticism began to mount that Welsh patients faced longer waiting times than in England. The Audit Commission, for example, reported that only “85% of patients in Wales are admitted for routine surgical treatment within a year of being placed on a waiting list, while 96% of patients in England are seen within the same period” (17). The Welsh Office responded by re-introducing targets in March 2001 (18).

On the basis of this reversal in Welsh policy alone, it could be argued that waiting time targets are an effective management tool. However, jumping to this conclusion may be premature, for two reasons. First, the differences in national waiting times may not be due to the target setting regimes in the two countries but to differences in such things as the health needs of the populations or hospital management structures.

Second, achievement of waiting time targets may have come at the expense of untargeted activities. In the remainder of this paper we address these problems.

4. Methods

Rather than comparing waiting times across English and Wales as a whole, we isolate the impact of different national priorities by analysing the performance of four NHS hospitals that serve both English and Welsh patients, three of which are located in England and one in Wales, these being:

- Royal Shrewsbury Hospitals NHS Trust
- Princess Royal NHS Trust
- Robert Jones and Agnes Hunt Orthopaedic Hospital & District NHS Trust
- North East Wales NHS Trust

(The former two hospitals were merged in 2003/04 to form the Shrewsbury and Telford Hospital NHS Trust.) Providers such as these, located close to the border, are likely to have to deal with both English and Welsh purchasers. As purchasers act as intermediaries in furthering national policy, such hospitals face two different “masters” and the supposition is that this may impact on the care delivered to patients from each country (19).

Rather than relying on official statistics, which may be subject to gaming and definitional differences (10), we analyse the Hospital Episode Statistics (HES), compiled by all English hospital trusts, and the equivalent for Welsh hospital trusts, the Patient Episode Database for Wales (PEDW). The HES/PEDW contains computerised patient-level information extracted from each patient’s medical record. Each observation comprises a Finished Consultant Episode (FCE), defined as the period when the patient is under the care of a specific consultant. An episode ends when the patient is discharged from hospital or into the care of another consultant. HES and PEDW contain a wealth of information about patients, including their purchasing authority, the reason for admission, diagnoses, procedures undertaken, discharge arrangements, and waiting times. Most of these variables are common across HES and PEDW. We analyse HES/PEDW data for six financial years, from

1997/98 (prior to devolution) to 2002/03. The database comprised a total of 882,112 observations.

We use these data to undertake two main types of analysis of these providers. First, we perform comparative longitudinal analysis to track the relationship between national policy and performance. The three English providers derive the majority of their income from English purchasing bodies, while the balance is reversed for the Welsh provider. This entails that the English providers will be judged predominantly against the more demanding English performance management regime than is the Welsh provider. If the performance management regimes matter, the English providers ought to display higher levels of output, perhaps coming at the expense of unmonitored aspects of care, than the Welsh provider. Trends at each provider are assessed in relation to:

- The amount and elective/non-elective mix of activity.
- Indicators of efficient behaviour, including day case rates, length of stay and waiting times.
- Patient characteristics, including age, gender and casemix. Casemix complexity is measured by the number of diagnoses that are coded, by the number of operations performed, and using an index constructed on the basis of the version 3.1 Healthcare Resource Group (HRG) to which each patient is assigned. Relative weights for each HRG were calculated using the 2003 National Reference Cost Schedule for English NHS Trusts.¹ The index is centred around 1, this value corresponding to the average Reference Cost reported for elective, non-elective and day case HRGs.
- Quality of care, proxied by in-hospital mortality rates.

Second, we compare waiting times faced by English and Welsh patients within each provider. In order to ascertain whether waiting time targets have a specific – rather than general – effect we assess whether English patients faced shorter waiting times or were more likely to be admitted within the maximum waiting time than Welsh patients. Further, we assess whether any differentials followed a similar pattern to the early relaxation and later strengthening of the target-setting regime in Wales.

We perform two sets of analysis of waiting time data. First, we estimate actual waiting times for the full period for which data are available. Second, we estimate the probability of being admitted before the maximum waiting time for the period in which targets applied. The purpose of these estimations is to identify whether the waiting time or probability of being admitted was associated with having an English or Welsh purchaser, after controlling for patient age, gender and various measures of their casemix complexity. A brief description of the variables is provided in table 2, and the specifications applied in each type of estimation are summarised in table 3.

Tables 2 and 3 around here

For both sets of estimation, we exclude episodes where admission occurred without waiting time, or where waiting time is not recorded. We also exclude episodes in rehabilitation or mental health, and episodes with missing information on one or more of the explanatory variables. This results in an annual sample of around 21,500 observations for Royal Shrewsbury, 11,000 for Princess Royal, 6,000 for Robert Jones, and 16,000 for North East Wales NHS Trust.

OLS estimation of individual waiting times

For each hospital, we assess whether English patients faced shorter waits than Welsh patients over the six year period. We analyse specialties and procedures with average waiting times over 50 days and where there are more than 50 FCEs per year purchased by both Welsh and English purchasers. As a dependent variable, we use waiting time in days. We control for various patient characteristics, as indicated in table 3, and our particular interest is in the dummy variable *Welsh* indicating whether the patient was the responsibility of a Welsh purchaser.

Given that national policy in the two countries changed over time, we expect to observe the following pattern:

¹http://www.dh.gov.uk/PublicationsAndStatistics/Publications/PublicationsPolicyAndGuidance/PublicationsPolicyAndGuidanceArticle/fs/en?CONTENT_ID=4070195&chk=UzhHA3

- 1997/98 – baseline: similar experiences, irrespective of whether patients are English or Welsh, because there was a common policy in the two countries. Hence, the coefficient on the *Welsh* variable should be non-significant.
- 1998/99-2000/01 – divergence: where the relaxing of targets and less stringent performance management regime in Wales leads to hospitals placing greater emphasis on meeting English targets, resulting in lower waiting times for English patients. The *Welsh* variable should, over this period, be increasingly positive and significant.
- Post-2001/02 – convergence: following the re-introduction of targets and strengthening of the performance management regime in Wales. The *Welsh* variable should remain positive and significant, but less so than during the period of divergence.

Probit estimations

We then analyse the probability of being admitted within the target waiting time during the three financial years covered by our data when such targets applied, this period being 2000/01-2002/03. This analysis is restricted to hospitals located in England, as the maximum waiting time targets did not apply to Welsh hospitals. We assess whether English patients were more likely than Welsh patients to be admitted within the maximum waiting time that applied at the time they were admitted. We estimate probit regressions for the probability of being admitted within target $E(\text{TARGET}/X) = P(\text{TARGET}-1/X)$, where TARGET is a binary variable indicating whether a patient is admitted within the target waiting time and X is the set of explanatory variables as listed in Table 3. For each of the three financial years, we estimate a joint model for all three English hospitals, and separate models for each hospital. We test for general misspecification of the models using a RESET test adapted to Maximum Likelihood Estimation, described in Jones (20).²

² In addition, we allow for a possible interrelation between age and casemix complexity by including interaction terms in an alternative model specification. We test for the validity of the interaction terms by conducting Likelihood Ratio tests between specifications with and without interaction terms (unrestricted and restricted models, respectively). The Likelihood Ratio tests favour the specification without interaction terms over the ones with for 7 out of the 11 models. For the sake of consistency, we do not include interaction terms in any of the models. Anyway, the alternative specification produces very similar results on our variable of interest, with the average effects of Welsh differing only in the 3rd or 4th decimal place for all models. The specification passes the RESET test for 6 of the 11 models, indicating that the models may suffer some misspecification. However, as the

We use STATA/SE 9.0 for the estimations and tests (21).

5. Results

Activity and waiting time trends

Tables 4-7 provide summary statistics for each of the four hospitals. The three English hospitals exhibit different temporal experience to the Welsh hospital in respect to the size and composition of their workload.

There was a substantial (17%) increase in activity at the Royal Shrewsbury over the six-year period (table 4, figure 2), the majority of which was elective work. Activity growth was achieved partly by reducing length of stay and doing more work on a day case basis. However, this failed to free up sufficient capacity to accommodate the additional workload, creating a backlog and forcing patients to wait longer prior to admission, as evidenced by a rise in mean waiting times from 81 in 1997/98 to 106 days in 2002/03.

There was also a dramatic (40%) rise in activity at the Princess Royal between 1997/98 and 2002/03, with elective and non-elective activity showing a similar trend (table 5). Capacity pressures were eased slightly by reducing length of stay (from 4.4 to 3.7 days), but waiting times increased from 74 to 89 days.

Little non-elective work is undertaken at the Robert Jones, but elective activity increased by 27% over the six years (table 6). Moreover, casemix complexity increased over time, as indicated by the casemix index, and there was also a declining proportion of patients recorded as having no operation. Despite increasing day case activity and reducing length of stay, mean waiting times increased from 135 to 156

joint model for all hospitals passes the RESET tests for 2 of the 3 years, we are reasonably confident that the problem will have only a slight impact on our results.

days. This suggests that, as at the other English hospitals, activity increased faster than available capacity.

In contrast to the English hospitals, activity at the North East Wales Trust remained constant over time, but the proportion of non-elective activity increased (table 7). Nevertheless there is no evidence that casemix became more complex over time. Without commensurate improvements in efficiency (average length of stay remained constant, and day case rates actually fell), the non-elective work crowded out the elective activity, resulting in an increase in waiting times from 71 to 97 days.

Table 4-7 around here

Figure 2 around here

Quality

The only routinely collected indicator of quality of care is in-hospital mortality, although this is not formally incorporated into contractual arrangements. Mortality rates at the three English hospitals are close to or below the English national average (of 2.3% in 2002/03 (22)). Over the period the rate averaged 1.5% at Royal Shrewsbury, 2.6% at the Princess Royal and 0.8% at Robert Jones, where a declining trend is evident (figure 3). Mortality rates were highest at the Welsh hospital and increased steadily from 2.9% in 1997/98 to 4.3% in 2002/03, the increase amounting to 664 additional deaths in this hospital than six years previously.

Figure 3 around here

Waiting times

Waiting times for each of the specialties and procedures listed in Table 1 were estimated according to the model specifications detailed in Table 3 on each year's worth of data. Although not intended as explanatory models, the majority of models estimating the waiting time for specific procedures pass the RESET test (83% - 52/63). The models estimating specialty-level waiting times appear misspecified, with only 12% (7/59) passing the RESET test.

Estimation results for the full set of variables are not reported. Rather, table 8 presents the coefficient estimates and their standard errors for the *Welsh* variable for each model. The coefficient shows the average difference in days that Welsh patients wait compared to their English counterparts after controlling for other patient characteristics. A positive coefficient implies that the Welsh wait longer, on average, than English patients.

Table 8 around here

In general, this provides weak evidence to suggest that English patients faced shorter waiting times than Welsh patients. This extent of the difference varies by provider and procedure. For example, at the start of the period, Welsh patients faced shorter waiting times for admission to trauma & orthopaedics at the Robert Jones, a situation that was reversed after 2000/01, in line with expectations (figure 4). The wait for admission for cataract removal at the Royal Shrewsbury also corresponds to the divergent temporal policy emphasis on waiting times in the two countries after 1998/99, although there is evidence that the English faced shorter waiting times prior to the divergence in policy (figure 5).

Figures 4 and 5 around here

After 2000, maximum waiting time targets were introduced for English hospitals, which became more demanding over time (3). Table 9 shows the number of elective patients treated at the three English hospitals according to whether or not they were admitted before the target waiting time that applied on the date that they were admitted. At the end of the first year that the maximum waiting time applied (2000/01) the vast majority (98.8%) of patients were seen within the target of 18 months. This proportion fell in later years, with only 96.6% being admitted within the target period that applied during 2002/03. The fall was partly due to the target becoming more demanding, but is also explained by an increase in activity, with 2.8% more elective patients being admitted during 2002/03 compared to 2000/01. As there was no compensating reduction in other types of activity (such as non-elective

admissions), the increase in elective activity placed greater pressure on existing capacity.

Table 9 around here

Throughout this period, patients who were the responsibility of English purchasers were more likely than their Welsh counterparts to be admitted within the target period. In 2000/01, 99% of English patients were admitted within 18 months, compared to only 97.5% of Welsh patients. Achievement of the target declined over time, as did the differential between the English and Welsh, with 97% of English patients being admitted within 12 months in 2002/03 compared to 93% of Welsh patients. The differential was most in evidence at the Robert Jones.

Table 10 presents the predicted probabilities of being treated within target, the average effects for the variable *Welsh* from the probit estimations, and their standard errors. The predicted probabilities are evaluated at the sample means of the explanatory variables. They show the probability of being treated within target for a typical observation, or the ‘reference patient’. The average effects measure the effect of being Welsh on being treated within target waiting time in units of probability (20, 23). The average effects are derived from the coefficient values using the formula:

$$\begin{aligned} &P(TARGET = 1/WELSH = 1) - P(TARGET = 1/WELSH = 0) \\ &= F(\bar{X}\beta | WELSH = 1) - F(\bar{X}\beta | WELSH = 0) \end{aligned}$$

The effect of the variable *Welsh* is evaluated at the sample mean of the other explanatory variables.

Table 10 around here

Table 10 shows that the average patient is likely to be treated within target waiting times with a probability of at least 95% in all hospitals and years. The predicted probability decreases slightly over the three years for all hospitals, however, and is lowest for the Robert Jones Trust. At this hospital the probability decreases from around 98% in 2000/01 to 95% in 2002/03. Welsh patients are significantly less likely to be treated within target waiting time than the average patient at Robert Jones Trust,

and at Royal Shrewsbury Trust in the years 2000/01 and 2001/02. In all hospitals, the probability of being treated within target decreases more rapidly for Welsh patients than for the average patient over the three years. At the Robert Jones Trust, in 2000/01, Welsh patients are around 1% less likely, whereas in 2002/03, they are around 7% less likely to be treated within target than the average patient. At less than 1%, the magnitudes of the average effects of Welsh are small for the Royal Shrewsbury Trust in both years where they reach significance.

6. Discussion

The contention of this paper is that differences observed between the English and Welsh hospitals are due - at least in part - to national policy, with the more demanding performance management regime that operated in England over most of the period having a more discernible impact on organisational culture and behaviour. However, there may be other reasons that might explain differences in performance. These include differences in contracting policies, local health needs, local health priorities and funding levels.

In this paper, the problem of controlling for institutional and contextual heterogeneity is avoided by examining routine data collected from four hospitals close to the English-Welsh border that have dealings with both English and Welsh purchasers to ascertain whether:

- There is evidence of differential performance over time that relates to the country where the hospital is located.
- English patients faced shorter waiting times than Welsh patients.
- English patients were more likely than Welsh patients to be admitted before the maximum time.

There is evidence that the English hospitals exerted more effort (e_u , e_v) than the Welsh hospital over the six year period. When comparing trends across hospitals, there are similarities among those located in England and differences to the North East Wales Trust. The English hospitals recorded increased levels of activity, reduced length of

stay and undertook proportionately more day case activity over the period. Activity levels remained constant at the Welsh hospital, the proportion of day case activity fell, and proportionately more non-elective patients were admitted.

There is no evidence that the English hospitals achieved activity increases by compromising on quality, q . Mortality rates at the English hospitals remained low or declined further over the period, but the high and rising hospital mortality rates at the North East Wales Trust are cause for concern. It may be that higher mortality rates at the North East Wales Trust are due partly to the proportionate increase in non-electives admitted to the hospital.

To evaluate the differential emphasis placed on waiting times in each country, $\phi_3^E > \phi_3^W$, we sought evidence that the waiting times faced by English and Welsh patients were related to the early relaxation and later strengthening of the target-setting regime in Wales. Differential waiting time was not a political objective (24) and the reward function described as “management by terror” (10) probably meant that $\gamma \approx 0$ in $P_r^E = \alpha + \gamma p[\bar{r}(v^E) - r(v^E)]$, dampening the incentive to ensure that the English patients faced significantly shorter waits. As such, strong evidence of differential access would not be expected.

Over the three year period in which specific waiting time targets were set for English hospitals, there is evidence that English patients were more likely than Welsh patients to be admitted within the target period. Moreover, over the six year period, a temporal pattern suggesting the English faced shorter waiting times is evident for hip replacements, knee replacements, cataract removals, and varicose vein procedures, with weaker evidence for removal of fixation devices, hernia repairs and tonsil removals. There is no evidence that the Welsh waited longer for examinations of the upper digestive system or bladder examinations. These observed differentials suggest that the emphasis placed on waiting times in the contracts with English purchasers had some, but not overwhelming, influence on internal hospital behaviour.

But consideration of the longer-term trends complicates matters. The hypothesis was that, because of the consistency in policy between the two countries prior to

devolution, there should have been no difference in the waiting times experienced by patients from the two countries prior to 1998. For the procedures that most closely match the expected temporal differentials between 1997/98 and 2002/03, there is evidence of a pre-existing differential. The Welsh waited significantly longer for cataract removals at the Royal Shrewsbury Trust in 1997/98. In contrast, at the Robert Jones Trust, the differential waiting time for hip and knee replacement favoured the Welsh in the period prior to the divergence in policy. Clearly, then, some reason other than the policy regime must be advanced to explain the differentials observed prior to 1997/98. By extension it would be inappropriate to disregard the possibility that this unidentified cause still had influence.

Furthermore, it may be inappropriate to attribute observed differences solely to waiting time policy because, despite holding many institutional and contextual features constant, it cannot be assumed the English and Welsh purchasers have similar arrangements for paying for the number of patients treated. If the formulation of $P^j(x)$ differs between the two purchasers, it may be this, not the waiting time policy, that drives discriminatory behaviour. Unfortunately it would be impossible to ascertain every actual specification of $P^j(x)$ given the wide range of purchasers that had dealings with each of the hospitals and the continuous re-organisation of these purchasing organisations over the period. But there is some recent evidence that inadequate funding from Powys Local Health Board in Wales might be the cause of longer waiting times for Welsh patients at the Royal Shrewsbury Hospital (http://news.bbc.co.uk/2/hi/uk_news/wales/mid_/2981761.stm). However this experience is not indicative of general differences in funding arrangements. NHS funding in Wales has long been more generous than in England, with per capita expenditure 16% higher in 1996/97 and 9% higher in 2002/03 (25, 26). This suggests that the stronger performance management regime may have acted as a counterweight to the lower level of funding in England.

Finally, it might be that longer waiting times are acceptable to the Welsh – if they get something else in return. Within each hospital, the utility loss associated with longer waiting times may have been offset by improvements in other dimensions of quality, q . This seems unlikely, though. The only measure we have of q is mortality rate, but

we found no evidence that the rate differed for English and Welsh patients within each hospital. “Compensation” may instead have come from other sources. For instance, commentators have suggested that in the immediate post-devolution period the Welsh Minister for Health and Social Services reduced attention to waiting time targets in Wales and instead “emphasised what has been termed ‘joined up working’ focusing on partnerships between health, local government and the voluntary sector” (2). It is unclear whether sufficient off-setting benefits were generated to justify the shift in policy focus, but the subsequent re-introduction of targets in Wales suggests not.

References

1. Hughes D, Griffiths L. On penalties and the Patient's Charter: centralism v de-centralised governance in the NHS. *Sociology of Health & Illness* 1999;21(1):71-94.
2. Burnett S. Finding a way through the forest: the Health and Social Services Committee. In: Jones J, Osmon J, editors. *Inclusive Government and Party Management: the National Assembly for Wales and the Work of its Committees*. Cardiff: Institute of Welsh Affairs and Welsh Governance Centre; 2001.
3. Secretary of State for Health. *The NHS Plan: a plan for investment, a plan for reform*. CM 4818-1. London: The Stationary Office; 2000.
4. Nove A. *The Soviet economic system*, second edition. London: Allen and Unwin; 1980.
5. Ellman M. *Socialist Planning*. Cambridge: Cambridge University Press; 1979.
6. Smith PC. Performance indicators and control in the public sector. In: Berry AJ, Broadbent J, Otley D, editors. *Management Control: Theories, Issues and Practices*. Basingstoke: Macmillan; 1995. p. 163-178.
7. Smith PC. Outcome-related performance indicators and organizational control in the public sector. *British Journal of Management* 1993;4:135-151.
8. Street A, Duckett S. Are waiting lists inevitable? *Health Policy* 1996;36:1-15.
9. Shifrin T. Milburn puts managers "on probation". *Health Service Journal* 2001;27 September.
10. Bevan G, Hood C. What's measured is what matters: targets and gaming in the English public health care system. mimeo 2005.
11. Siciliani L, Hurst J. Tackling excessive waiting times for elective surgery: a comparative analysis of policies in 12 OECD countries. *Health Policy* 2005;72:201-215.
12. Appleby J, Boyle S, Devlin N, Harley M, Harrison A, Thorlby R. Do English NHS waiting time targets distort treatment priorities in orthopaedic surgery? *Journal of Health Services Research and Policy* 2005;10(3):167-172.
13. Worthington DJ. Queuing models for hospital waiting lists. *Social Science and Medicine* 1987;83(5):413-22.
14. Department of Health. *The Patient's Charter and you*. London: HMSO; 1995.
15. Harrison A. The war on waiting. *Health Care UK* 2000;4:52-60.

16. McClelland S. Managing Health Services and Policy in Wales. In: The European Academy of Management 2nd Annual Conference, 'Innovative Research in Management'; 2002 9th-11th May 2002; Stockholm; 2002.
17. Audit Commission. Waiting for Elective Admission. London: HMSO; 1993.
18. McClelland S. Health policy in Wales - distinctive or derivative? *Social Policy & Society* 2002;1(4):325-333.
19. Commission for Health Improvement. Clinical Governance Report for Royal Shrewsbury Hospitals NHS Trust; 2002.
20. Jones AM. Applied Econometrics for Health Economists - A Practical Guide. London: Office of Health Economics; 2001.
21. STATA Corp. Stata Statistical Software: Release STATA/SE 9.0. College Station, Texas: STATA Corporation; 2005.
22. Dawson D, Gravelle H, O'Mahony M, Street A, Weale M, Castelli A, et al. Developing new approaches to measuring NHS outputs and productivity, Final Report. York: Centre for Health Economic research paper 6; 2005.
23. Greene WH. *Econometric Analysis*. 5th edition. Upper Saddle River, New Jersey: Prentice Hall; 2003.
24. Hutton J. Waiting times. In: House of Commons Hansard written answers for 11 Oct 2004. London: United Kingdom Parliament; 2004.
25. Dixon J, Inglis S, Klein R. Is the English NHS underfunded? *British Medical Journal* 1999;318:522-526.
26. Alvarez-Rosete A, Bevan G, Mays N, Dixon J. Effect of diverging policy across the NHS. *British Medical Journal* 2005;331:946-950.

Figure 1 {RevFuncTarget.xls}

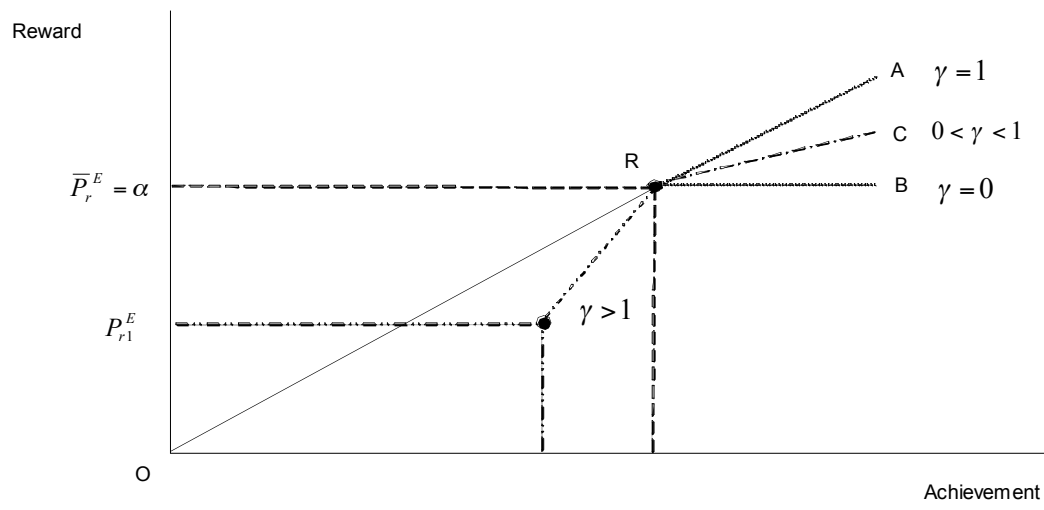


Table 1 English waiting time targets, applied to each Trust

Target No-one should be waiting more than ...	Date to be achieved
18 months	March 2000
15 months	March 2002
12 months	March 2003
9 months	March 2004
6 months	March 2005

Table 2: Variable descriptions

Variable name	Description	HES data field	PEDW data field
TARGET	Taking a value of 1 if patient is admitted within the target waiting time applicable at day of admission, 0 otherwise	Constructed from elec dur and admidate	-
ELECDUR	Waiting time in days	elec dur	'episode start date' minus 'date decided to admit'
ENGLISH	Purchaser is an English Health Authority, English GP fundholder, Primary Care Group, or Primary Care Trust	purcode	registered lhb
WELSH	Purchaser is a Welsh Health Authority, Welsh GP fundholder, or Local Health Board	purcode	registered lhb
OTHERPUR	Purchaser is private, Scottish or Irish NHS purchaser, Special Health Authority or Care Trust	purcode	registered lhb
DAYCASE	Taking a value of 1 if patient was treated as a daycase, 0 otherwise	classpat	patient class
AGE	Age in whole years at episode start date	age	age
AGE2	AGE ²	age	age
AGE3	AGE ³	age	age
FEMALE	Taking a value of 1 if patient is female, 0 otherwise	sex	sex
HRGW	Casemix complexity – hrg weights	hrglate	hrg drg submitted
NUMBERDIAG	Casemix complexity – number of diagnoses	diag_1–diag_3	diag1–diag6
NUMBEROP	Casemix complexity – number of operations	oper_1, oper_2	oper1–oper4
EXTRACAP	Type of procedure used in removal of cataract: extracapsular extraction	oper_1	oper1
PROSTHRE	Type of procedure used in removal of cataract: phacoemulsification	oper_1	oper1

Table 3: Model specifications and explanatory variables

+: included in model

-: excluded from model

	Explanatory variables										number diag	number op	daycase
	welsh	english	otherpur	female	age	age2	age3	hrgw	extracap	prosthre			
OLS Regressions													
Robert Jones and Agnes Hunt Orthopaedic Hospital NHS Trust													
Trauma & Ortho	+	-	+	+	+	+	+	+	-	-	+	+	+
Hip replacements	+	-	+	+	+	+	+	-	-	-	+	+	-
Knee replacements	+	-	+	+	+	+	+	-	-	-	+	+	-
Fixation removal	+	-	+	+	+	+	+	-	-	-	+	+	+
Royal Shrewsbury Hospital													
General Surgery	+	-	+	+	+	+	+	+	-	-	+	+	+
Hernia repairs	+	-	+	+	+	+	+	-	-	-	+	+	+
Varicose veins procedures	+	-	+	+	+	+	+	-	-	-	+	+	+
Digestive system exams	+	-	+	+	+	+	+	-	-	-	+	+	+
Bladder exams	+	-	+	+	+	+	+	-	-	-	+	+	+
Ophtamology	+	-	+	+	+	+	+	+	-	-	+	+	+
Cataract removals	+	-	+	+	+	+	+	-	+	+	+	+	+
Gynaecology	+	-	+	-	+	+	+	+	-	-	+	+	+
Ear, nose & throat	+	-	+	+	+	+	+	+	-	-	+	+	+
Tonsil removals	+	-	+	+	+	+	+	-	-	-	+	+	+
North East Wales NHS Trust													
General Surgery	+	-	+	+	+	+	+	+	-	-	+	+	+
Trauma & Ortho	+	-	+	+	+	+	+	+	-	-	+	+	+
Ophthalmology	+	-	+	+	+	+	+	+	-	-	+	+	+
Gynaecology	+	-	+	-	+	+	+	+	-	-	+	+	+
Probit Regressions													
All English trusts	+	-	+	+	+			+	-	-	+	+	+

excluded in some years due to insufficient number of observations or insufficient variation in the variable

Table 4: Descriptive statistics for Royal Shrewsbury Hospitals NHS Trust

	1997/98	1998/99	1999/2000	2000/01	2001/02	2002/03
Episodes	56,973	59,686	60,525	61,042	63,720	66,441
Elective Episodes (%)	38	40	40	41	43	43
Average Length of Stay	3.8	3.6	3.6	3.7	3.7	3.0
Daycases (%)	23	26	27	29	30	31
Mean waiting time (days)	81	93	75	88	110	106
Mortality Rates (%)	1.4%	1.6%	1.5	1.6	1.4	1.6
<i>Patient characteristics</i>						
English	49,613	52,293	53,700	53,802	55,069	56,725
Welsh	6,885	6,864	6,280	6,712	7,805	9,023
Mean age (years)	41.37	42.71	43.85	44.84	45.61	47.06
Female (%)	59	60	59	59	59	58
Casemix index	0.45	0.46	0.45	0.45	0.45	0.45

Table 5: Descriptive statistics for Princess Royal NHS Trust

	1997/98	1998/99	1999/2000	2000/01	2001/02	2002/03
Episodes	25,037	26,252	28,450	29,108	32,486	34,875
Elective Episodes (%)	74	72	74	73	74	73
Average Length of Stay	4.4	4.6	4.1	4.4	5.6	3.7
Daycases (%)	26	28	26	27	26	27
Mean waiting time (days)	74	87	72	85	78	89
Mortality Rates (%)	2.6	2.9	2.7	2.5	2.6	2.4
<i>Patient characteristics</i>						
English	24,720	25,639	27,822	28,462	31,744	34,074
Welsh	69	91	72	80	118	203
Mean age (years)	47.1	48.3	48.7	89.8	51.36	53.4
Female (%)	48	49	52	48	49	49
Casemix index	0.60	0.60	0.59	0.59	0.62	0.64

Table 6: Descriptive statistics for Robert Jones and Agnes Hunt Orthopaedic & District Hospital NHS Trust

	1997/98	1998/99	1999/2000	2000/01	2001/02	2002/03
Episodes	7,423	8,742	9,098	9,535	8,984	9,430
Elective Episodes (%)	74	79	79	83	85	88
Average Length of Stay	8.0	6.9	6.4	5.9	6.3	6.0
Daycases (%)	17	18	19	21	21	22
Mean waiting time (days)	135	141	142	150	155	156
Mortality Rates (%)	1.0	1.1	0.7	0.5	0.4	0.3
Patient characteristics						
English	4,073	5,032	5,493	5,608	5,038	5,404
Welsh	2,226	2,576	2,688	2,891	2,828	3,052
Mean age (years)	46.89	47.12	46.90	47.29	48.66	48.99
Female (%)	17	18	19	21	21	22
Casemix index	0.89	0.87	0.85	0.86	0.91	0.95

Table 7: Descriptive statistics for North East Wales NHS Trust

	1997/98	1998/99	1999/2000	2000/01	2001/02	2002/03
Episodes	46,947	46,898	46,463	47,937	49,137	47,563
Elective Episodes (%)	41	41	39	37	38	30
Average Length of Stay	5.7	5.5	5.2	5.6	5.6	5.4
Daycases (%)	22	22	20	20	20	13
Mean waiting time (days)	70.79	92.53	84.81	114.52	89.20	96.96
Mortality Rates (%)	2.9	3.1	3.1	3.4	3.6	4.3
Patient characteristics						
English	1,924	2,070	1,998	2,242	2,482	2,636
Welsh	44,634	44,752	44,408	45,654	46,584	44,865
Mean age (years)	49.87	50.85	51.35	51.83	52.02	52.82
Female (%)	57	57	58	58	57	55
Casemix index	0.54	0.53	0.53	0.55	0.55	0.57

Table 8: Estimated beta coefficients on *Welsh* variablenote: Coefficients in **bold** are significant at 5% confidence level

	1997/98		1998/99		1999/2000		2000/01		2001/02		2002/03	
	Coeff	SE	Coeff	SE	Coeff	SE	Coeff	SE	Coeff	SE	Coeff	SE
English specialist trust												
Trauma & Ortho	-37.21	5.17	-19.15	4.87	-17.83	4.46	8.66	4.63	35.97	4.23	39.56	3.73
Hip repl.	-65.96	14.15	-43.22	15.91	-15.02	18.83	49.35	17.52	102.66	16.27	90.25	13.31
Knee repl.	-43.73	17.67	-7.39	22.03	-13.64	21.86	44.73	20.27	60.10	19.09	85.47	14.91
Fixation removal	-38.71	20.66	-13.89	17.79	21.23	17.84	21.56	18.57	16.30	18.41	48.93	21.83
English general trust												
General Surg.	7.76	4.26	0.30	4.29	2.28	3.89	11.73	3.92	10.05	5.43	-4.48	6.47
Hernia repairs	30.50	18.45	-0.37	15.93	40.74	15.92	26.89	14.46	17.00	16.05	34.51	15.40
Var. veins proc.	20.27	28.52	7.27	19.70	20.46	24.01	78.21	23.26	68.16	30.94	64.31	32.32
Digest. sys. ex.	3.44	3.56	-9.02	4.04	-0.51	4.17	9.14	9.09	4.46	14.93	-22.13	11.59
Bladder exams	-8.08	7.85	-12.81	7.39	-16.16	8.56	0.76	7.54	1.71	13.44	-40.18	18.63
Ophtamology	25.40	4.98	8.80	4.93	14.35	5.10	42.12	4.82	41.08	4.94	7.25	4.54
Cataract removals	43.25	11.45	15.72	7.42	66.52	8.68	79.47	6.62	70.96	6.36	21.52	6.30
Gynaecology	4.94	5.47	3.41	6.41	-1.30	6.78	18.75	6.91	-11.07	7.28	-17.57	5.59
Ear, Nose & Thrt.	27.27	5.61	18.37	5.78	8.34	6.12	37.70	7.11	13.17	7.05	1.62	7.72
Tonsil removals	15.18	11.94	5.00	12.50	50.13	10.16	85.19	14.48	11.57	15.88	6.20	19.56
Welsh general trust												
General Surg.	1.24	11.88	5.15	13.21	-5.72	13.42	16.00	14.09	11.31	9.97	14.27	10.13
Trauma & Ortho	26.78	25.52	63.92	26.40	19.09	28.30	79.68	32.04	53.72	29.77	56.66	23.39
Ophtamology	-23.91	14.59	-21.31	20.37	-11.82	18.34	72.20	20.25	-8.85	20.95	-13.64	7.97
Gynaecology	-1.70	14.91	-17.99	16.81	13.59	19.12	-12.34	16.78	2.14	15.75	15.84	18.06

Table 9: English and Welsh patients admitted within and outside target, by English Trust³
(row percentages for every year)

		2000/01		2001/02		2002/03	
		English	Welsh	English	Welsh	English	Welsh
<i>All English trusts</i>							
Within target	n	31,908	4,763	31,276	4,693	31,792	5,095
	%	99.0	97.5	98.4	95.6	97.1	93.1
Outside target	n	327	121	514	218	938	377
	%	1.0	2.5	1.6	4.4	2.9	6.9
<i>Royal Shrewsbury</i>							
Within target	n	18,015	2,798	16,606	2,681	15,782	2,916
	%	99.4	99.1	98.0	97.2	96.8	96.7
Outside target	n	109	26	342	76	523	100
	%	0.6	0.9	2.0	2.8	3.2	3.3
<i>Princess Royal</i>							
Within target	n	10,039	45	11,166	57	12,063	134
	%	98.8	97.8	99.5	98.3	98.0	100
Outside target	n	117	1	58	1	249	0
	%	1.2	2.2	0.5	1.7	2.0	0
<i>Robert Jones</i>							
Within target	n	3,854	1,920	3,504	1,955	3,947	2,045
	%	97.4	95.3	96.8	93.3	96.0	88.1
Outside target	n	101	94	114	141	166	277
	%	2.6	4.7	3.2	6.7	4.0	11.9

³ The total of English and Welsh patients does not add up to the total sample size because the latter also includes episodes purchased by other purchasers (mainly private patients)

Table 10: Predicted probabilities and average effects of *Welsh*

	2000			2001			2002		
	Predicted P (at X)	average effect of <i>Welsh</i>	SE	predicted P (at X)	average effect of <i>Welsh</i>	SE	predicted P (at X)	average effect of <i>Welsh</i>	SE
All three trusts ¹	0.991	-0.009*	0.002	0.984	-0.019*	0.006	0.975	-0.031*	0.003
Robert Jones ²	0.979	-0.014*	0.004	0.968	-0.028*	0.005	0.946	-0.070*	0.007
Royal Shrews ³	0.995	-0.003*	0.001	0.981	-0.006*	0.003	0.976	-0.0002	0.003
Princess Royal ⁴	0.992	-0.012	0.019	0.996	-0.012	0.015	-	-	-

* indicates statistical significance at the 5% level

Results for the reset tests:

¹ LR chi2(1)=3.14, Prob>chi2=0.0766 in 2000; LR chi2(1)=1.21, Prob>chi2=0.2714 in 2001; LR chi2(1)=17.87, Prob>chi2=0.0000 in 2002; ² LR chi2(1)=15.83, Prob>chi2=0.0001 in 2000; LR chi2(1)=5.06, Prob>chi2=0.0245 in 2001; LR chi2(1)=5.51, Prob>chi2=0.0189 in 2002; ³ LR chi2(1)=1.38, Prob>chi2=0.2406 in 2000; LR chi2(1)=1.40, Prob>chi2=0.2366 in 2001; LR chi2(1)=88.99, Prob>chi2=0.0000 in 2002; ⁴ LR chi2(1)=1.13, Prob>chi2=0.2876 in 2000; LR chi2(1)=2.39, Prob>chi2=0.1222 in 2001;

Figure 2: activity rates {change over time.xls}

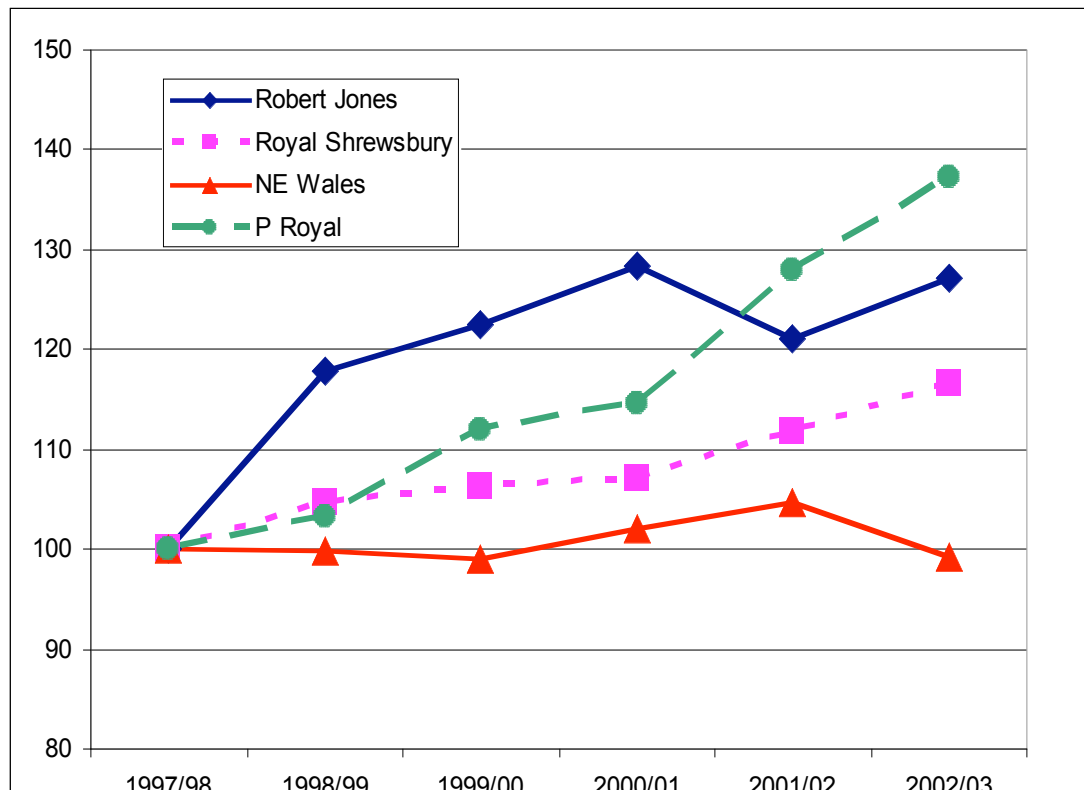


Figure 3: mortality rates {change over time.xls}

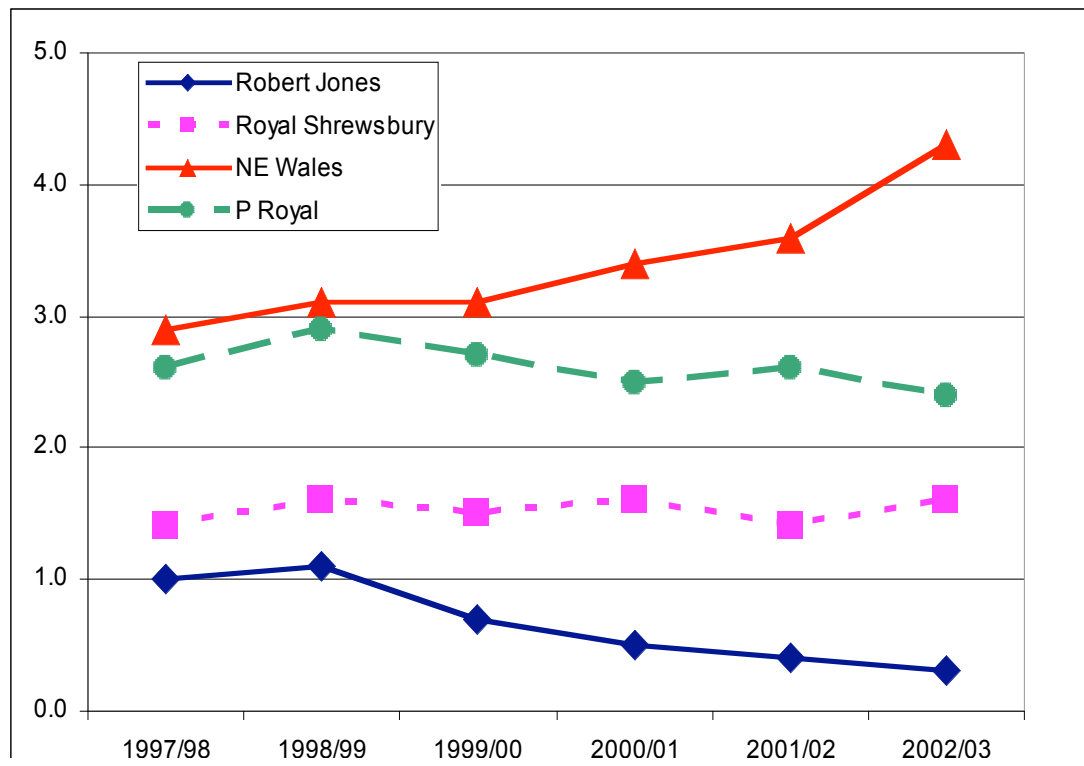


Figure 4

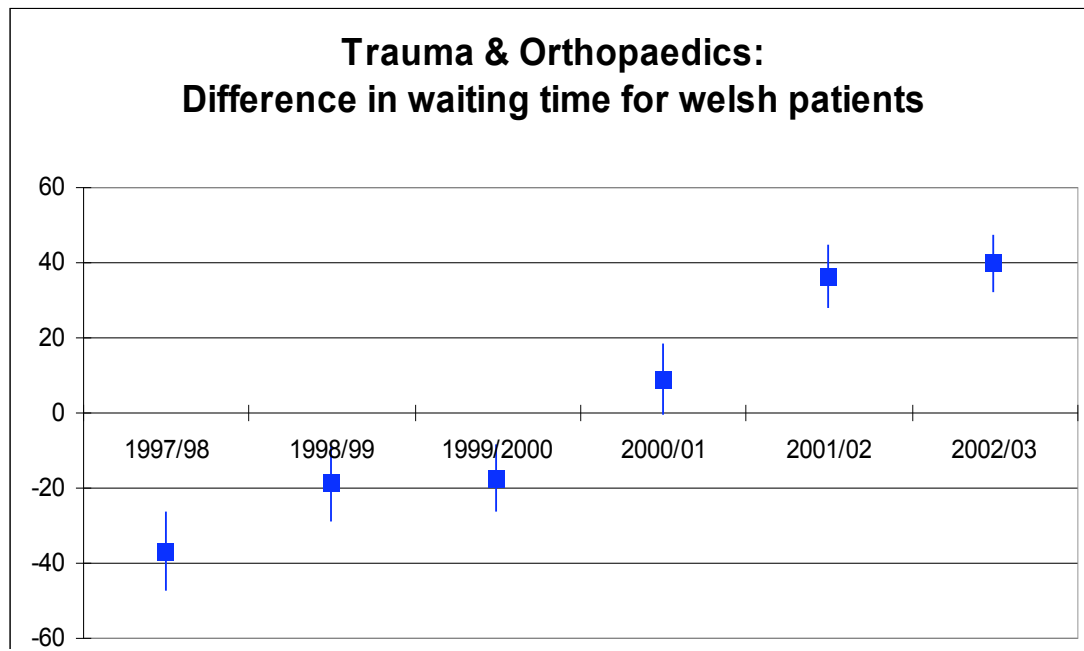


Figure 5

