

C5: Financial incentives: Pay-for- quality



Wilm Quentin

European Observatory at Department of Health Care Management
Berlin University of Technology

27 July 2018



Session outline



- What are the characteristics of the strategy?
- Overview of Pay for Quality initiatives in European countries?
- What do we know about the effectiveness and cost-effectiveness of pay-for-quality?
- How can pay-for-quality be implemented?
- Conclusions for policy-makers



European Observatory on Health Systems and Policies Series

Paying for Performance in Health Care

Implications for health system
performance and accountability

Edited by
Cheryl Cashin
Y-Ling Chi
Peter Smith
Michael Borowitz
and Sarah Thomson



OECD Health Policy Studies

Better Ways to Pay for Health Care





Underlying rationale of P4Q



- Traditional payment systems do not provide direct incentives for high quality care.
 - FFS versus capitation
 - DRGs versus global budgets
- Assumption: providers will improve quality if they have a direct financial interest to do so
- Proponents of P4Q:
 - Intrinsic motivation/non-financial incentives are not enough to motivate improvement
- Opponents of P4Q:
 - Financial incentives may crowd out intrinsic motivation → adverse consequences, e.g. exclusive focus on indicators



Defining Pay-for-Quality (P4Q)



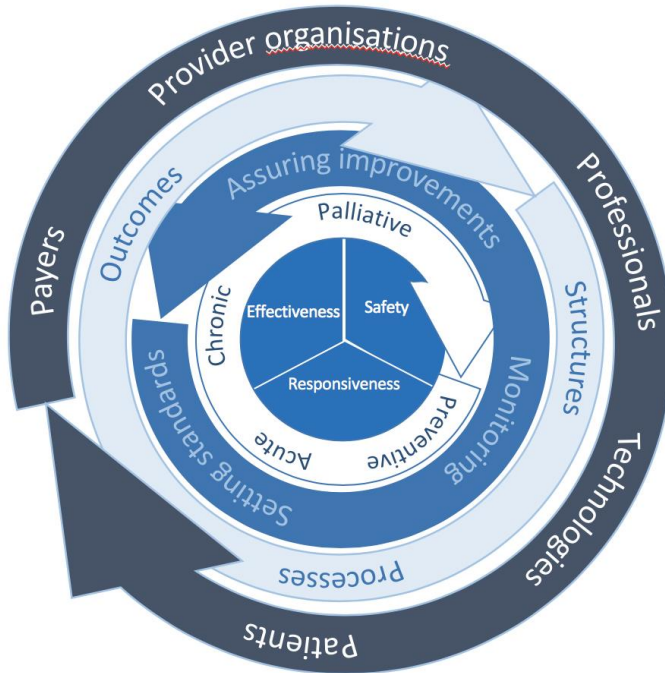
Two defining characteristics

- (1) performance of providers is monitored in relation to pre-specified quality indicators and
- (2) a monetary transfer is made conditional on the (achievement or improvement of) measured quality of care.

- P4Q is a more specific term than Pay-for-Performance (P4P)
- Many P4P programs include aspects of P4Q
- But most P4P programs target many other dimensions of performance



The 5-L Framework to analyse Pay for Quality in Europe



Dimensions of Quality

Areas of care

Activities of strategies

Donabedian's triad

Targets of strategies



Characteristics of P4P



In theory

- P4Q may aim at assuring or improving quality in terms of effectiveness, safety, and/or responsiveness.
- P4Q programmes can focus on different areas of care (preventive, acute, chronic, or palliative care).
- Quality can be measured by use of structure, process or outcome indicators of quality – also in P4Q.
- Target different types of professionals (e.g. physicians, nurses or social workers) and organisations (e.g. primary care practices, hospitals or departments).

In practice, P4Q focuses more on certain areas of care and characteristics differ across areas



P4Q Characteristics of incentives



- Voluntary or mandatory?
- Bonus or penalty?
- Size of incentive?
- Absolute measure or relative ranking?
- Attainment or improvement?
- Separate payment for each element or composite performance?



Pay for Quality in Primary Care in Europe

13 programmes in
12 countries



Country	Programme (nat./reg., volunt./mand.)	Start	Care area	Quality dimension	Type and number of indicators	Incentive structure	Size of FI as % of income	Type and number of provider
CZ	(NW, V)	-	PC, CC	EFFS	P, S	B, AM	-	IND
DE	DMP (R,V)	2001	PC, CC	EFFS	P, S	B, AM	-	IND, ORG
EE	PHC QBS (NW, V)	2005	PC, CC	EFFS	P, S - 45	B, AM, A+I	≤4.5 %	ORG (90%)
FR	ENMR (NW,V)	2009	PC, CC	EFFS	S	B, W	5%	ORG (300)
FR	ROSP (NW,V)	2009	PC, CC	EFFS	O, P, S - 29	B, AM, A+I	≤11 %	IND (97%)
HR	(NW, V)	2013	PC, CC	EFFS, RESP	P, S	B, AM	≤30%	IND
IT	Diabetes care program (R, M)	2003	CC	EEFS	P, S	B, AM, A	0.1 - 6%	IND (2,938)
LT	(NW, V)	2005	PC, CC	EFFS	O, P - 22	B, AM	-	ORG
LV	(NW, M)	2013	PC, CC	EFFS	O, P, S	B, AM	≤5%	ORG
PL	(NW, M)	2016	PC, CC	EFFS	P	B, AM	-	IND, ORG
PT	Model B (NW, V)	2006	PC, CC	EFFS, RESP	O, P, S - 22	B, AM, A	GPs: ≤30%, nurse: ≤10%	IND (181)
SE	Västra Götaland: M	2011	PC, CC	EFFS, RESP	O, P, S - 40	B, AM, A	≤4%	ORG (200)
UK (ENG)	QOF (NW,V)	2004	PC, CC	EFFS, RESP	O, P, S - 77	B, AM, A	≤15% (2013)	ORG (99%)



Overview of P4Q in Primary Care in Europe



- Programmes introduced since early 2000s
- Most programmes implemented at national level – but some (Germany, Italy, Sweden) regional programmes
- Half are mandatory – other half are voluntary
- All programmes award bonuses – almost always based on absolute performance
- Size of incentive: usually $< 5\%$ - but Croatia, France, Portugal and the UK pay $>10\%$



Overview of P4Q in Primary Care in Europe



- Area of care:
 - Focus on prevention and chronic care
 - No specific indicators for acute or palliative care
- Dimension of Quality:
 - Focus on effectiveness (e.g. provision of certain services, compliance with guidelines, improved blood pressure control).
 - Four programmes include indicators for responsiveness (patient experience/satisfaction)
- Donabedian's triad:
 - Focus on structures and processes
 - Six countries also incentivize quality in terms of outcomes
- Incentives paid to organisation or individual (depending on organisation of primary care in country)



Experiences from other countries?



Overview of P4Q in hospitals in Europe



13 programmes in
9 countries

Country	Programme	Start	Care area	Quality dimensions	Type and number of indicators	Incentive structure	Size of FI as % of total income	Type and number of provider
DK	Journalauditindikatorer (NW,M)	2009	AC	EFFS, RESP	O, P, S	B, PN, AM	<1%	IND (4 hospitals)
FR	IFAQ (NW,M+V)	2012	AC	EFFS	P, S	B, RR, TOP20P	0.4-0.6% - V; 0.2-0.5% - M; (15t- 500t€)	ORG – 460
HR	(NW, M)	2015	AC	EFFS	O, P, S	W, AM (RR)	10%	ORG
IT	PAFF (Lazio,M)	2009	AC	EFFS	P - 1	PN, AM	Reduced reimbursement	ORG
LU	Incitants qualité (NW,V)	1998	AC	EFFS, SFTY	O, P, S	B, AM	≤2.00%	ORG
NO	QBF (NW,V)	2014	AC	EFFS, RESP, SFTY	O, P – 33	W, RR	Redistribution of NOK 500M (10% of reg. bud)	ORG (4 regions)
PT	Hospital contract (NW,M)	2002	AC	EFFS, SFTY	O, P – 12	B, PN, RR	≤5%	ORG
SE	R, M (in 10 out of the 21 regions)	2004	AC	EFFS, RESP	O, P, S	W, AM	2-4%	
UK (ENG)	Advancing Quality (NW,V)	2008	AC	EFFS, RESP	O, P – 52	B, AM	2-4%	IND (teams), ORG (24)
UK (ENG)	CQUIN (NW,M)	2009	AC	EFFS, RESP, SFTY	O, P -	PN, AM	0.5 - 2.5% of the contract	ORG
UK (ENG)	BPT (NW,M+V)	2010	AC	EFFS, SFTY	P - 65	B, W, AM	<1% (5 - 43% of tariff)	ORG
UK (ENG)	Non-pay for never events (NW,M)	2009	AC	SFTY	O - 14	PN, AM	No reimbursement	ORG
UK (ENG)	Non-payment for ER (NW,M)	2011	AC	SFTY	O - 1	PN, AM	No reimbursement	ORG



Overview of P4Q in hospitals in Europe



- First programme introduced in 1998 in Luxembourg – others rather later than for primary care
- Almost all programmes are implemented nationwide and mandatory
- Type of incentive: Mostly penalties
- Size of incentive: usually relatively small (<2%), Portugal (<5%), Croatia (<10% - but broader P4P)
- Usually absolute measures – but France, Norway and Portugal use relative ranking



Overview of P4Q in hospitals in Europe



- Area of care: Focus on Acute Care
 - Acute Myocardial Infarction (AMI), acute stroke, renal failure, hip fracture and hip and knee replacement surgery
- Dimension of Quality:
 - Focus on effectiveness
 - Seven programmes with indicators for safety
 - Five programmes with indicators for responsiveness (Denmark, Norway, Sweden, UK)
- Donabedian's triad:
 - Focus on processes and outcomes
 - Only four countries with indicators for structures
- Incentives paid to hospital (only in DK/UK AQ to departments/teams)



Experiences from other countries?



Effectiveness of P4Q?



Effectiveness of P4Q in primary care – not only in Europe



Review	Reviews of P4Q in primary care	Included studies			
		Number*	Country of origin	Date range	Quality
Lin et al., 2016	Effect of P4Q on health care quality	44	FR, NL, TW, UK, US	1998-2013	M/H
Kondo et al., 2015	Effect of P4Q on health care quality (in ambulatory settings)	41	CA, FR, NL, TW, US, UK, TW	2006-2014	NR
Rashidian et al., 2015	Effects of FI-based drug policies on drug use, health care utilisation, health outcomes and costs	3 (18)	UK, NL	2007-2011	L
Damberg et al., 2014	Shared savings models (linked to quality of care)	6 (45)	US	2009-2013	L
Langdown and Peckham 2014	Efficacy of QOF in improving health outcomes, impact on non-incentivised activities and robustness of clinical targets	10 (11)	UK	2007-2012	L/M
Hamilton et al., 2013	Effectiveness of FI in provision of smoking cessation interventions and in health outcomes	18	UK, US, DE, TW	2003-2010	M
Huang et al., 2013	Effects of P4Q on management of diabetes (a meta-analysis)	11-MA (21)	UK, US	2003-2010	L/M
Houle et al., 2012	Effect of P4Q on health care quality	30	UK	1995-2012	L/M
Gillam et al., 2012	Impact of P4Q on quality of UK's primary care	53 (94)	US, TW	2004-2011	L
Scott et al., 2011	Effect of changes in the method and level of payment on the quality of care	6 (7)	UK, US, DE	2003-2009	L
De Bruin et al., 2011	Effects of P4Q on health care quality and costs of chronic care through disease management	5 (18)	US, AU	2003-2010	NA
Sabatino et al., 2008	Effectiveness of audit and feedback, and FI on cancer screening	3 (12)	US	1991-1998	M
Sorbero et al., 2006	Effect of P4Q on health care quality	15	US	1995-2006	NA
Petersen et al., 2006	Effect of explicit FI on health care quality	16 (17)	US	1992-2005	L/M
Dudley et al., 2004	Effects and potential of P4Q schemes to improve quality of care	9	US	1987-2003	NA
Kane et al., 2004	Effects of FI on preventive care delivery	10 (66)	UK, US	1992-2001	L/M
Giuffrida et al., 2000	Effects of target payments on cost, pattern, quantity and quality	2	UK, US	1992-1998	M
Achat et al., 1999	Effects of incentives on immunisation uptake	3(8)	UK, US	1992-1996	NA

18 reviews



Effectiveness of P4Q in primary care + hospitals – not only in Europe



Review	Review focus not limited to specific providers	Included studies			
		Number*	Country of origin	Date range	Study quality
Korenstein et al., 2016	Impact of system-level interventions on the value of US health care, defined as the balance between quality and cost	30	US	2009-2015	M
Ogundeji et al., 2016	To explore systematically the extent and sources of heterogeneity in the results of evaluations of P4Q schemes with the aim to identify features associated with success in P4Q schemes	96-MR, 37-MA	Div.	1998-2014	NA
Barreto et al., 2015	Effect of P4Q on health care quality	25 (27)	US, UK, SE, TW, IT	1991-2011	NR
Damberg et al., 2014 (P4Q)	Effects of P4Q on quality and resource use, efficiency, and costs	58 (89)	UK, US	2001-2013	L/M
Ivers et al., 2012	Effectiveness 12 reviews lack	3 (111)	US	1980-1999	L/M
Emmert et al., 2012	Analyse the existing literature regarding economic evaluation of P4Q	9	UK, US, DE	1992-2010	L
Van Herck et al., 2010	Effect of P4Q on health care quality	51 (128)	AU, DE, ES, IT, UK, US	1992-2009	M
French et al., 2010	Effects of interventions that aim to improve the appropriate use of imaging for people with musculoskeletal conditions	1 (28)	US	2007	M
Christianson et al., 2007, 2008	Effect of P4Q on health care quality	37	UK, US, TW, ES	1992-2007	NA
Armour et al., 2001	Effects of FI on physician resource use and the quality of medical care	5(7)	UK, US	1994-1998	NA

Reviews of P4Q in hospitals

Milstein and Schreyoegg, 2016	Impact of P4Q programmes in the inpatient sector	46	DK, CA, IT, KR, JP, TR, UK, US	2006-2015	NA
Kondo et al., 2015	Effect of P4Q on health care quality (in hospital setting)	7	IT, TW, UK, US	2010-2014	NR
Damberg et al., 2014 (BP)	Bundled payments (linked to quality)	1(3)	US	2007-2011	L
Damberg et al., 2007, Mehrotra et al., 2009	Impact of P4Q in inpatient or outpatient hospital services	9	US	2004-2007	L



Effectiveness of P4Q: Overview



- 32 reviews of P4Q between 1999 and 2016
- Reviews with a total of around 400 original studies
- 18 reviews only with studies from UK and/or US
- 14 reviews included studies from other countries – but rarely from other European countries



Effectiveness in **primary care**



Quality and Outcomes framework (QoF) (7 reviews with 71 studies)

- During first year of programme
 - significant improvements in process-of-care indicators (e.g. smoking cessation activities, diabetes management activities) (some improvements > 30%)
 - less improvement in intermediate health outcomes (e.g. blood pressure, cholesterol and blood glucose level under control).
- Subsequent years (2005-07)
 - Process-of-care indicators continued to slowly improve → similar to trend before implementation
 - For some health outcomes → improvement below pre-QoF trend
- Final health outcomes of targeted indications (incidence of AMI, stroke, renal failure and all-cause mortality) → no effect



Effectiveness in **primary care**



- Chronic care (124 studies)
 - Observational or uncontrolled quasi-experimental studies: → large positive effects on processes of care
 - Better designed studies (interrupted time series, controlled before after and other quasi-experimental designs with a comparison group + longer time periods) → no effect or small positive effects
- Preventive care (87 studies)
 - Higher quality studies: positive results for processes of care
 - Outcomes, e.g. quit rates, or smoking prevalence → no effects of P4Q



Effectiveness in **primary care**



- Effects on final health outcomes (two reviews of 11 studies)
 - Seven studies: positive effects on diabetes-related hospitalisation and long-term complications, reduced emergency department visits, depression treatment response and on neonatal ICU admissions
 - Two studies : no effects on 30-day mortality, readmissions, ED visits related to AMI etc.
 - Two studies: detrimental effects on ED visits related to asthma, diabetes and heart failure
- Effects on responsiveness (five reviews commented on effects)
 - Six observational studies: QOF did not improve communication, coordination, patient satisfaction → continuity deteriorated
 - Two reviews of several P4Q: no change of patient satisfaction
 - Two other reviews: positive effects – but could not be clearly attributed to P4Q



Effectiveness in **hospital care**



- Effects on processes of care
 - 30 studies of 15 P4Q programmes
 - Studies with comparison group:
 - small short-term and/or non-significant effects on composite of multiple processes of care measures
 - but positive effects for some individual indicators
 - Studies without comparison groups:
 - Positive effects on breast cancer, AMI
- Effects on health outcomes
 - 13 studies
 - Short term positive effects → but not in the long-term
 - Lower quality studies → sometimes positive effects



Effectiveness in **hospital care**



- Effects on patient safety:
 - 6 reviews including total of 7 studies of 7 programmes
 - Two studies: significantly positive effects
 - Four studies: small positive effects, not significant
- Effects on responsiveness
 - Five reviews reporting results of four studies → no reliable evidence for improvement



Health Policy. 2016 Oct;120(10):1141-1150. doi: 10.1016/j.healthpol.2016.09.002. Epub 2016 Sep 5.

The effectiveness of payment for performance in health care: A meta-analysis and exploration of variation in outcomes.

Oqundeji YK¹, Bland JM², Sheldon TA³.

+ Author information

Abstract

BACKGROUND: Pay for performance (P4P) is widely used in health care, but results of evaluations vary considerably.

METHODS: Evaluations of P4P schemes were identified in a search of health care schemes in health care in four bibliographic databases. Effect sizes (standardized mean differences) were estimated using meta-analysis. Subgroup analysis, meta-regression and sensitivity analysis were used. Random-effects models were used because of heterogeneity rather than just chance. Sensitivity analysis was used to assess the risk of bias.

FINDINGS: 96 primary studies were identified. The proportion of observed effect sizes was significantly greater than expected by regression. The proportion of observed effect sizes was significantly greater than expected by regression. Estimates of effect of P4P schemes were significantly greater than expected by regression compared to no controls (0.15; 95% CI: 0.09–0.21) compared to process measures (e.g., giving cessation advice) (0.18; 95% CI: 0.06–0.31). Schemes evaluated using less rigorous designs were 24 times more likely to have positive estimates of effect than those using randomized controlled trials (OR = 24; 95% CI: 6.3–92.8).

INTERPRETATION: Estimates of the effectiveness of incentive schemes on health outcomes are probably inflated due to poorly designed evaluations and a focus on process measures rather than health outcomes. Larger incentives and reducing the perceived risk of non-payment may increase the effect of these schemes on provider behavior.

Findings

- Effect of P4P schemes lower in evaluations using randomized controlled trials (SMD = 0.08; 95% CI: 0.01–0.15) compared to no controls (0.15; 95% CI: 0.09–0.21),
- Effects lower for those measuring outcomes (e.g., smoking cessation) (SMD = 0.0; 95% CI: –0.01 to 0.01) compared to process measures (e.g., giving cessation advice) (0.18; 95% CI: 0.06–0.31).
- Schemes evaluated using less rigorous designs were 24 times more likely to have positive estimates of effect than those using randomized controlled trials (OR = 24; 95% CI: 6.3–92.8).



Comparative effectiveness and Cost-effectiveness of P4Q



- P4Q versus audit & feedback:
 - One review identified one cluster-RCT → P4Q less effective in reducing test ordering
- Cost-effectiveness of P4Q:
 - Emmert identified two full economic evaluations and six partial economic evaluations
 - Study quality was generally low
 - Results were mixed
 - Effectiveness is uncertain → unlikely to be cost-effective



Implementation of P4Q



■ What to incentivize?

- Broad definition of quality
- Involving providers in programme design
- Aligning incentives with professional norms and values
- Consider risks of P4Q: risk-selection, focus on incentivized areas, gaming

■ How to measure?

- Include outcome measures if risk-adjustment is good and sample size sufficient
- Monitor also non-incentivized aspects (e.g. including patient satisfaction)

■ Who to incentivize?

- Group incentives are preferred over individual incentives
- Mandatory participation preferable
- Voluntary participation may secure acceptance



Implementation of P4Q



■ How to incentivize?

- Size should be sufficient but not too large.
- Short time lag.
- Absolute or relative measures?
- Penalty or bonus?
- Attainment or improvement?
- Individual indicator or composite performance?
- Blended systems, possibly with choice

■ How to implement?

- Involving all stakeholders (from start, through evaluation, and revision).
- Implement in a way that enables systematic evaluation.
- Sophisticated ICT is important
- Monitor potential unintended consequences.
- Periodic reevaluation of measures.



Conclusions for policy-makers



- P4Q is increasingly used in many European countries.
- There is very limited evidence from European P4Q programmes
- Studies report mostly positive effects – but
 - mostly for process measures (not for outcomes)
 - mostly short-term effects (no long-term improvements)
 - mostly based on low-quality evidence
- Implementation is politically and technically complex
- Programme implementation should be evaluated!