Pay for Performance for general practitioners: lessons from the UK

Hugh Gravelle

Centre for Health Economics
University of York

Hague, Wednesday 14 March 2012
© Hugh Gravelle
Quality and P4P schemes

• Problem: variations in quality across providers.
• Policy: link payment to quality measures
• Hospitals
  – England: CQUIN (Commissioning for Quality and Innovation)
    • 1.5% of payment linked to range of quality indicators
• Primary care
  – Focus on management of long term conditions
    • Australia, Canada, Germany, Italy, Netherlands, Spain, Taiwan, USA
    • UK: Quality and Outcomes Framework
      – Large set of quality indicators based on clinical evidence; powerful incentives
NHS

• NHS: tax financed
  – No charges for hospital care
  – No charges for general practice consultations
  – Charges for general practice prescribed drugs
    • Wide range of exemptions. 90% dispensed with no charge.

• General Practice list system
  – patients register with a general practice
  – do not pay for consultations
  – gatekeeping for elective (non-emergency) hospital care

• England: 8,200 practices
  – Private, limited liability partnerships
  – 36,000 GPs, 72,000 other staff
  – 4.4 GPs per practice
  – Average list per GP 1520
  – List per practice 6640
General practice contracts

- NHS contracts with general practices
- General Medical Services contract (55% of practices)
  - Nationally negotiated
  - Capitation, lump sum, quality incentive payments
- Personal Medical Service contract (45%)
  - Locally negotiated with PCT
  - Same services as GMS plus agreed services for particular groups. Payment: Total GMS plus markup
  - 10% higher income
- Dispensing practice (14%) (GMS and PMS)
  - Dispense as well prescribe medicines
  - 10% higher income
Quality and Outcomes Framework

- Introduced April 2004
- Voluntary: nearly all practices take part (including PMS)
- 146 quality indicators
- 1050 quality indicator points
## QOF indicators and points
### 2004/5-2005/6

<table>
<thead>
<tr>
<th>Category</th>
<th>Indicators</th>
<th>Max points</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Clinical quality</strong></td>
<td>76</td>
<td>550</td>
</tr>
<tr>
<td>Organisation (records, information for patients, education &amp; training, practice management)</td>
<td>56</td>
<td>184</td>
</tr>
<tr>
<td>Patient experience (consultation length, carry out patient surveys)</td>
<td>4</td>
<td>100</td>
</tr>
<tr>
<td>Additional services (screening etc)</td>
<td>10</td>
<td>36</td>
</tr>
<tr>
<td>Holistic care (3rd worst clinical perf)</td>
<td></td>
<td>100</td>
</tr>
<tr>
<td>Quality practice (3rd worst on other indices)</td>
<td></td>
<td>30</td>
</tr>
<tr>
<td>Access bonus (waits for appointment)</td>
<td></td>
<td>50</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>146</td>
<td>1050</td>
</tr>
</tbody>
</table>
## Clinical domains and points
2004/5-2005/6

<table>
<thead>
<tr>
<th>Clinical Domain</th>
<th>No. of Indicators</th>
<th>Max Points</th>
</tr>
</thead>
<tbody>
<tr>
<td>Asthma</td>
<td>7</td>
<td>72</td>
</tr>
<tr>
<td>Cancer</td>
<td>2</td>
<td>12</td>
</tr>
<tr>
<td>CHD</td>
<td>12</td>
<td>101</td>
</tr>
<tr>
<td>COPD</td>
<td>8</td>
<td>45</td>
</tr>
<tr>
<td>Diabetes</td>
<td>18</td>
<td>99</td>
</tr>
<tr>
<td>Epilepsy</td>
<td>4</td>
<td>72</td>
</tr>
<tr>
<td>Hypertension</td>
<td>5</td>
<td>105</td>
</tr>
<tr>
<td>Hypothyroidism</td>
<td>2</td>
<td>8</td>
</tr>
<tr>
<td>LVD</td>
<td>3</td>
<td>20</td>
</tr>
<tr>
<td>Mental health</td>
<td>5</td>
<td>41</td>
</tr>
<tr>
<td>Stroke and TIA</td>
<td>10</td>
<td>31</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>76</strong></td>
<td><strong>550</strong></td>
</tr>
</tbody>
</table>
Some CHD indicators  
2004/5-2005/6

<table>
<thead>
<tr>
<th>Indicator</th>
<th>Max points</th>
<th>Upper threshold</th>
</tr>
</thead>
<tbody>
<tr>
<td>CHD 1: Practice has register of patients with CHD</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>CHD 5: Percentage of CHD patients whose notes record BP in previous 15 months</td>
<td>7</td>
<td>90%</td>
</tr>
<tr>
<td>CHD 6: Percentage of CHD patients whose BP in previous 15 months is 150/90 or less</td>
<td>19</td>
<td>70%</td>
</tr>
</tbody>
</table>

Lower threshold all 65 ratio indicators: 25%
Quality and Outcomes Framework

- Introduced April 2004
- Voluntary: nearly all practices take part
- 146 quality indicators
- 1050 quality indicator points
- £76 per point for average practice in 2004/5
- payment per point varies with
  - relative practice disease prevalence
  - relative list size
- £80,000 pa potential additional gross income per average practice 2004/5

- 2005/6: payment per point increased to £125 (potential £130,000 pa per average practice) 2005/6

- UK expenditure on QOF £1,000,000M pa
Percent of maximum (1050) points achieved 2004/5

Predicted average percentage: 75%
Percent of maximum points achieved 2005/6
Quality and Outcomes Framework

- Introduced April 2004
- Voluntary: nearly all practices take part (including PMS)
- 146 quality indicators
- 1050 quality indicator points
- £76 per point for average practice in 2004/5
- payment per point varies with
  - relative practice disease prevalence
  - relative list size
- £80,000 pa potential additional gross income per average practice 2004/5

- 2005/6: payment per point increased to £125 (potential £130,000 pa per average practice) 2005/6

- UK expenditure on QOF £1,000,000M pa

- Subsequent changes from 2006/7 onwards
  - Indicators (definitions, upper and lower thresholds, points)
  - Disease areas
  - Prevalence adjustment
## QOF 2004/5 – 2012/13

<table>
<thead>
<tr>
<th>Year</th>
<th>Indicators</th>
<th>Clinical Domains</th>
<th>Points</th>
<th>Mean</th>
<th>% Practices at 100%</th>
<th>Other Changes</th>
</tr>
</thead>
<tbody>
<tr>
<td>2004/5</td>
<td>146</td>
<td>11</td>
<td>1050</td>
<td>91.3</td>
<td>2.6</td>
<td></td>
</tr>
<tr>
<td>2005/6</td>
<td>146</td>
<td>11</td>
<td>1050</td>
<td>96.2</td>
<td>9.7</td>
<td>Price per point increased</td>
</tr>
<tr>
<td>2006/7</td>
<td>135</td>
<td>19</td>
<td>1000</td>
<td>95.4</td>
<td>5.1</td>
<td>Thresholds increased</td>
</tr>
<tr>
<td>2007/8</td>
<td>135</td>
<td>19</td>
<td>1000</td>
<td>96.8</td>
<td>7.5</td>
<td></td>
</tr>
<tr>
<td>2008/9</td>
<td>129</td>
<td>19</td>
<td>1000</td>
<td>95.4</td>
<td>2.0</td>
<td>Access points based on pat surveys</td>
</tr>
<tr>
<td>2009/10</td>
<td>134</td>
<td>20</td>
<td>1000</td>
<td>93.7</td>
<td>1.0</td>
<td>Prevalence calc</td>
</tr>
<tr>
<td>2010/11</td>
<td>134</td>
<td>20</td>
<td>1000</td>
<td>94.7</td>
<td>1.3</td>
<td></td>
</tr>
<tr>
<td>2011/12</td>
<td>138</td>
<td>20</td>
<td>1000</td>
<td></td>
<td>96 points for new “Quality and Productivity indicators”</td>
<td></td>
</tr>
<tr>
<td>2012/13</td>
<td>142</td>
<td>22</td>
<td>1000</td>
<td></td>
<td>Thresholds increased</td>
<td></td>
</tr>
</tbody>
</table>
QOF questions

• Effects of QOF on
  – GPs
  – performance on incentivised activities
  – performance on unincentivised activities:
    – patient health
    – NHS costs
    – Cost effectiveness

• Lessons for pay for performance contract design
  – Appropriate incentives
  – Gaming
**Trends in GP hours, income, job satisfaction**

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Gross income (£000s)</td>
<td>183</td>
<td>202</td>
<td>230</td>
<td>245</td>
<td>247</td>
<td>252</td>
<td>274</td>
<td>278</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Expenses (£000s)</td>
<td>111</td>
<td>120</td>
<td>130</td>
<td>135</td>
<td>140</td>
<td>146</td>
<td>265</td>
<td>169</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Net Income (£000s)</td>
<td>72</td>
<td>72</td>
<td>82</td>
<td>100</td>
<td>110</td>
<td>108</td>
<td>106</td>
<td>110</td>
<td>109</td>
<td></td>
</tr>
<tr>
<td>hrs/week</td>
<td>45.7</td>
<td>44.4</td>
<td>38.9</td>
<td></td>
<td>40.1</td>
<td>41.4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Job satis(1-7)</td>
<td>4.0</td>
<td>4.7</td>
<td>5.2</td>
<td>4.7</td>
<td>4.7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Income data: Doctor and Dentists Review Body reports (sample tax records)
Hours, job satisfaction: National Primary Care Research and Development Centre GP Worklife Surveys
Did the QOF improve performance?

• No national data on QOF activities before QOF introduced
• No controls: simultaneous introduction in England, Scotland, Wales, Northern Ireland
• Use data on practices in GPRD, SPICE, QRESEARCH, THIN databases
  – Before/after studies (interrupted time series)
    • But not a random sample
  – Comparison of before/after for incentivised and unincentivised indicators (DID)
National trends in QOF indicators

Trends monitoring and control of blood pressure of hypertensive patients
Jan 2001 to July 2007

N = 470,725 general practice patients with hypertension
Trends in blood pressure of hypertensive patients Jan 2001 to July 2007

N = 470,725 general practice patients with hypertension
Trends in treatment intensity for hypertensive patients Jan 2001 to July 2007

N = 470,725 general practice patients with hypertension
DID design: trends for **incentivised** (BP, cholesterol, smoking status) and **unincentivised** (BMI, alcohol) monitoring of CHD patients

315 Scottish practices in SPICE. Sutton et al, Health Econ, 2010
DID: Incentivised and unincentivised quality indicators (Doran et al, BMJ 2010)

Figure 1. Mean achievement rate* by indicator group, 2000/1 to 2006/7

* Equal to the mean of the adjusted means for the individual indicators within the group
Effects on patient health?

- Direct evidence:
  - Change in patient health outcomes
- Indirect evidence
  - QOF improved performance on QOF indicators
  - Better QOF performance associated with better patient outcomes
ITS: trends in adverse outcomes for hypertensive patients Jan 2001 to July

N = 470,725 general practice patients with hypertension

DID: changes in unplanned admission rates for incentivised ACSCs relative to 2000/1 minus changes in non-incentivised ACSCs relative to 2000/1

Hatched blue line indicates the predicted difference between incentivised ACSCS and non-incentivised ACSCs assuming that the difference in pre QOF trends was continued. Vertical red lines indicate the announcement of the P4P scheme in 2003/4 and its introduction in 2004/5. Error bars indicate 95% confidence interval.

QOF performance and patient outcomes

• Better performance on process measures (monitoring etc) improves intermediate outcomes (blood pressure, HbA1c, etc)
  – DID
    • Ryan and Doran Medical Care, 2012
• Better control of blood sugar reduces emergency admissions for complications of diabetes
  – Cross sectional
  – Panel: within practices over time (2004/5 – 2008/9)
    • Dushieko et al HSR, 2010
• Patients in practices with better QOF performance have reduced mortality risk
  – Cross sectional
  – Panel: within practices over time (2004/5 - 2007/8)
    • Martin et al, Health Foundation, 2010
Effects of QOF on quality?

• Effect on quality performance indicators
  – Already trending upwards pre QOF
  – QOF had small additional effect
  – No clear evidence of effort diversion from unincentivised activities to incentivised

• Effects on health and other outcomes
  – Direct evidence
    • ITS: No change in CHD related outcomes?
    • DID: Reductions in ACSCs for incentivised conditions
  – Indirect evidence:
    • QOF improved performance on process indicators
    • Better performance on process indicators leads to
      – Better performance on intermediate outcomes
      – Fewer ACSC admissions
      – Reduced hospital costs
      – Lower mortality
Equity implications?

Improvements in QOF quality by level of area deprivation

But no evidence on deprivation and QOF quality before QOF

Source: Doran et al, Lancet, 2008
QOF performance and hospital costs

- Patients in practices with better QOF performance have lower hospital costs
  - Cross sectional and panel evidence
  - Reduced emergency admissions
  - Reduced outpatient attendances
  - No effect on elective admissions
  - No effect on costs once admitted
  - 10% improvement in QOF stroke performance between 2004/5 and 2007/8 reduced NHS hospital cost by £130M
    - Dusheiko et al, JHE 2012
Is the QOF cost effective?

- Does the incentivised activity improve outcomes or reduce costs?
- Did the QOF increase the incentivised activity?
- Was the incentive payment (for all patients treated) less than the value of the health gain (or cost reduction) to the additional patients treated

- Walker et al, BJGP 2010
  - Examined 9 QOF indicators
  - Required increase in activity for QOF to be cost effective generally very small
<table>
<thead>
<tr>
<th>Indicator</th>
<th>Description</th>
<th>Maximum points available for indicator</th>
<th>Mean annual payment received per treated patient, £</th>
<th>Mean utilisation level for indicator in 2004/2005, %</th>
<th>Evidence rating &amp; source</th>
<th>Incremental cost per QALY gained, £</th>
</tr>
</thead>
<tbody>
<tr>
<td>BP5</td>
<td>The percentage of patients with hypertension in whom the last blood pressure (measured in last 9 months) is 150/90 mmHg or less</td>
<td>56</td>
<td>8.35</td>
<td>71.3</td>
<td>Indicative&lt;sup&gt;36&lt;/sup&gt;</td>
<td>989</td>
</tr>
<tr>
<td>CHD9</td>
<td>The percentage of patients with CHD with a record in the last 15 months that aspirin, an alternative antiplatelet therapy, or an anticoagulant is being taken (unless a contraindication or side-effects are recorded)</td>
<td>7</td>
<td>2.64</td>
<td>90.0</td>
<td>Robust&lt;sup&gt;38&lt;/sup&gt; Dominant (less costly and more effective than comparator)&lt;sup&gt;39&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>CHD10</td>
<td>The percentage of patients with CHD who are currently treated with a beta-blocker (unless a contraindication or side-effects are recorded)</td>
<td>7</td>
<td>4.77</td>
<td>63.2</td>
<td>Indicative&lt;sup&gt;37&lt;/sup&gt;</td>
<td>58</td>
</tr>
<tr>
<td>CHD11</td>
<td>The percentage of patients with a history of myocardial infarction (diagnosed after 1 April 2003) who are currently treated with an ACE inhibitor</td>
<td>7</td>
<td>40.61</td>
<td>85.6</td>
<td>Robust&lt;sup&gt;35&lt;/sup&gt;</td>
<td>5623</td>
</tr>
<tr>
<td>CS1</td>
<td>The percentage of patients aged 25–64 years (in Scotland 25–60 years) whose notes record that a cervical smear has been performed in the last 3 to 5 years</td>
<td>11</td>
<td>0.63</td>
<td>80.2</td>
<td>Indicative&lt;sup&gt;37&lt;/sup&gt;</td>
<td>458</td>
</tr>
<tr>
<td>DM15</td>
<td>The percentage of patients with diabetes with proteinuria or micro-albuminuria who are treated with ACE inhibitors (or A2 antagonists)</td>
<td>3</td>
<td>17.86</td>
<td>82.1</td>
<td>Indicative&lt;sup&gt;36&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>DM21</td>
<td>The percentage of patients with diabetes who have a record of retinal screening in the previous 15 months</td>
<td>5</td>
<td>1.97</td>
<td>83.2</td>
<td>Robust&lt;sup&gt;35&lt;/sup&gt;</td>
<td>15 654</td>
</tr>
<tr>
<td>LVD3</td>
<td>The percentage of patients with a current diagnosis of heart failure due to LVD who are currently treated with an ACE inhibitor or an A2 antagonist, who can tolerate therapy and for whom there is no contraindication</td>
<td>10</td>
<td>36.25</td>
<td>82.1</td>
<td>Indicative&lt;sup&gt;37&lt;/sup&gt;</td>
<td>109</td>
</tr>
<tr>
<td>Stroke12</td>
<td>The percentage of patients with a stroke shown to be non-haemorrhagic, or a history of TIA, who have a record that an antiplatelet agent (aspirin, clopidogrel, dipyridamole or a combination), or an anticoagulant is being taken (unless a contraindication or side-effects are recorded)</td>
<td>4</td>
<td>6.00</td>
<td>89.3</td>
<td>Robust&lt;sup&gt;38&lt;/sup&gt;</td>
<td>2012</td>
</tr>
</tbody>
</table>
Proportionate change in achievement required for cost-effectiveness

BP5: percentage of patients in whom last recorded blood pressure is 150/90 or less
Proportionate change in achievement required for cost-effectiveness

CHD11: percentage of patients with CHD with a history of myocardial infarction who are currently treated with an ACE inhibitor
Proportionate change in achievement required for cost-effectiveness

DM15: percentage of patients with diabetes with proteinurial or micro-albuminuria who are treated with ACE inhibitors (or A2 antagonists)
QOF questions

• Effect of QOF on
  – GPs
    • Higher income, reduced hours
  – performance on incentivised activities:
    • small above trend increase
  – performance on unincentivised activities:
    • no evidence for effort diversion
  – patient health
    • small improvement?
  – NHS costs
    • Increased primary care costs
    • Reduced hospital cost
  – Cost effectiveness
    • Partial evidence of cost-effectiveness of QOF for some indicators

• Lessons for pay for performance contract design
Problematic features of QOF design

• Rewards and patient benefits
• Multiple payments for one activity
• Variation in incentives across practices
• Exception reporting
• Case finding and prevalence adjustment
Multiple rewards for same activity

- Multiple payment for activity undertaken in respect of same patient if
  - if undertaken in January-March (activities count if undertaken in 15 months prior to 31 March)
    - Activity needs to be achieved once every two years
  - If the outcome is controlled when initially measured (reward not contingent on treating the previously untreated)
  - If the patient has more than one QOF health condition (over 20% are multimorbid)
Alignment of rewards and patient benefit

Source: Fleetcroft and Cookson, JHSRP 2006

But: Gravelle & Siciliani (2012) - payment should be linked to costs and benefits unless these are uncorrelated across patients
QOF incentives: (i)

Practice revenue from clinical indicators depends on

– points achieved
  • varies with proportion of eligible patients for whom indicator achieved

– price per point
  • varies with practice characteristics:
    – reported disease prevalence
    – practice list size
QOF incentives: clinical indicators

Points vary with **achievement** ($\alpha$): proportion of relevant patients getting intervention

Number of points

$$\pi(\alpha)$$

Max points for indicator $\pi^0$

- $\alpha_L$: 25% lower threshold
- $\alpha_U$: 90% upper threshold
- 100% achievement
QOF incentives (iii): points

Reported achievement –
determines points

\[ \alpha = \frac{N}{D} = \frac{N}{P - E} \]

\(N\): number achieved (eg patients with CHD treated with beta blocker)

\(D\): number patients declared suitable for intervention

\(D = P - E\)

\(P\): reported prevalence (disease register eg number with CHD)

\(E\): exceptions
Grounds for exceptions

• indicator inappropriate eg terminal illness, extreme frailty
• patient has recently received a diagnosis or has recently registered with the practice.
• patient already on maximal tolerated dose of a medication
• adverse reaction or contraindication.
• required investigative service is unavailable

• patient has received at least three invitations for a review during the preceding 12 months but has not attended.
• patient does not agree to investigation or treatment.
QOF incentives (iii): points

Reported achievement – determines points

\[ \alpha = \frac{N}{D} = \frac{N}{P - E} \]

- \( N \): number achieved (e.g., patients with CHD treated with beta blocker)
- \( D \): number patients declared suitable for intervention
- \( P \): reported prevalence (disease register e.g., number with CHD)
- \( E \): exceptions

Incentives for

- treatment \( N \)
- gaming of reported prevalence \( P \)
- gaming of exception reporting \( E \)
Ratio indicators with exceptions and thresholds

• Exception reporting intended to reduce the risk that providers would inappropriately treat some patients
• Setting upper thresholds below 100% was justified on the same grounds

• The rationale is peculiar for those indicators where there is no risk to patients, for example the recording of blood pressure

• Marginal reward for achievement is zero above upper threshold
  – achievement above upper threshold in 91% cases in 2005/6
  – could have reduced number patients treated by 12.4% without reducing income
  – Altruism? Risk aversion? Rounding and small numbers?

• Ratio indicators imply marginal reward for exception reporting if below upper threshold - gaming incentive
QOF incentives (iv): indicator revenue

Revenue from indicator

\[ R = \pi(\alpha) v \left[ \frac{M}{\bar{M}} \right] F \]

Adjusted disease prevalence factor

\[ F = \sqrt{\frac{P}{M}} \left/ \left[ \frac{\text{Avg} \sqrt{\text{prev}}}{\text{Avg} \sqrt{\text{prev}}} \right] \right. \]

Indicator points

\[ \pi(\alpha) \]

Achievement

\[ \alpha = \frac{N}{D} = \frac{N}{(P - E)} \]

\( v \) national average price per point

\( M \) list size

\( \bar{M} \) Avg list size
Effects of practice list size

Revenue from indicator

\[ R = \pi(\alpha)[PM]^{\frac{1}{2}} v / \left[ \bar{M} \text{ Avg } \sqrt{\text{Prev}} \right] \]

Larger practices get higher revenue:

\[
\frac{\partial R}{\partial M}_{N,P,E} = \frac{1}{2} \frac{R}{M} > 0
\]

Larger practices have higher marginal revenue from treating patients

\[
\frac{\partial}{\partial M} \left( \frac{\partial R}{\partial N} \right)_{N,P,E} = \frac{1}{2} \frac{1}{M} \frac{\partial R}{\partial N} > 0
\]
Incentives for case finding/reporting?

\[ R = \pi(\alpha)[PM]^{1/2} \sqrt{\frac{\bar{M} \text{ Avg} \sqrt{\text{Prev}}}{\alpha}} \]

\[ \alpha = \frac{N}{P - E} \]

- Higher reported prevalence
  - Increases price per point and marginal rewards for achievement
  - Reduces reported achievement and hence points if below upper threshold
  - Increases revenue if above upper threshold
  - Reduces revenue if below upper threshold
Testing for gaming

• Compared to practices with $\alpha < \alpha_u$, practices with $\alpha > \alpha_u$ will have
  – lower exception reporting rates
  – higher reported prevalence
Test for gaming of 2005/6 exceptions

<table>
<thead>
<tr>
<th>Model</th>
<th>Effects allowed for</th>
<th>Coeff on $U^{04/05}$</th>
<th>t (or z)</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>Linear</td>
<td>Clustering</td>
<td>-0.129</td>
<td>-7.35</td>
<td>56980</td>
</tr>
<tr>
<td></td>
<td>FE Practice</td>
<td>-0.108</td>
<td>-8.78</td>
<td>56980</td>
</tr>
<tr>
<td></td>
<td>FE Practice-Disease</td>
<td>-0.139</td>
<td>-12.02</td>
<td>56899</td>
</tr>
<tr>
<td>NegBin</td>
<td>Clustering</td>
<td>-0.165</td>
<td>-11.75</td>
<td>56980</td>
</tr>
<tr>
<td></td>
<td>FE Practice</td>
<td>-0.080</td>
<td>-11.56</td>
<td>56980</td>
</tr>
<tr>
<td></td>
<td>FE Practice-disease</td>
<td>-0.115</td>
<td>-17.11</td>
<td>53500</td>
</tr>
</tbody>
</table>

$U^{04/05} = 1$ if practice above upper threshold for indicator in 2004/5

Gravelle, Sutton & Ma, EJ 2010
Tests of exception gaming: distance from upper threshold

<table>
<thead>
<tr>
<th></th>
<th>Coef</th>
<th>t (or z)</th>
<th>Coef</th>
<th>t (or z)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Effect of being above upper threshold $\alpha_{2004/5} \geq \alpha_U$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>versus being near below (0.8(\alpha_U \leq \alpha_{2004/5} &lt; \alpha_U))</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>versus being further below ((\alpha_{2004/5} &lt; 0.8\alpha_U))</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Method</th>
<th>Coef</th>
<th>t (or z)</th>
<th>Coef</th>
<th>t (or z)</th>
</tr>
</thead>
<tbody>
<tr>
<td>WLS</td>
<td>-0.075</td>
<td>-4.58</td>
<td>-0.230</td>
<td>-7.18</td>
</tr>
<tr>
<td></td>
<td>-0.064</td>
<td>-4.98</td>
<td>-0.180</td>
<td>-9.92</td>
</tr>
<tr>
<td>Practice-disease FE</td>
<td>-0.099</td>
<td>-7.64</td>
<td>-0.227</td>
<td>-12.61</td>
</tr>
<tr>
<td>NegBin</td>
<td>-0.103</td>
<td>-7.72</td>
<td>-0.254</td>
<td>-12.38</td>
</tr>
<tr>
<td></td>
<td>-0.049</td>
<td>-6.25</td>
<td>-0.138</td>
<td>-14.85</td>
</tr>
<tr>
<td>Practice-disease FE</td>
<td>-0.082</td>
<td>-11.10</td>
<td>-0.207</td>
<td>-21.23</td>
</tr>
</tbody>
</table>

$U_{04/05} = 1$ if practice above upper threshold for indicator in 2004/5

Gravelle, Sutton & Ma, EJ 2010
Magnitude of gaming effect

• 25% of 56,980 practice-indicator cases were below upper threshold in 2004/5
• estimated effect (NegBin pooled model) of exception gaming was to increase reported exception rate from 7.25% to 8.55%
Test for gaming of reported prevalence

• Gaming incentives: revenue increases if
  – reduce prevalence if below upper threshold
  – increase prevalence if above upper threshold
QOF incentives and prevalence reporting:

<table>
<thead>
<tr>
<th>Effects allowed for</th>
<th>Coef on $U^{04/05}$</th>
<th>t</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clustering</td>
<td>0.026</td>
<td>2.41</td>
<td>10044</td>
</tr>
<tr>
<td>FE practice</td>
<td>0.000</td>
<td>-0.02</td>
<td>10044</td>
</tr>
</tbody>
</table>

$U^{04/05}$ maximum points weighted proportion of indicators in disease domain for which practice is above the upper threshold for 11 diseases. Dependent variable practice 2005/06 reported prevalence relative to national mean for disease.
Prevalence adjustment and exception reporting

- Square root adjustment replaced with simple relative prevalence from 2009/10

\[ R = \pi(\alpha) v \left[ \frac{P}{M} \div \left( \frac{\bar{P}}{\bar{M}} \right) \right] \frac{M}{\bar{M}} = \pi(\alpha) v \left( \frac{P}{\bar{P}} \right) \]

- Next step: remove exception reporting: \( \alpha = N/P \)?
  - redundant because of upper thresholds
  - too easy to game

- So if practice below upper threshold:

\[ R = v \frac{N}{P} \left[ \frac{P}{M} \div A v \left( \frac{\bar{P}}{\bar{M}} \right) \right] \frac{M}{\bar{M}} = v N \left[ \frac{1}{\bar{P}} \right] \]

make payment proportional to patients treated - FFS
Do practices respond to small changes in QOF incentives?

Changes in flu vaccination indicators

<table>
<thead>
<tr>
<th>Condition</th>
<th>Attribute</th>
<th>2005/6</th>
<th>2006/7 onwards</th>
</tr>
</thead>
<tbody>
<tr>
<td>CHD</td>
<td>Lower threshold</td>
<td>25%</td>
<td>40%</td>
</tr>
<tr>
<td></td>
<td>Upper threshold</td>
<td>85%</td>
<td>90%</td>
</tr>
<tr>
<td>COPD</td>
<td>Lower threshold</td>
<td>25%</td>
<td>40%</td>
</tr>
<tr>
<td></td>
<td>Upper threshold</td>
<td>85%</td>
<td>85%</td>
</tr>
<tr>
<td>Diabetes</td>
<td>Lower threshold</td>
<td>25%</td>
<td>40%</td>
</tr>
<tr>
<td></td>
<td>Upper threshold</td>
<td>85%</td>
<td>85%</td>
</tr>
<tr>
<td>Stroke</td>
<td>Lower threshold</td>
<td>25%</td>
<td>40%</td>
</tr>
<tr>
<td></td>
<td>Upper threshold</td>
<td>85%</td>
<td>85%</td>
</tr>
<tr>
<td>Dependent variable</td>
<td>%Reported Achievement</td>
<td>%Population Achievement</td>
<td>%Exception Reporting</td>
</tr>
<tr>
<td>--------------------</td>
<td>-----------------------</td>
<td>-------------------------</td>
<td>----------------------</td>
</tr>
<tr>
<td></td>
<td>Coef % (95% CI)</td>
<td>Coefficient % (95% CI)</td>
<td>Coefficient % (95% CI)</td>
</tr>
<tr>
<td><strong>Model 1 DID</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Effect of change in CHD upper threshold on CHD outcome</td>
<td>0.69 (0.55, 0.82)</td>
<td>0.41 (0.25, 0.56)</td>
<td>0.26 (0.12, 0.40)</td>
</tr>
<tr>
<td><strong>Model 2 DID</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Effect of CHD upper threshold change on CHD outcome for practices with 2005/06 CHD RA ≥ 90%</td>
<td>0.49 (0.35, 0.63)</td>
<td>0.37 (0.21, 0.53)</td>
<td>0.13 (-0.02, 0.28)</td>
</tr>
<tr>
<td>2005/06 CHD RA in [85, 90]% range</td>
<td>0.98 (0.80, 1.16)</td>
<td>0.60 (0.40, 0.80)</td>
<td>0.29 (0.10, 0.47)</td>
</tr>
<tr>
<td>2005/06 CHD RA below 85%</td>
<td>1.47 (1.27, 1.68)</td>
<td>0.85 (0.62, 1.08)</td>
<td>0.50 (0.29, 0.72)</td>
</tr>
</tbody>
</table>

RA = 100*N/D      PA = 100*N/(D+E)      ER = 100*E/(D+E)

Kantopelis et al, *Health Services Research* (2011)
Lessons for design of PfP

• Measure activity to be incentivised before introducing incentives
  – Don’t pay for activity already achieved before P4P
• Trial major P4P schemes
• Cap total payments
• Use indicators with evidence base
• Use wide set of indicators
• Relative rewards should reflect health gain
• Include patient satisfaction measures
• No need for exceptions and upper thresholds to prevent over treatment
• Keep it simple: make payment proportional to numbers of treated
QOF effects

• Effect of QOF on
  – GPs
    • Higher income, reduced hours
  – performance on incentivised activities:
    • small above trend increase?
  – performance on unincentivised activities:
    • no evidence for effort diversion
  – patient health
    • small improvement?

  – NHS costs
    • Increased primary care costs
    • Reduced hospital cost

  – Cost effectiveness
    • Partial evidence of cost-effectiveness of QOF for some indicators
Latest version of the QOF
http://www.nhsemployers.org/PayAndContracts/GeneralMedicalServicesContract/QOF/Pages/QualityOutcomesFramework.aspx

Audit Body reports

Some Research studies
Doran, T. & Roland, M. Lessons from major initiatives to improve primary care in the United Kingdom. *Health Affairs.* 2010, 1023-1026