

Pay for performance and the management of hypertension

Rachel Foskett-Tharby¹, Nick Hex², Paramjit Gill¹

¹ Primary Care Clinical Sciences, Institute of Applied Health Research, University of Birmingham, Edgbaston, Birmingham, B15 2TT, United Kingdom;

² York Health Economics Consortium, Enterprise House, Innovation Way, University of York, York, YO10 5NQ, United Kingdom

INTRODUCTION

With the global epidemiological transition from communicable to non-communicable diseases, hypertension has become a major risk factor for burden of disease in many high, middle, and low income countries.^[1] While no absolute cut point exists for high blood pressure, persistent systolic blood pressure readings of > 140 and/or diastolic blood pressure readings of > 90 are commonly defined as hypertension.^[2] Hypertension is a risk factor for cardiovascular disease and its management is advocated to reduce burden to both individuals and health systems.

With the rise in long-term conditions, health care systems around the world are under pressure to curb health care costs while maintaining quality. In response, many countries have introduced pay for performance (P4P) programs that incentivize institutions and professionals to provide high quality care and to mitigate the potential weaknesses of other payment mechanisms such as fee for service.^[3] P4P programs have been widely adopted internationally in low, middle, and high income countries such as UK, US, Thailand, Germany, and Australia.

WHAT IS P4P?

Pay for performance is an over-arching term for a method of rewarding organizations and/or individuals based upon their performance against identified criteria, which, depending upon the scheme, may include measures of quality, reporting, efficiency, and/or value.^[4] Therefore, the aims, content, and structure of P4P

schemes are highly diverse. However, there is commonality in that schemes are focused upon modifying healthcare provider behaviors and that payments are linked to achievement of identified criteria, frequently quality indicators.^[5]

The theory underpinning this approach is that financial rewards are important in motivating healthcare providers, specifically financial incentives that focus on quality of care. This is important so that quality is not neglected in comparison with other measures such as volume of care provided.^[6] The size of the incentive is assumed to be key to how individuals respond. However, there is evidence to suggest that this response is more complex and also, affected by the design and implementation of the P4P scheme.^[7]

DESIGNING A P4P SCHEME

When designing a P4P scheme, there is an understandable tendency to focus upon the activities to be incentivized. While this is obviously important, it should not be to the detriment of consideration of design and implementation issues. These include definitions of quality and the scope of the scheme; identification of quality measures; measuring and rewarding performance; and data availability, reporting and verification.

Defining quality and the scope of the scheme

Quality is a multi-dimensional concept and its definition has changed over time (see Table 1). It includes concepts such as safety, effectiveness, being patient-centered, timeliness, efficiency, equity, value for money, access, and patient experience. And for

Address for Correspondence:
Paramjit Gill, FRCP, DM,
Primary Care Clinical Sciences,
Institute of Applied Health Research,
University of Birmingham, Edgbaston,
Birmingham, B15 2TT, United Kingdom.
Email: p.s.gill@bham.ac.uk

Access this article online

Website:

www.intern-med.com

DOI:

10.1515/jtim-2016-0004

Quick Response Code:



Table 1: Dimensions of quality of care

Donabedian 1990[9]	Maxwell 1992[8]	Campbell 2000[10]	Institute of Medicine 2001[11]	NHS Next Stage Review 2008[12]
	Access	Access		
Acceptability	Acceptability			
Effectiveness	Effectiveness	Effectiveness	Effectiveness	Effectiveness
Efficiency	Efficiency	Efficiency	Efficiency	
Equity	Equity	Equity	Equity	
Efficacy	Relevance		Patient-centeredness	Patient experience
Legitimacy			Safety	Safety
Optimality			Timeliness	

many health systems, the focus is shifting to patient safety and experience. Having a working definition of quality is important when developing a P4P scheme as it guides the selection and development of quality indicators.^[13]

Quality measures can then be categorized according to structure, process, and outcomes of care. Structural measures address the environment in which care is delivered and may reference, for example, the physical space, staffing, or available equipment. Process measures focus upon care activities undertaken by the healthcare provider such as recording of blood pressure or taking of blood tests. Outcome measures focus upon the ultimate impact of the care delivered in terms of the health outcomes experienced by patients. Outcome measures are challenging to incorporate into a P4P framework due to the question of attribution to an individual physician or organization.

Development and identification of quality measures

It is important to note that no international consensus exists as to the best approach to quality indicator development.^[13,14] Stelfox and Strauss^[13] identify two broad approaches to development: inductive or deductive. These are comparable to the classification offered by Campbell *et al.*^[15] of non-systematic and systematic approaches. Inductive or non-systematic approaches start with the available data or a clinical incident and then move towards defining the concept to be measured. Deductive or systematic approaches take the opposite approach in that the clinically important concepts are identified initially and used as the basis for indicator development. Deductive approaches aim to ensure a strong link between the scientific evidence and the resulting indicator, often having their roots in clinical guidelines. However, they have also been criticized for failing to consider issues of the importance of the care concept and being poorly specified from a patient's perspective.^[13] Rigid adherence to guideline recommendations also fails to acknowledge the many uncertainties and limits of scientific knowledge in relation to health care,^[16] which may be more pronounced

in different care settings, for example, family medicine. Incorporating expert opinion through the use of consensus techniques such as the RAND appropriateness method can be useful here.^[15]

Irrespective of the method used, there are a number of steps that need to be taken to move from a guideline recommendation to a quality measure. The first of these is to develop a quality indicator that specifies the clinical situation and the care that should or should not be given. It may be useful to write these in an IF-THEN format.^[14] Further detailed specification is then required to convert these statements to quality measures. We would agree with Shekelle^[14] that this requires input from a multi-disciplinary team composed of measurement experts alongside clinical experts. In our experience, this is also an iterative process as the implications of different approaches to measure wording and component specification are considered in conjunction with the available data sources. For example, when considering a family medicine indicator related to the monitoring of blood pressure in people with hypertension, it is first necessary to consider whether this will be measured from the patient's or clinician's perspective, then to define what constitutes a diagnosis of hypertension, what constitutes blood pressure monitoring, and the maximum reasonable time periods between monitoring.

Quality measures should also be subject to a period of field testing. This allows for an assessment of reliability, validity, feasibility, and acceptability of the measure to those being measured.^[17] As part of this process, the measures should be assessed to consider the extent to which their incentivization would constitute an efficient use of public funds and provide value for money. Cost-effectiveness analysis is one such approach, involving the calculation of costs per quality adjusted life year (QALY). Services with a lower cost per QALY can be considered cost-effective and those with a higher cost per QALY may not be considered cost-effective. This approach is attractive as it allows all measures to be assessed using the same metric, an incremental cost-effectiveness ratio, thus allowing the cost-effectiveness of different measures to be compared.^[18]

Measure development represents a significant undertaking and therefore consideration should always be given to whether there are existing measures that may be adapted. Both the National Quality Measures Clearinghouse (<http://www.qualitymeasures.ahrq.gov/>) in the US and the National Institute for Health and Care Excellence (<http://www.nice.org.uk/standards-and-indicators/qofindicators>) in the UK maintain a menu of quality measures that have been subject to initial assessment of reliability, validity, and acceptability. While quality measures identified in this way require an assessment to ensure that they are appropriate for adoption, previous work suggests that there are areas of commonality, even between quite differently funded health systems.^[19] As well as offering efficiencies in the development process, utilizing existing indicators supports international comparisons of care.

Measuring and rewarding performance

P4P schemes also need to detail the way in which performance will be measured and rewarded. This encompasses questions such as whether to reward absolute achievement against a measure, that is, achieving a pre-determined payment threshold or improvement above baseline measurement or a combination of both, the size of the incentive, when the reward should be given and to whom.

Data availability, reporting, and verification

Measure development and adoption will be influenced by the availability of data and how it is reported. Electronic medical records offer the potential for query specifications

to be developed centrally and anonymized data to be extracted and reported with minimal impact upon the organization. Manual reporting methods such as local audit, will require different levels of support and indicator specifications. Greater consideration will need to be given to sample size and selection criteria and ensuring inter-rater reliability. If this audit is to be completed by external personnel then this will add to the cost of the scheme.

USE OF P4P IN HYPERTENSION MANAGEMENT IN THE UK

In many respects, hypertension management lends itself to inclusion in P4P schemes. There are national and international guidelines for its identification and management, recommended care processes can be translated into measurable statements and it is possible to articulate and measure the desired outcomes of treatment.^[20] The care of these patients accounts for 17% of the available reward for general practices in the Quality and Outcomes Framework (QOF) for England.^[21] At present, five indicators are included (see Table 2), which focus upon hypertension as a discrete condition, six on the management of blood pressure in patients with other conditions such as diabetes, two upon modifiable risk factors associated with hypertension such as smoking and one upon population-based monitoring of blood pressure. Further measures have also been tested and are available for use via the indicator development process managed by NICE, although these have not been incorporated into the incentive structure. A similar range of measures are

Table 2: QOF indicators, payment thresholds, points and incentive value (per practice) 2015/16

Indicator	Payment threshold	Points value	Incentive value (per practice)
HYP001. The contractor establishes and maintains a register of patients with established hypertension	n/a	6	£961
HYP006. The percentage of patients with hypertension in whom the last blood pressure reading (measured in the preceding 12 months) is 150/90 mmHg or less	45–80%	20	£3203
CVD-PP001. In those patients with a new diagnosis of hypertension aged 30 or over and who have not attained the age of 75, recorded between the preceding 1 April to 31 March (excluding those with pre-existing CHD, diabetes, stroke and/or TIA), who have a recorded CVD risk assessment score (using an assessment tool agreed with the NHS Commissioning Board) of $\geq 20\%$ in the preceding 12 months: the percentage who are currently treated with statins	40–90%	10	£1602
SMOK002. The percentage of patients with any or any combination of the following conditions: CHD, PAD, stroke or TIA, hypertension, diabetes, COPD, CKD, asthma, schizophrenia, bipolar affective disorder or other psychoses whose notes record smoking status in the preceding 12 months	50–90%	25	£4004
SMOK005. The percentage of patients with any or any combination of the following conditions: CHD, PAD, stroke or TIA, hypertension, diabetes, COPD, CKD, asthma, schizophrenia, bipolar affective disorder or other psychoses who are recorded as current smokers who have a record of an offer of support and treatment within the preceding 12 months	56–96%	25	£4004

CHD: coronary heart disease, TIA: transient ischemic attack, CVD: cardiovascular disease, PAD: peripheral arterial disease, COPD: chronic obstructive pulmonary disease, CKD: chronic kidney disease.

available via the National Quality Measures Clearinghouse in the US.

IMPACT OF P4P UPON HYPERTENSION MANAGEMENT

So how has P4P impacted upon the management of hypertension? Available evidence is mixed and moderated by the design of the P4P scheme, which makes drawing definitive conclusions challenging. Between 2004 and 2014, average national achievement of the proportion of patients with hypertension in England whose latest recorded blood pressure (measured in the preceding 9 months) was 150/90 mmHg or less (an audit rather than an individual care standard) increased from 71.5 to 79.2%. A further slight increase to 80.4% was observed in 2015, when the time interval for measurement was increased to 12 months.^[22] Within England, therefore, the incentive has done little to improve the proportion of patients achieving blood pressure control at a national level, despite this being one of the most heavily incentivized indicators within the framework.

This lack of impact has also been reported by Serumaga *et al.*^[23] in their interrupted time series analysis of hypertension management before and after the introduction of P4P. Similarly, there were no significant changes to the numbers of patients being treated with combination therapy. This trend was observed prior to implementation of P4P and was subsequently sustained.

On a more positive note, they did not find evidence of gaming to achieve targets. Nationally reported data on the numbers of patients excluded from the indicator denominator through a process known as exception reporting has remained constant at approximately 3–4%, which would support this conclusion. However, potential gaming activity has been described by others, in particular the rounding down of recorded blood pressure by a few mmHg in order to meet the target.^[20] This may be perceived as being unlikely to have any significant clinical consequences for the patient, but may have significant financial impact on the physician or family practice.

Another contributing factor to this apparent lack of impact could be the threshold set for the payment of the incentive. Within the UK, the QOF rewards absolute achievement with an upper and lower threshold for minimum and maximum payment. In 2004/05, these were set at 25–70% and are currently set at 45–80%. This could well be below the level required for physicians to change their practice. Pilot testing of indicators, which did not commence until 2008, could have given an indication of current levels of achievement and been

used to inform threshold setting. Whether or not the incentives for hypertension demonstrate value for money would require an assessment of the benefits gained by the modest increase in the proportion of patients with hypertension with well controlled blood pressure compared with the costs of providing the required interventions, plus the incentive points awarded for threshold achievement.

Alternatively, it could be that the incentive was applied to the wrong organization. Petersen *et al.*^[24] undertook a cluster randomized controlled trial to evaluate the impact of P4P upon adherence to guideline-based hypertension care comparing incentives paid to individual physicians, to practices, mixed individual and practice payments and no payment. They observed no significant differences in blood pressure control between the intervention and control groups unless the incentive was aimed at the individual physician. In common with other studies, this effect was not sustained once the incentive was withdrawn.^[25]

While improvement against blood pressure control targets may be disappointing, it has also been suggested that this is not the most clinically meaningful measure, and that clinician response to a sub-optimal blood pressure recording is a better discriminator of quality. These clinical action measures place an equal emphasis upon achieving a control target or on taking appropriate clinical action such as modification of therapy in a timely manner. By taking this approach, Weiler *et al.*^[26] were able to identify 52% of patients with hypertension as receiving quality care as opposed to 20% when taking a target-based approach. Measures such as these, however, require a more sophisticated approach to data collection and analysis than control target measures alone.

Control target measures, in the absence of case-mix adjustment, may also promote over-treatment.^[27] This risk is becoming more acute given an aging population and an increase in the numbers of people with multi-morbidity who may require complex trade-offs in optimal single disease management to achieve individualized person-centered care. However, such case-mix adjustments are difficult to define. One option therefore is to allow clinicians to opt patients out of the care described in quality measures in a pre-determined set of circumstances. Within the UK QOF, this process is termed exception reporting. Recent analysis has suggested that the likelihood of being exception reported is strongly related to increasing age and numbers of co-morbid conditions.^[28] While this might suggest that it is being used to protect patients from the detrimental effects of over-treatment, further qualitative work is required to fully understand the process of deciding to exception report a patient.

CONCLUSION

P4P schemes are extremely diverse in their design and implementation with the potential for system-wide impact. Because of this, evaluation in one country may have limited transferability to other health systems and design structures. Many aspects of the management of hypertension appear amenable to quantitative measurement, although depending upon current levels of care, may or may not require incentivization. Evaluation of impact therefore is highly sensitive to the local context. In order to maximize the potential for shared learning, it is important that this is recognized and that attention is given not only to performing the evaluation itself, ideally within a trial setting, but also to describing the constituent parts of this complex intervention.

Previous Presentation

This editorial is based on talk given by PG at the Chinese Society of Cardiology & the 9th Oriental Congress of Cardiology, Shanghai 10–13 September 2015.

Sources of Financial Support

The authors are contracted to the National Institute for Health and Care excellence (NICE) to provide advice on piloting new indicators for the Quality and Outcomes Framework.

Conflict of Interest

All authors are fully independent of NICE and the Department of Health. NICE had no role in design and conduct of the study; collection, management, analysis, and interpretation of the data; and preparation, review, or approval of the manuscript; and the decision to submit the manuscript for publication.

REFERENCES

1. Lim S, Vos T, Flaxman AD, Danaei G, Shibuya K, Adair-Rohani H, *et al.* A comparative risk assessment of burden of disease and injury attributable to 67 risk factors and risk factor clusters in 21 regions, 1990–2010: a systematic analysis for the Global Burden of Disease Study 2010. *Lancet* 2012; 380:2224–60.
2. National Clinical Guideline Centre. Hypertension: the clinical management of primary hypertension in adults. Clinical Guideline 127. National Clinical Guideline Centre, London 2011.
3. Cashin C, Chi Y-Ling, Smith PC, Borowitz M, Thomson S. Health provider P4P and strategic health purchasing In: Cashin C, Chi Y-Ling, Smith PC, Borowitz M, Thomson S (eds.) *Paying for Performance in Health Care: Implications for health system performance and accountability*. Open University Press, Maidenhead, 2014, pp. 3–22.
4. Trisolini MG. Introduction to pay for performance In: Cromwell J, Trisolini MG, Pope GC, Mitchell JB, Greenwald LM (eds.) *Pay for Performance in Health Care: Methods and approaches*. RTI Press publication, Research Triangle Park, NC, 2011, pp. 7–32.
5. Witter S, Toonen J, Meessen B, Kagubare J, Fritsche G, Vaughan K. Performance-based financing as a health system reform: mapping the key dimensions for monitoring and evaluation. *BMC Health Services Research*. 2013; 13:367.
6. Trisolini MG. Theoretical Perspectives on Pay for Performance In: Cromwell J, Trisolini MG, Pope GC, Mitchell JB, Greenwald LM (eds.) *Pay for Performance in Health Care: Methods and approaches*. RTI Press publication, Research Triangle Park, NC, 2011, pp. 77–98.
7. Emanuel EJ, Ubel PA, Kessler JB, Meyer G, Muller RW, Navathe AS. Using behavioural economics to design physician incentives that deliver high-value care. *Ann Intern Med* 2016; 164:114–9.
8. Maxwell RJ. Dimensions of quality revisited: from thought to action. *Qual Health Care* 1992; 1:171–7.
9. Donabedian A. The 7 pillars of quality. *Arch Pathol Lab Med* 1990; 114:1115–8.
10. Campbell SM, Roland MO, Buetow SA. Defining quality of care. *Soc Sci Med* 2000;51:1611–25.
11. Institute of Medicine. *Crossing the Quality Chasm: A New Health System for the 21st Century*. National Academy Press, 2001.
12. Darzi A. *High quality care for all*. Department of Health, London, 2008.
13. Stelfox HT, Strauss SE. Measuring quality of care: considering conceptual approaches to quality indicator development and evaluation. *Journal Clin Epidemiol*. 2013; 66:1328–1337.
14. Shekelle PG. Quality indicators and performance measures: methods for development need more standardization. *Journal Clin Epidemiol* 2013; 66:1338–9.
15. Campbell SM, Braspenning J, Hutchinson A, Marshall M. Research methods used in developing and applying quality indicators in primary care. *Qual Saf Health Care* 2002; 11:358–64.
16. Naylor CD. Grey zones in clinical practice: some limits to evidence based medicine. *Lancet* 1995;345:840–2.
17. Campbell SM, Kontopantelis E, Hannon K, Burke M, Barber A, Lester HE. Framework and indicator testing protocol for developing and piloting quality indicators for the UK quality and outcomes framework. *BMC Family Practice* 2011;12:85.
18. Kautter J. Incorporating Efficiency Measures into Pay for Performance In: Cromwell J, Trisolini MG, Pope GC, Mitchell JB, Greenwald LM (eds.) *Pay for Performance in Health Care: Methods and approaches*. RTI Press publication, Research Triangle Park, NC, 2011, pp. 139–160.
19. Marshall MN, Shekelle PG, McGlynn EA, Campbell S, Brook RH, Roland MO. Can health care quality indicators be transferred between countries? *QualSaf Health Care*. 2003; 12:8–12.
20. Doran T, Fullwood C. Pay for performance: is it the best way to improve control of hypertension? *Curr Hypertens Rep* 2007; 9:360–7.
21. NHS Employers. 2015/16 General Medical Services (GMS) contract Quality and Outcomes Framework (QOF): guidance for GMS contract 2015/16. NHS Employers, London. 2015.
22. Health and Social Care Information Centre. Quality and Outcomes Framework <http://www.hscic.gov.uk/qof>. Last accessed on February 19, 2016.
23. Serumaga B, Ross-Degnan D, Avery AJ, Elliott RA, Majumdar SR, Zhang F, *et al.* Effect of pay for performance on the management and outcomes of hypertension in the United Kingdom: interrupted time series study. *BMJ* 2011; 342:d108.
24. Petersen LA, Simpson K, Pietz K, Urech TH, Hysong SJ, Profit J, *et al.* Effects of individual physician-level and practice-level financial incentives on hypertension care: a randomized trial. *JAMA* 2013;310:1042–50.
25. Lester H, Schmittiel J, Selby J, Fireman B, Campbell S, Lee J, *et al.* The impact of removing financial incentives from clinical quality indicators: longitudinal analysis of four Kaiser Permanente indicators. *BMJ* 2010;340:c1898.
26. Weiler S, Gemperli A, Collet TH, Bauer DC, Zimmerli L, Comuz J, *et al.* Clinically relevant quality measures for risk factor control in primary care: a retrospective cohort study. *BMC Health Serv Res* 2014;14:306.

27. Kerr EA, Lucatorto MA, Holleman R, Hogan MH, Klamerus ML, Hofer TP, *et al.* Monitoring performance for blood pressure management among diabetic patients: too much of a good thing? *Arch Intern Med* 2012;172:938-45.
28. Kontopantelis E, Springate DA, Ashcroft DM, Valderas JM, van der Veer SN, Reeves D, *et al.* Associations between exemption and survival outcomes in the UK's primary care pay-for-performance programme: a retrospective cohort study. *BMJ Qual Saf* 2015; 0:1-14.

How to cite this article: Foskett-Tharby R, Hex N, Gill P. Pay for performance and the management of hypertension. *J Transl Intern Med* 2016; 4: 14-9.