UK pay-for-performance scheme
Methods
Results
Conclusions

QOF impact on the quality of incentivised and non-incentivised aspects of primary care

Evangelos Kontopantelis¹  Tim Doran*¹
Stephen Campbell¹  Jose Valderas¹  Martin Roland²
Chris Salisbury³  David Reeves¹

¹National Primary Care Research and Development Centre
University of Manchester
²General Practice and Primary Care Research Unit, University of Cambridge
³Academic Unit of Primary Health Care, University of Bristol

HSRN, 3rd June 2010

Manchester! pronounced Manchestoh

- Where sunglasses last for a lifetime.
- 20 different accents in a 40 mile radius (and these are just the local ones).
- If you don’t speak football you die alone.
- Sarcasm is a way of life (Is it really? Nooooooooo!)
A P4P program kicked off in April 2004 with the introduction of a new GP contract.

- General practices are rewarded for achieving a set of quality targets for patients with chronic conditions.
- The aim was to increase overall quality of care and to reduce variation in quality between practices.

The incentive scheme for payment of GPs was named Quality and Outcomes Framework (QOF).

- Initial investment estimated at £1.8 bn for 3 years (increasing GP income by up to 25%).
- QOF is reviewed at least every two years.
Domains and indicators in year 1 (year 5):
- Clinical care for 10 (19) chronic diseases, with 76 (80) indicators.
- Organisation of care, with 56 (36) indicators.
- Additional services, with 10 (8) indicators.
- Patient experience, with 4 (5) indicators.

Implemented simultaneously in all practices (a control group was out of the question).

Into the 7th year now (01Mar10/31Apr11); cost for the first 6 years was well above the estimate at £5.8 bn approx.

Performance incentivised aspects keep improving but what about the non-incentivised ones?

- Quality scores for all QOF clinical indicators have been improving.
- Only a small proportion of all clinical care.
- Concerns that quality for non-incentivised aspects of care may have been neglected.
- How to measure performance on non-incentivised care?
Clinical indicators

- Two aspects to clinical indicators:
  - a disease condition (e.g. diabetes, CHD).
  - a care activity (e.g. influenza vaccination, BP control).
- Three indicator classes, in terms of incentivisation:
  - (A) Condition & process incentivised within QOF (28 ind)
  - (B) Condition or process incentivised (13 ind)
  - (C) Neither condition nor process incentivised (7 ind)

Example (Indicators)

- (A) DM11: Patients with diabetes in whom the last blood pressure (within 15m) is 145/85 or less
- (B) B4: Patients with peripheral arterial disease who have a record of total cholesterol in the last 15m
- (C) C4: Patients with back pain treated with strong analgesics (co-dydramol upwards) in the last 15m

Research questions
the obvious ones at least!

- We aimed to compare the three classes on changes in quality from pre-QOF (2000/01 - 2002/03) to post-QOF (2004/05 - 2006/07).
- Would class A indicators show most improvement?
- Would class B show some ‘halo’ effects since they involve either a QOF condition or activity?
- Has quality for class C indicators declined?
The General Practice Research Database (GPRD)

- Holds event data for more than 270 English practices, from 1999 (545 active practices in Apr10 and 11.2m patients).
- Database is broken down to numerous tables, because of the volume of data.
- Final sample was 600,000 patients from 150 nationally representative practices.
- Allows us to:
  - measure quality of care prior to the introduction of QOF.
  - construct non-incentivised indicators of quality.
- At the heart of the database lies the use of specific codes by GPs to record all information from a patient’s visit.

The process begins with a stab in the dark...

- Size of the tables prohibits looking at codes one by one.
- Instead we use search terms to identify potentially relevant codes in the lookup files and create draft lists.
- Clinicians go through the draft lists and select the relevant codes.
- Using the finalised code lists we search for events the main database files.
- For each indicator and year, we used the associated clinical rules and retrieved events to construct a % score:
  - (cases where indicator was met / eligible patients) * 100%.
Indicators classes are imbalanced.

Three different types of activities (x3 classes = 9 groups):

- Clinical processes related to measurement (PM/R).
  - A: 17  B: 9  C: 0
  - *e.g.* blood pressure measurement
- Clinical processes related to treatment (PT).
  - A: 6  B: 4  C: 7
  - *e.g.* influenza immunisation
- Intermediate outcome measures (I).
  - A: 5  B: 0  C: 0
  - *e.g.* control of HbA1c to 7.4 or below

Quality of care was already improving (prior to QOF).

The ceiling has been reached for certain ‘easy’ indicators.

The main analysis used logit-transformed scores, due to the ceiling effect.

Untransformed scores were used in a sensitivity analysis.

The six available indicator groups (of a possible nine) were compared, on performance above expectation.

FE model selected; controlling for RTTM, denominator, patient age and gender.

All analyses performed in STATA v10.1

Interrupted Time Series methods employed...
The design

Methods

Results

Conclusions

The approach

Interrupted Time Series

- ITS multi-level multivariate regression analyses, allowed us to compare the six indicator groups on the basis of two outcome measures:
  - The difference between observed and expected achievement, in 2004/05.
  - The difference between observed and expected achievement, in 2006/07.

Kontopantelis

The QOF impact

Trends

Indicator performance
- Short term QOF effect - by indicator group
- Long term QOF effect - by indicator group

Trends by indicator group
untransformed scores

**Indicator group performance**

Percentage scores

using group means of indicator means (by practice) in brackets, the number of indicators in each group
**Trends by indicator group**

**logit transformed scores**

---

**Indicator group performance**

Logit transformed scores

---

**Observed performance compared to expectation**

**logit transformed scores**

---

<table>
<thead>
<tr>
<th>Indicator group</th>
<th>Difference in 2004/05</th>
<th>Difference in 2006/07</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>sig positive</td>
<td>-</td>
</tr>
<tr>
<td>A-I</td>
<td>5</td>
<td>0</td>
</tr>
<tr>
<td>A-PM/R</td>
<td>17</td>
<td>0</td>
</tr>
<tr>
<td>B-PM/R</td>
<td>3</td>
<td>5</td>
</tr>
<tr>
<td>A-PT</td>
<td>5</td>
<td>1</td>
</tr>
<tr>
<td>B-PT</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>C-PT</td>
<td>1</td>
<td>4</td>
</tr>
</tbody>
</table>
**Difference in 2004/05 of observed performance compared to expectation**

- All three indicator groups of class A significantly increased in level above expectation post-QOF.
- Treatment indicators of class B significantly decreased in level below expectation post-QOF.

**Kontopantelis**

The QOF impact

---

**Difference in 2004/05 comparing ordered means across groups, using the Newman-Keuls procedure**

<table>
<thead>
<tr>
<th>Uplift in 2004/5</th>
<th>Group</th>
<th>Fully incentivized measurement</th>
<th>Fully incentivized outcome</th>
<th>Fully incentivized treatment</th>
<th>Partially incentivized measurement</th>
<th>Partially incentivized treatment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>14.5%</td>
<td>8.2%</td>
<td>4.2%</td>
<td>0.8%</td>
<td>-0.7%</td>
<td>-1.5%</td>
</tr>
<tr>
<td>95% confidence interval</td>
<td>(14.0, 15.0)</td>
<td>(7.3, 9.2)</td>
<td>(3.2, 5.3)</td>
<td>(-0.2, 1.8)</td>
<td>(-1.8, 0.5)</td>
<td>(-3.0, -0.2)</td>
</tr>
<tr>
<td>P value</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
<td>0.128</td>
<td>0.257</td>
<td>0.03</td>
</tr>
<tr>
<td>Difference between means**</td>
<td>[------------------------]</td>
<td>[------------------------]</td>
<td>[------------------------]</td>
<td>[------------------------]</td>
<td>[------------------------]</td>
<td>[------------------------]</td>
</tr>
</tbody>
</table>

* Group means based on logit-transformed data, back-transformed to percentage scores.
** Neuman-Keuls tests. Means connected by a dashed line were not significantly different (p > 0.05).
All three indicator groups of class A significantly increased in level above expectation post-QOF.

All partially incentivised and non-incentivised indicator groups significantly decreased in level below expectation post-QOF.
Summary

- Short term:
  - Overall, all three groups of fully incentivised indicators exhibited performance above expectation. The uplift varied from 1.1% to 38.2% with 4 indicators (smoking) having uplifts of over 30%.
  - Partially incentivised Measurement/Recording indicators demonstrated significantly lower than expected gains (on average).

- Long term:
  - Overall, the three fully incentivised groups continued to have positive uplifts, although none exceeded 4%.
  - The three partially incentivised and non-incentivised groups displayed significantly negative uplifts, on average.

Discussion

- Improvement in quality pre-QOF suggests that the quality initiatives at that time had an effect.
- The rate of improvement was not sustained and improvement appeared to plateau after 04/05.
- No such plateau effect for non-QOF indicators with similar rates of achievement. Practices may have relaxed their efforts for the incentivized indicators after the first year?
- QOF did not generate positive spill-overs to other activities & appears to have had a negative impact on non-inc ones for otherwise inc conditions.
  - Focus on the activities for which incentives applied when managing particular conditions?
  - Performed particular activities preferentially for patients for whom rewards applied?
UK pay-for-performance scheme

Methods

Results

Conclusions

Something goes around something but that’s as far as I’ve got...

Comments and questions:
e.kontopantelis@manchester.ac.uk

Kontopantelis  The QOF impact