

Advantages, disadvantages and feasibility of the introduction of 'Pay for Quality' programmes in Belgium - Supplement

KCE reports 1185

The Belgian Health Care Knowledge Centre

Introduction : The Belgian Health Care Knowledge Centre (KCE) is an organization of public interest, created on the 24th of December 2002 under the supervision of the Minister of Public Health and Social Affairs.
KCE is in charge of conducting studies that support the political decision making on health care and health insurance.

Administrative Council

Actual Members : Pierre Gillet (President), Dirk Cuypers (Vice-president), Jo De Cock (Vice-president), Frank Van Massenhove (Vice-president), Yolande Avondtroodt, Jean-Pierre Baeyens, Ri de Ridder, Olivier De Stexhe, Peter Degadt, Daniel Devos, Jean-Noël Godin, Floris Goyens, Jef Maes, Pascal Mertens, Raf Mertens, Marc Moens, François Perl, Marco Schetgen, Yves Smeets, Patrick Verertbruggen, Michel Foulon, Myriam Hubinon

Substitute Members : Rita Cuypers, Christiaan De Coster, Benoît Collin, Lambert Stamatakis, Karel Vermeyen, Katrien Kesteloot, Bart Ooghe, Frederic Lernoux, Anne Vanderstappen, Paul Palsterman, Geert Messiaen, Anne Remacle, Roland Lemeye, Annick Poncé, Pierre Smiets, Jan Bertels, Catherine Lucet, Ludo Meyers, Olivier Thonon.

Government commissioner : Roger Yves

Management

Chief Executive Officer a.i. : Jean-Pierre Closon

Information

Federaal Kenniscentrum voor de gezondheidszorg - Centre fédéral d'expertise des soins de santé – Belgian Health Care Knowledge Centre.

Centre Administratif Botanique, Doorbuilding (10th floor)

Boulevard du Jardin Botanique 55

B-1000 Brussels

Belgium

Tel: +32 [0]2 287 33 88

Fax: +32 [0]2 287 33 85

Email : info@kce.fgov.be

Web : <http://www.kce.fgov.be>

Advantages, disadvantages and
feasibility of the introduction
of 'Pay for Quality'
programmes in Belgium -
Supplement

KCE reports 118S

LIEVEN ANNEMANS, PAULINE BOECKSTAENS, LIESBETH BORGERMANS, DELPHINE DE SMEDT,
CHRISTIANE DUCHESNES, JAN HEYRMAN, ROY REMMEN, WALTER SERMEUS,
CARINE VAN DEN BROEKE, PIETER VAN HERCK, MARC VANMEERBEEK,
SARA WILLEMS, KRISTEL DE GAUQUIER

KCE reports I I8S

- Title :** Advantages, disadvantages and feasibility of the introduction of 'Pay for Quality' programmes in Belgium - Supplement
- Authors :** Lieven Annemans (UGent), Pauline Boeckxstaens (UGent), Liesbeth Borgermans (KULeuven), Delphine De Smedt (UGent), Christiane Duchesnes (Université de Liège), Jan Heyrman (KULeuven), Roy Remmen (Universiteit Antwerpen), Walter Sermeus (KULeuven), Carine Van Den Broeke (KULeuven), Pieter Van Herck (KULeuven), Marc Vanmeerbeek (Université de Liège), Sara Willems (UGent), Kristel De Gauquier (KCE)
- Reviewers:** Christian Léonard (KCE), Imgard Vinck (KCE)
- External experts:** Hugo Casteleyn (AZ Sint-Blasius Dendermonde), Pierre Chevalier (UCL), Xavier de Béthune (Alliance Nationale des Mutualités Chrétiennes), Pierre Gillet (Université de Liège), Margareta Haelterman (Federal Public Service Health, Food Chain Safety and Environment), Annelies Van Linden (Domus Medica)
- Acknowledgements:** Belgium: Bert Aertgeerts (CEBAM, KULeuven), Dirk Broeckx (APB), Geneviève Bruwier (ULg, SSMG, Forum des Associations de Médecins Généralistes), Piet Calcoen (DKV), Xavier de Béthune (Alliance Nationale des Mutualités Chrétiennes), Jo De Cock (RIZIV/INAMI), Christiaan Decoster (Federal Public Service Health, Food Chain Safety and Environment), Peter Degadt (Zorgnet Vlaanderen), Didier de Laminne de Bex (DKV), Jan De Maeseneer (UGent), Ri De Ridder (RIZIV/INAMI), Daniel Désir (CHU Brugmann), Jos Desmedt (Domus Medica), Jacques De Toeuf (CHIREC, ABSYM), Alain De Wever (ULB), Pierre Drielsma (Maisons Médicales), Guy Durant (UCL), Micky Fierens (LUSS), Pierre Gillet (ULg), Johan Hellings (Ziekenhuis Oost-Limburg), Marc Justaert (Landsbond der Christelijke Mutualiteiten), Johan Kips (UZLeuven), Jean-Marc Laasman (Union nationale des mutualités socialistes), Roger Lonfils (Communauté française Wallonie-Bruxelles, Direction générale de la Santé), Catherine Lucet (Union nationale des mutualités socialistes), Pascal Meeus (RIZIV/INAMI), Marc Moens (BVAS), Dominique Pestiaux (UCL), Michel Roland (ULB, CEBAM), Robert Rutsaert (ASGB), Marco Schetgen (ULB), Erik Schokkaert (KULeuven), Carine Serano (Ligue des Usagers des Services de Santé), Piet Vandenbussche (UGent), Walter Van Den Eede (Agentschap Inspectie Welzijn, Volkgezondheid en Gezin), Chris Vander Auwera (Flemish Agency for Care and Health), Philippe Vandermeeren (Groupement Belge des Omnipraticiens), Michel Van Hoegaarden (Federal Public Service Health, Food Chain Safety and Environment), Bernard Vercruysse (UCL, Forum des Associations de Médecins Généralistes), Michel Vermeylen (Association des Médecins de Famille, ABSYM), Arthur Vleugels (KULeuven), Ilse Weeghman (Vlaams Patiëntenplatform)
- International: Mark Ashworth (King's College, London, UK), Jozé Braspenning (Scientific Institute for Quality of Health Care, the Netherlands), Alyn Chien (University of Chicago, US), Tim Coleman (University of Leeds, UK), Sheryl Damberg (RAND, US), Stephen Duckett (Queensland Health Government, Australia & Alberta Health Services, Canada), Adams Dudley (University of California, US), Robert Fleetcroft (University of East Anglia, UK), Christine Gardel (Haute Autorité de Santé, France), Ruth McDonald (University of Manchester, UK), Gary McLean (University of Glasgow, Scotland, UK), Peggy McNamara (AHRQ, US), Christopher Millett (Imperial College, London, UK), Karen Murphy (Temple University, Philadelphia, US), Martin Roland (University of Manchester/ Cambridge, UK), Meredith Rosenthal (Harvard University, US), Ian Scott (Princess Alexandra Hospital, Queensland, Australia), Colin

Simpson (University of Aberdeen, Scotland, UK), Abd Tahrani (University of Birmingham, UK)

External validators: Marc Jegers (VUB), Matthew Sutton (University of Manchester, UK), Cordula Wagner (Nivel & Free University Amsterdam medical centre, the Netherlands)

Conflict of interest: Hugo Casteleyn regularly speeches at seminars or at training sessions for health care institutions. Xavier de Béthune works for an organisation that may gain or lose financially when P4Q programmes are implemented.

Disclaimer: The external experts collaborated on the scientific report that was subsequently submitted to the validators. The validation of the report results from a consensus or a voting process between the validators. Only the KCE is responsible for errors or omissions that could persist. The policy recommendations are also under the full responsibility of the KCE.

Layout: Ine Verhulst

Brussels, 16th November 2009

Study nr 2008-12

Domain: Health Services Research (HSR)

MeSH: Salaries and Fringe Benefits; Reimbursement, Incentive; Fees and Charges; Quality Assurance, Health Care; Quality Control; Health Services Accessibility

NLM classification: W 84.4

Language: English

Format: Adobe® PDF™ (A4)

Legal depot: D/2009/10.273/53

Any partial reproduction of this document is allowed if the source is indicated. This document is available on the website of the Belgian Health Care Knowledge Centre.

How to refer to this document?

Annemans L, Boeckxstaens P, Borgermans L, De Smedt D, Duchesnes C, Heyrman J, Remmen R, Sermeus W, Van Den Broecke C, Van Herck P, Vanmeerbeek M, Willems S, De Gauquier K. Advantages, disadvantages and feasibility of the introduction of 'Pay for Quality' programmes in Belgium. Health Services Research (HSR). Brussels: Belgian Health Care Knowledge Centre (KCE). 2009. KCE Reports I 18S. D/2009/10.273/53

Supplement

Table of contents

APPENDICES WITH CHAPTER 2	2
APPENDIX 1 SYSTEMATIC REVIEWS.....	3
Search strategy in electronic databases.....	3
APPENDIX 2 PRIMARY EVALUATION STUDIES	8
Search strategy in electronic databases.....	8
APPENDIX 3 SYSTEMATIC REVIEWS.....	12
Relevance screening	12
APPENDIX 4 PRIMARY EVALUATION STUDIES	14
Relevance screening	14
APPENDIX 5 SYSTEMATIC REVIEWS.....	16
Quality appraisal.....	16
APPENDIX 6 PRIMARY EVALUATION STUDIES	19
Quality appraisal.....	19
APPENDIX 7 COST-EFFECTIVENESS AND MODELING STUDIES	30
Quality appraisal.....	30
APPENDIX 8 SYSTEMATIC REVIEWS.....	31
Citations included in full text analysis.....	31
APPENDIX 9 PRIMARY EVALUATION STUDIES	36
Citations included in full text analysis.....	36
APPENDIX 10 GRID EQUITY RELATED CONCEPTS	53
GENERAL INFORMATION QOF TABLE	53
METHODS	56
EQUITY ASPECTS FINDINGS	64
GENERAL INFORMATION NON QOF	74
METHODS	75
EQUITY ASPECTS FINDINGS	77
APPENDIX 11 OVERVIEW INTERVIEW CONTENT INTERNATIONAL EXPERTS	79
SEMI STRUCTURED INTERVIEW OF P4Q EXPERT COUNTRY REPRESENTATIVES	79
APPENDIX 12 FEASIBILITY ASSESSMENT CURRENT QUALITY CYCLES : SEE SEPARATE PDF DOCUMENT	81
APPENDIX 13 LIST OF STAKEHOLDERS.....	86
APPENDIX 14 STAKEHOLDERS QUESTIONNAIRE	86
APPENDICES WITH CHAPTER 4 – EVIDENCE BASE FOR P4Q	95
APPENDIX 15 DESCRIPTION OF STUDIES.....	95
APPENDIX 16A DETAILED DESCRIPTION	100
APPENDIX 16B EVIDENCE TABLES.....	107
APPENDIX 16C SYSTEMATIC REVIEW RESULTS PERIOD JANUARY – JULY 2009	140
REFERENCES	141

Appendices with Chapter 2

APPENDIX I SYSTEMATIC REVIEWS

SEARCH STRATEGY IN ELECTRONIC DATABASES

Database	Medline
Host	http://www.ncbi.nlm.nih.gov/sites/entrez (Pubmed)
Date of search	31/12/2008
Years covered	2000-2008
Search Strategy	(("Salaries and Fringe Benefits"[Majr] OR "Reimbursement, Incentive"[Majr] OR "Fees and Charges"[Majr] OR p4q OR p4p OR pay* OR incentive* OR compensation* OR reimbursement* OR financ* OR bonus* OR remunerat*) AND ("Treatment Outcome"[Majr] OR "Medical Errors"[Majr] OR "Quality Control"[Majr] OR "Cost-Benefit Analysis"[Majr] OR "Safety"[Majr] OR "Health Services Accessibility"[Majr] OR quality OR outcome* OR performance OR error* OR safety* OR access* OR equity OR effectiveness) AND ("Hospitals"[Majr] OR "Physicians"[Majr] OR hospital* OR physician* OR practitioner*)) AND systematic[sb] Limits: only items with abstracts, Humans (("Salaries and Fringe Benefits"[Majr] OR "Reimbursement, Incentive"[Majr] OR "Fees and Charges"[Majr] OR p4q OR p4p OR pay* OR incentive* OR compensation* OR reimbursement* OR financ* OR bonus* OR remunerat*) AND ("Treatment Outcome"[Majr] OR "Medical Errors"[Majr] OR "Quality Control"[Majr] OR "Cost-Benefit Analysis"[Majr] OR "Safety"[Majr] OR "Health Services Accessibility"[Majr] OR quality OR outcome* OR performance OR error* OR safety* OR access* OR equity OR effectiveness) AND ("Hospitals"[Majr] OR "Physicians"[Majr] OR hospital* OR physician* OR practitioner*)) Limits: only items with abstracts, Humans, Meta-Analysis, Practice Guideline, Review
Language restrictions	none
Number of citations	1549

Database	Embase
Host	http://www.embase.com/home
Date of search	31/12/2008
Years covered	2000-2008
Search Strategy	('salary and fringe benefit'/exp/mj OR 'reimbursement'/exp/mj OR 'fee'/exp/mj OR p4q OR p4p OR pay* OR incentive* OR compensation* OR reimbursement* OR financ* OR bonus* OR remunerat*) AND ('clinical effectiveness'/exp/mj OR 'medical error'/exp/mj OR 'treatment outcome'/exp/mj OR 'quality control'/exp/mj OR 'cost effectiveness analysis'/exp/mj OR 'patient safety'/exp/mj OR 'health care access'/exp/mj OR quality OR outcome* OR 'performance'/exp/mj OR error* OR safety* OR access* OR equity OR 'cost effectiveness'/exp/mj OR effectiveness) AND ('hospital'/exp/mj OR 'physician'/exp/mj OR hospital* OR physician* OR practitioner*) AND [humans]/lim AND [abstracts]/lim AND [review]/lim AND [embase]/lim ('salary and fringe benefit'/exp/mj OR 'reimbursement'/exp/mj OR 'fee'/exp/mj OR p4q OR p4p OR pay* OR incentive* OR compensation* OR reimbursement* OR financ* OR bonus* OR remunerat*) AND ('clinical effectiveness'/exp/mj OR 'medical error'/exp/mj OR 'treatment outcome'/exp/mj OR 'quality control'/exp/mj OR 'cost effectiveness analysis'/exp/mj OR 'patient safety'/exp/mj OR 'health care access'/exp/mj OR quality OR outcome* OR 'performance'/exp/mj OR error* OR safety* OR access* OR equity OR 'cost effectiveness'/exp/mj OR effectiveness)

	AND ('hospital'/exp/mj OR 'physician'/exp/mj OR hospital* OR physician* OR practitioner*) AND [humans]/lim AND [abstracts]/lim AND ([cochrane review]/lim OR [meta analysis]/lim OR [systematic review]/lim) AND [embase]/lim
Language restrictions	none
Number of citations	1865

Database	Web of science
Host	http://apps.isiknowledge.com/
Date of search	31/12/2008
Years covered	2000-2008
Search Strategy	TS=(salary and fringe benefit OR reimbursement OR fee OR p4q OR p4p OR pay* OR incentive* OR compensation* OR reimbursement* OR financ* OR bonus* OR remunerat*) AND TS=(clinical effectiveness OR medical error OR treatment outcome OR quality control OR cost effectiveness analysis OR patient safety OR health care access OR quality OR outcome* OR performance OR error* OR safety* OR access* OR equity OR cost effectiveness OR effectiveness) AND TS=(hospital OR physician OR hospital* OR physician* OR practitioner*) AND Document Type=(Review)
Language restrictions	none
Number of citations	423

Database	Centre for Research and Dissemination
Host	http://www.york.ac.uk/inst/crd/
Date of search	31/12/2008
Years covered	2000-2008
Search Strategy	TS=(salary and fringe benefit OR reimbursement OR fee OR p4q OR p4p OR pay* OR incentive* OR compensation* OR reimbursement* OR financ* OR bonus* OR remunerat*) AND TS=(clinical effectiveness OR medical error OR treatment outcome OR quality control OR cost effectiveness analysis OR patient safety OR health care access OR quality OR outcome* OR performance OR error* OR safety* OR access* OR equity OR cost effectiveness OR effectiveness) AND TS=(hospital OR physician OR hospital* OR physician* OR practitioner*) AND Document Type=(Review) Further selection of DARE and HTA results.
Language restrictions	none
Number of citations	409

Database	Cochrane Library																																																																																							
Host	http://www3.interscience.wiley.com/cgi-bin/mrwhome/106568753/HOME?CRETRY=1&SRETRY=0																																																																																							
Date of search	31/12/2008																																																																																							
Years covered	2000-2008																																																																																							
Search Strategy	<table border="0"> <tr><td># 1</td><td>MeSH Salaries and Fringe Benefits EXPLODE I 2</td><td>305</td></tr> <tr><td># 2</td><td>MeSH Reimbursement, Incentive EXPLODE I</td><td>23</td></tr> <tr><td># 3</td><td>MeSH Fees and Charges EXPLODE I</td><td>852</td></tr> <tr><td># 4</td><td>p4q</td><td>0</td></tr> <tr><td># 5</td><td>p4p</td><td>0</td></tr> <tr><td># 6</td><td>pay*</td><td>2480</td></tr> <tr><td># 7</td><td>incentive*</td><td>172</td></tr> <tr><td># 8</td><td>compensation*</td><td>127</td></tr> <tr><td># 9</td><td>reimbursement*</td><td>873</td></tr> <tr><td># 10</td><td>financ*</td><td>1495</td></tr> <tr><td># 11</td><td>bonus*</td><td>4</td></tr> <tr><td># 12</td><td>remunerat*</td><td>26</td></tr> <tr><td># 13</td><td>#1 OR #2 OR #3 OR #4 OR #5 OR #6 OR #7 OR #8 OR #9 OR #10 OR #11 OR #12</td><td>5327</td></tr> <tr><td># 14</td><td>MeSH Treatment Outcome EXPLODE I 2 3</td><td>5912</td></tr> <tr><td># 15</td><td>MeSH Medical Errors EXPLODE I</td><td>161</td></tr> <tr><td># 16</td><td>MeSH Quality Control EXPLODE I</td><td>65</td></tr> <tr><td># 17</td><td>MeSH Cost-Benefit Analysis EXPLODE I</td><td>10851</td></tr> <tr><td># 18</td><td>MeSH Safety EXPLODE I</td><td>284</td></tr> <tr><td># 19</td><td>MeSH Health Services Accessibility EXPLODE I 2</td><td>566</td></tr> <tr><td># 20</td><td>quality</td><td>16751</td></tr> <tr><td># 21</td><td>outcome*</td><td>17295</td></tr> <tr><td># 22</td><td>performance</td><td>1511</td></tr> <tr><td># 23</td><td>safety*</td><td>3431</td></tr> <tr><td># 24</td><td>access*</td><td>2203</td></tr> <tr><td># 25</td><td>equity</td><td>60</td></tr> <tr><td># 26</td><td>effectiveness</td><td>15571</td></tr> <tr><td># 27</td><td>#14 OR #15 OR #16 OR #17 OR #18 OR #19 OR #20 OR #21 OR #22 OR #23 OR #24 OR #25 OR #26</td><td>32615</td></tr> <tr><td># 28</td><td>#13 AND #27</td><td>4251</td></tr> <tr><td># 29</td><td>#13 AND #27 RESTRICT YR 2000 2008</td><td>2988</td></tr> </table> <p>Further selection of Cochrane reviews and Other review results</p>	# 1	MeSH Salaries and Fringe Benefits EXPLODE I 2	305	# 2	MeSH Reimbursement, Incentive EXPLODE I	23	# 3	MeSH Fees and Charges EXPLODE I	852	# 4	p4q	0	# 5	p4p	0	# 6	pay*	2480	# 7	incentive*	172	# 8	compensation*	127	# 9	reimbursement*	873	# 10	financ*	1495	# 11	bonus*	4	# 12	remunerat*	26	# 13	#1 OR #2 OR #3 OR #4 OR #5 OR #6 OR #7 OR #8 OR #9 OR #10 OR #11 OR #12	5327	# 14	MeSH Treatment Outcome EXPLODE I 2 3	5912	# 15	MeSH Medical Errors EXPLODE I	161	# 16	MeSH Quality Control EXPLODE I	65	# 17	MeSH Cost-Benefit Analysis EXPLODE I	10851	# 18	MeSH Safety EXPLODE I	284	# 19	MeSH Health Services Accessibility EXPLODE I 2	566	# 20	quality	16751	# 21	outcome*	17295	# 22	performance	1511	# 23	safety*	3431	# 24	access*	2203	# 25	equity	60	# 26	effectiveness	15571	# 27	#14 OR #15 OR #16 OR #17 OR #18 OR #19 OR #20 OR #21 OR #22 OR #23 OR #24 OR #25 OR #26	32615	# 28	#13 AND #27	4251	# 29	#13 AND #27 RESTRICT YR 2000 2008	2988
# 1	MeSH Salaries and Fringe Benefits EXPLODE I 2	305																																																																																						
# 2	MeSH Reimbursement, Incentive EXPLODE I	23																																																																																						
# 3	MeSH Fees and Charges EXPLODE I	852																																																																																						
# 4	p4q	0																																																																																						
# 5	p4p	0																																																																																						
# 6	pay*	2480																																																																																						
# 7	incentive*	172																																																																																						
# 8	compensation*	127																																																																																						
# 9	reimbursement*	873																																																																																						
# 10	financ*	1495																																																																																						
# 11	bonus*	4																																																																																						
# 12	remunerat*	26																																																																																						
# 13	#1 OR #2 OR #3 OR #4 OR #5 OR #6 OR #7 OR #8 OR #9 OR #10 OR #11 OR #12	5327																																																																																						
# 14	MeSH Treatment Outcome EXPLODE I 2 3	5912																																																																																						
# 15	MeSH Medical Errors EXPLODE I	161																																																																																						
# 16	MeSH Quality Control EXPLODE I	65																																																																																						
# 17	MeSH Cost-Benefit Analysis EXPLODE I	10851																																																																																						
# 18	MeSH Safety EXPLODE I	284																																																																																						
# 19	MeSH Health Services Accessibility EXPLODE I 2	566																																																																																						
# 20	quality	16751																																																																																						
# 21	outcome*	17295																																																																																						
# 22	performance	1511																																																																																						
# 23	safety*	3431																																																																																						
# 24	access*	2203																																																																																						
# 25	equity	60																																																																																						
# 26	effectiveness	15571																																																																																						
# 27	#14 OR #15 OR #16 OR #17 OR #18 OR #19 OR #20 OR #21 OR #22 OR #23 OR #24 OR #25 OR #26	32615																																																																																						
# 28	#13 AND #27	4251																																																																																						
# 29	#13 AND #27 RESTRICT YR 2000 2008	2988																																																																																						
Language restrictions	none																																																																																							
Number of citations	912																																																																																							

Database	Psycinfo
Host	http://ovidsp.tx.ovid.com/spa/ovidweb.cgi?T=JS&D=psych&PAGE=main
Date of search	31/12/2008
Years covered	2000-2008
Search Strategy	<ol style="list-style-type: none"> 1. exp salaries/ or exp bonuses/ or exp employee benefits/ or exp professional fees/ 2. incentives/ or exp monetary incentives/ or exp "awards (merit)"/ or exp rewards/ 3. p4p.mp. [mp=title, abstract, heading word, table of contents, key concepts] 4. p4q.mp. [mp=title, abstract, heading word, table of contents, key concepts] 5. pay*.mp. [mp=title, abstract, heading word, table of contents, key concepts] 6. incentive*.mp. [mp=title, abstract, heading word, table of contents, key concepts] 7. compensation*.mp. [mp=title, abstract, heading word, table of contents, key concepts] 8. reimbursement*.mp. [mp=title, abstract, heading word, table of contents, key concepts] 9. bonus*.mp. [mp=title, abstract, heading word, table of contents, key concepts] 10. remunerat*.mp. [mp=title, abstract, heading word, table of contents, key concepts] 11. exp Treatment Effectiveness Evaluation/ 12. exp "Quality of Care"/ 13. exp Errors/ 14. exp Treatment Outcomes/ 15. exp Quality Control/ 16. exp "Costs and Cost Analysis"/ 17. exp Safety/ 18. exp health disparities/ 19. quality.mp. [mp=title, abstract, heading word, table of contents, key concepts] 20. outcome*.mp. [mp=title, abstract, heading word, table of contents, key concepts] 21. error*.mp. [mp=title, abstract, heading word, table of contents, key concepts] 22. safety*.mp. [mp=title, abstract, heading word, table of contents, key concepts] 23. access*.mp. [mp=title, abstract, heading word, table of contents, key concepts] 24. equity.mp. [mp=title, abstract, heading word, table of contents, key concepts] 25. effectiveness.mp. [mp=title, abstract, heading word, table of contents, key concepts] 26. exp Hospitals/ 27. exp Physicians/ 28. exp Health Care Services/ 29. hospital*.mp. [mp=title, abstract, heading word, table of contents, key concepts] 30. physician*.mp. [mp=title, abstract, heading word, table of contents, key concepts] 31. practitioner*.mp. [mp=title, abstract, heading word, table of contents, key concepts] 32. healthcare.mp. [mp=title, abstract, heading word, table of contents, key concepts] 33. 6 or 3 or 7 or 9 or 2 or 8 or 1 or 4 or 10 or 5 34. 25 or 11 or 21 or 17 or 12 or 20 or 15 or 14 or 22 or 18 or 24 or 23 or 13 or 16 or 19 35. 27 or 32 or 28 or 30 or 26 or 31 or 29 36. 35 and 33 and 34 37. limit 36 to (human and abstracts and yr="2000 - 2008") 38. limit 37 to ("0800 literature review" or "0830 systematic review" or 1200 meta analysis)

Language restrictions	none
Number of citations	72

Database	Econlit
Host	http://csaweb110v.csa.com/
Date of search	31/12/2008
Years covered	2000-2008
Search Strategy	((KW=(p4q or p4p or pay*) or KW=(incentive* or compensation* or reimbursement*) or KW=(bonus* or remunerat*)) or (DE=("pay" or "salary")) or (DE=("incentive compatibility" or "bonuses"))) and ((DE=("cost effectiveness" or "quality" or "safety")) or (KW=(quality or outcome* or error*) or KW=(effectiveness or safety or access*) or KW=(disparit* or equit*))) and ((DE=("healthcare" or "hospital" or "physician")) or (KW=(hospital* or physician* or practitioner*) or KW=healthcare))
Language restrictions	none
Number of citations	224

APPENDIX 2 PRIMARY EVALUATION STUDIES

SEARCH STRATEGY IN ELECTRONIC DATABASES

Database	Medline
Host	http://www.ncbi.nlm.nih.gov/sites/entrez (Pubmed)
Date of search	31/12/2008
Years covered	2005-2008. The most recent, Frolich et al (2007) searched till June 2005. Custers et al (2008) and Schatz (2008) have some methodological drawbacks.
Search Strategy	("Salaries and Fringe Benefits"[Majr] OR "Reimbursement, Incentive"[Majr] OR "Fees and Charges"[Majr] OR p4q OR p4p OR pay* OR incentive* OR bonus*) AND ("Treatment Outcome"[Majr] OR "Medical Errors"[Majr] OR "Quality Control"[Majr] OR "Cost-Benefit Analysis"[Majr] OR "Safety"[Majr] OR "Health Services Accessibility"[Majr] OR quality OR outcome* OR performance OR error* OR safety* OR access* OR equity OR effectiveness) AND ("Hospitals"[Majr] OR "Physicians"[Majr] OR hospital* OR physician* OR practitioner*) AND (hasabstract[text] AND ("2005/06/01"[EDat]:"2008/12/30"[EDat]) AND (Humans[Mesh]) AND (Clinical Trial[ptyp] OR Randomized Controlled Trial[ptyp] OR Case Reports[ptyp] OR Clinical Trial, Phase I[ptyp] OR Clinical Trial, Phase II[ptyp] OR Clinical Trial, Phase III[ptyp] OR Clinical Trial, Phase IV[ptyp] OR Comparative Study[ptyp] OR Controlled Clinical Trial[ptyp] OR Evaluation Studies[ptyp] OR Technical Report[ptyp] OR Validation Studies[ptyp]))
Language restrictions	none
Number of citations	491

Database	Embase
Host	http://www.embase.com/home
Date of search	31/12/2008
Years covered	2004-2009. Embase was previously searched in three systematic reviews: Chaix-Couturier et al, 2000 (broad, till 1999), Giuffrida et al, 2000 (target payments, till 1997) and Stone et al, 2002 (prevention, till 1999). The last five years are selected as a standard approach.
Search Strategy	('salary and fringe benefit'/exp/mj OR 'reimbursement'/exp/mj OR 'fee'/exp/mj OR p4q OR p4p OR pay* OR incentive* OR bonus*) AND ('clinical effectiveness'/exp/mj OR 'medical error'/exp/mj OR 'treatment outcome'/exp/mj OR 'quality control'/exp/mj OR 'cost effectiveness analysis'/exp/mj OR 'patient safety'/exp/mj OR 'health care access'/exp/mj OR quality OR outcome* OR 'performance'/exp/mj OR error* OR safety* OR access* OR equity OR 'cost effectiveness'/exp/mj OR effectiveness) AND ('hospital'/exp/mj OR 'physician'/exp/mj OR hospital* OR physician* OR practitioner*) AND [humans]/lim AND [abstracts]/lim AND [embase]/lim AND [article]/lim AND [2004-2009]/py
Language restrictions	none
Number of citations	1555

Database	Web of science
Host	http://apps.isiknowledge.com/
Date of search	31/12/2008
Years covered	2004-2008. This database was most recently partially covered by Rosenthal & Frank (2006), who searched till 2003.
Search Strategy	TS=(salary and fringe benefit OR reimbursement OR fee OR p4q OR p4p OR pay* OR incentive* OR compensation* OR reimbursement* OR financ* OR bonus* OR remunerat*) AND TS=(clinical effectiveness OR medical error OR treatment outcome OR quality control OR cost effectiveness analysis OR patient safety OR health care access OR quality OR outcome* OR performance OR error* OR safety* OR access* OR equity OR cost effectiveness OR effectiveness) AND TS=(hospital OR physician OR hospital* OR physician* OR practitioner*)
Language restrictions	none
Number of citations	1661

Database	Cochrane Library	
Host	http://www3.interscience.wiley.com/cgi-bin/mrwhome/106568753/HOME?CRETRY=1&SRETRY=0	
Date of search	31/12/2008	
Years covered	2005-2008. Frolich et al (2007) searched till June 2005.	
Search Strategy	#1 MeSH descriptor Salaries and Fringe Benefits explode all trees	530
	#2 MeSH descriptor Reimbursement, Incentive explode all trees	41
	#3 MeSH descriptor Fees and Charges explode all trees	1020
	#4 p4q	0
	#5 p4p	1
	#6 pay*	4658
	#7 incentive*	1075
	#8 bonus*	49
	#9 MeSH descriptor Treatment Outcome explode all trees	54602
	#10 MeSH descriptor Medical Errors explode all trees	1657
	#11 MeSH descriptor Quality Control explode all trees	360
	#12 MeSH descriptor Cost-Benefit Analysis explode all trees	13414
	#13 MeSH descriptor Safety explode all trees	2676
	#14 MeSH descriptor Health Services Accessibility explode all trees	884
	#15 quality	46778
	#16 outcome*	116160
	#17 performance	28230
	#18 error*	9569

	#19 safety*	39786
	#20 access*	379154
	#21 equity	165
	#22 cost effectiveness	17832
	#23 effectiveness	44156
	(#9 OR #10 OR #11 OR #12 OR #13 OR #14 OR #15 #24 OR #16 OR #17 OR #18 OR #19 OR #20 OR #21 OR #22 OR #23)	436796
	#25 MeSH descriptor Hospitals explode all trees	2738
	#26 MeSH descriptor Physicians explode all trees	799
	#27 hospital*	118973
	#28 physician*	16578
	#29 practitioner*	5943
	#30 (#25 OR #26 OR #27 OR #28 OR #29)	131648
	#31 (#1 OR #2 OR #3 OR #4 OR #5 OR #6 OR #7 OR #8)	6983
	#32 (#24 AND #30 AND #31), from 2005 to 2008	1028
	Further exclusion of economical evaluation database.	
Language restrictions	none	
Number of citations	217	

Database	Psycinfo
Host	http://ovidsp.tx.ovid.com/spa/ovidweb.cgi?T=J&D=psych&PAGE=main
Date of search	31/12/2008
Years covered	2004-2008. This database was most recently covered by Rosenthal & Frank (2006), who searched till 2003.
Search Strategy	<ol style="list-style-type: none"> 1. exp salaries/ or exp bonuses/ or exp employee benefits/ or exp professional fees/ 2. incentives/ or exp monetary incentives/ or exp "awards (merit)"/ or exp rewards/ 3. p4p.mp. [mp=title, abstract, heading word, table of contents, key concepts] 4. p4q.mp. [mp=title, abstract, heading word, table of contents, key concepts] 5. pay*.mp. [mp=title, abstract, heading word, table of contents, key concepts] 6. incentive*.mp. [mp=title, abstract, heading word, table of contents, key concepts] 7. bonus*.mp. [mp=title, abstract, heading word, table of contents, key concepts] 8. exp Treatment Effectiveness Evaluation/ 9. exp "Quality of Care"/ 10. exp Errors/ 11. exp Treatment Outcomes/ 12. exp Quality Control/ 13. exp "Costs and Cost Analysis"/ 14. exp Safety/ 15. exp health disparities/ 16. quality.mp. [mp=title, abstract, heading word, table of contents, key concepts] 17. outcome*.mp. [mp=title, abstract, heading word, table of contents, key

	<p>concepts]</p> <p>18. error*.mp. [mp=title, abstract, heading word, table of contents, key concepts]</p> <p>19. safety*.mp. [mp=title, abstract, heading word, table of contents, key concepts]</p> <p>20. access*.mp. [mp=title, abstract, heading word, table of contents, key concepts]</p> <p>21. equity.mp. [mp=title, abstract, heading word, table of contents, key concepts]</p> <p>22. effectiveness.mp. [mp=title, abstract, heading word, table of contents, key concepts]</p> <p>23. exp Hospitals/</p> <p>24. exp Physicians/</p> <p>25. exp Health Care Services/</p> <p>26. hospital*.mp. [mp=title, abstract, heading word, table of contents, key concepts]</p> <p>27. physician*.mp. [mp=title, abstract, heading word, table of contents, key concepts]</p> <p>28. practitioner*.mp. [mp=title, abstract, heading word, table of contents, key concepts]</p> <p>29. healthcare.mp. [mp=title, abstract, heading word, table of contents, key concepts]</p> <p>30. 22 or 8 or 18 or 14 or 9 or 17 or 12 or 11 or 19 or 15 or 21 or 20 or 10 or 13 or 16</p> <p>31. 24 or 29 or 25 or 27 or 23 or 28 or 26</p> <p>32. 6 or 4 or 1 or 3 or 7 or 2 or 5</p> <p>33. 32 and 30 and 31</p> <p>34. limit 33 to (human and abstracts and "0110 peer-reviewed journal" and journal article and yr="2004 - 2008")</p>
Language restrictions	none
Number of citations	612

Database	Econlit
Host	http://csaweb110v.csa.com/
Date of search	31/12/2008
Years covered	2004-2008. This database was most recently covered by Rosenthal & Frank (2006), who searched till 2003.
Search Strategy	((KW=(p4q or p4p or pay*) or KW=(incentive* or compensation* or reimbursement*) or KW=(bonus* or remunerat*)) or(DE=("pay" or "salary")) or(DE=("incentive compatibility" or "bonuses"))) and((DE=("cost effectiveness" or "quality" or "safety")) or(KW=(quality or outcome* or error*) or KW=(effectiveness or safety or access*) or KW=(disparit* or equit*))) and((DE=("healthcare" or "hospital" or "physician")) or(KW=(hospital* or physician* or practitioner*) or KW=healthcare))
Language restrictions	none
Number of citations	213

APPENDIX 3 SYSTEMATIC REVIEWS

RELEVANCE SCREENING

Table 1: Rationale for exclusion based on full text review

Citation	Patient/ population	Intervention	Outcome	Design
Chen & Feldman, 2000		no explicit fin incentive for quality		
Gosden et al, 2000		only implicit incentives in payments systems		
Baker, 2002		only implicit incentives in payment systems	only cost and productivity related managed care measures, no other quality measures	
Armour & Pitts, 2003			costs and productivity	
Havranek et al, 2003				no review
Yavroff et al, 2003		no financial incentive		
Baily, 2004		only implicit incentives in payment systems		
Borenstein et al, 2004				no review
Mojica et al, 2004		Financial incentives aimed at patients		
Arnold & Straus, 2005		about reimbursement (yes/no) as such		
Harris et al, 2005		about patient compensation status		
Shortell et al, 2005		Financing and incentives for quality are addressed, but not related as an intervention. Remaining two factors next to each other.		
Veloski et al, 2006		about feedback		
Yen et al, 2006				reviews only reviews
Khunti et al, 2007				Reviews two periods for comparison as in a primary study set up

Citation	Patient/ population	Intervention	Outcome	Design
Sood et al, 2007			only quantity focused, without an evidence base	
Akbari et al, 2008		only implicit incentives in payment systems		
Chan et al, 2008	pharmacy only	no explicit quality goal		
Glickman et al, 2008		specific indicator use in P4Q		
Kaestner & Guardado, 2008		about reimbursement level as such		
Lu et al, 2008		modification of patient copayments		
Raftery et al, 2008			aimed at research participation	
Welton et al, 2008		no explicit fin incentive for quality goal		no review

APPENDIX 4 PRIMARY EVALUATION STUDIES

RELEVANCE SCREENING

Table 2: Rationale for exclusion based on full text review

Citation	Patient/ population	Intervention	Outcome	Design
Hemenway et al, 1995		No P4Q evaluation		
Shen, 2003	Mental health			
Dixon et al, 2004		Managed care, no P4Q		
Eggleston & Hsieh, 2004		Cost sharing		
Feldman et al, 2004	Mental health			Panel discussion
Forsberg et al, 2004			Only efficiency focused	
Laurence et al, 2004		No P4Q		
Mullen, 2004		Indicator use as such		
St Jacques et al, 2004			Only productivity focused	
Zivin & Pfaff, 2004		No P4Q		
Beersen et al, 2005		No P4Q		
Feldman et al, 2005	Mental health			
Gandjour & Lauterbach, 2005		No P4Q		
Grembowski et al, 2005		Managed care		
Jack, 2005		Cost sharing		
Koffman et al, 2005		Patient incentives		
Mentari et al, 2005		Reimbursement elevation as such		
Saitto et al, 2005		Prospective payment		
Schneider et al, 2005	Mental health			No evaluation
Shepard et al, 2005	Mental health			

Citation	Patient/ population	Intervention	Outcome	Design
Spertus et al, 2005		No P4Q		
Bachman, 2006	Behavioral health			No evaluation
Bloche, 2006		Tax exemption, no P4Q		
Ginsburg, 2006		No P4Q		
Paleologou et al, 2006		Development of HRM attitude measurement instrument, including wages		
Thomas et al, 2006	Mental health			
Biai et al, 2007		Cost reimbursement of additional administration		
Blue ribbon panel, 2007				No evaluation
Cotter, 2007		No P4Q		
Harrison et al, 2007		Cost quality combination, no P4Q		
Huddle, 2007		No P4Q		
Ittner et al, 2007			Only productivity focused	
Pelonero et al, 2007	Mental health			
Pronovost et al, 2007		No P4Q		
Bottle et al, 2008		No P4Q		
Boucai & Zonszein, 2008		No P4Q		
Chung et al, 2008		No P4Q		
Greene et al, 2008		Not EBP based	Cost reduction, variability reduction	
Quimbo et al, 2008		Payment vs no payment, no target or intervention handled		
Sinsky et al, 2008		How to present measures in guidelines		
Tuerk et al, 2008		No P4Q		
Wright et al, 2008		No P4Q		

APPENDIX 5 SYSTEMATIC REVIEWS

QUALITY APPRAISAL

Citation	Research question (PICOD)	Search strategy			Relevance selection	Quality appraisal	Data extraction	Studies description	Heterogeneity and pooling	Validity rating
		Databases	Entry terms	Period						
Sarnoff & Rundall, 1998	Clear (O = prevention)	Medline only, English only	6 terms, unclear string	1975-1997	clear	unclear	unclear	clear	clear	Insuf-ficient
Achat et al, 1999	Clear (O = prevention)	Medline only, English only	2 terms	1966-1998	unclear	unclear	unclear	clear	NA	Insuf-ficient
Buchan et al, 2000	clear	broad (7), English only	11 terms, but unclear string	1989-1999	clear	unclear	unclear	unclear	NA	Insuf-ficient
Chaix-Couturier et al, 2000	clear (broad fin incentives)	broad (6), Engl, French	very broad	1993-1999	clear	clear (EPOC)	clear	clear	NA	sufficient
Giuffrida et al, 2000	clear (target payment)	broad (7), plus grey	strong string	x-1997	clear	clear (EPOC)	clear	clear	NA	sufficient
Armour et al, 2001	clear	refers to Cochrane collaboration handbook					clear	clear	NA	sufficient
Heffner, 2001	clear	narrative, unsystematic review, no criteria fulfilled								Insuf-ficient
Flynn et al, 2002	clear	no methods specified								Insuf-ficient
Stone et al, 2002	clear (O = immunization)	three	unclear	x-1999	clear	clear	clear	clear	meta regression, ok	sufficient
Grol & Grimshaw, 2003	clear	listing of other reviews, no methods specified								Insuf-ficient
Vittorio et al, 2003	clear	two, plus grey, English only	22 terms, but unclear string	1995-2002	unclear	unclear	unclear	unclear	NA	Insuf-ficient
Alper, 2004	clear	narrative, unsystematic review, no criteria fulfilled								Insuf-ficient
Conrad & Christianson, 2004	clear	no methods specified								Insuf-ficient
Czubak et al, 2004	clear	no methods specified								Insuf-ficient
Dudley et al,	clear (RCT's)	two, grey,	very broad	1980-2003	clear	clear	clear	clear	NA	sufficient

Citation	Research question (PICOD)	Search strategy			Relevance selection	Quality appraisal	Data extraction	Studies description	Heterogeneity and pooling	Validity rating
		Databases	Entry terms	Period						
2004		ongoing								
Kane et al, 2004	clear (O = prevention)	five	broad	1966-2002	clear	clear	clear	clear	NA	sufficient
Shortell, 2004	clear	narrative, unsystematic review, no criteria fulfilled								Insuf-ficient
Smellie & Roy, 2005		no methods specified								Insuf-ficient
Sturm et al, 2007	clear (P = prescribers)	fifteen	strong string	x-2003/4/5	clear	clear	clear	clear	NA	sufficient
Town et al, 2005	clear (O = prevention, RCT's)	four, 17nglish only	broad	1966-2002	clear	clear	clear	clear	NA	sufficient
Fenter & Lewis, 2006		no methods specified								Insuf-ficient
Freed & Uren, 2006	clear	no methods specified								Insuf-ficient
McNamara, 2006	clear	narrative, unsystematic review, no criteria fulfilled								Insuf-ficient
Petersen et al, 2006	clear	medline only, 17nglish only	broad	1980- nov 2005	clear	clear	clear	clear	NA	sufficient
Pink et al, 2006	clear	narrative, unsystematic review, no criteria fulfilled								Insuf-ficient
Rosenthal & Frank, 2006	clear	five	narrow P4Q terms	x-2003	clear	clear	clear	clear	NA	sufficient
Chien et al, 2007	clear	narrative, unsystematic review, no criteria fulfilled								Insuf-ficient
Doran & Fullwood, 2007	clear	narrative, unsystematic review, no criteria fulfilled								Insuf-ficient
Ensor & Weinzierl, 2007	Clear (p = low income countries)	Medline, econlit, grey	unclear	unclear	unclear	unclear	clear	clear	NA	Insuf-ficient
Frolich et al, 2007	clear (RCT's)	medline, cochrane	broad	1980- june 2005	clear	clear	clear	clear	NA	sufficient
Gonzalez et al, 2007	clear	no methods specified								Insuf-ficient
Hartig & Allison, 2007	clear	narrative, unsystematic review, no criteria fulfilled								Insuf-ficient

Citation	Research question (PICOD)	Search strategy			Relevance selection	Quality appraisal	Data extraction	Studies description	Heterogeneity and pooling	Validity rating
		Databases	Entry terms	Period						
Pierce et al, 2007	clear	medline only	narrow P4Q terms	x- nov 2005	unclear	unclear	unclear	unclear	NA	Insuf-ficient
Scott, 2007	clear	narrative, unsystematic review, no criteria fulfilled								Insuf-ficient
Sikka, 2007	clear	narrative, unsystematic review, no criteria fulfilled								Insuf-ficient
Varela, 2007	clear	narrative, unsystematic review, no criteria fulfilled								Insuf-ficient
Chopra et al, 2008	clear (SYST REV ONLY))	four, English only	unclear	1979- sept 2006	clear	clear	clear	clear	NA	Insuf-ficient
Christianson et al, 2008	clear	five	broad	x- june 2007	clear	unclear	clear	clear	NA	Sufficient
Curry et al, 2008	clear	narrative, unsystematic review, no criteria fulfilled								Insuf-ficient
Custers et al, 2008	clear	medline, proquest, grey, English only	broad	1995- 1 may 2006	clear	unclear	unclear	clear	NA	sufficient
Hamblin, 2008	clear	No methods specified								Insuf-ficient
Hart-Hester et al, 2008	clear	No methods specified								Insuf-ficient
Mason, 2008	clear	nine	unclear	1992- march 2006	clear	unclear	clear	clear	NA	Insuf-ficient
Mcdonald et al, 2008	clear (P = aust, engl, new-zealand, O = access)	four	unclear	1995- mid 2007	unclear	Not done and no reporting by design or quality subgroup	unclear	unclear	NA	Insuf-ficient
Sabatino et al, 2008	Clear (O = prevention)	five	clear	X – 2004	clear	clear	clear	clear	NA	sufficient
Schatz, 2008	clear (ambulatory)	medline only	very narrow terms	2006-2007	clear	unclear	clear	clear	NA	sufficient
Thomas & Rosenthal, 2008	clear	Medline and grey	Very narrow terms	unclear	clear	clear	unclear	unclear	NA	Insuf-ficient
Sloan & Kasper, 2008	clear	narrative, unsystematic review, no criteria fulfilled								Insuf-ficient

	Research question	Patient population and setting	Intervention	Comparison	Outcome	Design	Sample size	Statistics	Generalizability	Confounders addressed	Randomization	Blinding	Clustering effect	Nr. data points	Total score
Hillman et al, 1999													0	0	12
Hopkins, 1999				0	-			-	-		0	0	0	0	3
Lebaron et al, 1999								-			0	0	0	0	8
Schauffler et al, 1999	Purely descriptive study of % targets attained, no statistical analysis present.										0	0	0	0	Over-rule
Woodson, 1999	Two descriptive case studies, only narrative results reporting.										0	0	0	0	Over-rule
Safran et al, 2000				0							0	0	0	0	9
Sussman et al, 2002	No reported P4Q description and outcomes, which is one of the intervention components										0	0	0	0	Over-rule
Cattaneo et al, 2001				0							0	0	0	0	9
Coleman et al, 2001 ¹								-		-		0	-	0	6
Fairbrother et al, 2001												0		0	12
Shortell et al, 2001				0							0	0	0	0	9
Ashworth et al, 2002b	primary outcome results (effect on prescribing) not shown, no figures, only narrative										0	0	0	0	Over-rule
Ashworth et al, 2002a	descriptive survey design, without any result related relationship testing										0	0	0	0	Over-rule
Bond et al, 2002				0							0	0	0	0	9
Amundson et al, 2003				0							0	0	0	0	9

	Research question	Patient population and setting	Intervention	Comparison	Outcome	Design	Sample size	Statistics	Generalisability	Confounders addressed	Randomization	Blinding	Clustering effect	Nr. data points	Total score
Casalino et al, 2003				0							0	0	0	0	9
Chung et al, 2003				0							0	0	0	0	9
Larsen et al, 2003				0							0	0	0	0	9
McMenamin et al, 2003				0							0	0	0	0	9
Roski et al, 2003												0		0	12
Amour et al, 2004				0							0	0	0	0	9
Ashworth et al, 2004	Although presented as quality – incentive effects the author acknowledges the lack of reporting this primary outcome.										0	0	0	0	Overrule
Berthiaume et al, 2004	Purely descriptive, no results testing, no outcome reporting.										0	0	0	0	Overrule
Borenstein et al, 2004	Selection of accreditation applicants as participants. Too big threat to generalisability,										0	0	0	0	Overrule
Greene et al, 2004				0							0	0	0	0	9
Hippisley-cox et al, 2004				0							0	0	0	0	9
Keating et al, 2004	Although very limited P4Q results in univariate analysis, these are not further addressed (confounders, generalisability,,,))										0	0	0	0	Overrule
Li et al, 2004				0							0	0	0	0	9
McMenamin et al, 2004				0							0	0	0	0	9
Schmittziel et al, 2004				0							0	0	0	0	9

	Research question	Patient population and setting	Intervention	Comparison	Outcome	Design	Sample size	Statistics	Generalizability	Confounders addressed	Randomization	Blinding	Clustering effect	Nr. data points	Total score
Wickizer et al, 2004	No statistical testing for time trends in fig 1 and 2, the only P4Q congruent process measures reported.										0	0	0	0	Overrule
Ashworth et al, 2005				0							0	0	0	0	9
Beaulieu & Horrigan, 2005									-		0	0	0	0	8
Harries et al, 2005	no statistical testing.										0	0	0	0	Overrule
Majeed et al, 2005	no statistical testing.										0	0	0	0	Overrule
May, 2005	descriptive, no testing.										0	0	0	0	Overrule
Pourat et al, 2005				0							0	0	0	0	9
Qual letter, 2005	narrative, descriptive, no statistical analysis										0	0	0	0	Overrule
Qual letter, 2005b	narrative, descriptive, no statistical analysis										0	0	0	0	Overrule
Rosenthal et al, 2005											0	0	0	0	10
Shenkman et al, 2005				0							0	0	0	0	9
Sperl-hillen & O'Connor, 2005	P4Q only used in last year of 10 year study, was not statistically tested.										0	0	0	0	Overrule
Bailit, 2006	descriptive case studies, no further evaluation										0	0	0	0	Overrule
Doran et al, 2006				0							0	0	0	0	9
Ettner et al, 2006				0							0			0	11

	Research question	Patient population and setting	Intervention	Comparison	Outcome	Design	Sample size	Statistics	Generalizability	Confounders addressed	Randomization	Blinding	Clustering effect	Nr. data points	Total score
Francis et al, 2006			-	0							0	0	0	0	7
Grossbart, 2006									-		0	0	0	0	8
Healy et al, 2006	descriptive Australia profile, no scientific evaluation										0	0	0	0	Over-rule
Jaiveer et al, 2006				0							0	0	0	0	8
Levin-Scherz et al, 2006				0							0	0	0	0	9
McLean et al, 2006				0							0	0	0	0	9
Pines, 2006	case studies and comment only, no evaluation										0	0	0	0	Over-rule
Ramsay et al, 2006	no statistical analysis										0	0	0	0	Over-rule
Reiter et al, 2006				0							0	0	0	0	9
Reschovsky et al, 2006				0							0	0	0	0	9
Rittenhouse & Robinson, 2006				0							0	0	0	0	9
Sigfrid et al, 2006				0							0	0	0	0	9
Simpson et al, 2006				0							0	0	0	0	9
Srilangalingam et al, 2006				0							0	0	0	0	9

	Research question	Patient population and setting	Intervention	Comparison	Outcome	Design	Sample size	Statistics	Generalizability	Confounders addressed	Randomization	Blinding	Clustering effect	Nr. data points	Total score
Strong et al, 2006				0							0	0	0	0	9
Sutton & Mclean, 2006				0							0	0	0	0	9
Trisolini et al, 2006	No statistical analysis										0	0	0	0	Over-rule
Wang et al, 2006				0							0	0	0	0	9
Whalley et al, 2006	testing only right before P4Q implementation										0	0	0	0	Over-rule
Williams et al, 2006 ²								-			0	0	0	0	8
Williams et al, 2006 ³											0	0	0	0	10
Wright et al, 2006				0							0	0	0	0	9
Ashworth et al, 2007				0							0	0	0	0	9
Ashworth et al, 2007b				0							0	0	0	0	9
Campbell et al, 2007				0							0	0	0	-	8
Carey et al, 2007	no relationship or effect testing										0	0	0	0	Over-rule
Casale et al, 2007				0							0	0	0	0	9
Coleman et al, 2007 ⁴				0							0	0	0	0	9
Coleman et al, 2007 ⁵				0							0	0	0	0	9

	Research question	Patient population and setting	Intervention	Comparison	Outcome	Design	Sample size	Statistics	Generalisability	Confounders addressed	Randomization	Blinding	Clustering effect	Nr. data points	Total score	
Downing et al, 2007				0							0	0	0	0	9	
Felt-Lisk et al, 2007	no significance testing										0	0	0	0	0	Over-rule
Gene badia et al, 2007				0							0	0	0	0	9	
Gilmore et al, 2007											0	0	0	0	10	
Glickman et al, 2007											0	0	1	0	11	
Gray et al, 2007				0							0	0	0	1	10	
Gulliford et al, 2007				0							0	0	0	0	9	
Halanych et al, 2007	no testing or adjustment for P4Q										0	0	0	0	0	Over-rule
Helm & Tortorella, 2007	some response rates below 50%										0	0	0	0	0	Over-rule
Heneghan et al, 2007	only 50% response rate										0	0	0	0	0	Over-rule
Hughes, 2007	no referencing, no testing										0	0	0	0	0	Over-rule
Kautter et al, 2007	descriptive case study findings, no further evaluation or testing										0	0	0	0	0	Over-rule
Lindenauer et al, 2007											0	0	0	0	10	
Mandel & Kotagal, 2007	no statistical testing										0	0	0	0	0	Over-rule

	Research question	Patient population and setting	Intervention	Comparison	Outcome	Design	Sample size	Statistics	Generalizability	Confounders addressed	Randomization	Blinding	Clustering effect	Nr. data points	Total score
McCarlie et al, 2007	no statistical testing										0	0	0	0	Over-rule
McLean et al, 2007a				0							0	0	0	0	9
McLean et al, 2007b				0							0	0	0	0	9
Mehrota et al, 2007				0							0	0	0	0	9
Millett et al, 2007 ⁶				0							0	0		0	10
Millett et al, 2007 ⁷				0							0	0		0	10
Millett et al, 2007 ⁸				0							0	0	0	0	9
Nalli et al, 2007	no statistical testing, significance analysis										0	0	0	0	Over-rule
O'Malley et al, 2007				0							0	0	0	0	9
Patel et al, 2007	no statistical analysis										0	0	0	0	Over-rule
Rosenthal & Camillus, 2007	descriptive, no testing, no references										0	0	0	0	Over-rule
Saxena et al, 2007				0							0	0	0	0	9
Shohet et al, 2007				0							0	0	0	0	9
Simon et al, 2007				0							0	0	0	0	9
Simpson et al, 2007				0							0	0	0	0	9

	Research question	Patient population and setting	Intervention	Comparison	Outcome	Design	Sample size	Statistics	Generalizability	Confounders addressed	Randomization	Blinding	Clustering effect	Nr. data points	Total score
Smith, 2007	no statistical analysis										0	0	0	0	Over-rule
Steel et al, 2007				0							0	0	0	0	9
Tahrani et al, 2007				0							0	0	0	0	9
Ting et al, 2007	no statistical analysis										0	0	0	0	Over-rule
Twardella & Brenner, 2007												0	0	0	11
Young et al, 2007				0							0	0	0	1	10
An et al, 2008 ⁹												0	0	0	11
Ashworth et al, 2008				0							0	0	0	0	9
Bhattacharyya et al, 2008				0							0	0	0	0	9
Chang et al, 2008	too high non response, attrition (> 50%)										0	0	0	0	Over-rule
Cupples et al, 2008											0	0	1	0	11
Doran et al, 2008a				0							0	0	0	0	9
Doran et al, 2008b				0							0	0	0	0	9
Fleetcroft et al, 2008a	no significance testing or relationship analysis										0	0	0	0	Over-rule

	Research question	Patient population and setting	Intervention	Comparison	Outcome	Design	Sample size	Statistics	Generalizability	Confounders addressed	Randomization	Blinding	Clustering effect	Nr. data points	Total score
Gravelle et al, 2008											0	0	0	0	10
Greenberg et al, 2008				0							0	0	0	0	9
Gross et al, 2008	no statistical analysis										0	0	0	0	Over-rule
Herrin et al, 2008											0	0		0	11
Karve et al, 2008				0							0	0	0	0	9
Kirschner et al, 2008	too small sample (n=11 practices), no significance testing										0	0	0	0	Over-rule
McBride-Stewart et al, 2008				0							0	0	0		10
McGovern et al, 2008 ¹⁰				0							0	0		0	10
McGovern et al, 2008 ¹¹				0							0	0		0	10
Millett et al, 2008 ¹²				0							0	0		0	10
Millett et al, 2008 ¹³				0							0	0		0	10
Millett et al, 2008 ¹⁴				0							0	0	0	0	9
Pearson et al, 2008									-		0	0	0	0	8
Rosenthal et al, 2008											0	0	0	0	10
Steel et al, 2008				0							0	0	0	0	9

	Research question	Patient population and setting	Intervention	Comparison	Outcome	Design	Sample size	Statistics	Generalizability	Confounders addressed	Randomization	Blinding	Clustering effect	Nr. data points	Total score
Tahrani et al, 2008				0							0	0	0	0	9
Vaghela et al, 2008				0							0	0	0	0	9
Weber et al, 2008				0							0	0	0	0	9
Yao et al, 2008	unclear and small sampling, discordance between text and tables in results interpretation										0	0	0	0	Over-rule

APPENDIX 7 COST-EFFECTIVENESS AND MODELING STUDIES

QUALITY APPRAISAL

Citation	according to farmaco-economic guidelines KCE: guideline 10: Modelling + ISPOR guidelines						
	model design (as simple as possible)	assumptions tested in sensitivity analysis/scenario analysis	original data set provided	sources used presented and described in detail (high quality)	scenarios (for models that extrapolate to longer time periods)	calibration (results should be logically consistent with real-life observations and data)	face validity (the results of the model should be intuitively correct) and cross-validation (transparent enough to allow an explanation of the differences with other models for the same intervention)
Kahn et al, 2006	+/-	-	+	+	NA	+	+
Fleetcroft & Cookson, 2006	+	-	+	+	-	+	+
McElduff et al, 2004	+	+	+	+	-	+ (if proposed targets are achieved)	+
Averill et al, 2006	+	+/-	+/-	+/-	NA	+	+
fleetcroft et al, 2008	+	+	+/-	+	-	+	+

Citation	according to farmaco-economic guidelines KCE													
	literature review	perspective of the evaluation	target population	comparator	analytic technique	study design	calculation of costs	valuation of outcomes	data source	incremental cost-effectiveness	time horizon	modelling	handling uncertainty	discount rate
Nahra et al, 2006	+/-	+	+	NA (new method already in use)	cost-utility	+	+/- (kost behandelings niet opgenomen)	+	+	NA (old method no costs)	NA	+	+	+ (5%)
Curtin et al, 2006	-	+	+	+	cost-benefit	+/-	+	+	+	NA (old method no costs)	NA	NA	NA	NA
Mason et al, 2008	+/-	+	+	+/-	cost-utility	+	+/-	+	+/-	NA (old method no costs)	NA	+	+	+/-

APPENDIX 8 SYSTEMATIC REVIEWS

CITATIONS INCLUDED IN FULL TEXT ANALYSIS

- 1 Akbari, A., Mayhew, A., Al Alawi, M. A., Grimshaw, J., Winkens, R., Glidewell, E. et al. (2008). Interventions to improve outpatient referrals from primary care to secondary care. *Cochrane.Database.Syst.Rev.*, CD005471.
- 2 Alper, P. R. (2004). Pitfalls of financial incentives for improving diabetes outcomes. *Diabetes trends*, 16S, 36-40.
- 3 Armour, B. S., Pitts, M. M., Maclean, R., Cangialose, C., Kishel, M., Imai, H. et al. (2001). The effect of explicit financial incentives on physician behavior. *Arch.Intern.Med.*, 161, 1261-1266.
- 4 Armour, B. S. & Pitts, M. M. (2003). Physician financial incentives in managed care - Resource use, quality and cost implications. *Disease Management & Health Outcomes*, 11, 139-147.
- 5 Arnold, S. R. & Straus, S. E. (2005). Interventions to improve antibiotic prescribing practices in ambulatory care. *Cochrane.Database.Syst.Rev.*, CD003539.
- 6 Baily, M. A. (2004). Ethics, economics, and physician reimbursement. *Mt.Sinai J Med.*, 71, 231-235.
- 7 Baker, L. C. (2003). Managed care spillover effects. *Annual Review of Public Health*, 24, 435-456.
- 8 Berven, S., Smith, A., Bozic, K., & Bradford, D. S. (2007). Pay-for-performance: considerations in application to the management of spinal disorders. *Spine*, 32, S33-S38.
- 9 Bordley, W. C., Chelminski, A., Margolis, P. A., Kraus, R., Szilagyi, P. G., & Vann, J. J. (2000). The effect of audit and feedback on immunization delivery: a systematic review. *Am.J.Prev.Med.*, 18, 343-350.
- 10 Borenstein, J., Badamgarav, E., Henning, J. M., Gano, A. D., Jr., & Weingarten, S. R. (2004). The association between quality improvement activities performed by managed care organizations and quality of care. *Am.J Med.*, 117, 297-304.
- 11 Buchan, J., Thompson, M., & O'May, F. (2000). *Issues in health service delivery. Discussion Paper Nr. 4. Incentive and remuneration strategies. Health workforce incentives and remuneration strategies. A research review.* (Rep. No. WHO/EIP/OSD/00.14). Evidence and information for policy. Department of organization of healthservices delivery. World Health Organization. Geneva..
- 12 Casalino, L. P. (2003). Markets and medicine: barriers to creating a "business case for quality". *Perspect.Biol.Med.*, 46, 38-51.
- 13 Chaillet, N., Dube, E., Dugas, M., Audibert, F., Tourigny, C., Fraser, W. D. et al. (2006). Evidence-based strategies for implementing guidelines in obstetrics: a systematic review. *Obstet.Gynecol.*, 108, 1234-1245.
- 14 Chaix-Couturier, C., Durand-Zaleski, I., Jolly, D., & Durieux, P. (2000). Effects of financial incentives on medical practice: results from a systematic review of the literature and methodological issues. *Int J Qual.Health Care*, 12, 133-142.
- 15 Chan, P., Grindrod, K., Bougher, D., Passuto, F. M., Wilgosh, C., Eberhart, G. et al. (2008). A systematic review of remuneration systems for clinical pharmacy care services. *Can.Pharm.J.*, 141, 102-112.
- 16 Chauhan, D. & Mason, A. (2008). Factors affecting the uptake of new medicines in secondary care - a literature review. *Journal of Clinical Pharmacy and Therapeutics*, 33, 339-348.
- 17 Chen, G. J. & Feldman, S. R. (2000). Economic aspect of health care systems. Advantage and disadvantage incentives in different systems. *Dermatol.Clin.*, 18, 211-214.
- 18 Chien, A. T., Conti, R. M., & Pollack, H. A. (2007). A pediatric-focused review of the performance incentive literature. *Curr.Opin.Pediatr.*, 19, 719-725.
- 19 Chopra, M., Munro, S., Lavis, J. N., Vist, G., & Bennett, S. (2008). Effects of policy options for human resources for health: an analysis of systematic reviews. *Lancet*, 371, 668-674.
- 20 Conrad, D. A. & Christianson, J. B. (2004). Penetrating the "black box": financial incentives for enhancing the quality of physician services. *Med.Care Res.Rev.*, 61, 37S-68S.

- 21 Curry, S. J., Keller, P. A., Orleans, C. T., & Fiore, M. C. (2008). The role of health care systems in increased tobacco cessation. *Annu.Rev.Public Health*, 29, 411-428.
- 22 Custers, T., Hurley, J., Klazinga, N. S., & Brown, A. D. (2008). Selecting effective incentive structures in health care: A decision framework to support health care purchasers in finding the right incentives to drive performance. *BMC.Health Serv.Res.*, 8, 66.
- 23 Czubak, R., Tucker, J., & Zarowitz, B. J. (2004). Optimizing drug prescribing in managed care populations - Improving clinical and economic outcomes. *Disease Management & Health Outcomes*, 12, 147-167.
- 24 de Brantes, F., Wickland, P. S., & Williams, J. P. (2008). The value of ambulatory care measures: A review of clinical and financial impact from an employer/payer perspective. *American Journal of Managed Care*, 14, 360-368.
- 25 Doran, T. & Fullwood, C. (2007). Pay for performance: is it the best way to improve control of hypertension? *Curr.Hypertens.Rep.*, 9, 360-367.
- 26 Dudley, R. A., Frolich, A., Robinowitz, D. L., Talavera, J. A., Broadhead, P., & Luft, H. S. (2004). *To support quality-based purchasing: A review of the evidence* (Rep. No. Technical review 10. AHRQ Publication No. 04-0057.). Rockville, MD: Prepared by the Stanford-University of California San Francisco Evidence-based Practice Center under Contract No. 290-02-0017.
- 27 Ensor, T. & Weinzierl, S. (2007). Regulating health care in low- and middle-income countries: Broadening the policy response in resource constrained environments. *Social Science & Medicine*, 65, 355-366.
- 28 Faulkner, A., Mills, N., Bainton, D., Baxter, K., Kinnersley, P., Peters, T. J. et al. (2003). A systematic review of the effect of primary care-based service innovations on quality and patterns of referral to specialist secondary care. *Br.J Gen.Pract.*, 53, 878-884.
- 29 Fenter, T. C. & Lewis, S. J. (2008). Pay-for-performance initiatives. *Journal of Managed Care Pharmacy*, 14, S12-S15.
- 30 Figueras, J., Robinson, R., & Jakubowski, E. (2005). *Purchasing to improve health systems performance* Open University Press.
- 31 Flynn, K. E., Smith, M. A., & Davis, M. K. (2002). From physician to consumer: The effectiveness of strategies to manage health care utilization. *Medical Care Research and Review*, 59, 455-481.
- 32 Freed, G. L. & Uren, R. L. (2006). Pay-for-performance: an overview for pediatrics. *J Pediatr.*, 149, 120-124.
- 33 Frolich, A., Talavera, J. A., Broadhead, P., & Dudley, R. A. (2007). A behavioral model of clinician responses to incentives to improve quality. *Health Policy*, 80, 179-193.
- 34 Garcia-Prado, A. (2005). *Sweetening the carrot. Motivating public physicians for better performance* (Rep. No. 3772).
- 35 Giuffrida, A., Gosden, T., Forland, F., Kristiansen, I. S., Sergison, M., Leese, B. et al. (2000). Target payments in primary care: effects on professional practice and health care outcomes. *Cochrane.Database.Syst.Rev.*, CD000531.
- 36 Glickman, S. W., Schulman, K. A., Peterson, E. D., Hocker, M. B., & Cairns, C. B. (2008). Evidence-based perspectives on pay for performance and quality of patient care and outcomes in emergency medicine. *Ann.Emerg.Med.*, 51, 622-631.
- 37 Gonzalez, C. M., Penson, D., Kosiak, B., Dupree, J., & Clemens, J. Q. (2007). Pay for performance: rationale and potential implications for urology. *J Urol.*, 178, 402-408.
- 38 Gosden, T., Forland, F., Kristiansen, I. S., Sutton, M., Leese, B., Giuffrida, A. et al. (2000). Capitation, salary, fee-for-service and mixed systems of payment: effects on the behaviour of primary care physicians. *Cochrane.Database.Syst.Rev.*, CD002215.
- 39 Gosden, T., Forland, F., Kristiansen, I. S., Sutton, M., Leese, B., Giuffrida, A. et al. (2001). Impact of payment method on behaviour of primary care physicians: a systematic review. *J Health Serv.Res.Policy*, 6, 44-55.

- 40 Greene, S. E. & Nash, D. B. (2008). Pay for Performance: An Overview of the Literature. *Am.J Med.Qual.*
- 41 Grol, R. & Grimshaw, J. (2003). From best evidence to best practice: effective implementation of change in patients' care. *Lancet*, 362, 1225-1230.
- 42 Grol, R. P., Bosch, M. C., Hulscher, M. E., Eccles, M. P., & Wensing, M. (2007). Planning and studying improvement in patient care: the use of theoretical perspectives. *Milbank Q.*, 85, 93-138.
- 43 Harris, I., Mulford, J., Solomon, M., van Gelder, J. M., & Young, J. (2005). Association between compensation status and outcome after surgery: a meta-analysis. *JAMA*, 293, 1644-1652.
- 44 Hartig, J. R. & Allison, J. (2007). Physician performance improvement: an overview of methodologies. *Clin.Exp.Rheumatol.*, 25, 50-54.
- 45 Havranek, E. P., Krumholz, H. M., Dudley, R. A., Adams, K., Gregory, D., Lampert, S. et al. (2003). Aligning quality and payment for heart failure care: defining the challenges. *J Card Fail.*, 9, 251-254.
- 46 Heffner, J. E. (2001). Altering physician behavior to improve clinical performance. *Top.Health Inf.Manage.*, 22, 1-9.
- 47 Hicks, V. & Adams, O. (2000). *Issues in health services delivery. Discussion paper nr. 5. Economic and policy incentives. The effect of economic and policy incentives on provider practice.* (Rep. No. WHO/EIP/OSD/00.8). Evidence and information for policy. Department and organization of health services delivery. World Health Organization. Geneva.
- 48 Hwang, R. W. & Herndon, J. H. (2007). The business case for patient safety. *Clinical Orthopaedics and Related Research*, 21-34.
- 49 Institute Of Medicine (2007). *Rewarding provider performance. Aligning incentives in Medicare.* Washington, D.C.: the National Academies Press.
- 50 Kaestner, R. & Guardado, J. (2008). Medicare reimbursement, nurse staffing, and patient outcomes. *J Health Econ.*, 27, 339-361.
- 51 Kane, R. L., Johnson, P. E., Town, R. J., & Butler, M. (2004). Economic incentives for preventive care. *Evid.Rep.Technol.Assess.(Summ.)*, 1-7.
- 52 Khunti, K., Gadsby, R., Millett, C., Majeed, A., & Davies, M. (2007). Quality of diabetes care in the UK: comparison of published quality-of-care reports with results of the Quality and Outcomes Framework for Diabetes. *Diabetic Medicine*, 24, 1436-1441.
- 53 Lanier, D. C., Roland, M., Burstin, H., & Knottnerus, J. A. (2003). Doctor performance and public accountability. *Lancet*, 362, 1404-1408.
- 54 Lewin, S., Lavis, J. N., Oxman, A. D., Bastias, G., Chopra, M., Ciapponi, A. et al. (2008). Supporting the delivery of cost-effective interventions in primary health-care systems in low-income and middle-income countries: an overview of systematic reviews. *Lancet*, 372, 928-939.
- 55 Lu, C. Y., Ross-Degnan, D., Soumerai, S. B., & Pearson, S. A. (2008). Interventions designed to improve the quality and efficiency of medication use in managed care: A critical review of the literature - 2001-2007. *Bmc Health Services Research*, 8.
- 56 Maio, V., Goldfarb, N. I., Carter, C., & Nash, D. B. (2003). *Value-based purchasing: a review of the literature* (Rep. No. #636). The Commonwealth Fund.
- 57 Mason, A. (2008). New medicines in primary care: a review of influences on general practitioner prescribing. *Journal of Clinical Pharmacy and Therapeutics*, 33, 1-10.
- 58 McDonald, J., Harris, M. F., Cumming, J., Davies, G. P., & Burns, P. (2008). The implementation and impact of different funding initiatives on access to multidisciplinary primary health care and policy implications. *Medical Journal of Australia*, 188, S69-S72.
- 59 McNamara, P. (2006). Purchaser strategies to influence quality of care: from rhetoric to global applications. *Qual.Saf Health Care*, 15, 171-173.
- 60 Miller, D. C., Wei, J. T., Montie, J. E., & Hollenbeck, B. K. (2006). Quality of care and performance-based reimbursement: the contemporary landscape and implications for urologists. *Urology*, 67, 1117-1125.

- 61 Mojica, W. A., Suttorp, M. J., Sherman, S. E., Morton, S. C., Roth, E. A., Maglione, M. A. et al. (2004). Smoking-cessation interventions by type of provider - A meta-analysis. *American Journal of Preventive Medicine*, 26, 391-401.
- 62 Mossialos, E., Walley, T., & Rudisill, C. (2005). Provider incentives and prescribing behavior in Europe. *Expert Rev.Pharmacoeconomics Outcomes Res.*, 5, 81-93.
- 63 Petersen, L. A., Woodard, L. D., Urech, T., Daw, C., & Sookanan, S. (2006). Does pay-for-performance improve the quality of health care? *Ann.Intern.Med.*, 145, 265-272.
- 64 Pierce, R. G., Bozic, K. J., & Bradford, D. S. (2007). Pay for performance in orthopaedic surgery. *Clin.Orthop.Relat Res.*, 457, 87-95.
- 65 Pink, G. H., Brown, A. D., Studer, M. L., Reiter, K. L., & Leatt, P. (2006). Pay-for-performance in publicly financed healthcare: some international experience and considerations for Canada. *Healthc.Pap.*, 6, 8-26.
- 66 Profit, J., Zupancic, J. A., Gould, J. B., & Petersen, L. A. (2007). Implementing pay-for-performance in the neonatal intensive care unit. *Pediatrics*, 119, 975-982.
- 67 Raftery, J., Bryant, J., Powell, J., Kerr, C., & Hawker, S. (2008). Payment to healthcare professionals for patient recruitment to trials: systematic review and qualitative study. *Health Technol.Assess.*, 12, 1-128, iii.
- 68 Renders, C. M., Valk, G. D., Griffin, S., Wagner, E. H., Eijk, J. T., & Assendelft, W. J. (2001). Interventions to improve the management of diabetes mellitus in primary care, outpatient and community settings. *Cochrane.Database.Syst.Rev.*, CD001481.
- 69 Rosenthal, M. B., Fernandopulle, R., Song, H. R., & Landon, B. (2004). Paying for quality: providers' incentives for quality improvement. *Health Aff.(Millwood.)*, 23, 127-141.
- 70 Rosenthal, M. B. & Frank, R. G. (2006). What is the empirical basis for paying for quality in health care? *Med.Care Res.Rev.*, 63, 135-157.
- 71 Rosenthal, M. B. & Dudley, R. A. (2007). Pay-for-performance: will the latest payment trend improve care? *JAMA*, 297, 740-744.
- 72 Rowe, J. W. (2006). Pay-for-performance and accountability: related themes in improving health care. *Ann.Intern.Med.*, 145, 695-699.
- 73 Sabatino, S. A., Habarta, N., Baron, R. C., Coates, R. J., Rimer, B. K., Kerner, J. et al. (2008). Interventions to increase recommendation and delivery of screening for breast, cervical, and colorectal cancers by healthcare providers - Systematic reviews of provider assessment and feedback and provider incentives. *American Journal of Preventive Medicine*, 35, S67-S74.
- 74 Scalise, D. (2006). Just rewards. *Hospitals & Health Networks*, 80, 40-+.
- 75 Schatz, M. (2008). Does pay-for-performance influence the quality of care? *Curr.Opin.Allergy Clin.Immunol.*, 8, 213-221.
- 76 Scott, I. A. (2007). Pay for performance in health care: strategic issues for Australian experiments. *Med.J Aust.*, 187, 31-35.
- 77 Seidel, R. L. & Nash, D. B. (2004). Paying for performance in diagnostic imaging: current challenges and future prospects. *J Am.Coll.Radiol.*, 1, 952-956.
- 78 Shojania, K. G., Ranji, S. R., McDonald, K. M., Grimshaw, J. M., Sundaram, V., Rushakoff, R. J. et al. (2006). Effects of quality improvement strategies for type 2 diabetes on glycemic control: a meta-regression analysis. *JAMA*, 296, 427-440.
- 79 Shojania, K. G., McDonald, K. M., Wachter, R. M., & Owens, D. K. (2007). *Closing the quality gap: A critical analysis of quality improvement strategies. Vol. 6. Prevention of healthcare-associated infections* (Rep. No. Technical Review 9, AHRQ Publication No. 04(07)-0051-6). Rockville, MD: Prepared by the Stanford University-UCSF Evidence-based Practice Center under Contract No. 290-02-0017 for the Agency for Healthcare Research and Quality.
- 80 Shortell, S. M. (2004). Increasing value: a research agenda for addressing the managerial and organizational challenges facing health care delivery in the United States. *Med.Care Res.Rev.*, 61, 12S-30S.

- 81 Shortell, S. M., Schmittiel, J., Wang, M. C., Li, R., Gillies, R. R., Casalino, L. P. et al. (2005). An empirical assessment of high-performing medical groups: Results from a national study. *Medical Care Research and Review*, 62, 407-434.
- 82 Sikka, R. (2007). Pay for performance in emergency medicine. *Ann.Emerg.Med.*, 49, 756-761.
- 83 Sloan, F. A. & Kasper, H. (2008). *Incentives and choice in health care*. Cambridge, Massachusetts London, England: MIT Press.
- 84 Smellie, W. S. A. & Roy, D. V. (2005). Impact of the new General Medical Services contract on the clinical laboratory. *Annals of Clinical Biochemistry*, 42, 4-10.
- 85 Sood, R., Sood, A., & Ghosh, A. K. (2007). Non-evidence-based variables affecting physicians' test-ordering tendencies: a systematic review. *Netherlands Journal of Medicine*, 65, 167-177.
- 86 Stone, E. G., Morton, S. C., Hulscher, M. E., Maglione, M. A., Roth, E. A., Grimshaw, J. M. et al. (2002). Interventions that increase use of adult immunization and cancer screening services: a meta-analysis. *Ann.Intern.Med.*, 136, 641-651.
- 87 Sturm, H., Austvoll-Dahlgren, A., Aaserud, M., Oxman, A. D., Ramsay, C., Vernby, A. et al. (2007). Pharmaceutical policies: effects of financial incentives for prescribers. *Cochrane.Database.Syst.Rev.*, CD006731.
- 88 Swayne, L. C. (2005). Pay for performance: pay more or pay less? *J Am.Coll.Radiol.*, 2, 777-781.
- 89 Thomas, C. & Rosenthal, M. B. (2008). The medical home: Growing evidence to support a new approach to primary care. *Journal of the American Board of Family Medicine*, 21, 427-440.
- 90 Town, R., Wholey, D. R., Kralewski, J., & Dowd, B. (2004). Assessing the influence of incentives on physicians and medical groups. *Med.Care Res.Rev.*, 61, 80S-118S.
- 91 Town, R., Kane, R., Johnson, P., & Butler, M. (2005). Economic incentives and physicians' delivery of preventive care: a systematic review. *Am.J Prev.Med.*, 28, 234-240.
- 92 Unruh, L., Lugo, N. R., White, S. V., & Byers, J. F. (2005). Managed care and patient safety: risks and opportunities. *Health Care Manag.(Frederick.)*, 24, 245-256.
- 93 Varela, G. (2007). Pay for performance in thoracic surgery. *Thorac.Surg.Clin.*, 17, 431-435.
- 94 Veloski, J., Boex, J. R., Grasberger, M. J., Evans, A., & Wolfson, D. B. (2006). Systematic review of the literature on assessment, feedback and physicians' clinical performance: BEME Guide No. 7. *Medical Teacher*, 28, 117-128.
- 95 Walsh, J. M. E., McDonald, K. M., Shojania, K. G., Sundaram, V., Nayak, S., Lewis, R. et al. (2006). Quality improvement strategies for hypertension management - A systematic review. *Medical Care*, 44, 646-657.
- 96 Welton, J. M. (2008). Implications of Medicare reimbursement changes related to inpatient nursing care quality. *J Nurs.Adm*, 38, 325-330.
- 97 Woodburn, J., Branson, R., Pavoloni, G., Lin, C. C., Fritts, M., & Goertz, C. (2006). Assessment of chiropractic outcomes for low back pain and neck pain: a health plan quality incentive model. *J Healthc.Qual.*, 28, 32-39.
- 98 Yabroff, K. R., Mangan, P., & Mandelblatt, J. (2003). Effectiveness of interventions to increase Papanicolaou smear use. *J Am.Board Fam.Pract.*, 16, 188-203.
- 99 Yen, B. M. (2006). Engaging physicians to change practice. *Journal of clinical outcomes management*, 13, 103-110.

APPENDIX 9 PRIMARY EVALUATION STUDIES

CITATIONS INCLUDED IN FULL TEXT ANALYSIS

- 1 Looking at lessons on quality from the Medicare pay-for-performance hospital demonstration (2005). *Qual.Lett.Healthc.Lead.*, 17, 2-13, 1.
- 2 Aligning incentives to promote quality care: lessons from pay-for-performance initiatives (2005). *Qual.Lett.Healthc.Lead.*, 17, 2-7, 1.
- 3 Ahmann, A. J. (2007). Guidelines and performance measures for diabetes. *American Journal of Managed Care*, 13, S41-S46.
- 4 Allen, K., Hazelett, S., Jarjoura, D., Wright, K., Clough, L., & Weinhardt, J. (2004). Improving stroke outcomes: Implementation of a postdischarge care management model. *Journal of Clinical Outcomes Management* 11[11], 707-714.
Ref Type: Journal (Full)
- 5 An, L. C., Bluhm, J. H., Foldes, S. S., Alesci, N. L., Klatt, C. M., Center, B. A. et al. (2008). A randomized trial of a pay-for-performance program targeting clinician referral to a state tobacco quitline. *Archives of Internal Medicine*, 168, 1993-1999.
- 6 Anderson, J., Hackman, M., Burnich, J., & Gurgio, T. R. (2007). Determining hospital performance based on rank ordering: Is it appropriate? *American Journal of Medical Quality*, 22, 177-185.
- 7 Anderson, K. K., Sebaldt, R. J., Lohfeld, L., Burgess, K., Donald, F. C., & Kaczorowski, J. (2006). Views of family physicians in southwestern Ontario on preventive care services and performance incentives. *Family Practice*, 23, 469-471.
- 8 Angus, D. C. & Black, N. (2004). Improving care of the critically ill: institutional and health-care system approaches. *Lancet*, 363, 1314-1320.
- 9 Armour, B. S., Friedman, C., Pitts, M. M., Wike, J., Alley, L., & Etchason, J. (2004). The influence of year-end bonuses on colorectal cancer screening. *American Journal of Managed Care*, 10, 617-624.
- 10 Ashworth, M., Golding, S., & Majeed, A. (2002). Prescribing indicators and their use by primary care groups to influence prescribing. *Journal of Clinical Pharmacy and Therapeutics*, 27, 197-204.
- 11 Ashworth, M., Golding, S., Shephard, L., & Majeed, A. (2002). Prescribing incentive schemes in two NHS regions: cross sectional survey. *British Medical Journal*, 324, 1187-1188.
- 12 Ashworth, M., Lea, R., Gray, H., Rowlands, G., Gravelle, H., & Majeed, A. (2004). How are primary care organizations using financial incentives to influence prescribing? *Journal of Public Health*, 26, 48-51.
- 13 Ashworth, M., Armstrong, D., de Freitas, J., Boullier, G., Garforth, J., & Virji, A. (2005). The relationship between income and performance indicators in general practice: a cross-sectional study. *Health Serv.Manage.Res.*, 18, 258-264.
- 14 Ashworth, M., Medina, J., & Morgan, M. (2008). Effect of social deprivation on blood pressure monitoring and control in England: a survey of data from the quality and outcomes framework. *BMJ*, 337, a2030.
- 15 Atun, R. A., Menabde, N., Saluvere, K., Jesse, M., & Habicht, J. (2006). Introducing a complex health innovation - Primary health care reforms in Estonia (multimethods evaluation). *Health Policy*, 79, 79-91.
- 16 Averill, R. F., Vertrees, J. C., McCullough, E. C., Hughes, J. S., & Goldfield, N. I. (2006). Redesigning medicare inpatient PPS to adjust payment for post-admission complications. *Health Care Financing Review*, 27, 83-93.
- 17 Babbott, S. F., Boby, J., Day, S. C., Dugdale, D. C., Fihn, S. D., Kopoor, W. N. et al. (2007). Redesigning the practice model for general internal medicine. A proposal for coordinated care - A policy monograph of the society of general internal medicine. *Journal of General Internal Medicine*, 22, 400-409.
- 18 Bachman, J. (2006). Pay for performance in primary and specialty behavioral health care: Two "concept" proposals. *Professional Psychology-Research and Practice*, 37, 384-388.

- 19 Bailit Health Purchasing, L. L. C. (2006). *Incentives and rewards best practices primer: Lessons learned from early pilots* The Leapfrog Group.
- 20 Bailit Health Purchasing, L. L. C. (2007). *Value-driven health care. A purchaser guide. Version 2.0.*
- 21 Baumann, M. H. & Dellert, E. (2006). Performance measures and pay for performance. *Chest*, 129, 188-191.
- 22 Beaulieu, N. D. & Horrigan, D. R. (2005). Putting smart money to work for quality improvement. *Health Services Research*, 40, 1318-1334.
- 23 Beckman, H., Suchman, A. L., Curtin, K., & Greene, R. A. (2006). Physician reactions to quantitative individual performance reports. *American Journal of Medical Quality*, 21, 192-199.
- 24 Beersen, N., Bart de Bruijn, J. H., Dekkers, M. A., Ten Have, P., Hekster, G. B., Redekop, W. K. et al. (2004). Developing a national continuous quality improvement system for neuromodulation treatment in The Netherlands. *Jt.Comm J Qual.Saf*, 30, 310-321.
- 25 Beersen, N., Redekop, W. K., de Bruijn, J. H. B., Theuvenet, P. J., Berg, M., & Klazinga, N. S. (2005). Quality based social insurance coverage and payment of the application of a high cost medical therapy: the case of spinal cord stimulation for chronic non-oncologic pain in The Netherlands. *Health Policy*, 71, 107-115.
- 26 Beich, J., Scanlon, D. P., Ulbrecht, J., Ford, E. W., & Ibrahim, I. A. (2006). The role of disease management in pay-for-performance programs for improving the care of chronically ill patients. *Medical Care Research and Review*, 63, 96S-116S.
- 27 Bennett, N. M., Lewis, B., Doniger, A. S., Bell, K., Kouides, R., LaForce, F. M. et al. (1994). A Coordinated, Community-Wide Program in Monroe County, New-York, to Increase Influenza Immunization Rates in the Elderly. *Archives of Internal Medicine*, 154, 1741-1745.
- 28 Berthiaume, J. T., Tyler, P. A., Ng-Osorio, J., & LaBresh, K. A. (2004). Aligning financial incentives with "get with the guidelines" to improve cardiovascular care. *American Journal of Managed Care*, 10, 501-504.
- 29 Bhattacharyya, T., Mehta, P., & Freiberg, A. A. (2008). Hospital characteristics associated with success in a pay-for-performance program in orthopaedic surgery. *Journal of Bone and Joint Surgery-American Volume*, 90A, 1240-1243.
- 30 Biai, S., Rodrigues, A., Gomes, M., Ribeiro, I., Sodemann, M., Alves, F. et al. (2007). Reduced in-hospital mortality after improved management of children under 5 years admitted to hospital with malaria: randomised trial. *BMJ*, 335, 862.
- 31 Bloche, M. G. (2006). Perspective - Tax preferences for nonprofits: From per se exemption to pay-for-performance. *Health Affairs*, 25, W304-W307.
- 32 Bokhour, B. G., Burgess, J. F., Hook, J. M., White, B., Berlowitz, D., Guldin, M. R. et al. (2006). Incentive implementation in physician practices: A qualitative study of practice executive perspectives on pay for performance. *Medical Care Research and Review*, 63, 73S-95S.
- 33 Bond, L., Davie, G., Carlin, J. B., Lester, R., & Nolan, T. (2002). Increases in vaccination coverage for children in child care, 1997 to 2000: an evaluation of the impact of government incentives and initiatives. *Australian and New Zealand Journal of Public Health*, 26, 58-64.
- 34 Borenstein, J., Badamgarav, E., Henning, J. M., Gano, A. D., & Weingarten, S. R. (2004). The association between quality improvement activities performed by managed care organizations and quality of care. *American Journal of Medicine*, 117, 297-304.
- 35 Bottle, A., Gnani, S., Saxena, S., Aylin, P., Mainous, A. G., & Majeed, A. (2008). Association between quality of primary care and hospitalization for coronary heart disease in England: National cross-sectional study. *Journal of General Internal Medicine*, 23, 135-141.
- 36 Boucai, L. & Zonszein, J. (2007). Effects of quality improvement strategies for type 2 diabetes in Bronx, N.Y. *Clinical Diabetes* 25[4], 155-159.
Ref Type: Journal (Full)
- 37 Boyd, C. M., Darer, J., Boulton, C., Fried, L. P., Boulton, L., & Wu, A. W. (2005). Clinical practice guidelines and quality of care for older patients with multiple comorbid diseases - Implications for pay for performance. *Jama-Journal of the American Medical Association*, 294, 716-724.

- 38 Bregm, K. (2008). Management of the new pay systems in the public sector - some implications of insights gained from experiments. *International Review of Administrative Sciences*, 74, 79-93.
- 39 Brush, J. E., Krumholz, H. M., Wright, J. S., Brindis, R. G., Cacchione, J. G., Drozda, J. P. et al. (2006). American college of cardiology principles to guide physician pay-for-performance programs - A report of the American College of Cardiology Work Group on Pay for Performance (A joint working group of the ACC quality strategic direction committee and the ACC advocacy committee). *Journal of the American College of Cardiology*, 48, 2603-2609.
- 40 Buetow, S. (2008). Pay-for-performance in New Zealand primary health care. *J Health Organ Manag.*, 22, 36-47.
- 41 Cabana, M. D., Dombkowski, K. J., Yoon, E. Y., & Clark, S. J. (2005). Variation in pediatric asthma quality improvement programs by managed care plans. *American Journal of Medical Quality*, 20, 204-209.
- 42 Campbell, S., Reeves, D., Kontopantelis, E., Middleton, E., Sibbald, B., & Roland, M. (2007). Quality of primary care in England with the introduction of pay for performance. *New England Journal of Medicine*, 357, 181-190.
- 43 Campbell, S. A., McDonald, R., & Lester, H. (2008). The experience of pay for performance in english family practice: A qualitative study. *Annals of Family Medicine*, 6, 228-234.
- 44 Campbell, S. M., Roland, M. O., Middleton, E., & Reeves, D. (2005). Improvements in quality of clinical care in English general practice 1998-2003: longitudinal observational study. *British Medical Journal*, 331, 1121-1123.
- 45 Carey, I. M., DeWilde, S., Harris, T., Whincup, P. H., & Cook, D. G. (2007). Spurious trends in coronary heart disease incidence: unintended consequences of the new GP contract? *British Journal of General Practice*, 57, 486-489.
- 46 Carlsen, B. & Norheim, O. F. (2005). "Saying no is no easy matter" - A qualitative study of competing concerns in rationing decisions in general practice. *Bmc Health Services Research*, 5.
- 47 Casale, A. S., Paulus, R. A., Selna, M. J., Doll, M. C., Bothe, A. E., McKinley, K. E. et al. (2007). "ProvenCare(SM)" a provider-driven pay-for-performance program for acute episodic cardiac surgical core. *Annals of Surgery*, 246, 613-623.
- 48 Casalino, L., Gillies, R. R., Shortell, S. M., Schmittiel, J. A., Bodenheimer, T., Robinson, J. C. et al. (2003). External incentives, information technology, and organized processes to improve health care quality for patients with chronic diseases. *Jama-Journal of the American Medical Association*, 289, 434-441.
- 49 Casalino, L. P., Elster, A., Eisenberg, A., Lewis, E., Montgomery, J., & Ramos, D. (2007). Will pay-for-performance and quality reporting affect health care disparities? *Health Aff.(Millwood.)*, 26, w405-w414.
- 50 Cattaneo, A., Borgnolo, G., & Simon, G. (2001). Breastfeeding by objectives. *European Journal of Public Health*, 11, 397-401.
- 51 Cebul, R. D., Rebitzer, J. B., Taylor, L. J., & Votruba, M. (2008). *Organizational Fragmentation and Care Quality in the U.S. Health Care System* (Rep. No. Working Paper 14212). NATIONAL BUREAU OF ECONOMIC RESEARCH.
- 52 Chang, F. C., Hu, T. W., Lin, M., Yu, P. T., & Chao, K. Y. (2008). Effects of financing smoking cessation outpatient services in Taiwan. *Tobacco Control*, 17, 183-189.
- 53 Checkland, K., Harrison, S., McDonald, R., Grant, S., Campbell, S., & Guthrie, B. (2008). Biomedicine, holism and general medical practice: responses to the 2004 General Practitioner contract. *Sociology of Health & Illness*, 30, 788-803.
- 54 Chien, A. T., Chin, M. H., Davis, A. M., & Casalino, L. P. (2007). Pay for performance, public reporting, and racial disparities in health care - How are programs being designed? *Medical Care Research and Review*, 64, 283S-304S.
- 55 Choné, P. & Ma, C. A. (2006). *Assymetric information from physician agency: optimal payment and healthcare quality* Boston: Department of Economics, Boston University.

- 56 Choné, P. & Ma, C. A. (2007). *Optimal health care contracts under physician agency* Boston: Department of Economics, Boston University.
- 57 Chung, E. S., Guo, L., Casey, D. E., Bartone, C., Menon, S., Saghir, S. et al. (2008). Relationship of a quality measure composite to clinical outcomes for patients with heart failure. *American Journal of Medical Quality*, 23, 168-175.
- 58 Chung, K. P., Lai, M. S., Cheng, S. H., Tang, S. T., Huang, C. C., Cheng, A. L. et al. (2008). Organization-based performance measures of cancer care quality: core measure development for breast cancer in Taiwan. *European Journal of Cancer Care*, 17, 5-18.
- 59 Chung, R. S., Chernicoff, H. O., Nakao, K. A., Nickel, R. C., & Legorreta, A. P. (2003). A quality-driven physician compensation model: four-year follow-up study. *J Healthc. Qual.*, 25, 31-37.
- 60 Clarkson, J. E., Turner, S., Grimshaw, J. M., Ramsay, C. R., Johnston, M., Scott, A. et al. (2008). Changing clinicians' behavior: A randomized controlled trial of fees and education. *Journal of Dental Research*, 87, 640-644.
- 61 Coleman, K., Reiter, K. L., & Fulwiler, D. (2007). The impact of pay-for-performance on diabetes care in a large network of community health centers. *Journal of Health Care for the Poor and Underserved*, 18, 966-983.
- 62 Coleman, T., Wynn, A. T., Stevenson, K., & Cheater, F. (2001). Qualitative study of pilot payment aimed at increasing general practitioners' antismoking advice to smokers. *British Medical Journal*, 323, 432-435.
- 63 Coleman, T., Lewis, S., Hubbard, R., & Smith, C. (2007). Impact of contractual financial incentives on the ascertainment and management of smoking in primary care. *Addiction*, 102, 803-808.
- 64 Cornell, S. A. (2007). Clinical case study: Achieving long-term control of insulin resistance. *Journal of Managed Care Pharmacy*, 13, S11-S15.
- 65 Cotter, C. M. (2007). Making the case for a clinical information system: The chief information officer view. *Journal of Critical Care*, 22, 56-65.
- 66 Crinson, I., Shaw, A., Durrant, R., De Lusignan, S., & Williams, B. (2007). Coronary heart disease and the management of risk: patient perspectives of outcomes associated with the clinical implementation of the National Service Framework targets. *Health, risk & society* 9[4], 359-373. Ref Type: Journal (Full)
- 67 Cronenwett, J. L., Likosky, D. S., Russell, M. T., Eldrup-Jorgensen, J., Stanley, A. C., & Nolan, B. W. (2007). A regional registry for quality assurance and improvement: The Vascular Study Group of Northern New England (VSGNNE). *Journal of Vascular Surgery*, 46, 1093-+.
- 68 Cupples, M. E., Smith, S. M., & Murphy, A. W. (2008). How effective is prevention in coronary heart disease? *Heart*, 94, 1370-1371.
- 69 Curtin, K., Beckman, H., Pankow, G., Milillo, Y., & Greene, R. A. (2006). Return on investment in pay for performance: A diabetes case study. *Journal of Healthcare Management*, 51, 365-374.
- 70 Custers, T., Arah, O. A., & Klazinga, N. S. (2007). Is there a business case for quality in The Netherlands? A critical analysis of the recent reforms of the health care system. *Health Policy*, 82, 226-239.
- 71 Cutler, T. W., Palmieri, J., Khalsa, M., & Stebbins, M. (2007). Evaluation of the relationship between a chronic disease care management program and california pay-for-performance diabetes care cholesterol measures in one medical group. *Journal of Managed Care Pharmacy*, 13, 578-588.
- 72 Davidson, G., Moscovice, I., & Remus, D. (2007). Hospital size, uncertainty, and pay-for-performance. *Health Care Financing Review*, 29, 45-57.
- 73 Davis, K., Schoenbaum, S. C., & Audet, A. M. (2005). A 2020 vision of patient-centered primary care. *Journal of General Internal Medicine*, 20, 953-957.
- 74 Davis, K. & Guterman, S. (2007). Rewarding excellence and efficiency in medicare payments. *Milbank Quarterly*, 85, 449-468.
- 75 Dennis, S. M., Zwar, N., Griffiths, R., Roland, M., Hasan, I., Davies, G. P. et al. (2008). Chronic disease management in primary care: from evidence to policy. *Medical Journal of Australia*, 188, S53-S56.

- 76 Desai, A. A., Garber, A. M., & Chertow, G. M. (2007). Rise of pay for performance: Implications for care of people with chronic kidney disease. *Clinical Journal of the American Society of Nephrology*, 2, 1087-1095.
- 77 Dimick, J. B., Weeks, W. B., Karia, R. J., Das, S., & Campbell, D. A. (2006). Who pays for poor surgical quality? Building a business case for quality improvement. *Journal of the American College of Surgeons*, 202, 933-937.
- 78 DiPiero, A. & Sanders, D. (2005). Condition based payment: improving care of chronic illness. *British Medical Journal*, 330, 654-657.
- 79 Dixon, J., Lewis, R., Rosen, R., Finlayson, B., & Gray, D. (2004). Can the NHS learn from US managed care organisations? *British Medical Journal*, 328, 223-225.
- 80 Dixon, J., Chantler, C., & Billings, J. (2007). Competition on outcomes and physician leadership are not enough to reform health care. *Jama-Journal of the American Medical Association*, 298, 1445-1447.
- 81 Dixon, J. (2007). Improving management of chronic illness in the National Health Service: better incentives are the key. *Chronic.Illn.*, 3, 181-193.
- 82 Dobson, R. T., Lepnurm, R., & Struening, E. (2005). Developing a scale for measuring professional equity among Canadian physicians. *Social Science & Medicine*, 61, 263-266.
- 83 Doran, T., Fullwood, C., Gravelle, H., Reeves, D., Kontopantelis, E., Hiroeh, U. et al. (2006). Pay-for-performance programs in family practices in the United Kingdom. *New England Journal of Medicine*, 355, 375-384.
- 84 Doran, T., Fullwood, C., Reeves, D., Gravelle, H., & Roland, M. (2008). Exclusion of patients from pay-for-performance targets by english physicians. *New England Journal of Medicine*, 359, 274-284.
- 85 Doran, T. (2008). Lessons from early experience with pay for performance. *Disease Management & Health Outcomes*, 16, 69-77.
- 86 Doran, T., Fullwood, C., Kontopantelis, E., & Reeves, D. (2008). Effect of financial incentives on inequalities in the delivery of primary clinical care in England: analysis of clinical activity indicators for the quality and outcomes framework. *Lancet*, 372, 728-736.
- 87 Downing, A., Rudge, G., Cheng, Y., Tu, Y. K., Keen, J., & Gilthorpe, M. S. (2007). Do the UK government's new Quality and Outcomes Framework (QOF) scores adequately measure primary care performance? A cross-sectional survey of routine healthcare data. *Bmc Health Services Research*, 7.
- 88 Duckett, S., Daniels, S., Kamp, M., Stockwell, A., Walker, G., & Ward, M. (2008). Pay for performance in Australia: Queensland's new Clinical Practice Improvement Payment. *Journal of Health Services Research & Policy*, 13, 174-177.
- 89 Dunn, J. D. (2007). Pharmacy management approach: How do we align all the incentives? *Journal of Managed Care Pharmacy*, 13, S16-S19.
- 90 Eggleston, K. & Hsieh, C. R. (2004). Healthcare payment incentives : a comparative analysis of reforms in taiwan, South Korea and china. *Appl.Health Econ Health Policy*, 3, 47-56.
- 91 Englesbe, M. J., Dimick, J. B., Sonnenday, C. J., Share, D. A., & Campbell, D. A. (2007). The Michigan surgical quality collaborative - Will a statewide quality improvement initiative pay for itself? *Annals of Surgery*, 246, 1100-1103.
- 92 Ettner, S. L., Thompson, T. J., Stevens, M. R., Mangione, C. M., Kim, C., Steers, W. N. et al. (2006). Are physician reimbursement strategies associated with processes of care and patient satisfaction for patients with diabetes in managed care? *Health Services Research*, 41, 1221-1241.
- 93 Evans, J. H., Kim, K., & Nagarajan, N. J. (2006). Uncertainty, legal liability, and incentive contracts. *Accounting Review*, 81, 1045-1071.
- 94 Fairbrother, G., Friedman, S., Hanson, K. L., & Butts, G. C. (1997). Effect of the vaccines for children program on inner-city neighborhood physicians. *Archives of Pediatrics & Adolescent Medicine*, 151, 1229-1235.
- 95 Fairbrother, G., Hanson, K. L., Friedman, S., & Butts, G. C. (1999). The impact of physician bonuses, enhanced fees, and feedback on childhood immunization coverage rates. *American Journal of Public Health*, 89, 171-175.

- 96 Fairbrother, G., Hanson, K. L., Butts, G. C., & Friedman, S. (2001). Comparison of preventive care in Medicaid managed care and medicaid fee for service in institutions and private practices. *Ambulatory Pediatrics, 1*, 294-301.
- 97 Feess, E. & Ossig, S. (2007). Reimbursement schemes for hospitals, malpractice liability, and intrinsic motivation. *International Review of Law and Economics, 27*, 423-441.
- 98 Feldman, M. D., Areal, P. A., Ong, M. K., Lee, D. L., & Feldmann, S. (2005). Incentives for primary care providers to participate in a collaborative care program for depression. *Psychiatric Services, 56*, 1344-1346.
- 99 Feldman, S. (2004). Rewarding results: Improving the quality of treatment for people with alcohol and drug problems - Join together, a national policy panel. *Administration and Policy in Mental Health, 31*, 283-312.
- 100 Felt-Lisk, S., Gimm, G., & Peterson, S. (2007). Making pay-for-performance work in Medicaid. *Health Affairs, 26*, W516-W527.
- 101 Fendrick, A. M. & Shapiro, N. L. (2008). A commentary on the potential of value-based insurance design (VBID) to contain costs and preserve quality. *Journal of Managed Care Pharmacy, 14*, S11-S15.
- 102 Fisher, E. S. (2006). Paying for performance - Risks and recommendations. *New England Journal of Medicine, 355*, 1845-1847.
- 103 Fisher, E. S. (2007). 2007 Robert and Alma Moreton lecture: pay for performance: more than rearranging the deck chairs? *J Am.Coll.Radiol., 4*, 879-885.
- 104 Fleetcroft, R. & Cookson, R. (2006). Do the incentive payments in the new NHS contract for primary care reflect likely population health gains? *J Health Serv.Res.Policy, 11*, 27-31.
- 105 Fleetcroft, R., Steel, N., Cookson, R., & Howe, A. (2008). "Mind the gap!" Evaluation of the performance gap attributable to exception reporting and target thresholds in the new GMS contract: National database analysis. *Bmc Health Services Research, 8*.
- 106 Forest, V. (2008). Performance-related pay and work motivation: theoretical and empirical perspectives for the French civil service. *International Review of Administrative Sciences, 74*, 325-339.
- 107 Forsberg, E., Axelsson, R., & Arnetz, B. (2004). The relative importance of leadership and payment system - Effects on quality of care and work environment. *Health Policy, 69*, 73-82.
- 108 Francis, D. O., Beckman, H., Chamberlain, J., Partridge, G., & Greene, R. A. (2006). Introducing a multifaceted intervention to improve the management of otitis media: How do pediatricians, internists, and family physicians respond? *American Journal of Medical Quality, 21*, 134-143.
- 109 Gallagher, T. H. & Levinson, W. (2004). A prescription for protecting the doctor-patient relationship. *American Journal of Managed Care, 10*, 61-68.
- 110 Gandjour, A. & Lauterbach, K. W. (2005). How much does it cost to change the behavior of health professionals? A mathematical model and an application to academic detailing. *Medical Decision Making, 25*, 341-347.
- 111 Garson, A. (2004). US healthcare: The intertwined caduceus of physicians, coverage, quality, and cost. *Journal of the American College of Cardiology, 43*, 1-5.
- 112 Gene-Badia, J., Escaramis-Babiano, G., Sans-Corrales, M., Sampietro-Colom, L., Aguado-Menguy, F., Cabezas-Pena, C. et al. (2007). Impact of economic incentives on quality of professional life and on end-user satisfaction in primary care. *Health Policy, 80*, 2-10.
- 113 Gilmore, A. S., Zhao, Y. X., Kang, N., Ryskina, K. L., Legorreta, A. P., Taira, D. A. et al. (2007). Patient outcomes and evidence-based medicine in a preferred provider organization setting: A six-year evaluation of a physician pay-for-performance program. *Health Services Research, 42*, 2140-2159.
- 114 Ginsburg, P. B. (2006). Recalibrating Medicare payments for inpatient care. *New England Journal of Medicine, 355*, 2061-2064.
- 115 Glickman, S. W., Ou, F. S., DeLong, E. R., Roe, M. T., Lytle, B. L., Mulgund, J. et al. (2007). Pay for performance, quality of care, and outcomes in acute myocardial infarction. *Jama-Journal of the American Medical Association, 297*, 2373-2380.

- 116 Glickman, S. W., Baggett, K. A., Krubert, C. G., Peterson, E. D., & Schulman, K. A. (2007). Promoting quality: the health-care organization from a management perspective. *International Journal for Quality in Health Care*, 19, 341-348.
- 117 Glickman, S. W., Schulman, K. A., Peterson, E. D., Hocker, M. B., & Cairns, C. B. (2008). Evidence-based perspectives on pay for performance and quality of patient care and outcomes in emergency medicine. *Annals of Emergency Medicine*, 51, 622-631.
- 118 Goldman, L. E., Henderson, S., Dohan, D. P., Talavera, J. A., & Dudley, R. A. (2007). Public reporting and pay-for-performance: Safety-net hospital executives' concerns and policy suggestions. *Inquiry-the Journal of Health Care Organization Provision and Financing*, 44, 137-145.
- 119 Goroll, A. H., Berenson, R. A., Schoenbaum, S. C., & Gardner, L. B. (2007). Fundamental reform of payment for adult primary care: Comprehensive payment for comprehensive care. *Journal of General Internal Medicine*, 22, 410-415.
- 120 Grady, K. E., Lemkau, J. P., Lee, N. R., & Caddell, C. (1997). Enhancing mammography referral in primary care. *Preventive Medicine*, 26, 791-800.
- 121 Grant, S., Huby, G., Watkins, F., Checkland, K., McDonald, R., Davies, H. et al. (2008). The impact of pay-for-performance on professional boundaries in UK general practice: an ethnographic study. *Sociol.Health Illn.*
- 122 Greenberg, M. R., Weinstock, M., Fenimore, D. G., & Sierzega, G. M. (2008). Emergency department tobacco cessation program: staff participation and intervention success among patients. *J Am.Osteopath.Assoc.*, 108, 391-396.
- 123 Greene, R. A., Beckman, H., Chamberlain, J., Partridge, G., Miller, M., Burden, D. et al. (2004). Increasing adherence to a community-based guideline for acute sinusitis through education, physician profiling, and financial incentives. *American Journal of Managed Care*, 10, 670-678.
- 124 Greene, R. A., Beckman, H. B., & Mahoney, T. (2008). Beyond the efficiency index: Finding a better way to reduce overuse and increase efficiency in physician care. *Health Affairs*, 27, W250-W259.
- 125 Grembowski, D., Paschane, D., Diehr, P., Katon, W., Martin, D., & Patrick, D. L. (2005). Managed care, physician job satisfaction, and the quality of primary care. *Journal of General Internal Medicine*, 20, 271-277.
- 126 Gress, S., Focke, A., Hessel, F., & Wasem, J. (2006). Financial incentives for disease management programmes and integrated care in German social health insurance. *Health Policy*, 78, 295-305.
- 127 Gross, R., Elhaynay, A., Friedman, N., & Buetow, S. (2008). Pay-for-performance programs in P4P programs Israeli sick funds. *J Health Organ Manag.*, 22, 23-35.
- 128 Grossbart, S. R. (2006). What's the return? Assessing the effect of "pay-for-performance" initiatives on the quality of care delivery. *Medical Care Research and Review*, 63, 295-485.
- 129 Gulliford, M. C., Ashworth, M., Robotham, D., & Mohiddin, A. (2007). Achievement of metabolic targets for diabetes by English primary care practices under a new system of incentives. *Diabetic Medicine*, 24, 505-511.
- 130 Guthrie, B., McLean, G., & Sutton, M. (2006). Workload and reward in the Quality and Outcomes Framework of the 2004 general practice contract. *British Journal of General Practice*, 56, 836-841.
- 131 Hackbarth, G. & Milgate, K. (2005). Using quality incentives to drive physician adoption of health information technology. *Health Affairs*, 24, 1147-1149.
- 132 Halanych, J. H., Safford, M. M., Keys, W. C., Person, S. D., Shikany, J. M., Kim, Y. I. et al. (2007). Burden of comorbid medical conditions and quality of diabetes care. *Diabetes Care*, 30, 2999-3004.
- 133 Harries, A. D., Salaniponi, F. M., Nunn, R. R., & Raviglione, M. (2005). Performance-related allowances within the Malawi National Tuberculosis Control Programme. *International Journal of Tuberculosis and Lung Disease*, 9, 138-144.
- 134 Harrison, J. P. & Coppola, M. N. (2007). Is the quality of hospital care a function of leadership? *Health Care Manag.(Frederick)*, 26, 263-272.
- 135 Healy, J., Sharman, E., & Lokuge, B. (2006). Australia: Health system review. *Health Systems in Transition* 8[5], 1-158.
Ref Type: Journal (Full)

- 136 Helm, C., Holladay, C. L., & Tortorella, F. R. (2007). The performance management system: Applying and evaluating a pay-for-performance initiative. *Journal of Healthcare Management*, 52, 49-62.
- 137 Hemenway, D. (1995). Financial Incentives for Childhood Immunization. *Journal of Policy Analysis and Management*, 14, 133-139.
- 138 Hemmila, M. R., Jakubus, J. L., Maggio, P. M., Wahl, W. L., Dimick, J. B., Campbell, D. A. et al. (2008). Real money: Complications and hospital costs in trauma patients. *Surgery*, 144, 307-316.
- 139 Heneghan, C., Perera, R., Mant, D., & Glasziou, P. (2007). Hypertension guideline recommendations in general practice: awareness, agreement, adoption, and adherence. *British Journal of General Practice*, 57, 948-952.
- 140 Herrin, J., Nicewander, D., & Ballard, D. J. (2008). The effect of health care system administrator pay-for-performance on quality of care. *Jt.Comm J Qual.Patient Saf*, 34, 646-654.
- 141 Hillman, A. L., Ripley, K., Goldfarb, N., Nuamah, I., Weiner, J., & Lusk, E. (1998). Physician financial incentives and feedback: Failure to increase cancer screening in Medicaid managed care. *American Journal of Public Health*, 88, 1699-1701.
- 142 Hillman, A. L., Ripley, K., Goldfarb, N., Weiner, J., Nuamah, I., & Lusk, E. (1999). The use of physician financial incentives and feedback to improve pediatric preventive care in Medicaid managed care. *Pediatrics*, 104, 931-935.
- 143 Hindle, D. & Kalanj, K. (2004). New general practitioner payment formula in Croatia: Is it consistent with worldwide trends? *Croatian Medical Journal*, 45, 604-610.
- 144 Hippisley-Cox, J., Pringle, M., Cater, R., Coupland, C., & Meal, A. (2005). Coronary heart disease prevention and age inequalities: the first year of the National Service Framework for CHD. *Br.J Gen.Pract.*, 55, 369-375.
- 145 Hollingsworth, J. M., Krein, S. L., Miller, D. C., DeMonner, S., & Hollenbeck, B. K. (2007). Payer leverage and hospital compliance with a benchmark: a population-based observational study. *Bmc Health Services Research*, 7.
- 146 Honore, P. A., Simoes, E. J., Moonesinghe, R., Kirbey, H. C., & Renner, M. (2004). Applying principles for outcomes-based contracting in a public health program. *J Public Health Manag.Pract.*, 10, 451-457.
- 147 Hopkins, J. R. (1999). Financial incentives for ambulatory care performance improvement. *Jt.Comm J Qual.Improv.*, 25, 223-238.
- 148 Huby, G., Guthrie, B., Grant, S., Watkins, F., Checkland, K., McDonald, R. et al. (2008). Whither British general practice after the 2004 GMS contract? Stories and realities of change in four UK general practices. *J Health Organ Manag.*, 22, 63-78.
- 149 Huddle, T. S. (2007). The limits of objective assessment of medical practice. *Theoretical Medicine and Bioethics*, 28, 487-496.
- 150 Hudon, E., Beaulieu, M. D., & Roberge, D. (2004). Integration of the recommendations of the Canadian Task Force on Preventive Health Care - Obstacles perceived by a group of family physicians. *Family Practice*, 21, 11-17.
- 151 Hughes, E. (2007). Payment by results--a model for other diabetes healthcare systems? *Prim.Care Diabetes*, 1, 111-113.
- 152 Ittner, C. D., Larcker, D. F., & Pizzini, M. (2007). Performance-based compensation in member-owned firms: An examination of medical group practices. *Journal of Accounting & Economics*, 44, 300-327.
- 153 Jack, W. (2005). Purchasing health care services from providers with unknown altruism. *Journal of Health Economics*, 24, 73-93.
- 154 Jaiveer, P. K., Jaiveer, S., Jujavarapu, S. B., Morrissey, J., White, J., Gadsby, R. et al. (2006). Improvements in clinical diabetes care in the first year of the new General Medical Services contract in the UK. *The British Journal of Diabetes and Vascular Disease* 6, 45-48. Ref Type: Journal (Full)
- 155 Johnson, D. A., Shaheen, N., Sarles, H., Cattau, E., Deal, S., Flax, I. et al. (2007). Pay for performance: ACG guide for physicians. *American Journal of Gastroenterology*, 102, 2119-2122.

- 156 Kahn, C. N., Ault, T., Isenstein, H., Potetz, L., & Van Gelder, S. (2006). Snapshot of hospital quality reporting and pay-for-performance under Medicare. *Health Affairs*, 25, 148-162.
- 157 Karve, A. M., Ou, F. S., Lytle, B. L., & Peterson, E. D. (2008). Potential unintended financial consequences of pay-for-performance on the quality of care for minority patients. *American Heart Journal*, 155, 571-576.
- 158 Kautter, J., Pope, G. C., Trisolini, M., & Grund, S. (2007). Medicare Physician Group Practice demonstration design: Quality and efficiency pay-for-performance. *Health Care Financing Review*, 29, 15-29.
- 159 Keating, N. L., Landon, B. E., Ayanian, J. Z., Borbas, C., & Guadagnoli, E. (2004). Practice, clinical management, and financial arrangements of practicing generalists. *Journal of General Internal Medicine*, 19, 410-418.
- 160 Keating, N. L., Landrum, M. B., Landon, B. E., Ayanian, J. Z., Borbas, C., Robert, W. F. et al. (2004). The influence of physicians' practice management strategies and financial arrangements on quality of care among patients with diabetes. *Medical Care*, 42, 829-839.
- 161 Kennerly, S. (2007). The impending reimbursement revolution how to prepare for future APN reimbursement. *Nursing Economics*, 25, 81-84.
- 162 Kilbridge, P. M. & Classen, D. C. (2008). The informatics opportunities at the intersection of patient safety and clinical informatics. *Journal of the American Medical Informatics Association*, 15, 397-407.
- 163 Kim, C., Steers, W. N., Herman, W. H., Mangione, C. M., Narayan, K. M. V., & Ettner, S. L. (2007). Physician compensation from salary and quality of diabetes care. *Journal of General Internal Medicine*, 22, 448-452.
- 164 Kirschner, K., Braspenning, J., Batenburg, J., Van de Rijt, D., Muijers, P., Van Everdingen, C. et al. (2008). *Value for money: een model voor honoreren van kwaliteit in de huisartsenpraktijk. Project Transparantie Huisartsenzorg (Fase 2)* Nijmegen: Wetenschappelijk Instituut Kwaliteit Gezondheidszorg, UMC St Radboud.
- 165 Klebe, B., Farmer, C., Cooley, R., De Lusignan, S., Middleton, R., O'Donoghue, D. et al. (2007). Kidney disease management in UK primary care: guidelines, incentives and Information Technology. *Family Practice*, 24, 330-335.
- 166 Koffman, D. M. M., Goetzel, R. Z., Anwuri, V. V., Shore, K. K., Orenstein, D., & LaPier, T. (2005). Heart healthy and stroke free - Successful business strategies to prevent cardiovascular disease. *American Journal of Preventive Medicine*, 29, 113-121.
- 167 Kouides, R. W., Lewis, B., Bennett, N. M., Bell, K. M., Barker, W. H., Black, E. R. et al. (1993). A Performance-Based Incentive Program for Influenza Immunization in the Elderly. *American Journal of Preventive Medicine*, 9, 250-255.
- 168 Kouides, R. W., Bennett, N. M., Lewis, B., Cappuccio, J. D., Barker, W. H., & LaForce, F. M. (1998). Performance-based physician reimbursement and influenza immunization rates in the elderly. *American Journal of Preventive Medicine*, 14, 89-95.
- 169 Larriviere, D. G. & Bernat, J. L. (2008). Invited article: Threats to physician autonomy in a performance-based reimbursement system. *Neurology*, 70, 2338-2342.
- 170 Larson, E. B., Fihn, S. D., Kirk, L. M., Levinson, W., Loge, R. V., Reynolds, E. et al. (2004). The future of general internal medicine - Report and recommendations from the Society of General Internal Medicine (SGIM) task force on the domain of general internal medicine. *Journal of General Internal Medicine*, 19, 69-77.
- 171 Laurence, C. O., Beilby, J., Campbell, S., Campbell, J., Ponte, L., & Woodward, G. (2004). Process for improving the integration of care across the primary and acute care settings in rural South Australia: asthma as a case study. *Aust.J Rural.Health*, 12, 264-268.
- 172 Leape, L. L. & Berwick, D. M. (2005). Five years after to err is human - What have we learned? *Jama-Journal of the American Medical Association*, 293, 2384-2390.
- 173 LeBaron, C. W., Mercer, J. T., Massoudi, M. S., Dini, E., Stevenson, J., Fischer, W. M. et al. (1999). Changes in clinic vaccination coverage after institution of measurement and feedback in 4 states and 2 cities. *Archives of Pediatrics & Adolescent Medicine*, 153, 879-886.

- 174 Leonard, K. L. & Zivin, J. G. (2005). Outcome versus service based payments in health care: lessons from African traditional healers. *Health Economics*, 14, 575-593.
- 175 Lester, H. & Hobbs, F. D. R. (2007). Major policy changes for primary care: Potential lessons for the US new model of family medicine from the quality and outcomes framework in the United Kingdom. *Family Medicine*, 39, 96-102.
- 176 Levin-Scherz, J., DeVita, N., & Timbie, J. (2006). Impact of pay-for-performance contracts and network registry on diabetes and asthma HEDIS (R) measures in an integrated delivery network. *Medical Care Research and Review*, 63, 14S-28S.
- 177 Li, R., Simon, J., Bodenheimer, T., Gillies, R. R., Casalino, L., Schmittiel, J. et al. (2004). Organizational factors affecting the adoption of diabetes care management process in physician organizations. *Diabetes Care*, 27, 2312-2316.
- 178 Lilford, R., Mohammed, M. A., Spiegelhalter, D., & Thomson, R. (2004). Use and misuse of process and outcome data in managing performance of acute medical care: avoiding institutional stigma. *Lancet*, 363, 1147-1154.
- 179 Lindenauer, P. K., Remus, D., Roman, S., Rothberg, M. B., Benjamin, E. M., Ma, A. et al. (2007). Public reporting and pay for performance in hospital quality improvement. *New England Journal of Medicine*, 356, 486-496.
- 180 Locke, R. G. & Srinivasan, M. (2008). Attitudes toward pay-for-performance initiatives among primary care osteopathic physicians in small group practices. *J Am.Osteopath.Assoc.*, 108, 21-24.
- 181 Lynch, M. (1995). Effect of Practice and Patient Population Characteristics on the Uptake of Childhood Immunizations. *British Journal of General Practice*, 45, 205-208.
- 182 MacBride-Stewart, S. P., Elton, R., & Walley, T. (2008). Do quality incentives change prescribing patterns in primary care? An observational study in Scotland. *Family Practice*, 25, 27-32.
- 183 Maisey, S., Steel, N., Marsh, R., Gillam, S., Fleetcroft, R., & Howe, A. (2008). Effects of payment for performance in primary care: qualitative interview study. *Journal of Health Services Research & Policy*, 13, 133-139.
- 184 Majeed, A., Williams, J., De Lusignan, S., & Chan, T. (2005). Management of heart failure in primary care after implementation of the National Service Framework for Coronary Heart Disease: a cross-sectional study. *Public Health*, 119, 105-111.
- 185 Mandel, K. E. & Kotagal, U. R. (2007). Pay for performance alone cannot drive quality. *Archives of Pediatrics & Adolescent Medicine*, 161, 650-655.
- 186 Mannion, R., Goddard, M., Kuhn, M., & Bate, A. (2005). Decentralization strategies and provider incentives in healthcare: evidence from the english national health service. *Appl.Health Econ Health Policy*, 4, 47-54.
- 187 Marang-van de Mheen, P. J., Stadlander, M. C., & Kievit, J. (2006). Adverse outcomes in surgical patients: implementation of a nationwide reporting system. *Quality & Safety in Health Care*, 15, 320-324.
- 188 Martens, J. D., Werkhoven, M. J., Severens, J. L., & Winkens, R. A. G. (2007). Effects of a behaviour independent financial incentive on prescribing behaviour of general practitioners. *Journal of Evaluation in Clinical Practice*, 13, 369-373.
- 189 Mathauer, I. & Imhoff, I. (2006). Health worker motivation in Africa: the role of non-financial incentives and human resource management tools. *Hum.Resour.Health*, 4, 24.
- 190 May, E. L. (2005). Take the lead or take your chances: engaging physicians in pay-for-performance. *Healthc.Exec.*, 20, 24-28.
- 191 McCarlie, J., Reid, E., & Brady, A. J. B. (2007). Audit of the new GMS contract Quality and Outcomes Framework: Raising standards in CHD. *The British Journal of Cardiology* 14, 117-120. Ref Type: Journal (Full)
- 192 McDermott, S. & Williams, T. (2006). *Advancing quality through collaboration: the California pay for performance program* Integrated Healthcare Association.
- 193 McDonald, R., Harrison, S., Checkland, K., Campbell, S. M., & Roland, M. (2007). Impact of financial incentives on clinical autonomy and internal motivation in primary care: ethnographic study. *British Medical Journal*, 334, 1357-1359.

- 194 McElduff, P., Lyratzopoulos, G., Edwards, R., Heller, R. F., Shekelle, P., & Roland, M. (2004). Will changes in primary care improve health outcomes? Modelling the impact of financial incentives introduced to improve quality of care in the UK. *Quality & Safety in Health Care*, 13, 191-197.
- 195 McGovern, M. P., Boroujerdi, M. A., Taylor, M. W., Williams, D. J., Hannaford, P. C., Lefevre, K. E. et al. (2008). The effect of the UK incentive-based contract on the management of patients with coronary heart disease in primary care. *Family Practice*, 25, 33-39.
- 196 McGovern, M. P., Williams, D. J., Hannaford, P. C., Taylor, M. W., Lefevre, K. E., Boroujerdi, M. A. et al. (2008). Introduction of a new incentive and target-based contract for family physicians in the UK: good for older patients with diabetes but less good for women? *Diabetic Medicine*, 25, 1083-1089.
- 197 McLean, G., Sutton, M., & Guthrie, B. (2006). Deprivation and quality of primary care services: evidence for persistence of the inverse care law from the UK Quality and Outcomes Framework. *Journal of Epidemiology and Community Health*, 60, 917-922.
- 198 McLean, G., Guthrie, B., & Sutton, M. (2007). Differences in the quality of primary medical care for CVD and diabetes across the NHS: evidence from the quality and outcomes framework. *Bmc Health Services Research*, 7.
- 199 McMenamin, S. B., Schauffler, H. H., Shortell, S. M., Rundall, T. G., & Gillies, R. R. (2003). Support for smoking cessation interventions in physician organizations - Results from a National Study. *Medical Care*, 41, 1396-1406.
- 200 McMenamin, S. B., Schmittiel, J., Halpin, H. A., Gillies, R., Rundall, T. G., & Shortell, S. A. (2004). Health promotion in physician organizations - Results from a national study. *American Journal of Preventive Medicine*, 26, 259-264.
- 201 McNamara, P. (2005). Quality-based payment: six case examples. *International Journal for Quality in Health Care*, 17, 357-362.
- 202 Mead, N., Bower, P., & Roland, M. (2008). The General Practice Assessment Questionnaire (GPAQ) - Development and psychometric characteristics. *Bmc Family Practice*, 9.
- 203 Meddings, J. A. & McMahon, L. F. (2008). Measuring quality in pay-for-performance programs - From 'one-size-fits-all' measures to individual patient risk-reduction scores. *Disease Management & Health Outcomes*, 16, 205-216.
- 204 Mehrotra, A., Pearson, S. D., Coltin, K. L., Kleinman, K. P., Singer, J. A., Rabson, B. et al. (2007). The response of physician groups to P4P incentives. *Am.J Manag.Care*, 13, 249-255.
- 205 Mehta, R. H., Liang, L., Karve, A. M., Hernandez, A. F., Rumsfeld, J. S., Fonarow, G. C. et al. (2008). Association of patient case-mix adjustment, hospital process performance rankings, and eligibility for financial incentives. *Jama-Journal of the American Medical Association*, 300, 1897-1903.
- 206 Mentari, E. K., Deoreo, P. B., O'Connor, A. S., Love, T. E., Ricanati, E. S., & Sehgal, A. R. (2005). Changes in Medicare reimbursement and patient-nephrologist visits, quality of care, and health-related quality of life. *American Journal of Kidney Diseases*, 46, 621-627.
- 207 Meterko, M., Young, G. J., White, B., Bokhour, B. G., Burgess, J. F., Berlowitz, D. et al. (2006). Provider attitudes toward pay-for-performance programs: Development and validation of a measurement instrument. *Health Services Research*, 41, 1959-1978.
- 208 Milgate, K. & Cheng, S. B. (2006). Pay-for-performance: The MedPAC perspective. *Health Affairs*, 25, 413-419.
- 209 Millett, C., Gray, J., Saxena, S., Netuveli, G., Khunti, K., & Majeed, A. (2007). Ethnic disparities in diabetes management and pay-for-performance in the UK: The Wandsworth prospective diabetes study. *Plos Medicine*, 4, 1087-1093.
- 210 Millett, C., Gray, J., Saxena, S., Netuveli, G., & Majeed, A. (2007). Impact of a pay-for-performance incentive on support for smoking cessation and on smoking prevalence among people with diabetes. *Canadian Medical Association Journal*, 176, 1705-1710.
- 211 Millett, C., Gray, J., Bottle, A., & Majeed, A. (2008). Ethnic disparities in blood pressure management in patients with hypertension after the introduction of pay for performance. *Ann.Fam.Med.*, 6, 490-496.

- 212 Millett, C., Gray, J., Wall, M., & Majeed, A. (2008). Ethnic Disparities in Coronary Heart Disease Management and Pay for Performance in the UK. *J Gen.Intern.Med.*
- 213 Mullen, P. M. (2004). Using performance indicators to improve performance. *Health Services Management Research* 17, 217-228.
Ref Type: Journal (Full)
- 214 Murphy, K. M. & Nash, D. B. (2008). Nonprimary care physicians' views on office-based quality incentive and improvement programs. *Am.J Med.Qual.*, 23, 427-439.
- 215 Nahra, T. A., Reiter, K. L., Hirth, R. A., Shermer, J. E., & Wheeler, J. R. C. (2006). Cost-effectiveness of hospital pay-for-performance incentives. *Medical Care Research and Review*, 63, 49S-72S.
- 216 Nalli, G. A., Scanlon, D. P., & Libby, D. (2007). Developing a performance-based incentive program for hospitals: a case study from Maine. *Health Aff.(Millwood.)*, 26, 817-824.
- 217 Nassiri, A. & Rochaix, L. (2006). Revisiting physicians' financial incentives in Quebec: a panel system approach. *Health Economics*, 15, 49-64.
- 218 Newhouse, J. P. (2005). Medicare's challenges in paying providers. *Health Care Financing Review*, 27, 35-44.
- 219 Nicholson, S., Pauly, M. V., Wu, A. Y. J., Murray, J. F., Teutsch, S. M., & Berger, M. L. (2008). Getting real performance out of pay-for-performance. *Milbank Quarterly*, 86, 435-457.
- 220 Normand, S. L. T., Wolf, R. E., Ayanian, J. Z., & Mcneil, B. J. (2007). Assessing the accuracy of hospital clinical performance measures. *Medical Decision Making*, 27, 9-20.
- 221 O'Brien, S. M., DeLong, E. R., Dokholyan, R. S., Edwards, F. H., & Peterson, E. D. (2007). Exploring the behavior of hospital composite performance measures - An example from coronary artery bypass surgery. *Circulation*, 116, 2969-2975.
- 222 O'Kane, M. E. (2007). Performance-based measures: The early results are in. *Journal of Managed Care Pharmacy*, 13, S3-S6.
- 223 O'Malley, A. S., Pham, H. H., & Reschovsky, J. D. (2007). Predictors of the growing influence of clinical practice guidelines. *Journal of General Internal Medicine*, 22, 742-748.
- 224 O'Shea, J. S. (2008). Medicare physician payment reform: Changing incentives to maintain access to quality surgical services. *Journal of the American College of Surgeons*, 206, 165-170.
- 225 Paleologou, V., Kontodimopoulos, N., Stamouli, A., Aletras, V., & Niakas, D. (2006). Developing and testing an instrument for identifying performance incentives in the Greek health care sector. *Bmc Health Services Research*, 6.
- 226 Patel, M. M., Eisenberg, L., Witsell, D., & Schulz, K. A. (2008). Assessment of acute otitis externa and otitis media with effusion performance measures in otolaryngology practices. *Otolaryngology-Head and Neck Surgery*, 139, 490-494.
- 227 Patel, P. H., Siemons, D., & Shields, M. C. (2007). Proven methods to achieve high payment for performance. *J Med.Pract.Manage.*, 23, 5-11.
- 228 Patmas, M. A., Rosenblum, R., & Hatch, C. (2006). Pay-for-performance program focuses on Web-based quality data. *Physician Exec.*, 32, 58-61.
- 229 Pearson, S. D., Schneider, E. C., Kleinman, K. P., Coltin, K. L., & Singer, J. A. (2008). The impact of pay-for-performance on health care quality in Massachusetts, 2001-2003. *Health Affairs*, 27, 1167-1176.
- 230 Pelonero, A. L. & Johnson, R. L. (2007). A pay-for-performance program for behavioral health care practitioners. *Psychiatric Services*, 58, 442-444.
- 231 Perdue, W. C., Mensah, G. A., Goodman, R. A., & Moulton, A. D. (2005). A legal framework for preventing cardiovascular diseases. *American Journal of Preventive Medicine*, 29, 139-145.
- 232 Pham, H. H., Schrag, D., O'Malley, A. S., Wu, B. N., & Bach, P. B. (2007). Care patterns in Medicare and their implications for pay for performance. *New England Journal of Medicine*, 356, 1130-1139.
- 233 Pierdon, S. B. & Eckrote, B. (2004). Changing compensation plans. Moving beyond last year's, this year's and next year's. *Physician Exec.*, 30, 26-29.

- 234 Pines, J. M. (2006). Profiles in patient safety: Antibiotic timing in pneumonia and pay-for-performance. *Academic Emergency Medicine*, 13, 787-790.
- 235 Pines, J. M., Hollander, J. E., Lee, H., Everett, W. W., Uscher-Pines, L., & Metlay, J. P. (2007). Emergency department operational changes in response to pay-for-performance and antibiotic timing in pneumonia. *Academic Emergency Medicine*, 14, 545-548.
- 236 Pit, S. W., Byles, J. E., Henry, D. A., Holt, L., Hansen, V., & Bowman, D. A. (2007). A Quality Use of Medicines program for general practitioners and older people: a cluster randomised controlled trial. *Medical Journal of Australia*, 187, 23-30.
- 237 Polk, H. C. (2006). Renewal of surgical quality and safety initiatives: A multispecialty challenge. *Mayo Clinic Proceedings*, 81, 345-352.
- 238 Poon, E. G., Blumenthal, D., Jaggi, T., Honour, M. M., Bates, D. W., & Kaushal, R. (2004). Overcoming barriers to adopting and implementing computerized physician order entry systems in US hospitals. *Health Affairs*, 23, 184-190.
- 239 Pourat, N., Rice, T., Tai-Seale, M., Bolan, G., & Nihalani, J. (2005). Association between physician compensation methods and delivery of guideline-concordant STD care: Is there a link? *American Journal of Managed Care*, 11, 426-432.
- 240 Preker, A. S. & Langenbrunner, J. C. (2005). The role of purchasing in hospital performance. *World hospitals and health services* 41[4], 22-9. Ref Type: Journal (Full)
- 241 Pronovost, P., Thompson, D. A., Holzmüller, C. G., Dorman, T., & Morlock, L. L. (2007). Impact of the Leapfrog Group's intensive care unit physician staffing standard. *Journal of Critical Care*, 22, 89-96.
- 242 Quimbo, S. A., Peabody, J. W., Shimkhada, R., Woo, K., & Solon, O. (2008). Should we have confidence if a physician is accredited? A study of the relative impacts of accreditation and insurance payments on quality of care in the Philippines. *Social Science & Medicine*, 67, 505-510.
- 243 Ramsay, S. E., Morris, R. W., Papacosta, O., Lennon, L. T., Thomas, M. C., & Whincup, P. H. (2005). Secondary prevention of coronary heart disease in older British men: extent of inequalities before and after implementation of the National Service Framework. *Journal of Public Health*, 27, 338-343.
- 244 Ramsay, S. E., Whincup, P. H., Lawlor, D. A., Papacosta, O., Lennon, L. T., Thomas, M. C. et al. (2006). Secondary prevention of coronary heart disease in older patients after the national service framework: population based study. *British Medical Journal*, 332, 144-145.
- 245 Reid, G. S., Robertson, A. J., Bissett, C., Smith, J., Waugh, N., & Halkerston, R. (1991). Cervical Screening in Perth and Kinross Since Introduction of the New Contract. *British Medical Journal*, 303, 447-450.
- 246 Reiter, K. L., Nahra, T. A., Alexander, J. A., & Wheeler, J. R. (2006). Hospital responses to pay-for-performance incentives. *Health Serv.Manage.Res.*, 19, 123-134.
- 247 Reschovsky, J. D., Hadley, J., & Landon, B. E. (2006). Effects of compensation methods and physician group structure on physicians' perceived incentives to alter services to patients. *Health Services Research*, 41, 1200-1220.
- 248 Richardson, J. R. (2005). Priorities of health policy: cost shifting or population health. *Aust.New Zealand Health Policy*, 2, 1.
- 249 Ritchie, L. D., Bisset, A. F., Russell, D., Leslie, V., & Thomson, I. (1992). Primary and Preschool Immunization in Grampian - Progress and the 1990 Contract. *British Medical Journal*, 304, 816-819.
- 250 Rittenhouse, D. R., Grumbach, K., O'Neil, E. H., Dower, C., & Bindman, A. (2004). Physician organization and care management in California: From cottage to kaiser. *Health Affairs*, 23, 51-62.
- 251 Rittenhouse, D. R. & Robinson, J. C. (2006). Improving quality in Medicaid - The use of care management processes for chronic illness and preventive care. *Medical Care*, 44, 47-54.
- 252 Rosenbaum, P., Shortt, S. E. D., & Walker, D. M. C. (2004). Alternative funding for academic medicine: Experience at a Canadian health sciences center. *Academic Medicine*, 79, 197-204.
- 253 Rosenthal, M. B., Frank, R. G., Li, Z. H., & Epstein, A. M. (2005). Early experience with pay-for-performance - From concept to practice. *Jama-Journal of the American Medical Association*, 294, 1788-1793.

- 254 Rosenthal, M. B., Landon, B. E., Normand, S. T., Frank, R. G., & Epstein, A. M. (2006). Pay for performance in commercial HMOs. *New England Journal of Medicine*, 355, 1895-1902.
- 255 Rosenthal, M. B., Landon, B. E., Normand, S. L. T., Frank, R. G., Ahmad, T. S., & Epstein, A. M. (2007). Employers' use of value-based purchasing strategies. *Jama-Journal of the American Medical Association*, 298, 2281-2288.
- 256 Rosenthal, M. B. & Camillus, J. (2007). *How four purchasers designed and implemented quality-based purchasing activities* Agency for Healthcare Research and Quality.
- 257 Rosenthal, M. B., de Brantes, F. S., Sinaiko, A. D., Frankel, M., Robbins, R. D., & Young, S. (2008). Bridges to Excellence - Recognizing High-Quality Care: Analysis of Physician Quality and Resource Use. *American Journal of Managed Care*, 14, 670-677.
- 258 Roski, J., Jeddelloh, R., An, L., Lando, H., Hannan, P., Hall, C. et al. (2003). The impact of financial incentives and a patient registry on preventive care quality: increasing provider adherence to evidence-based smoking cessation practice guidelines. *Preventive Medicine*, 36, 291-299.
- 259 Rust, G., Strothers, H. S., & Zimmerman, R. K. (2005). Re-engineering the primary care practice to eliminate adult immunization disparities. *Ethnicity & Disease*, 15, 21-26.
- 260 Safran, D. G., Rogers, W. H., Tarlov, A. R., Inui, T., Taira, D. A., Montgomery, J. E. et al. (2000). Organizational and financial characteristics of health plans - Are they related to primary care performance? *Archives of Internal Medicine*, 160, 69-76.
- 261 Saitto, C., Marino, C., Fusco, D., Arca, M., & Perucci, C. A. (2005). Toward a new payment system for inpatient rehabilitation - Part II: Reimbursing providers. *Medical Care*, 43, 856-864.
- 262 Saunders, M., Schattner, P., & Mathews, M. (2008). Diabetes 'cycles of care' in general practice - Do government incentives help? *Australian Family Physician*, 37, 781-784.
- 263 Sautter, K. M., Bokhour, B. G., White, B., Young, G. J., Burgess, J. F., Berlowitz, D. et al. (2007). The early experience of a hospital-based pay-for-performance program. *Journal of Healthcare Management*, 52, 95-107.
- 264 Saxena, S., Car, J., Eldred, D., Soljak, M., & Majeed, A. (2007). Practice size, caseload, deprivation and quality of care of patients with coronary heart disease, hypertension and stroke in primary care: national cross-sectional study. *Bmc Health Services Research*, 7.
- 265 Schaubroeck, J., Shaw, J. D., Duffy, M. K., & Mitra, A. (2008). An under-met and over-met expectations model of employee reactions to merit raises. *Journal of Applied Psychology*, 93, 424-434.
- 266 Schauffler, H. H., Brown, C., & Milstein, A. (1999). Raising the bar: The use of performance guarantees by the Pacific Business Group on Health. *Health Affairs*, 18, 134-142.
- 267 Schmittiel, J., McMenamin, S. B., Halpin, H. A., Gillies, R. R., Bodenheimer, T., Shortell, S. M. et al. (2004). The use of patient and physician reminders for preventive services: results from a National Study of Physician Organizations. *Preventive Medicine*, 39, 1000-1006.
- 268 Schneider, F., Menke, R., Harter, M., Salize, H. J., Janssen, B., Bergmann, F. et al. (2005). Are bonus systems applicable to guideline-oriented depression treatment provided by general practitioners and psychiatrists? *Nervenarzt*, 76, 308-+.
- 269 Schoen, C., Osborn, R., Huynh, P. T., Doty, M., Peugh, J., & Zapert, K. (2006). On the front lines of care: Primary care doctors' office systems, experiences, and views in seven countries. *Health Affairs*, 25, W555-W571.
- 270 Shen, Y. J. (2003). Selection incentives in a performance-based contracting system. *Health Services Research*, 38, 535-552.
- 271 Shenkman, E., Tian, L. L., Nackashi, J., & Schatz, D. (2005). Managed care organization characteristics and outpatient specialty care use among children with chronic illness. *Pediatrics*, 115, 1547-1554.
- 272 Shepard, D. S., Daley, M. C., Beinecke, R. H., & Hurley, C. L. (2005). Managed behavioral health care: Lessons from Massachusetts. *Administration and Policy in Mental Health*, 32, 311-319.
- 273 Shohet, C., Yelloly, J., Bingham, P., & Lyratzopoulos, G. (2007). The association between the quality of epilepsy management in primary care, general practice population deprivation status and epilepsy-related emergency hospitalisations. *Seizure-European Journal of Epilepsy*, 16, 351-355.

- 274 Sigfrid, L. A., Turner, C., Crook, D., & Ray, S. (2006). Using the UK primary care Quality and Outcomes Framework to audit health care equity: preliminary data on diabetes management. *Journal of Public Health, 28*, 221-225.
- 275 Sikka, R. (2007). Pay for performance in emergency medicine. *Annals of Emergency Medicine, 49*, 756-761.
- 276 Simon, J. S., Rundall, T. G., & Shortell, S. M. (2007). Adoption of order entry with decision support for chronic care by physician organizations. *Journal of the American Medical Informatics Association, 14*, 432-439.
- 277 Simpson, C. R., Hannaford, P. C., Lefevre, K., & Williams, D. (2006). Effect of the UK incentive-based contract on the management of patients with stroke in primary care. *Stroke, 37*, 2354-2360.
- 278 Simpson, C. R., Hannaford, P. C., McGovern, M., Taylor, M. W., Green, P. N., Lefevre, K. et al. (2007). Are different groups of patients with stroke more likely to be excluded from the new UK general medical services contract? A cross-sectional retrospective analysis of a large primary care population. *Bmc Family Practice, 8*.
- 279 Sinsky, C. A., Foreman-Hoffman, V., & Cram, P. (2008). The impact of expressions of treatment efficacy and out-of-pocket expenses on patient and physician interest in osteoporosis treatment: Implications for pay-for-performance programs. *Journal of General Internal Medicine, 23*, 164-168.
- 280 Smith, A. L. (2007). Merging P4P and disease management: How do you know which one is working? *Journal of Managed Care Pharmacy, 13*, S7-S10.
- 281 Snyder, L. & Neubauer, R. L. (2007). Pay-for-performance principles that promote patient-centered care: An ethics manifesto. *Annals of Internal Medicine, 147*, 792-233.
- 282 Soeters, R., Habineza, C., & Peerenboom, P. B. (2006). Performance-based financing and changing the district health system: experience from Rwanda. *Bulletin of the World Health Organization, 84*, 884-889.
- 283 Sperl-Hillen, J. M. & O'Connor, P. J. (2005). Factors driving diabetes care improvement in a large medical group: Ten years of progress. *American Journal of Managed Care, 11*, S177-S185.
- 284 Spertus, J. A., Eagle, K. A., Krumholz, H. M., Mitchell, K. R., & Normand, S. L. T. (2005). American College of Cardiology and American Heart Association methodology for the selection and creation of performance measures for quantifying the quality of cardiovascular care. *Journal of the American College of Cardiology, 45*, 1147-1156.
- 285 Spinelli, R. J. & Fromknecht, J. M. (2007). Pay for performance: improving quality care. *Health Care Manag.(Frederick), 26*, 128-137.
- 286 St Jacques, P. J., Patel, N., & Higgins, M. S. (2004). Improving anesthesiologist performance through profiling and incentives. *Journal of Clinical Anesthesia, 16*, 523-528.
- 287 Steel, N., Maisey, S., Clark, A., Fleetcroft, R., & Howe, A. (2007). Quality of clinical primary care and targeted incentive payments: an observational study. *Br J Gen.Pract., 57*, 449-454.
- 288 Steel, N., Bachmann, M., Maisey, S., Shekelle, P., Breeze, E., Marmot, M. et al. (2008). Self reported receipt of care consistent with 32 quality indicators: national population survey of adults aged 50 or more in England. *British Medical Journal, 337*.
- 289 Strong, M., Maheswaran, R., & Radford, J. (2006). Socioeconomic deprivation, coronary heart disease prevalence and quality of care: a practice-level analysis in Rotherham using data from the new UK general practitioner Quality and Outcomes Framework. *Journal of Public Health, 28*, 39-42.
- 290 Sussman, A. J., Fairchild, D. G., Coblyn, J., & Brennan, T. A. (2001). Primary care compensation at an academic medical center: A model for the mixed-payer environment. *Academic Medicine, 76*, 693-699.
- 291 Sutton, M. & McLean, G. (2006). Determinants of primary medical care quality measured under the new UK contract: cross sectional study. *British Medical Journal, 332*, 389-390.
- 292 Swerissen, H. & Taylor, M. J. (2008). Reforming funding for chronic illness: Medicare-MM. *Australian Health Review, 32*, 76-84.
- 293 Tabak, Y. P., Johannes, R. S., & Silber, J. H. (2007). Using automated clinical data for risk adjustment - Development and validation of six disease-specific mortality predictive models for pay-for-performance. *Medical Care, 45*, 789-805.

- 294 Tacke, M., Braspenning, J., Plas, L., Muijers, J., Winckels, S., Batenburg, J. et al. (2005). *Naar een kwaliteitsbonus voor huisartsenzorg, Fase 1. Project Transparantie Huisartsenzorg* Nijmegen: Centre for Quality of Care Research (WOK), UMC St Radboud.
- 295 Tahrani, A. A., McCarthy, M., Godson, J., Taylor, S., Slater, H., Capps, N. et al. (2007). Diabetes care and the new GMS contract: the evidence for a whole county. *British Journal of General Practice*, 57, 483-485.
- 296 Tahrani, A. A., McCarthy, M., Godson, J., Taylor, S., Slater, H., Capps, N. et al. (2008). Impact of practice size on delivery of diabetes care before and after the Quality and Outcomes Framework implementation. *British Journal of General Practice*, 58, 576-579.
- 297 Teleki, S. S., Damberg, C. L., Pham, C., & Berry, S. H. (2006). Will financial incentives stimulate quality improvement? Reactions from frontline physicians. *American Journal of Medical Quality*, 21, 367-374.
- 298 Thomas, M. R., Waxmonsky, J. A., McGinnis, G. F., & Barry, C. L. (2006). Realignment clinical and economic incentives to support depression management within a Medicaid population: The Colorado Access experience. *Administration and Policy in Mental Health*, 33, 26-33.
- 299 Timbie, J. W., Newhouse, J. P., Rosenthal, M. B., & Normand, S. L. T. (2008). A cost-effectiveness framework for profiling the value of hospital care. *Medical Decision Making*, 28, 419-434.
- 300 Ting, H. H., Galvin, R. S., Krumholz, H. M., Petersen, L. A., & Block, P. C. (2007). Do economic incentives improve quality of health care? Implications for pay-for-performance. *ACC CardioSource Review Journal* 16[7], 22-25.
Ref Type: Journal (Full)
- 301 Tisnado, D. M., Rose-Ash, D. E., Malin, J. L., Adams, J. L., Ganz, P. A., & Kahn, K. L. (2008). Financial incentives for quality in breast cancer care. *American Journal of Managed Care*, 14, 457-466.
- 302 Trisolini, M., Pope, G., Kautter, J., & Aggarwal, J. (2006). *Medicare physician group practices: innovations in quality and efficiency* (Rep. No. 971). The Commonwealth Fund.
- 303 Trude, S., Au, M., & Christianson, J. B. (2006). Health plan pay-for-performance strategies. *American Journal of Managed Care*, 12, 537-542.
- 304 Tuerk, P. W., Mueller, M., & Egede, L. E. (2008). Estimating physician effects on glycemic control in the treatment of diabetes - Methods, effects sizes, and implications for treatment policy. *Diabetes Care*, 31, 869-873.
- 305 Turenne, M. N., Hirth, R. A., Pan, Q., Wolfe, R. A., Messana, J. M., & Wheeler, J. R. C. (2008). Using knowledge of multiple levels of variation in care to target performance incentives to providers. *Medical Care*, 46, 120-126.
- 306 Twardella, D. & Brenner, H. (2007). Effects of practitioner education, practitioner payment and reimbursement of patients' drug costs on smoking cessation in primary care: a cluster randomised trial. *Tobacco Control*, 16, 15-21.
- 307 Twomey, P. J. & Pledger, D. R. (2008). Different DCCT-aligned HbA1c methods and the GMS contract. *International Journal of Clinical Practice*, 62, 202-205.
- 308 Waldman, J. D. & Schargel, F. P. (2006). Twins in trouble (II): Systems thinking in healthcare and education. *Total Quality Management & Business Excellence*, 17, 117-130.
- 309 Walker, O., Strong, M., Atchinson, R., Saunders, J., & Abbott, J. (2007). A qualitative study of primary care clinicians' views of treating childhood obesity. *Bmc Family Practice*, 8.
- 310 Wang, Y. Y., O'Donnell, C. A., Mackay, D. F., & Watt, G. C. M. (2006). Practice size and quality attainment under the new GMS contract: a cross-sectional analysis. *British Journal of General Practice*, 56, 830-835.
- 311 Waters, H. R., Morlock, L. L., & Hatt, L. (2004). Quality-based purchasing in health care. *International Journal of Health Planning and Management*, 19, 365-381.
- 312 Watts, I. T. & Wenck, B. (2007). Financing and the quality framework. *Australian Family Physician*, 36, 32-34.
- 313 Weber, V., Bloom, F., Pierdon, S., & Wood, C. (2008). Employing the electronic health record to improve diabetes care: A multifaceted intervention in an integrated delivery system. *Journal of General Internal Medicine*, 23, 379-382.

- 314 Welke, K. F., Diggs, B. S., Karamlou, T., & Ungerleider, R. M. (2008). Measurement of quality in pediatric cardiac surgery: Understanding the threats to validity. *Asaio Journal*, 54, 447-450.
- 315 Wennberg, J. E. (2004). Practice variations and health care reform: Connecting the dots. *Health Affairs*, 23, VAR140-VAR144.
- 316 Whalley, D., Bojke, C., Gravelle, H., & Sibbald, B. (2006). GP job satisfaction in view of contract reform: a national survey. *Br J Gen Pract*, 56, 87-92.
- 317 Whalley, D., Gravelle, H., & Sibbald, B. (2008). Effect of the new contract on GPs' working lives and perceptions of quality of care: a longitudinal study. *British Journal of General Practice*, 58, 8-14.
- 318 Wickizer, T. M., Franklin, G., Gluck, J. V., & Fulton-Kehoe, D. (2004). Improving quality through identifying inappropriate care: The use of guideline-based utilization review protocols in the Washington state workers' compensation system. *Journal of Occupational and Environmental Medicine*, 46, 198-204.
- 319 Willcox, S. (2005). Buying best value health care: Evolution of purchasing among Australian private health insurers. *Aust. New Zealand Health Policy*, 2, 6.
- 320 Williams, P. H. & De Lusignan, S. (2006). Does a higher 'quality points' score mean better care in stroke? An audit of general practice medical records. *Inform. Prim. Care*, 14, 29-40.
- 321 Williams, T. R., Raube, K., Damberg, C. L., & Mardon, R. E. (2006). Pay for performance: its influence on the use of IT in physician organizations. *J Med. Pract. Manage.*, 21, 301-306.
- 322 Woodson, S. B. (1999). Making the connection between physician performance and pay. *Healthc. Financ. Manage.*, 53, 39-42, 44.
- 323 Wright, S. W., Trott, A., Lindsell, C. J., Smith, C., & Gibler, W. B. (2008). Evidence-based emergency medicine. Creating a system to facilitate translation of evidence into standardized clinical practice: a preliminary report. *Ann. Emerg. Med.*, 51, 80-6, 86.
- 324 Wynn, B. O. & Sorbero, M. E. (2007). *Pay-for-performance in California's workers' compensation medical treatment system. An assessment of options, challenges and potential benefits.*
- 325 Yao, H., Wei, X., Liu, J., Zhao, J., Hu, D., & Walley, J. D. (2008). Evaluating the effects of providing financial incentives to tuberculosis patients and health providers in China. *International Journal of Tuberculosis and Lung Disease*, 12, 1166-1172.
- 326 Young, G. J., Meterko, M., White, B., Bokhour, B. G., Sautter, K. M., Berlowitz, D. et al. (2007). Physician attitudes toward pay-for-quality programs - Perspectives from the front line. *Medical Care Research and Review*, 64, 331-343.
- 327 Young, G. J. & Conrad, D. A. (2007). Practical issues in the design and implementation of pay-for-quality programs. *Journal of Healthcare Management*, 52, 10-18.
- 328 Young, G. J., Burgess, J. F., & White, B. (2007). Pioneering pay-for-quality: Lessons from the rewarding results demonstrations. *Health Care Financing Review*, 29, 59-70.
- 329 Young, G. J., Meterko, M., Beckman, H., Baker, E., White, B., Sautter, K. M. et al. (2007). Effects of paying physicians based on their relative performance for quality. *Journal of General Internal Medicine*, 22, 872-876.
- 330 Zivin, J. G. & Pfaff, A. S. P. (2004). To err on humans is not benign - Incentives for adoption of medical error-reporting systems. *Journal of Health Economics*, 23, 935-949.
- 331 Zwar, N. A., Comino, E. J., Hasan, I., & Harris, M. F. (2005). General practitioner views on barriers and facilitators to implementation of the Asthma 3+Visit Plan. *Medical Journal of Australia*, 183, 64-67.

APPENDIX 10 GRID EQUITY RELATED CONCEPTS

GENERAL INFORMATION QOF TABLE

First author	Affiliation authors	Country (area)	Healthcare setting	Journal	Title	Aim
Millett 2007 ⁷	Primary health care	UK (England)	General practice	CMAJ	Impact of a pay-for-performance incentive on support for smoking cessation and on smoking prevalence among people with diabetes.	To examine the impact of a pay-for-performance incentive in the UK introduced in 2004 as part of the new general practitioner contract to improve support for smoking cessation and to reduce the prevalence of smoking among people with chronic diseases such as diabetes.
Millett 2008 ¹⁴	Primary health care	UK (England, Wandsworth, London)	General practice	Diabetes Care	Impact of pay for performance on ethnic disparities in intermediate outcomes for diabetes: longitudinal study.	To examine the impact of a major pay for performance incentive on trends in the quality of diabetes care in white, black and South-Asian ethnic groups.
Millett 2007 ⁶	Primary health care Health services research	UK (England, Wandsworth, London)	General practice	PLoS Medicine	Ethnic Disparities in Diabetes Management and Pay-for-Performance in the UK: The Wandsworth Prospective Diabetes Study.	To study whether the introduction of pay-for-performance management in general medical practice in the UK leads to a reduction in ethnic disparities in the quality of diabetic care.
Ashworth 2007a	Primary health care	UK (England)	General Practice	British Journal of General Practice	The relationship between social deprivation and the quality of primary care: a national survey using indicators from the UK Quality and Outcomes Framework.	To use Quality and Outcomes Framework (QOF) Indicators to explore the characteristics of primary care in deprived communities.
Ashworth 2008	Primary health care	UK (England)	General Practice	BMJ	Effect of social deprivation on blood pressure monitoring and control in England: a survey of data from the quality and outcomes framework.	To determine levels of blood pressure monitoring and control in primary care and to determine the effect of social deprivation on these levels.
Doran 2008a	Primary health care	UK (England)	General Practice	The Lancet	Effect of financial incentives on inequalities in the delivery of primary clinical care in England: analysis of clinical activity indicators for the quality and outcomes framework.	To examine the relationship between socioeconomic inequalities and delivered quality of care in the first three years of the QOF.
Millett 2008 ¹³	Primary health care	UK (England, Wandsworth, London)	General practice	J Gen Intern Med	Ethnic Disparities in Coronary Heart Disease Management and Pay for Performance in the UK.	To study whether the introduction of pay for performance management in general medical practice in the UK leads to a reduction in ethnic disparities in

						coronary health disease management and outcomes.
Simpson 2006	Primary health care Pharmacology	UK (Scotland)	General Practice	Stroke	Effect of the UK Incentive-Based Contract on the Management of Patients With Stroke in Primary Care.	To examine whether the introduction of the new GMS contract improved recording of quality indicators for patients with stroke and whether there is a difference in change recording to sex, age, and deprivation?
McGovern 2008 ¹¹	Primary health care Public health Pharmacology	UK (Aberdeen)	General practice	Family Practice	The effect of the UK incentive-based contract on the management of patients with coronary heart disease in primary care.	To determine whether the recording of CHD related health indicators and prescribing of medicines have increased following the introduction of the nGMS contract and whether differences in treatment of patients of differing age, gender and deprivation have been affected.
McGovern 2008 ¹⁰	Primary health care	UK (Scotland)	General Practice	Diabetic Medicine	Introduction of a new incentive and target-based contract for family physicians in the UK: good for older patients with diabetes but less good for women?	To determine whether the recording of diabetes related health indicators has increased following the introduction of the nGMS contract and whether differences between age gender and deprivation groups have been affected.
Ashworth 2007b	Primary health care	UK (England)	General Practice	Journal of Public Health	Social deprivation and statin prescribing: a cross-sectional analysis using data from the new UK general practitioner 'Quality and Outcomes Framework'.	To study the relationship between the prescribing of lipid-lowering medication, social deprivation and other general practice characteristics.
Doran 2006	Primary health care	UK (England)	General Practice	The New England Journal of Medicine	Pay-for-Performance Programs in Family Practices in the United Kingdom.	To examine the effects of patients and practice characteristics on performance and to assess the impact of exception reporting on reported achievement in a pay for performance program.
Downing 2007	Epidemiology Health Services Research Biostatistics	UK (England)	General practice	BioMed Central Health Services Research	Do the UK government's new Quality and Outcomes Framework (QOF) scores adequately measure primary care performance? A cross-sectional survey of routine healthcare data.	To assess the extent to which measures of health observed in practice populations are correlated with their QOF scores after accounting for the established associations between health outcomes and socio demographics
Gray 2007	Primary health care Health sciences research	UK (England)	General Practice	JGIM	Ethnicity and Quality of Diabetes Care in a Health System with Universal Coverage: Population-Based Cross-sectional Survey in Primary Care.	To assess the quality of diabetes care and intermediate clinical outcomes within a multiethnic population after a sustained period of investment in quality improvement.

Gulliford 2007	Health services research	UK (England)	General Practice	Diabetic Medicine	Achievement of metabolic targets for diabetes by English primary care practices under a new system of incentives.	To analyze achievement of metabolic targets for diabetes by English general practices following the introduction of a new system of incentives.
Millett 2007 ^a	Primary health care Health services research	UK (England, Wandsworth, London)	General practice	Journal of the Royal Society of Medicine	Diabetes prevalence, process of care and outcomes in relation to practice size, caseload and deprivation: national cross-sectional study in primary care.	To examine the association between practice list size, deprivation and the quality of care of patients with diabetes.
Saxena 2007	Primary health care	UK (England and Scotland)	General Practice	BMC Health Services Research	Practice size, caseload, deprivation and quality of care of patients with coronary heart disease, hypertension and stroke in primary care: national cross-sectional study.	To study the association between quality of care for cardiovascular disease by general practice caseload, practice size and area based deprivation measures.
Strong 2006	Primary health care Public health	UK (England)	General Practice	Journal of Public Health	Socioeconomic deprivation, coronary heart disease prevalence and quality of care: a practice-level analysis in Rotherham using data from the new UK general practitioner Quality and Outcomes Framework.	To examine whether GP practice-level CHD prevalence and quality of care are associated with area-level socioeconomic deprivation.
Sutton 2006	Health economics Primary health care	UK (Scotland, NSH Ayrshire and Arran area)	General Practice	BMJ	Determinants of primary medical care quality measured under the new UK contract: cross sectional study.	To identify factors associated with the quality of primary medical care incentivized under the new UK general medical services contract;
Wang et al 2006	Health services research Primary health care	UK (Scotland mainland)	General Practice	British Journal of General Practice	Practice size and quality attainment under the new GMS contract: a cross-sectional analysis.	To explore the relationship between practice size and points attained in the QOF.
Sigfrid 2006	Public health Primary health care	UK (England, Brighton and Hove area)	General Practice	Journal of Public Health	Using the UK primary care Quality and Outcomes Framework to audit health care equity: preliminary data on diabetes management.	To explore whether exception reporting in the QOF is linked to socioeconomic deprivation.
Wright et al 2006	Geography	UK (England)	General Practice	British Journal of General Practice	Overall Quality of Outcomes Framework scores lower in practices in deprived areas.	To assess the relationships between deprivation, rurality and the number of overall QOF points achieved by general practices.
Mc Lean 2006	Primary health care	UK (Scotland)	General Practice	Journal of epidemiology and community health	Deprivation and quality of primary care services : evidence for persistence of the inverse care law from the UK quality and outcomes framework.	To examine whether the quality of primary care measured by the 2004 contract varies with socioeconomic deprivation
Doran 2008b	Primary health care Health economics	UK (England)	General Practice	The New England Journal of Medicine	Exclusion of Patients from Pay-for-Performance Targets by English Physicians.	To analyze determining factors in the rate of exception reporting by English Physicians

Millett 2008 ¹²	Primary health care	UK (England)	General practice	Annals of Family Medicine	Ethnic Disparities in Blood Pressure Management in Patients With Hypertension After the Introduction of Pay for Performance.	To examine the impact of a major pay for performance incentive on trends in blood pressure management in white, black and South-Asian groups
Simpson 2006	Primary health care Pharmacology	UK (Scotland)	General practice	BMC Family Practice	Are different groups of patients with stroke more likely to be excluded from the new UK general medical services contract? A cross-sectional retrospective analysis of a large primary care population.	To examine whether there is a difference in exception reporting in the GMS contract recording to patient' characteristics?
Hippisley-Cox 2004	Primary health care	UK (England)	General practice	BMJ	Association of deprivation, ethnicity, and sex with quality indicators for diabetes: population based survey of 53 000 patients in primary care.	To determine the effect of deprivation and ethnicity on the achievement of quality indicators for patients with diabetes.

METHODS

First author	Study design ^a Timeframe ^b	Study population / study group	Data source	Outcome variable
Millett 2007	Longitudinal (repeated measurements for each patient) Pre- and post-contract: 2003-2004 2005-2006	All patients with diabetes type I or II in 32 general practices contracting with the Wandsworth Primary Care Trust. Wandsworth is an ethnic diverse region with higher levels of socioeconomic deprivation relative to elsewhere in England.	Wandsworth Primary Care Based Registers - Diagnosis and treatment targets of diabetes for all diabetic patients (using the Wandsworth primary care-based diabetes registers, based on the practices' electronic records) - Patient's self-rated ethnic origin - Patient's socioeconomic status based on the postal area where the patient lives (using the 2004 Index of Multiple Deprivation)	- Smoking status - Smoking cessation advice
Millett 2008	Longitudinal (repeated measurements for each patient) Pre- and post-contract: 2000 till 2005-2006	All patients with diabetes type I or II in 16 general practices in the Battersea area contracting with the Wandsworth Primary Care Trust (n=1968). Wandsworth is an ethnic diverse region.	Wandsworth Primary Care Based Registers - Diabetes related clinical outcome measures (HbA1c and blood pressure) using the Wandsworth primary care-based diabetes registers which are based on the practices' electronic records - Patient's characteristics (age, gender, duration of	- Diabetes related clinical outcome measures

^a Definitions used in this context:

Cross-sectional: measurement at one point in time

Serial cross-sectional: measurements at two or more points in time, the data from the same study subject (e.g. patient) is not linked over time

Longitudinal: measurements at two or more points in time, the data from the same study subject (e.g. patient) is linked over time

^b Post-contract : after the introduction of the GMS contract for GPs in march 2004

			<p>diabetes) using the Wandsworth primary care-based diabetes registers which are based on the practices' electronic record</p> <ul style="list-style-type: none"> - Patient's self-rated ethnic origin - Patient's socioeconomic status based on the postal area where the patient lives (using the 2004 Index of Multiple Deprivation) - Family practice characteristics (list size, number of full-time GPs and neighborhood SES) obtained from the National Primary Care Research and Development Centre. 	
Millett 2007	<p>Longitudinal (repeated measurements for each patient)</p> <p>Pre- and post-contract: 2003-2004 2005-2006</p>	<p>All patients with diabetes type I or II in 32 general practices contracting with the Wandsworth Primary Care Trust (n=4284). Wandsworth is an ethnic diverse region.</p>	<p>Wandsworth Primary Care Based Registers</p> <ul style="list-style-type: none"> - Diagnosis and treatment targets of diabetes for all diabetic patients (using the Wandsworth primary care-based diabetes registers, based on the practices' electronic records) - Patient's self-rated ethnic origin - Patient's socioeconomic status based on the postal area where the patient lives (using the 2004 Index of Multiple Deprivation) 	<ul style="list-style-type: none"> - Prescribing levels - Intermediate clinical diabetes outcome measures
Ashworth 2007a	<p>Serial cross-sectional</p> <p>Post contract: 2004-2005 2005-2006</p>	<p>All General Practices in England</p> <p>Practices with a list size of under 750 patients or under 500 patients per full-time equivalent GP were excluded.</p> <p>Complete data were available for 8480 practices in England in 2004-2005 and for 8264 practices in England in 2005-2006.</p>	<ul style="list-style-type: none"> - QOF data for each general practice in the UK - Practice characteristics incl. the SOA^c in which the practice is located - Social deprivation data (Index of Multiple Deprivation Scores 2004) for all lower layer SOAs - National urbanicity scores for all SOAs 	<ul style="list-style-type: none"> - Total QOF points - QOF points per domain - Detailed QOF scores
Ashworth 2008	<p>Serial cross-sectional</p> <p>Post contract: April 2004 till March 2007</p>	<p>All General Practices in England.</p> <p>Practices with a list size of under 750 patients or under 500 patients per full-time equivalent GP were excluded.</p>	<ul style="list-style-type: none"> - QOF data for specific indicators relevant for BP monitoring and control - Practice characteristics including the SOA in which the practice is located, which forms the basis for calculating the index of multiple deprivation, 2004. - Ethnicity data available from the 2001 Census, again aggregated at the level of SOA, at practice 	<ul style="list-style-type: none"> - Achievement levels of BP monitoring - Achievement levels of BP control

^c Super Output Area (SOA): geographical socially homogeneous areas with a population of around 1500

			level.	
Doran 2008a	Serial cross-sectional Post-contract: 2004 till 2007	All General Practices in England for which achievement data were available (n=8277) Practices were excluded when <ul style="list-style-type: none"> - List s size < 1000 - One or more disease registers missing - Relocation of a practice to a more or less affluent area during the study period - Incomplete availability of exclusion data - Change of over 25% in practice population Subanalyses were undertaken for excluded practices Complete data were available for 7637 practices	<ul style="list-style-type: none"> - QOF achievement data - Level of deprivation of a practice based on the level of area deprivation in the census super output area where the practice is located- with data from the Index of Deprivation 2004 - Information pf practice and patient characteristics from the 2006 general medical statistics database (dept of Health) 	Practices median reported overall achievement rate for 48 clinical indicators
Millett 2008	Serial cross-sectional Pre- and post contract: 2003-2004 2005-2006	All patients with coronary heart disease in 32 general practices contracting with the Wandsworth Primary Care Trust (n=2891). Wandsworth is an ethnic diverse region.	Wandsworth Primary Care Based Registers <ul style="list-style-type: none"> - Diagnosis and treatment targets of CHD for all CHD patients (using the Wandsworth primary care-based CHD registers, based on the practices' electronic records) - Patient's self-rated ethnic origin - Patient's socioeconomic status based on the postal area where the patient lives (using the 2004 Index of Multiple Deprivation). 	10 quality indicators: <ul style="list-style-type: none"> - Process of care measures - Prescribing measures - Intermediate clinical CHD outcome measures
Simpson 2006	Serial cross-sectional Post- contract: 2004 2005	All patients with a computer record of transient ischemic attack or stroke from all general practices in Scotland using the General Practice Administrative Software System and that participated in the Scottish Program for Improving Clinical Effectiveness (SPICE) (n practices= 310, n patients in 2004= 21901, n patients in 2005=32 401).	SPICE database ^d <ul style="list-style-type: none"> - patient characteristics (age, sex, stroke-related co-morbidity, deprivation status based on postal code) - stroke related QOF indicators 	<ul style="list-style-type: none"> - Recording of QOF scores - Stroke/TIA prevalence

^d SPICE : As part of the SPICE program (Scottish Program for Improving Clinical Effectiveness), data entry templates were developed for use by clinicians to systematically record data about a number of chronic conditions. From 2003 onwards these templates were modified to include all information required for the new GMS contract resulting in a database containing the same variables as in the QOF database but on patient level (Mc Govern, Simpson).

McGovern 2008	Serial cross-sectional Pre- and post contract: 2000 till 2005	Patients with CHD as defined by their GP, included in 310 general practices in Scotland 58406 patients pre-contract 75495 patients post-contract	SPICE database - Gender, age, comorbidities - Deprivation status (deprivation quintiles based on Carstairs's DEPCAT postcode categorization) - number of patients registered in the practice of the patient)	Recording of 11 CHD related quality indicators
McGovern 2008	Serial cross-sectional Pre- and post contract: 2000 till 2005	Patients with DM as defined by their GP, included in 310 general practices in Scotland 37329 patients pre-contract 56561 patients post-contract	SPICE database - Gender, age, comorbidities - Deprivation status (deprivation quintiles based on Carstairs's DEPCAT postcode categorization)	Recording of 8 DM related quality indicators
Ashworth 2007b	Cross-sectional Post-contract: 2004-2005	All General Practices in England Practices were excluded when they had a list size of under 750 patients or under 500 per full-time GP. Complete data were available for 8430 practices	- QOF data for each general practice in the UK - Census based variables - Prescribing data. were collected from the National Prescribing Analysis and Cost (PACT) data (prescribed + over the counter medication) - Social deprivation data (Index of Multiple Deprivation 2004)	Prescribing volume for statins
Doran 2006	Cross-sectional Post-contract: 2004-2005	General Practices out of QMAS dataset (n = 8576) Practices with a list size of under 1000 patients or with the reported register missing or who included less than half the patients subsequently reported for individual indicators were excluded. 8105 practices remained (94.5%)	- QOF data for 8105 general practices in England - Reported achievement out of QMAS database - Population achievement for 30 of the 76 indicators (only for those were the indicator is based on all patients with that condition (eg without age limits)) - Exception reporting for 30 of the 76 indicators (only for those were the indicator is based on all patients with that condition (eg without age limits)) - Socioeconomic characteristics attributed to each practice based on data of the UK 2001 Census and indices of deprivation (SOA) - Information on practice characteristics from the 2004 General Medical Statistics database maintained by the department of Health	- Summary outcome scores for each condition - Global scores - Rates of exception reporting - Reported achievement - Population achievement
Downing 2007	Cross-sectional Post-contract:	All general practices in two English Primary Care Trusts	- QOF data for all general practices in two English Primary Care Trusts (April 2004-March 2005) - Data for emergency hospital admissions (for	- QOF scores - Admission rates - Overall mortality

	2004-2005		<p>asthma, cancer, COPD, coronary heart disease, diabetes, stroke and all other conditions)</p> <ul style="list-style-type: none"> - All cause mortality (September 2004 to August 2005) - a geographical measure of socioeconomic deprivation was obtained by matching the income domain scores of the Index of Multiple deprivation 2004 to the census based super output area of residence. 	
Gray 2007	Cross-sectional Post-contract: 2005-2006	Identification of 7605 diabetes patients in 32/36 primary care practices in Wandsworth.	Wandsworth Prospective Diabetes Study : <ul style="list-style-type: none"> - Quality indicators were registered out of clinical information recorded on the practice computer. - Neighborhood socioeconomic status was assigned to individual patients based on their postcode using the index of multiple deprivation 2004. 	Achievement rate of the quality indicators for diabetes
Gulliford 2007	Cross-sectional Post-contract: 2005	<p>Patients :</p> <ul style="list-style-type: none"> - Clinical data of 1441 patients who had diabetes since 2000 or before - In this analysis deprivation or ethnicity are not considered <p>Practices</p> <ul style="list-style-type: none"> - All English General Practices - Practices with a list size of under 750 patients or under 500 patients per full-time equivalent GP were excluded. - Comparisons were made between tertiles of deprivation/ethnicity 	<ul style="list-style-type: none"> - QOF data for all English practices (2005) - Clinical data of 1441 patients out of 26 practices in South London who agreed to take part in a local diabetes care project (2000-2004) - Deprivation scores were linked to practices using the practice postcode (based on 2001 Census – SOA – IMD 2004) - Ethnicity was calculated as 100 – percentage of white subjects in the SOA of the practice. 	Achievement targets for HbA1C, BP and cholesterol
Millett 2007	Cross-sectional Post-contract: Exact year not specified	9411 general practices in England and Scotland	<ul style="list-style-type: none"> - QOF data - Practices socioeconomic status based on the index of multiple deprivation 	Achievement rates for 18 diabetes related QOF indicators
Saxena 2007	Cross-sectional Post-contract: 2004-2005	<p>All General Practices in England and Scotland returning QOF data (n = 9411).</p> <p>Practices were excluded if they could not be matched to deprivation data via their postcode (n=441)</p>	<ul style="list-style-type: none"> - QOF data on 26 cardiovascular disease related indicators - Social deprivation data of super output areas (Index of Multiple Deprivation Scores 2004) 	<ul style="list-style-type: none"> - Cardiovascular disease prevalence - Target achievement for 26 cardiovascular disease QOF indicators

Strong 2006	Cross-sectional Post-contract: 2004-2005	All general practices contracting with the Rotherham Primary Care Trust and taking part in the QOF (n=38).	<ul style="list-style-type: none"> - QOF data: total points achieved on the 11 CHD indicators and number of CHD patients in the QOF register (sex and age standardized) - Per practice the mean index of multiple deprivation (IMS) score using the 2004 Super Output Area level IMD scores, weighted for the proportion of patients living in each super output area. 	<ul style="list-style-type: none"> - Sex and age standardized CHD prevalence - The proportion of practices achieving each of the 11 CHD targets.
Sutton 2006	Cross-sectional Post-contract: 2004-2005	60 of the 61 general practices in the NHS Ayrshire and Arran area.	<ul style="list-style-type: none"> - QOF data: total points achieved on the 10 clinical domains and holistic care - Practice and GP characteristics (clinical team size and composition, mean age GPs, proportion female GPs, training practices, accreditation data, data on salaried contract, ex-funding practice, income from other sources) - Material deprivation of the population based on the Scottish Index of Multiple Deprivation 2004 - Standardized chronic illness rate of the population based on Scottish Census 2001 - Urbanicity categories based on the Scottish Executive Urban-Rural Classification (SEURC) 	Total QOF points achieved on the 10 clinical domains and holistic care
Wang 2006	Cross-sectional Post-contract: 2005	<p>All urban General Practices in mainland Scotland returning QOF data (636 practices).</p> <p>No exclusion criteria.</p> <p>Only bivariate analyses were performed with “being a small or single-handed practice” as outcome variable.</p> <p>Comparisons were made between small and single-handed urban practices (n=286) and medium or large urban practices (n=350).</p>	<ul style="list-style-type: none"> - QOF data (2005) - Practice and GP characteristics (list size, nb of GPs, proportion females and South-Asian GPs, personal medical services and training practices) for each general practice in Scotland (2002) - Percentage of Indian, Pakistani and South Asian patients in the practice using output area (2) level data from the 2001 census - Information whether the practice received Practice Accreditation (PA) or the Quality Practice Award (QPA) or participated in the Scottish Programme to Improve Clinical Effectiveness (SPICE) - Social deprivation data in the practice population based on the modified Scottish Index of Multiple Deprivation - Urbanicity categories based on the Scottish Executive Urban-Rural Classification (SEURC) 	<ul style="list-style-type: none"> - Total QOF points - QOF points per domain (median)

			<ul style="list-style-type: none"> - Patient's healthcare need, captured using the 2001 census based indicator of limiting long-term illness for those aged under 64 years - Data on coronary heart disease mortality for under-70s, standardized for age and sex of the practice population 	
Sigfrid 2006	Cross-sectional Post-contract: 2004-2005	All General Practices contracting with the Brighton and Hove City Primary Care Trust (n=52). Practices without electronic records (n=2) and practices with an atypical population of exclusively homeless people (n=1) were excluded.	<ul style="list-style-type: none"> - QOF data: 15 diabetes related QOF indicators - Per practice the mean index of multiple deprivation (IMS) score using the 2004 Super Output Area level IMD scores, weighted for the proportion of patients living in each super output area. - Diabetes prevalence standardized for sex and age bands (using the Diabetes UK data and the Exeter system) 	<ul style="list-style-type: none"> - Diabetes (type 1 +type 2) prevalence - Target achievement for 15 diabetes indicators - Exception reporting rates for each of the 15 indicators
Wright 2006	Cross-sectional Post-contract: 2004	All General Practices in England with a complete set of data (8569 practices).	<ul style="list-style-type: none"> - QOF data for (all? no more detailed information available) general practices in England - Practice characteristics: postcode and address - Social deprivation data of super output areas (Index of Multiple Deprivation Scores 2004) - Urban/Rural Classification of output areas (Office for National Statistics) 	Total QOF score
Mc Lean 2006	Cross-sectional Post-contract: 2005	1024 general practices in Scotland	<ul style="list-style-type: none"> - QOF achievement data for which payment quality (based on payment denominators) and delivered quality (based on total register size) are calculated - Practice deprivation is derived from the income domain of the Scottish Index of Multiple Deprivation 2004 	Regression coefficients summarizing the relationships between deprivation and payment and delivered quality. Where the coefficient on delivered quality minus the coefficient on payment quality is negative the implied exclusion rates are higher in more deprived practices.
Doran 2008b	Cross-sectional Post-contract: 2005-2006	General Practices in QMAS database Exclusion : - Practices with < 1000 patients	<ul style="list-style-type: none"> - QMAS data on 65 clinical indicators (concerning diagnosis and referral, measurement and review, offer of treatment, provision of treatment, intermediate outcomes) for 10 diseases. 	Rate of exception reporting

		<ul style="list-style-type: none"> - Practices with one or more disease registers missing - Practices with missing or incomplete data regarding exception reporting <p>Data were collected for 8409 practices.</p>	<ul style="list-style-type: none"> - Rates of exception reporting : number of patients who were excluded for each indicator as a proportion of the number of patients who were eligible for the target - Information on characteristics of medical practices from the 2004 General Medical Statistics Database - Attribution of socio-economic characteristics to each practice on the basis of the electoral district in which the practice is located using Census data and the Index of Multiple Deprivation 	
Millett 2008 ¹²	Cross-sectional Post-contract: 2005-2006	All patients with essential hypertension in 16 primary care practices in Wandsworth.	Wandsworth Primary Care Based Registers <ul style="list-style-type: none"> - Cardiovascular comorbidities - Blood pressure values - Currently prescribed antihypertensive medications - Patient's self-rated ethnic origin - Patient's socioeconomic status based on the postal area where the patient lives (using the 2004 Index of Multiple Deprivation) 	Achievement of blood pressure control
Simpson 2006	Cross-sectional Post-contract: 2005	All patients with a computer record of transient ischemic attack or stroke and an exception report* from all general practices in Scotland and that participated in the Scottish Program for Improving Clinical Effectiveness (n practices= 310, n patients =1749).	SPICE database <ul style="list-style-type: none"> - patient characteristics (age, sex, stroke-related comorbidity, dementia, deprivation status based on postal code) - stroke related QOF indicators - 'top level' exception reporting codes 	<ul style="list-style-type: none"> - Recording of exception reporting - Stroke/TIA prevalence
Hippisley-Cox 2004	Cross-sectional Pre-contract	Patients with diabetes (n=54180) included in the new general practices database (QSEARCH	QSEARCH database linking the following data <ul style="list-style-type: none"> - Clinical data - Townsend scores (derived from the 2001 Census) as a proxy for material deprivation- based on SOA - Ethnicity – based on SOA 	<ul style="list-style-type: none"> - Interpractice variation in achievement rate of quality indicators for diabetes

EQUITY ASPECTS FINDINGS

First author	Analysis: <ul style="list-style-type: none"> - Level of analysis - Groups taken into consideration in the analysis - Extent of exception reporting analysed 	Type of equity in health care			Findings related to equity
		access	treatment	outcomes	
Millett 2007	<p>Level of analysis: patients</p> <p>Groups:</p> <ul style="list-style-type: none"> - Socioeconomic groups based on the IMD score of the patients' SOA: patients were grouped into 5 groups - Ethnic groups based on the patients' self-rated ethnicity <p>Exception reporting is not explicitly mentioned by the author.</p>		X		<ul style="list-style-type: none"> - Significantly more patients with diabetes had their smoking status ever recorded in 2005 than in 2003 (98.8% vs 90.0%). The proportion of patients with documented smoking cessation advice also increased significantly over this period (from 48.0% to 83.5%) - The prevalence of smoking decreased significantly from 20% to 16.2%. This reduction was lower among women (OR 0.71) but was not significantly different in the most and least affluent groups. - In 2005 smoking rates continued to differ significantly with age, sex and ethnic background.
Millett 2008	<p>Level of analysis : patients</p> <p>Groups:</p> <p>Ethnic groups based on the patients' self-rated ethnicity</p> <p>Exception reporting is not explicitly mentioned by the author.</p>			X	<ul style="list-style-type: none"> - The introduction of the pay for performance was associated with reductions in blood pressure and in HbA1c for all ethnic groups. - However the magnitude of the improvement appeared to differ between ethnic groups: after adjusting for the effects of age, gender, years since the diagnosis, practice size and deprivation of the area where the patient lives and the area where the practice is located, the average reductions in blood pressure were lower in the black patients than in the white patients. No sign. difference between south Asian patients and white patients was found. A sign. reduction of HbA1c was found for the white patients but not for the black and south Asian patients. - The introduction of the pay for performance seems to widen the existing inequalities in diabetes control. However, the differences were generally modest and the associated clinical impact likely to be small. Nevertheless this widening remain a concern.
Millett 2007	<p>Level of analysis : patients</p>		X	X	<ul style="list-style-type: none"> - The proportion of patients reaching treatment targets for HbA1c, blood pressure and total cholesterol increased sign. after the

	<p>Groups: Ethnic groups based on the patients' self-rated ethnicity</p> <p>Exception reporting is not explicitly mentioned by the author.</p>				<p>introduction of the GP contract. These increases were broadly uniform across ethnic groups, except for the black Caribbean group, which had improvements in HbA1c and BP control that were sign. lower than in the British group.</p> <ul style="list-style-type: none"> - The number of patients who met the targets improved but the gap between ethnic groups remained about the same. - The new GP contract fails to address known disparities in diabetes management and outcomes between ethnic groups.
Ashworth 2007a	<p>Level of analysis: practices</p> <p>Groups: Socioeconomic groups based on the IMD score of the practices' SOA: comparisons were made between practices located in the least and most deprived quintile SOAs in England.</p> <p>Exception reporting has not been included in the study. The "raw" QOF database is used which excludes all patients with an exception report code.</p>		X	X	<ul style="list-style-type: none"> - More practices, more fte GP's, more single handed practices and less training practices in the most deprived quintiles. - Differences between primary care quality indicators in deprived and prosperous communities were small. - For 22 of the 147 specific indicators (both clinical and non-clinical) a difference of more than 5% between groups was reported: <ul style="list-style-type: none"> - Medicines: identify and follow up SMI patients who do not attend their injectable neuroleptic appointment - Medicines: medication review in last 15 months for all patients on repeat medication - Information: surgery open > or = 45 hours/week - Education: practice has conducted > or = 12 significant event audits in the last 3 years - Education: practice nurses have personal learning plan - Education: practice has conducted > or = 6 significant event audits in last 3 years - Education: all practice nurses have annual appraisal - Patient experience: practice has discussed patient survey with patient group or non-executive director of PCT, changes proposed and some evidence that changes enacted - Records: case notes have clinical summary in > or = 80% - Records: case notes have clinical summary in > or = 60% - Child health surveillance: practice offers child health surveillance checks - Epilepsy: seizure free for > or = 12 months - Mental health: on lithium and serum level in therapeutic range - Mental health: on lithium and creatinine level function checked - CHD: % new angina diagnosis confirmed by exercise test - LVD: % left ventricular disease patients with diagnosis confirmed by ECHO test - COPD: % new cases with diagnosis confirmed by spirometry - COPD: % all cases who have had spirometry testing - COPD: FEV1.0 in all patients diagnosed with COPD

					<ul style="list-style-type: none"> - Stroke: % new cases referred for confirmation of diagnosis by CT/MRI scan - Cervical screening: % women aged 25–65 years who have had a smear in the last 5 years - If adjusting for exception reporting, the existing differences between the practices increased and a significant difference was found for three additional indicators: COPD (% given flu vaccination), diabetes (% with record of neuropathy testing), diabetes (% with HbA1c < 10 mmol/L) - Existing differences between the practices in the most and those in the least deprived areas narrowed between 2004-2005 and 2005-2006.
Ashworth 2008	<p>Level of analysis: practices</p> <p>Groups: Socioeconomic groups based on the IMD score of the practices' SOA: comparisons were made between practices located in the least and most deprived SOAs in England.</p> <p>Accounted for exception reporting: All data presented are the values reported by general practitioners before they had excluded any patients using the mechanism of exception reporting.</p>			X	<ul style="list-style-type: none"> - Practice characteristics : although there are about twice as many practices in deprived areas they have larger list sizes per full time equivalent GP and are less likely to be training or group practices. - Blood pressure recording in the adult population : the small discrepancy between achievement of BP monitoring in the least and most deprived areas has all but disappeared by 2007 - Prevalence of five chronic conditions : recorded disease prevalence has increased over the observation period but the differences between least and most deprived areas are small. - Achievement of BP targets for five chronic conditions : Modest shortfalls in blood pressure control by practices in more deprived areas have largely disappeared by the third year of the QOF even though the small residual differences were significant.
Doran 2008a	<p>Level of analysis : practices</p> <p>Groups: Socioeconomic groups based on the IMD score of the practices' SOA: practices were grouped into equal sized quintiles based on their SOA's IMD score</p> <p>Exception reporting is not explicitly mentioned by the author. However the QOF database is used which excludes all patients with an exception report code.</p>		X	X	<ul style="list-style-type: none"> - In year 1 area deprivation was associated with lower levels of achievement. Quintile 1 (least deprived) 86.8% - Quintile 5 (most deprived) 82.8% - Greater deprivation was associated with marginally higher exclusion rates (6.29% Q1 – 6.80 Q5 in year 2 and 7.21% Q1-759% Q5 in year 3) The association between area deprivation and reported exclusion rates remained significant after regression analysis with practices serving the most deprived population having a modelled exclusion rate that was 0.55% higher than did those serving the least deprived in year 2 and 0.67% higher in year 3. - A 1% higher rate of exclusion was associated with a 0.35% higher rate of achievement in year 2 and a 0.16% higher rate in year 3. - Between year 1 and 3 the gap in median achievement narrowed from 4.0% to 0.8% - The lower the achievement in the previous year the greater the increase in achievement. More rapid improvement in achievement in

					more deprived quintiles was therefore attributable to poorer initial performance and not location in deprived area per se.
Millett 2008	<p>Level of analysis : patients</p> <p>Groups: Ethnic groups based on the patients' self-rated ethnicity</p> <p>Accounted for exception reporting: All data presented are the values reported by general practitioners before they had excluded any patients using the mechanism of exception reporting.</p>		X	X	<ul style="list-style-type: none"> - Significantly more patients achieved established quality indicators for CHD after the implementation of the pay for performance programme in the UK (e.i. blood pressure control and total cholesterol). - Improvements in the blood pressure control were greater in the black group compared to whites, with disparities evident at baseline attenuated. - Disparities in the blood pressure between the south Asian group and the white group, attenuated. - Statin prescribing remained sign. lower in the black group compared with the south Asian and the white groups after the implementation of pay for performance.
Simpson 2006	<p>Level of analysis: patients</p> <p>Groups: - Males/females - Age - Socioeconomic groups based on the Carstairs' DEPCAT score of the patient's postcode.</p> <p>Patients with an exception code for any of the measured indicators were excluded.</p>		X	X	<ul style="list-style-type: none"> - The recording of stroke related QOF indicators increased after the introduction of the contract. - Large increases in the recording of risk factors in the oldest patients tended to attenuate age differences. - Women had larger increases in recording of quality indicators over time than men, however sex differences persisted in some components of care. - More affluent patients tended to have larger increases in recording of quality indicators than did the most deprived. This resulted in increasing deprivation differences in certain aspects over time: the recording of a magnetic resonance imaging/computed tomography scan, smoking, cholesterol, antiplatelet or anticoagulant therapy, and influenza vaccination. - A significant difference between the most and least deprived patients emerged after the contract, with the most deprived stroke patients being less likely to have a record of smoking status and blood pressure. - "... inequitable care exists, which may have important implications for female, older, and more deprived subgroups in terms of stroke recurrence and mortality."
McGovern 2008	<p>Level of analysis : patients</p> <p>Groups: - Males/females - Age - Socioeconomic groups based on the Carstairs' DEPCAT score of the patient's postcode: patients were grouped into 5 quintiles</p>		X	X	<ul style="list-style-type: none"> - Introduction QOF : dramatic rise in the recording of CHD related quality indicators. - Not all the population benefitted equally. - Women, older patients and the most deprived were less likely to have a record than men, the youngest and least deprived respectively. - Post contract, women with a history of CHD were less likely than men to be referred for an exercise test and or specialist assessment

	<p>Patients with an exception code for any of the measured indicators were excluded.</p>				<p>after angina is first diagnosed, have blood pressure recorded and controlled, be prescribed an anticoagulant, betablocker or ace-inhibitor therapy or receive an influenzavaccination (even after adjustment for age, number of comorbidities and deprivation) However women who smoke were more likely to receive smoking cessation advice.</p> <ul style="list-style-type: none"> - Post contract, the most deprived patients were more likely to receive antiplatelet or anticoagulant or aceinhibitor therapy than the least deprived. However the most deprived patients were less likely to have smoking status recorded. The most deprived were also less likely to have their blood pressure measured, receive betablocker therapy or an influenza vaccination
McGovern 2008	<p>Level of analysis : patients</p> <p>Groups:</p> <ul style="list-style-type: none"> - Males/females - Age - Socioeconomic groups based on the Carstairs DEPCAT score of the patient's postcode: patients were grouped into 5 quintiles <p>Patients with an exception code for any of the measured indicators were excluded.</p>		X	X	<ul style="list-style-type: none"> - Introduction QOF : rise in the recording of patients with diabetes and the recording of the relevant quality indicators. - Women have not benefitted equally. - Pre-contract women were as likely as men to have recording of HbA1c, blood pressure, serumcreatinine and cholesterol. Post contract women were less likely to have HbA1c, serumcreatinine and cholesterol recorded (even after adjustment for age, number of diabetes related comorbidities and deprivation). - Few statistically significant differences were found between deprivation groups. - Differences between the oldest and youngest age groups in the pre contract dataset disappeared in the post contract dataset suggesting that older patients benefited most from contract changes
Ashworth 2007b	<p>Level of analysis : practices</p> <p>Groups:</p> <ul style="list-style-type: none"> - Age groups - Socioeconomic groups based on the IMD score of the practices' SOA - Ethnic groups based on the estimation of the proportion of south Asian and Afro- Caribbean patients in the practice's SOA <p>The data were analysed for each of the chronic disease indicators using both raw QOF data and, where possible, data adjusted for exception reporting following the method of Doran, 2006.¹⁵</p>		X		<ul style="list-style-type: none"> - This study found higher prescription rates for practices serving more deprived populations even after adjustment for other factors such as the increased prevalence of cardiovascular disease and diabetes - Patients over 75 are being prescribed proportionally less statins. Even after controlling for factors such as social deprivation or reported prevalence this association remained relatively strong. - Practices in areas with higher proportions of Afro Caribbean or south-Asian had lower volumes of statin prescribing even though these patients have a higher need for coronary health care.
Doran 2006	<p>Level of analysis : practices</p> <p>Groups:</p> <ul style="list-style-type: none"> - Males/females 		X	X	<ul style="list-style-type: none"> - Sociodemographic characteristics of the patients had moderate but significant effects on performance (living in income deprived household, long term unemployed, living in social housing, living in 1 parent household, member of racial or ethnic minority)

	<ul style="list-style-type: none"> - Age groups (≤ 15 Yr of age and ≥ 65 Yr of age) - Socioeconomic groups based on IMD scores and Census data of the practices' electoral district: the proportion of the population living in income-deprived households, the proportion without any educational qualifications, the proportion long-term unemployed, the proportion living in social housing and the proportion living in 1-parent households - Socioeconomic groups based on the number of patients from a racial or ethnic minority (no clear information on how this data is obtained) <p>Analyzing exception reporting was one of the aims of this study.</p>				<ul style="list-style-type: none"> - For reported achievement the factor with the greatest effect was exception reporting. An increase of 1% in the estimated proportion of patients excluded was associated with an increase of 0.31% for every additional 1000 patients on the practice list
Downing 2007	<p>Level of analysis : practices</p> <p>Groups:</p> <ul style="list-style-type: none"> - Males/females - Age - Socioeconomic groups based on the income domain of the IMD score of the practices' SOA <p>Exception reporting is not explicitly mentioned by the author. However the QOF database is used which excludes all patients with an exception report code.</p>		X	X	<ul style="list-style-type: none"> - The associations between QOF scores and emergency admissions and mortality were small and inconsistent, whilst the impact of socioeconomic deprivation on the outcomes was much stronger. These results have implications for the use of target based remuneration of general practitioners and emphasise the need to tackle inequalities and improve the health of disadvantaged groups and the population as a whole.
Gray 2007	<p>Level of analysis : patients</p> <p>Groups:</p> <ul style="list-style-type: none"> - Socioeconomic groups based on the IMD score of the patients' SOA - Ethnic groups based on the patients' self-rated ethnicity <p>Exception reporting is not explicitly mentioned by the author.</p>		X	X	<ul style="list-style-type: none"> - Recording of process measures varied only minimally between ethnic groups. No significant differences in recording blood pressure, HbA1c, cholesterol, microalbuminuria, creatinine or retinopathy screening. Blacks were significantly more asked about smoking status and to have their BMI and peripheral pulses measured than whites but less likely to be offered smoking cessation advice. South Asians were more likely to be asked about their smoking status and to have their peripheral pulses measured than whites. - The black and south asian groups were significantly less likely to meet all three treatment targets (for BP, HbA1c and cholesterol control) than the white group. The black group had significantly lower BP and HbA1c control than the white group. The south asian group had significantly poorer HbA1c control but better cholesterol control than the white group. These disparities were present after controlling for age, gender, and neighborhood socioeconomic status.
Gulliford 2007	<p>Level of analysis : practices (for the part of the study considering equity)</p> <p>Groups:</p>			X	<ul style="list-style-type: none"> - Comparing the highest and lowest tertiles of deprivation the percentage achieving HbA1c below 7.4% was 2.69% lower in the most deprived areas. - In areas with the highest proportion of ethnic minorities the

	<ul style="list-style-type: none"> - Socioeconomic groups based on the IMD score of the practices' SOA: practices were grouped into equal sized tertiles - Ethnic groups based on the proportion of ethnic minorities in the practice's SOA: practices were grouped into equal sized tertiles <p>Accounted for exception reporting.</p>				percentage achieving HbA1c below 7.4% was 2.73% lower than where there were few ethnic minorities
Millett 2007	<p>Level of analysis : practices</p> <p>Groups:</p> <ul style="list-style-type: none"> - Socioeconomic groups based on the IMD score of the practices' SOA: practices were grouped into 3 groups <p>The authors mention that they were unable to adjust for exception reporting in their analysis.</p>		X	X	<ul style="list-style-type: none"> - Smaller and more deprived practices had a higher mean prevalence than larger and more affluent practices (3.8% vs 2.8%) Deprivation had a negative effect on the achieved scores and this was more pronounced for smaller practices.
Saxena 2007	<p>Level of analysis: practices</p> <p>Groups:</p> <p>Socioeconomic groups based on the IMD score of the practices' SOA</p> <p>Accounted for exception reporting : Level of exception reporting was monitored for each individual practice and if it would be unusually high or low the data would need to be verified. The median rate of exception reporting after the first year was small <6% and we did not make any adjustment for this</p>		X	X	<ul style="list-style-type: none"> - Prevalence of CHD was consistent across all areas (from least deprived to most deprived). - Despite wide variations in practice size and deprivation levels, little variation in achieving quality outcome indicators were found. However, some exceptions were found: practices in more affluent areas have a higher achievement of indicators requiring referral for further investigation. For the other indicators, no significant associations were found.
Strong 2006	<p>Level of analysis: practices</p> <p>Groups:</p> <p>Socioeconomic groups based on the IMD score of the patients' SOA (IMD ratings for practice populations were calculated proportionately according to number of patients from their list represented in each SOA)</p> <p>Exception reporting is not explicitly mentioned by the author. However the QOF database is used which excludes all patients with an exception report code.</p>		X	X	<ul style="list-style-type: none"> - Practice-level CHD prevalence has a positive correlation with deprivation. - A relationship was found between the level of deprivation of the practice's patient population and the achievement of one of the 11 CHD related QOF targets. For the other targets, no significant association was found.
Sutton 2006	<p>Level of analysis: practices</p> <p>Groups:</p>		X	X	<ul style="list-style-type: none"> - In a multivariate analysis, quality of care is higher for deprived areas. quality of care is higher for deprived areas. 53% of the variation in

	<p>Socioeconomic groups based on the IMD score of the practice's SOA</p> <p>Exception reporting is not explicitly mentioned by the author. However the QOF database is used which excludes all patients with an exception report code.</p>				<p>quality scores was explained by a multivariate model which included measures of deprivation, clinical team size and composition and financial incentives. No significant effects are found for accreditation, training status, and average age of the general practitioner.</p> <ul style="list-style-type: none"> - No univariate relation was found between the urban/rural character of the area and the quality of care.
Wang 2006	<p>Level of analysis : practices</p> <p>Socioeconomic groups based on the IMD score of the practice's SOA: bivariate analyses show a significant relationship between being a small or single handed practice and being located in lower socioeconomic area</p> <p>Ethnic minority groups (unclear how the ethnicity of the patients was measured): Bivariate analyses show a significant relationship between being a small or single-handed practice and having higher numbers of patients with an minority ethnic background.</p> <p>However, socioeconomic groups or ethnic patient groups were not included when analyzing the relationship between practice size and QOF points.</p> <p>Exception reporting was not taken into account as no data on exception reporting for practices were available to the authors.</p>		X	X	<ul style="list-style-type: none"> - Single-handed and smaller practices were more likely to be located in areas of greater socioeconomic deprivation (bivariate). - Single-handed and smaller practices were more likely to have patients with poorer health (bivariate). - Single-handed and smaller practices were more likely to have patients from minority ethnic groups (bivariate). - Single-handed and smaller practices obtain a significant lower number of QOF-points, due to lower point attainment in the organizational domain (bivariate). - Within the clinical domain, single-handed and smaller practices achieved slightly (but significant) less median points for COPD and CHD. - After controlling for socioeconomic deprivation, single-handed and small practices perform as well as larger practices in the clinical care, the patient experience domains, holistic care, additional services and quality practice payments. They score lower on the organizational domain.
Sigfrid 2006	<p>Level of analysis: practices</p> <p>Groups: Socioeconomic groups based on the IMD score of the practices' patients' SOA (IMD ratings for practice populations were calculated proportionately according to number of patients from their list represented in each SOA)</p> <p>Analysing exception reporting was the aim of this study.</p>		X	X	<ul style="list-style-type: none"> - Patients with diabetes living in deprived areas are more likely to be 'exception reported' from QOF clinical indicators. - Correlations between exception reporting and deprivation were seen for 10 of the 15 diabetes indicators: for these indicators practices with a more deprived patient population were more likely to report 'exceptions'. For the other 5 indicators correlations were weaker but in the same direction. - Deprivation accounted for 9-16% of the exception reporting. - No relationship between the deprivation of the patient population and the achievement of QOF targets was found. - Since the level of achievement of targets is similar between practices with different levels of deprivation, they receive equal resource allocation, regardless of exception reporting. So high levels of

					exception reporting, particularly in practices with deprived populations, may be disguising unmet need in those populations. More work is needed to detect diabetes, prevent complications and target hard-to-reach populations. So analyzing exception reporting should be used to allocate health care resources.
Wright 2006	<p>Level of analysis : practices</p> <p>Socioeconomic groups based on the IMD score of the practices' SOA: comparisons were made between practices located in the least and most deprived SOAs.</p> <p>Exception reporting is not explicitly mentioned by the author. However the QOF database is used which excludes all patients with an exception report code.</p>		X	X	<ul style="list-style-type: none"> - Multiple deprivation is significantly inversely related to quality points achieved. - Deprivation affects a practice's ability to score quality points. This translates into 8400 BP less income for the most deprived compared to the least deprived practices in England.
Mc Lean 2006	<p>Level of analysis : practices</p> <p>Socioeconomic groups based on the income domain of the IMD score of the practices' SOA.</p> <p>Exception reporting was taken into account by analyzing both payment quality (that allows for exclusion of patients) and delivered quality (based on the care delivered to all patients)</p>				<ul style="list-style-type: none"> - Little systematic association is found between payment quality and deprivation but for 17/33 indicators examined delivered quality falls with increasing deprivation. - Absolute differences in delivered quality are small for most simpler process measures, such as recording of smoking status or blood pressure. - Greater inequalities are seen for more complex process measures such as diagnostic procedures, some intermediate outcome measures such as glycemic control in diabetes and measures of treatment such as influenza vaccination.
Doran 2008b	<p>Level of analysis : practices</p> <p>Groups:</p> <ul style="list-style-type: none"> - Socioeconomic groups based on the IMD score of the practices' SOA - Ethnic groups defined by Census data on the practices' SOA <p>Analyzing exception reporting was the aim of this study.</p>				<p>The characteristics of patients and practices explained only 2.7% in the variance of exception reporting.</p> <ul style="list-style-type: none"> - Living in income deprived households total 0.04 (beta coefficient) - Member of racial or ethnic minority : total -0.04 (beta coefficient)
Millett 2008	<p>Level of analysis : patients</p> <p>Groups:</p> <p>Ethnic groups based on the patients' self-rated ethnicity</p> <p>Exception reporting is not explicitly mentioned by the author.</p>			X	<ul style="list-style-type: none"> - Black patients with hypertension are significantly less likely to achieve treatment targets for BP than white or South Asian patients (OR 0.86) - Prevalence of cardiovascular comorbidities was higher among SA patients than among their white or black counterparts (41% vs 28.5% vs 28.8%) - The presence of 2 or more cardiovascular comorbidities was associated with significantly improved BP control among white

					<p>patients but not among black or South Asian patients. (-9.4 mm Hg, - 0.6 mm Hg, - 1.8 mm Hg)</p> <ul style="list-style-type: none"> - SA patients were prescribed fewer antihypertensive medications with poorly controlled BP when compared with their black or white peers (OR 0.66)
Simpson 2006	<p>Level of analysis: patients</p> <p>Groups:</p> <ul style="list-style-type: none"> - Males/females - Age - Socioeconomic groups based on the Carstairs DEPCAT score of the patient's postcode. <p>Analyzing exception reporting was one of the aims of the study</p>				<ul style="list-style-type: none"> - There is no significant association between the practice's exception reporting and the practice having proportionately more female, older or deprived stroke/TIA patients. - Stroke/TIA patients with the 'top level' exclusion code 'patient unsuitable for inclusion' were more likely to be female, older, and have a diagnosis with dementia when compared to those patients without such a code. - The youngest and patients from more deprived parts of Scotland were more likely to have the exception codes: 'informed dissent' or 'no response to letters'. - Females were more likely to be excluded from the specific quality indicators of achieving blood pressure or cholesterol control. More deprived patients were not likely to be excluded from these quality indicators. - Younger and more deprived patients were more likely to be recorded as having refused to attend for review or not replying to letters asking for attendance at primary care clinics. It is important to identify and monitor these individuals so that all patients fully benefit from the implementation of an incentive based contract and receive appropriate clinical care to prevent stroke recurrence, further disability and mortality.
Hippisley-Cox 2004	<p>Level of analysis : patients</p> <p>Socioeconomic groups : Patients from the most deprived fifth compared with those from the most affluent fifth – based Townsend score of the patient (SOA based)</p> <p>Ethnicity groups : patients of the fifth with highest ethnicity compared with that of lowest ethnicity (based on SOA of the patient)</p> <p>gender</p>		X	X	<ul style="list-style-type: none"> - Compared with patients from affluent areas those from deprived areas were less likely to have BMI and smoking status recorded. They were also less likely to have records for HbA1c, an HbA1c value < 7.5% of < 10%, retinal screening, blood pressure, testing for neuropathy or microalbuminuria, or flu vaccination. - Compared with patients from areas of low ethnicity those from areas of high ethnicity were less likely to have many measures recorded. - Women were significantly less likely to have records for BMI, pulses, BP below 145/85, testing for microalbuminuria, serum cholesterol concentration, serum cholesterol values < 5mmol/l and ACE inhibitors given in the presence of proteinuria or microalbuminuria. - Of the 17 quality indicators 10 were adversely associated with deprivation and nine were adversely associated with ethnicity.

GENERAL INFORMATION NON QOF

First author	Affiliation authors	Country (area)	Healthcare setting	Journal	Title	Aim
Karve 2008	Clinical research	USA (Durham, NC)	Hospital Care	American Heart journal	Potential unintended financial consequences of pay-for-performance on the quality of care for minority patients.	To determine whether pay for performance increases existing racial care disparities.
Langham 1995	Health Services Research Public health	UK (England)	General Practice	British Journal of General Practice	The carrot, the stick and the general practitioner: how have changes in financial incentives affected health promotion activity in general practice?	To evaluate the effect of the change in June 1993, in financial incentives for health promotion activity in primary care on the distribution of health promotion payments in two family health services authorities
Lynch 1995	Public Health	UK (Scotland)	General practice	British Journal of General Practice	Effect of practice and patient population characteristics on the uptake of childhood immunizations.	To examine the relationship between the factors which provide a broad profile of practices and general practitioners performance in terms of primary childhood immunization targets
Safran 2000	Primary health care Public health Health services research	USA (Massachusetts)	General Practice	Archives of Internal Medicine	Organizational and Financial Characteristics of Health Plans. Are They Related to Primary Care Performance?	To compare the primary care received by patients in each of 5 models of managed care and identify specific characteristics of health plans associated with performance differences.
Shenkman 2005	Epidemiology Health services research	USA (Gainesville, Florida)	Specialist outpatient care	Pediatrics Official Journal of the American Academy of Pediatrics	Managed Care Organization Characteristics and Outpatient Specialty Care Use Among Children With Chronic Illness.	To examine the association between managed care organization characteristics in which primary care providers serve as gatekeeper, and outpatient physician specialist use among children with chronic conditions and who are publically insured (meaning they require the same benefit package and the same copayment structure).

METHODS

First author	Study design ^e Timeframe ^f	Study population / study group	Data source	Outcome variable
Karve 2008	Cross sectional Post contract	Hospitals (n=3449) Exclusion of hospitals <ul style="list-style-type: none"> - with less than 10 patients - without performance data - without hospital characteristics data - with < 30 cases eligible for any given measure for AMI, CAP and HF respectively 	Cross-sectional databank (Data from quarter 2 2004 to quarter 1 2005) composed for this study linking the following data <ul style="list-style-type: none"> - Hospital Compare Database 2006 provided hospital Performance data on 3 conditions AMI, CAP and heart failure. - Racial demographics were collected from a 5% sample of Medicare Claims data. - Hospital characteristics were obtained from the American hospital association annual survey database (2003) To determine the percentage AA treated by a center data from Hospital Care were matched with claims data.	Hospital performance rates for AMI, CAP and heart failure defined as the number of times a hospital delivered a guideline based therapy divided by the number of opportunities to administer that therapy in that hospital.
Langham 1995	Longitudinal at pre and post contract time points.	General Practices <ul style="list-style-type: none"> - Rural area of Bedfordshire - Inner city London encompassing large deprived areas : Kensington, Chelsea and Westminster 	Serial cross-sectional databank (1992-1993) composed for this study linking the following data <ul style="list-style-type: none"> - Health promotion payment data from the family health service authorities - The Jarman underprivileged area score which is a measure of population deprivation factors likely to affect general practitioner workload. 	change in remuneration pre- and post contract (from clinic activity payment to target payment for health promotion)
Lynch 1995	Cross sectional Post contract	208 general practices in Greater Glasgow Health Board	Cross sectional databank (1991-1992) <ul style="list-style-type: none"> - anonymized information on the uptake of childhood immunizations for 208 general practices which grouped practices according to their immunization targets. - characteristics of the practices - characteristics of the practices patient populations - deprivation data according to the percentage 	Immunization targets (high target, low target or neither) (consistent achievers, occasional achievers and non achievers)

^e Definitions used in this context:

Cross-sectional: measurement at one point in time

Serial cross-sectional: measurements at two or more points in time, the data from the same study subject (e.g. patient) is not linked over time

Longitudinal: measurements at two or more points in time, the data from the same study subject (e.g. patient) is linked over time

^f Post-contract : after the introduction of the GMS contract for GPs in march 2004

			of patients attracting deprivation payments based on the Jarman score and the neighbourhood type of each practice location (census data)	
Safran 2000	Cross sectional Post contract	<p>Employees of the Commonwealth of Massachusetts enrolled in any of 12 health plans offered to state workers.</p> <p>A random sample of 10733 employees was taken. Only adults who reported having a regular family physician and for whom plan-type was known were included. Employees unlocatable by mail, who were deceased or no longer working as an employee were excluded. This resulted in a study sample of 6018.</p>	<p>Cross-sectional databank (1996) including data from two sources:</p> <ul style="list-style-type: none"> - the Primary Care Assessment Survey (PCAS): a validated, 51-item, patient-completed questionnaire designed to measure the essential elements of primary care. - Survey of health plan executives: this organizational survey elicited information about the plan's physician recruitment, selection, and deselection criteria, compensation and financial incentives; and nonfinancial influences on care. 	<p>Primary care performance on the following characteristics:</p> <ul style="list-style-type: none"> - accessibility - continuity - comprehensiveness - integration - clinical interaction - humane interpersonal treatment - patient trust
Shenkman 2005	Cross-sectional Post contract	All children between 5 and 18 years enrolled in the Florida's State Child's Health Insurance Program who have been diagnosed with a chronic condition and had functional limitations, an increased need for or use of health care services beyond what children normally use, and/or dependence on medications or home medical equipment. (n=2333)	<p>Cross-sectional databank (data gathered between 1999 and 2003) composed for this study, linking the following data:</p> <ul style="list-style-type: none"> - Child-level enrollment files: age, gender, family income and nb of years enrolled in the program - Child-level health care claims and encounter files: child's diagnosis, prior specialty care use - Parent telephone survey: consequences of the child's health problem e.g. on daily functioning - MCO administrator interviews: MCO characteristics - Area resource files: County specific data (provider availability in the MCO service delivery area) - Census: number of children under 18 	<ul style="list-style-type: none"> - Odds of an outpatient physician specialist visit 1 year after study entry

EQUITY ASPECTS FINDINGS

First author	Analysis: <ul style="list-style-type: none"> - Level of analysis - Groups taken into consideration in the analysis 	Type of equity in health care			Findings related to equity
		access	Treatmen	outcomes	
Karve 2008	<p>Level of analysis : hospitals</p> <p>Ethnicity : Centers treating large African American populations (> 20%) versus centres treating less than 20% African Americans.</p>		X		<ul style="list-style-type: none"> - The percentage of AA patients treated by a center was inversely associated with performance for AMI and CAP but not HF. - Relative to hospitals with < 20% AA, those with > 20% AA were less likely eligible for financial bonuses and more likely to face penalties.
Langham 1995	<p>Level of analysis : practices</p> <p>Practices were divided into two groups : high or low need according to the Jarman underprivileged area score.</p> <p>High relative need was defined as more than 25% of the practice population living in electoral wards with a Jarman score of over 30.</p>				<ul style="list-style-type: none"> - The new arrangements for health promotion activity have resulted in a more even spread of financial resources. - This has been achieved with a disproportionate financial loss to single handed practices and to practices situated in areas of high relative need. - Mean changes in payments in practices for Bedfordshire -179£ for high Jarman score + 174 for low Jarman score. For Kensington, Chelsea and Westminster -1797£ for high Jarman score, -968 £ for low Jarman score.
Lynch 1995	<p>Level of analysis : practices</p> <p>Socioeconomic data according to the percentage of patients attracting deprivation payments based on the Jarman score and the neighbourhood type of each practice location (census data)</p>		X		<ul style="list-style-type: none"> - A disproportionate number of practices reaching the high target were located in the more affluent areas, whereas a higher than expected proportion of those which either achieved the low target or missed both targets was located in the more deprived areas. Similar results were obtained when the consecutive reaching of targets (consistent achievers, inconsistent achievers, non achievers) was considered.

Safran 2000	<p>Level of analysis: individuals</p> <ul style="list-style-type: none"> - Socioeconomic groups: household income and years of education - Age - Sex - Ethnic minority groups: white vs not-white <p>All data collected using a self-administered survey.</p>	×	×		<ul style="list-style-type: none"> - Patients in the indemnity insurance (which scores most favorable on performance) were older, sicker and disproportionately male and white compared with patients in the other models of care. - Patients in the IPA (which perform equally to the indemnity system on many attributes and at intermediate levels for most others) and staff-model HMOs (which score least favorable on performance) were of average younger than those in each of the other 3 models. Patients in staff-model HMOs had more non-white patients. - The proportion of employees from low-income households did not differ across plan-types.
Shenkman 2005	<p>Level of analysis: patients</p> <ul style="list-style-type: none"> - Socioeconomic groups: family income - Girls/boys - Race: whites, blacks, others - Ethnicity: hispanic, non-hispanic 	×			<ul style="list-style-type: none"> - Black children were 55% less likely than white children to receive an outpatient physician specialist visit, even after consideration of other covariates in the model (such as MCO characteristics, the child's socioeconomic characteristics, the child's condition and consequences of the condition). - Gender was marginally significant in a reduced model (girls being less likely to receive care). When adding more covariates in the model gender was no longer significant. - Children's ethnicities (Hispanic versus non-Hispanic) and family incomes were not significant in the full model including covariates such as MCO characteristics, ethnicity, race, the child's socioeconomic characteristics, the child's condition and consequences of the condition).

APPENDIX II OVERVIEW INTERVIEW CONTENT INTERNATIONAL EXPERTS

SEMI STRUCTURED INTERVIEW OF P4Q EXPERT COUNTRY REPRESENTATIVES

1. Is there any form of ‘performance based payment’ (e.g. pay for performance, pay for quality) present in the health care system of your country, as you are aware of?

Please consider thereby pay for quality as Policies, including laws, rules, financial or administrative orders, made by governments, non-governmental organizations (health funds, provider organizations,...), public or private insurers, that specifically intend to affect the quality of care, by means of financial incentives. The intervention can be combined with other interventions. It can comprise a financial incentive directed at a person’s income or directed at further investment in quality improvement. The financial incentive can be either positive or negative. Target payments, being the practice of paying professionals only if they provide a minimum level of care, is considered a form of P4Q. Implicit financial incentives, which might influence quality of care, but are not specifically intended as such to promote quality explicitly, nor are directly related to quality goals, are not considered to be a form of P4Q.

2. How did ‘pay for quality’ arise in your country?
 - a. Who initiated it?
 - b. Based on which rationale?
 - c. Are there specific initiatives, abroad or domestic, that are considered as key examples for inspiring your country in developing P4Q?
3. What are/ were necessary cornerstones for the implementation of pay for quality, in the culture of your health care system?
4. How is ‘pay for quality’ developed and implemented in your country?
 - a. Which are/should be the goals? (quality domains addressed)
 - b. Which are/should be the targets? (patient group, provider setting, indicator selection)
 - c. Who is/should be involved in setting goals and targets?
 - d. What kind of incentive was/should be developed? (reward/withhold, continuous or threshold, size, competitive or not, on income or investment budget)
 - e. Which level does/should the incentive target? (individual provider, team, provider organization, insurer, patient trajectory through settings)
 - f. How was/should it be implemented? (level of local involvement, communication, phased approach or not, provision of quality improvement support)
 - g. Is/should participation be mandatory or voluntary? (advantages and disadvantages of both options)
 - h. How is/should the quality be measured? (by whom, risk adjustment for outcomes, exception reporting, data availability and quality)
 - i. What kind of indicator selection criteria are/should be used?
 - j. Is/should it be linked with feedback? How is/should it be organized? (level of data, benchmarking, time delay before data availability, with/without comments, with/without suggestions for improvement)
 - k. Who organizes/should organize the program? (state, payer, provider organizations, independent specific organization)
 - l. Who provides/should provide the P4Q budget? Which percentage of the national budget? Where do these resources come from?
5. What are the reported or likely effects of ‘pay for quality’ in your country?

- a. Effects on clinical effectiveness, timeliness and patient safety?
 - b. Effects on care access and equity?
 - c. Effects on continuity and coordination?
 - d. Effects on the care relationship and patient centeredness?
 - e. Effects on efficiency and cost effectiveness?
 - f. Which unintended consequences should be expected?
6. What is your opinion about P4Q programs? (desirability and feasibility)
 - a. Which advantages could be gained (for the patient, for the providers, for policy makers (e.g. hospital directors), for the state/payer, for the society)
 - b. Do you quote disadvantages? (for the patient, for the providers, for policy makers (e.g. hospital directors), for the state/payer, for the society)
 - c. What kind of barriers were experienced or do you expect?
 7. Which specific health system characteristics have largely influenced or will influence P4Q design, implementation and effects in your country? (general payment system incentives, level of competition, level of fragmentation, medical culture, state influence, etc.)
 8. What does the future hold for 'pay for quality' in your country? Are there plans to set up further or new P4Q programmes? If yes, in which way will they be different from initiatives in other countries? Which lessons were learned from previous experiences nationally and/or in other countries?

If there are no further P4Q plans, what are the reasons not to use P4Q interventions?

9. What is your view on the current P4Q research status and its future evolution? Which recommendations do you formulate for the next five years national and international research agenda?
10. Which key recommendations do you formulate as an advice to a country at the first initial stage of considering the implementation of P4Q?

APPENDIX 12 FEASIBILITY ASSESSMENT CURRENT QUALITY CYCLES

	quality dimensions		quality targets measured											data co
	(safety: SA, access: AC, effectiveness: EFFECT, patient centeredness: PC, timeliness: TI, equity: EQ, efficiency: EFFIC, continuity: CO, provider experience: PE)	setting, type of care (preventive, acute, chronic), medical conditions/generic	structure	proces	outcome		N	appropriate (A)/ inappropriate (I) care focus	Selection criteria				static/dynamical target selection	method
					intermediate	long term			clinical evidence base	SMART	room for improvement	cost effectiveness		
Starting from scratch	first phase: EFFECT + monitoring EQ and PE, incentivizing other dimensions later on	primary and hospital care, mix of care types, condition specific + generic	x	x	x	Related using intermediate measures	What is a sufficient but not overwhelming number? (phased, but not starting too low)	first phase: A, later on A + I	Only high level of evidence targets	Specific, measurable, actionable, realistic, timely	based on local baseline and continuous measurement	based on target specific health gain per unit of expense	dynamical approach	automatic extraction wherever possible + additionally sampled approach?
Starting from existing initiatives														
Care itineraries	EFFECT + CO	primary + hospital care, chronic care focus, medical condition specific	/	x	x	/	four and five measures	A	x	x	/	unclear	dynamical	automatic extraction planned
Clinical pathways	all dimensions possible (hospital or practice specific)	primary + hospital care, acute + chronic care, medical condition specific	/	x	x	x	Limited number of targets (clinical pathway specific)	A + I	pathway specific	x	x	unclear	dynamical	sampled approach
Providers' accreditation	EFFECT + EFFIC + CO	Primary and hospital care, mix of care types, generic	/	/	/	/	/	A + I	/	/	/	/	/	/
Breast cancer screening prevention bonus	EFFECT + TI + CO	primary care (GPs and gynaecologists), preventive, medical condition specific	/	x	/	/	Very limited number of targets (one)	A	yes	x	no	unclear	static	secondary data analysis
Capitation funding reevaluation in the medical houses	EFFIC	primary, mix of care types, generic	/	x	/	/	three target groups: medical imaging, clinical biology, hospitalization	I	no	no	no	unclear	static	secondary data analysis
EPA tool	SA + AC + EFFECT + PC + TI + EFFIC + CO + PE	primary care, mix of care types, generic	x	x	/	/	199 indicators	A	no	x	yes	unclear	dynamical	combined approach
Global medical record	CO + PC	primary care, mix of care types, generic	/	x	/	/	having one registered contact per patient per year as the only indicator	NA	NA	NA	NA	unclear	static	secondary data analysis
Prescription feedback	EFFECT + EFFIC	primary care, mix of care types, medical condition specific + generic	/	x	/	/	various classes of target drugs	I	yes	yes	no	unclear	dynamical	secondary data analysis
Preventive module in global medical record	EFFECT + CO + PC	primary care, preventive care, medical condition specific	/	x	/	/	limited number of targets	A	yes	x	no	unclear	static	sampled approach
Centres of reference	SA + EFFECT + PC + TI	hospital care, acute + chronic care, generic + medical condition specific	x	x	/	x	limited number of targets	A	partial	x	no	unclear	static	unclear
Hospital accreditation	all dimensions possible (program specific)	hospital care, acute + chronic care, medical condition specific + generic	x	x	/	/	limited number of targets (program specific)	A + I	program specific	x	program specific	unclear	dynamical	combined approach
Hospital benchmarking	all dimensions possible (hospital specific)	hospital care, acute + chronic care, medical condition specific + generic	/	x	x	x	large sets, hospital specific	A + I	hospital specific	x	hospital specific	unclear	dynamical	combined approach
Quality and patient safety in hospitals	all dimensions possible (hospital specific)	hospital care, acute + chronic care, medical condition specific + generic	x	x	x	x	limited number of targets (hospital specific)	A + I	hospital specific	x	hospital specific	unclear	dynamical	hospital specific
Reference payment hospitals	EFFIC	hospital care, acute + chronic care, medical condition specific	/	x	/	/	three target groups: medical imaging, clinical biology, technical services	I	no	no	no	unclear	static	secondary data analysis

	lection	quality measurement			P4Q incentive							
		validity	case mix adjustment for outcome measures	exception reporting	monitoring unintended consequences	Financing related	Revenue and/or cost (positive and/or negative within the financial structure)	Explicit/implicit	Performance related or independent; measuring and/or reporting related	Reward/penalty	Incentive level	Relative/absolute
Starting fr	Health care record as most valid source? Existing databases?	risk adjustment or specific subgroup comparisons only for outcome measures; Structure and process measures are not casemix dependent,	using a standardized or variable approach? Additional specific data collection and analysis	Monitoring patient equity, monitoring provider equity, monitoring quality target equity	Yes, combined with non financial incentives	Incentive as a positive revenue, with cost coverage as a minimum	Explicit	Phased: measuring related (1), reporting related (2), performance related (3)	Reward	individual + team	Absolute (no ranking competition)	threshold + improvement
Starting fr initiatives												
Care itinal	based on health care record	yes	no	no	yes	fixed fee independent of costs	explicit	measuring and reporting related	reward	individual/team	absolute	NA
Clinical pa	based on health care record	yes	yes, using a variable approach (variance analysis, based on pathway deviations)	no	no	only costs, no revenues	only implicit effects	independent	NA	NA	NA	NA
Providers /	/	/	/	/	yes	FFS increase + fixed fee	explicit	independent	reward	individual	absolute	NA
Breast can prevention	based on existing clinical database	NA	no	no	yes	bonus leading to higher revenues	explicit	performance related	reward	individual, but partly based on regional performance	absolute	improvement
Capitation revaluation houses	based on administrative data	NA	no	no	yes	10% of capitation increase	explicit	global performance related	reward	collective (all providers involved)	relative + absolute	improvement
EPA tool	based on self assessment, surveys, inspection, interview	NA	no	no	no	only costs, no revenues	only implicit effects	independent	NA	NA	NA	NA
Global me	based on administrative data/ health care record	NA	no	no	yes	fixed fee independent of costs	explicit	volume related	reward	individual	absolute	improvement
Prescriptio	based on administrative data	NA	no	no	yes	FFS increase	explicit one time incentive	global performance related	reward	individual	relative + absolute	improvement
Preventive medical re	based on health care record	NA	no	no	yes	fixed fee independent of costs	explicit	reporting related	reward	individual	absolute	improvement
Centres of	unclear	unclear	no	no	no	no direct revenue or cost effect	only implicit effects	independent	NA	NA	NA	NA
Hospital a	accreditation body specific	NA	no	no	no	only costs, no revenues	only implicit effects	independent	NA	NA	NA	NA
Hospital b	hospital specific	yes	no, but often exclusions applied	no	no	only costs, no revenues	only implicit effects	independent	NA	NA	NA	NA
Quality an hospitals	hospital specific	hospital specific	no	no	yes	fixed fee independent of costs	explicit	measuring related	reward	hospital	absolute	NA
Reference hospitals	based on administrative data	NA	no	no	yes	refunding of revenues	explicit	performance related	penalty	hospital	relative	threshold

Starting from initiatives	Incentive structure						Implementing and communicating the program					Evaluation of the program	
	Weighting, composite, all or none	Size	Frequency	Stability	Simplicity vs. Complexity	Provider involvement in setting goals	Provider communication and awareness + methods	Mandatory/voluntary	Staged approach	Stand alone / embedded program	Sustainability of change	Validation of the program	
Starting from initiatives	Weighting based on workload and health gain; initially no use of composite or all or none approach	10% of total revenues	Each 3 to 6 months a performance feedback with related incentive	Targets: As long there is locally sufficient room for improvement; Incentive: no intermediate incentive changes	Front office simplicity, back office accuracy	Sufficient representation; scientific + professional group input	Direct and intensive provider communication of the program	Voluntary	Beforehand modelling and pilot testing	Embedded in a program providing quality improvement tools and support	Long term follow up, from quality improvement to quality maintenance	Evaluation of effects, using scientifically sound methods	
Care itin	NA	Yearly 80 euro per patient	Yearly	ongoing, premature at present	High level of simplicity	Sufficient	ongoing, premature at present	voluntary	Staged	Embedded	ongoing, premature at present	ongoing, premature at present	
Clinical pa	NA	NA	NA	NA	NA	sufficient	high level	Voluntary	staged	embedded	addressed	regular	
Providers	NA	limited as compared to total provider revenues (< 10%)	per service + yearly	sufficient	High level of simplicity	sufficient	high level	Voluntary	no?	embedded	unclear	exceptional	
Breast can prevention	NA	limited as compared to total provider revenues (< 10%), but with volume component	two years interval	sufficient	High level of simplicity	unclear	limited	Voluntary	no	stand alone	ongoing, premature at present	ongoing, premature at present	
Capitation revaluatio	NA	sufficient	yearly	sufficient	High level of simplicity	no involvement	unclear	mandatory	no	stand alone	addressed	sporadic	
EPA tool	NA	NA	NA	NA	NA	sufficient	unclear	Voluntary	no	embedded	ongoing, premature at present	ongoing, premature at present	
Global me	NA	limited as compared to total provider revenues (< 10%)	yearly	sufficient	High level of simplicity	unclear	high level	Voluntary	yes	stand alone	addressed	regular	
Prescriptio	NA	limited as compared to total provider revenues (< 10%)	one time	one time	High level of simplicity	unclear	high level	mandatory	no	embedded	ongoing, premature at present	ongoing, premature at present	
Preventive medical re	NA	limited as compared to total provider revenues (< 10%)	yearly	sufficient	to be determined	ongoing, premature at present	ongoing, premature at present	Voluntary	yes	embedded	ongoing, premature at present	ongoing, premature at present	
Centres of	NA	NA	NA	NA	NA	unclear	unclear	NA	no	stand alone	ongoing, premature at present	ongoing, premature at present	
Hospital a	NA	NA	NA	NA	NA	sufficient	high level	Voluntary	staged	embedded	addressed	regular	
Hospital b	NA	NA	NA	NA	NA	sufficient	unclear	Voluntary	no	stand alone	addressed	regular	
Quality an	NA	limited as compared to total hospital revenues (< 10%)	yearly	ongoing, premature at present	High level of simplicity	sufficient	high level	voluntary	yes	embedded	ongoing, premature at present	regular	
Reference hospitals	NA	limited as compared to total hospital revenues (< 10%)	yearly, with intermediate feedback	ongoing, premature at present	lacking	lacking	limited	mandatory	no	stand alone	ongoing, premature at present	ongoing, premature at present	

APPENDIX 13 LIST OF STAKEHOLDERS

GROUP	NAME	
PHARMACY	Dirk Broeckx	Secretaris-Generaal van APB
PRIMARY CARE	Jos Desmedt	Huisarts, Voorzitter directiecomité van Domus Medica
	Geneviève Bruwier	Médecin généraliste, Professeur DUMG-ULg, Vice-présidente de la SSMG, Vice présidente du FAG
	Marco Schetgen	Médecin généraliste, Professeur CUMG-ULB, Conseiller de la Ministre Onkelinx
	Bernard Vercruysse	Médecin généraliste, Titulaire de la Chaire de Médecine générale UCL, Ancien conseiller du ministre Demotte, Vice-président du FAG
	Piet Vandebussche	Huisarts, Professor UGent
SECONDARY CARE	Daniel Désir	Endocrinologue, Directeur général médical / Médecin-chef du CHU Brugmann
	Johan Kips	Internist, Algemeen directeur van UZ Leuven
	Jacques De Toeuf	Chirurgien, Directeur Général Médical du CHIREC, vice président de l'ABSYM
	Robert Rutsaert	Nefroloog, Voorzitter van ASGB
	Guy Durant	Professeur UCL, Administrateur général des Cliniques Saint-Luc (UCL)
	Johan Hellings	Verpleegkundige, Professor UHasselt, Algemeen directeur van Ziekenhuis Oost-Limburg
	Alain De Wever	Professeur ULB, Directeur de "De Wever Health Care Consulting" (DWHCC), Ancien Directeur médical des hôpitaux Brugmann et Erasme, Ancien conseiller des ministres Busquin et Moureaux
	Peter Degadt	Economist, Gedelegeerd bestuurder Zorgnet Vlaanderen
FEDERAL GOVERNMENT HEALTH CARE	Michel Van Hoegaarden	FOD Volksgezondheid, Veiligheid van de Voedselketen en Leefmilieu, Directeur Generaal van Basisgezondheidszorg en Crisisbeheer (DG2)
	Christiaan Decoster	FOD Volksgezondheid, Veiligheid van de Voedselketen en Leefmilieu, Directeur Generaal van Organisatie Gezondheidszorgvoorzieningen (DG1)
REGIONAL AND COMMUNITY GOVERNMENTS	Chris Vander Auwera	Vlaams Agentschap Zorg en Gezondheid, Administrateur Generaal
	Walter Van Den Eede	Vlaams Agentschap Inspectie, Welzijn, Volksgezondheid en Gezin, Voormalig Administrateur Generaal
	Roger Lonfils	Communauté française de Belgique, Direction générale de la Santé, Direction de la Promotion de la Santé
INSURERS	Catherine Lucet	Médecin spécialiste en gestion des systèmes de santé, Attachée à la Direction Etudes des Mutualités Socialistes
	Jean-Marc Laasman	Economiste, Directeur du service d'études de l'Union Nationale des Mutualités Socialistes
	Xavier De Bethune	Médecin généraliste, Responsable des initiatives qualité, Alliance Nationale des Mutualités Chrétiennes, membre du conseil de direction du CEBAM, Département de Santé publique Institut de médecine tropicale Anvers, Maître de conférences invité Ecole de santé publique UCL
	Marc Justaert	Voorzitter Landsbond der Christelijke Mutualiteiten

	Didier de Laminne de Bex	Administrateur-directeur du DKV Belgium
	Piet Calcoen	Geneesheer, Advocaat, Medisch Directeur van DKV Belgium
RIZIV/INAMI	Ri De Ridder	Directeur Generaal van de Dienst voor Geneeskundige Verzorging, RIZIV
	Jo De Cock	Administrateur Generaal RIZIV
	Pascal Meeus	Médecin Généraliste, INAMI, Direction Recherche, Développement et Promotion de la Qualité
SCIENTISTS	Bert Aertgeerts	Huisarts, Directeur CEBAM, Professor en Afdelingshoofd Academisch Centrum voor Huisartsgeneeskunde , KULeuven
	Jan De Maeseneer	Huisarts, Professor UGent
	Dominique Pestiaux	Médecin Généraliste, Professeur CAMG-UCL
	Michel Roland	Médecin Généraliste, Professeur CUMG-ULB, Service de promotion de la santé et de la qualité Fédération des Maisons médicales, Membre du conseil de direction du CEBAM
	Arthur Vleugels	Geneesheer, Afdelingshoofd Centrum voor Ziekenhuis- en Verplegingswetenschap,, KULeuven
	Pierre Gillet	Médecin Généraliste, Professeur ULg, Médecin-chef adjoint du CHU de Liège, Président du conseil d'administration du KCE
	Erik Schokkaert	Economist, Professor KULeuven
TRADE UNIONS	Michel Vermeylen	Médecin Généraliste, Président de l'Association des Médecins de Famille (AMF)
	Philippe Vandermeeren	Médecin Généraliste, Président du Groupement belge des omnipraticiens (GBO)
	Pierre Drielsma	Médecin Généraliste, Responsable du Service d'études et de recherches, Fédération des Maisons médicales
	Marc Moens	Geneesheer, Ondervoorzitter van de Belgische Vereniging van Artsensyndicaten (BVAS)
PATIENT ORGANISATIONS	Micky Fierens	Présidente de la Ligue des Usagers des Services de Santé (LUSS)
	Carine Serano	Chargée de communication, Ligue des Usagers des Services de Santé (LUSS)
	Ilse Weeghmans	Coördinator van het Vlaamse Patiëntenplatform

APPENDIX 14 STAKEHOLDERS QUESTIONNAIRE

Questionnaire for stakeholders

Before the beginning of the interview, the interviewer proposes a small synthesis of the principles of a P4Q programme in order to be sure the interviewee is sufficiently informed on the subject in terms of P4Q definition, P4Q development, implementation and evaluation, and international examples of dissemination.

Questions	Answers
Global understanding of P4Q and its context	
1. Do you understand the subject of the project to a sufficient degree?	
Knowledge about / involvement in P4Q programmes	
2. Have you already heard of P4Q programmes?	
2.1. Which one?	
2.2. In which country was it developed?	
2.3. Have you got a report about it? <i>If yes, make sure that all items listed below are included in the report. We want a first impression on their knowledge of the subject. So, some further questions should be asked (maybe not all, depending on what the participant shares spontaneously).</i>	
2.4. Can you describe it?	
2.4.1. Which were the goals? <i>Quality dimensions addressed?</i>	
2.4.2. Which were the targets? <i>Patient groups, professionals, care setting?</i>	
2.4.3. Who was involved in defining targets and goals? <i>State, payer, provider, patient, academic...</i>	
2.4.4. What kind of incentives was proposed? <i>Reward / withhold, size, receiver level?</i>	
2.4.5. Was it implemented? How? <i>Level of local involvement, communication?</i>	
2.4.6. Was it assessed? How? <i>By whom, data collection method?</i>	
2.4.7. How was the quality measured? <i>Indicator set, risk adjustment for outcomes, exception reporting, data availability and quality?</i>	
2.4.8. What were the reported effects? <i>Effectiveness, timeliness and safety, care access and equity, continuity and coordination, patient centeredness and satisfaction, efficiency, productivity and cost effectiveness.</i> <i>If a participant focuses on programs only focused on efficiency</i>	

<p>or productivity, without a relation with EB indicators or outcomes on the other quality domains, then it should be made clear to the participant that this doesn't comply with the study definition of P4Q.</p>	
<p>3. Have you already been involved in the development of a P4Q programme (locally, regionally, nationally or internationally)? If no, skip to question 4.</p>	
<p>3.1. Can you tell me more about it?</p>	
<p>3.2. Have you got a report about it?</p>	
<p>3.3. Can you describe it? If not addressed sufficiently in question 3.1</p>	
<p>3.3.1. Which were the goals? Quality dimensions addressed?</p>	
<p>3.3.2. Which were the targets? Patient groups, professionals, care setting?</p>	
<p>3.3.3. Who was involved in defining targets and goals? State, payer, provider, patient, academic...</p>	
<p>3.3.4. What kind of incentives was proposed? Reward/withhold, size, receiver level?</p>	
<p>3.3.5. Was it implemented? How? Level of local involvement, communication?</p>	
<p>3.3.6. Was it assessed? How? By whom, data collection method?</p>	
<p>3.3.7. How was the quality measured? Indicator set, risk adjustment for outcomes, exception reporting, data availability and quality?</p>	
<p>3.3.8. What were the reported effects? Effectiveness, timeliness and safety, care access and equity, continuity and coordination, patient centeredness and satisfaction, efficiency, productivity and cost effectiveness.</p>	
<p>Now, imagine that a P4Q programme is proposed in Belgium</p>	
<p>Personal perceptions about P4Q programmes</p>	
<p>4. What is your opinion about the introduction of such kind of programme in Belgium? Desirability, feasibility. This part will already induce many answers depending on how the program is developed and implemented. This is the case for all potential advantages and disadvantages. If this is indicated by a participant further clarification of conditions should be sought, wherever possible, also when this comes only later in the proposed structure.</p>	
<p>5. Suppose this kind of programme is implemented in Belgium, which advantages could we gain?</p>	
<p>5.1. For the patients? Clinical safety, patient, access / equity and continuity / coordination domain depending on goals and targets.</p>	
<p>5.2. For the providers? Work experience, commitment and satisfaction; recognition,</p>	

<p><i>fairness and equity in payment, refocusing on core values (quality as a higher weight than productivity), inter provider cooperation and teamwork, higher provider inflow e.g. In primary care due to more balanced incentivized workload. Let each respondent answer from his/her professional point of view.</i></p>	
5.2.1. In primary care?	
5.2.2. For specialists in ambulatory care?	
5.2.3. For specialists in a hospital setting?	
5.2.4. For other professional groups? <i>Nursing, dieticians, psychologists, etc.</i>	
5.2.5. For provider organization policy makers such as hospital directors? <i>Alignment of incentives with professional providers, more predictable and focused steering capacity.</i>	
5.3. For the state and payers? <i>National Institute for Sickness and Invalid Insurance (NISII) (INAMI/RIZIV), insurance funds? (alignment of incentives with professional providers, more predictable and focused steering capacity)</i>	
5.4. For the society? <i>All above, cost effectiveness?</i>	
6. Do you quote disadvantages?	
6.1. For the patients? <i>Negative effects on clinical domains, patient selection in service delivery, admission, length of stay,... as an effect on equity, less attention for patient preferences, less attention for holistic and individualized care, ethics of internal vs. external motivation, negligence of high co morbidity patients, negligence of local quality priorities, negligence of unincentivized quality targets.</i>	
6.2. For the providers? <i>Withhold effects if negative incentives or ranking. Mattheus effect: rich richer, poor poorer depending on quality scores, therapeutic freedom, inappropriate responsibility allocation between providers, overly control between providers, workload of administration/paperwork. Let each respondent answer from his/her professional point of view.</i>	
6.2.1. In primary care?	
6.2.2. For specialists in ambulatory care?	
6.2.3. For specialists in a hospital setting?	
6.2.4. For other professional groups? <i>Nursing, dieticians, psychologists, etc.</i>	
6.3. For provider organization policy makers such as hospital directors?	
6.4. For the state and payers <i>National Institute for Sickness and Invalid Insurance (NISII) (INAMI/RIZIV), insurance funds? (budget equilibrium)</i>	
6.5. For the society? <i>All above, cost effectiveness?</i>	
7. Which perverse effects should be expected? <i>Data gaming, documentation improvement only.</i>	
8. What kind of barriers do you expect? <i>Shortcomings of quality measurement (what is quality issues, how to measure issues, IT issues), lobbying of stakeholders</i>	

depending on own interests in the current system.	
Goals	
9. What kind of authorities should define the goals of P4Q in Belgium?	
9.1. Political level: <i>Guard that this discussion doesn't take too much of the interview time. This isn't the main subject.</i> <i>Levels to be considered are Federal, Regional, Community and Province.</i>	
9.2. Scientific level?	
9.2.1. Institute of Public Health? (ISP/WIV)	
9.2.2. Health council? (CSS/HGR)	
9.2.3. Scientific societies of specialities?	
9.2.4. Universities?	
9.2.5. Other?	
9.3. Providers' representatives?	
9.3.1. Trade unions?	
9.3.2. Scientific societies?	
9.3.3. Hospitals associations?	
9.3.4. Local associations?	
9.4. Payers' representatives?	
9.4.1. Public? <i>State, National Institute for Sickness and Invalid Insurance (NISII) (INAMI/RIZIV)</i>	
9.4.2. Sick funds?	
9.4.3. Private insurers?	
9.5. Patients' representatives?	
10. How should this kind of programme define goals?	
10.1. How to choose priorities in terms of patient groups and measures? <i>Volume, cost, identified quality gaps, epidemiological evolution, availability of measures.</i>	
10.2. From which data could the goals be derived?	
10.2.1. Crossroads Bank for Social Security? (BCSS/KSZ)	
10.2.2. Data from the <i>National Institute for Sickness and Invalid Insurance (NISII) (INAMI/RIZIV)</i> ? <i>Claims data, Financial data (MFG/RFM)</i>	
10.2.3. Minimal medical (MKG/RCM) and nursing (MVG/RIM) datasets of hospitals?	
10.2.4. Pharmanet? (Pharmaceutical Policy Management Unit – Healthcare Department)	
10.2.5. Data to be collected locally? <i>Patient's electronic records.</i>	
10.2.6. Other?	
10.3. Should the goals be defined from the current situation or from levels defined by the professionals or the authorities?	
10.4. Should the goals take into account setting specific characteristics (e.g. primary care vs. hospital care)? Can these goals complement each other	

between settings?	
10.5. Can you quote some examples of goals to focus upon in a P4Q program?	
Targets	
11. Which parts of the healthcare system should be targeted first?	
11.1. General and/or mental health care?	
11.2. Nursing homes, community health networks, etc.?	
11.3. Physician professional providers?	
11.4. Non physician professional providers? (Psychologists, dieticians...)	
12. Should it be implemented in primary/ambulatory care or in hospital care or both?	
12.1. For primary/ambulatory care: solo practices, group practices, both?	
12.2. For hospitals: by department, by specialty, by interdisciplinary team, as a whole, or a combination?	
13. Should it target incentives to :	
13.1. Individuals?	
13.2. Teams (group practice in primary care, interdisciplinary team in hospital care)	
13.3. Provider organizations as a whole (hospital)	
13.4. A patient trajectory (multiple settings involved)?	
14. Do you think patients should also receive some incentives?	
15. How should you rank the following dimensions of quality?	
15.1. Effectiveness of clinical/preventive care?	
15.2. Effectiveness of interpersonal relationship?	
15.3. Continuity and coordination of care?	
15.4. Patient safety?	
15.5. Access to care?	
15.6. Equity?	
15.7. Patient centeredness? <i>Respect of rights and preferences</i>	
15.8. Efficiency and Cost-effectiveness?	
16. Should it target specific diseases?	
16.1. Which ones? <i>Prevention, acute/chronic care.</i>	
16.2. Why?	
17. Do you think about some priorities?	
P4Q incentives	
18. What kind of incentives would you prefer?	
18.1. Should P4Q make use of rewards and/or withholds?	
18.2. Should providers' performance be compared in the P4Q design (competitive ranking) and/or be assessed separately?	

18.3. Should P4Q be directed at personal income, at practice investment stimulation or at both?	
18.4. Non financial next to financial incentives? <i>Recognition, public reporting, accreditation, training time, less control, etc.</i>	
19. Should P4Q be mixed with other quality incentives? <i>Development of practice assessment, benchmarking, etc.</i>	
20. What percentage part of the total income of the care providers could be ensured by P4Q? <i>A low part can be inefficient; a high part can be unaccepted. About 5, 10 and 25% are used abroad</i>	
20.1. A minor part?	
20.2. A balanced part?	
20.3. A major part?	
Implementing and communicating the programme	
21. What kind of structures should organize and manage P4Q in Belgium?	
21.1. Existing structures?	
21.1.1. State/government?	
21.1.2. National Institute for Sickness and Invalid Insurance (NISII)? (INAMI/RIZIV)	
21.1.3. National Council for Quality Promotion? (CNPQ/NRK)	
21.1.4. Insurance funds?	
21.1.5. Local organizations (e.g. hospital)?	
21.1.6. Other?	
21.2. Structures to be created, developed?	
21.3. At which institutional level? <i>Federal, Regional, Community, Province?</i>	
22. Where should it be discussed before implementation?	
22.1. With which partners?	
23. Should the introduction be phased? How? <i>Demonstration projects, geographically staged, pay for participation/reporting.</i>	
24. Should P4Q be mandatory or introduced on a voluntary basis?	
24.1. Can you cite advantages/disadvantages of each proposal?	
25. Should the recipients be supported in this process?	
25.1. In which setting? (primary care/hospitals)	
25.2. By whom?	
25.3. For which tasks?	
25.3.1. Data collection?	
25.3.2. Data management?	
25.3.3. Defining goals?	
25.3.4. Other?	
26. How should the feedback be organized?	
26.1. Personally?	
26.2. By group practice / hospital ward /	

specialty/team?	
26.3. Live? (as soon as it is possible)	
26.4. Periodically?	
26.5. With/without comments?	
26.6. With/without benchmarking?	
26.7. With/without suggestions for improvement?	
26.8. Other?	
Quality assessment	
27. Who would be responsible for measurements?	
27.1.1. State/government?	
27.1.2. National Institute for Sickness and Invalid Insurance (NISII)? (INAMI/RIZIV)	
27.1.3. Insurance funds?	
27.1.4. National Council for Quality Promotion? (CNPQ/NRK)	
27.1.5. Scientific societies?	
27.1.6. Local associations?	
27.1.7. Local providers or provider organizations?	
27.1.8. Other?	
28. How could the recipients be assessed?	
28.1. Structure assessment?	
28.2. Process assessment?	
28.3. Outcomes?	
28.4. Continuous quality improvement or quality threshold?	
29. Can you quote some examples of measures?	
Evaluation of the programme	
30. How could we know if the P4Q programme is working?	
31. How to make indicators acceptable for the recipients?	
Budgetary impact	
32. At which financial level do you think the Social Security should support P4Q?	
32.1. A minor part of its budget?	
32.2. A major part of its budget?	
33. Do you imagine other sources of funding? Which ones?	
33.1. Direct funding by the Ministry of Health? (SPF/FOD)	
33.2. Care providers?	
33.3. New payroll tax for employees?	
33.4. Return on money saving by quality of care improvement?	
33.5. Other?	
Making sure we have interviewed the major stakeholders of the field	
34. Could you list 3 persons who you think that should be interviewed about this subject?	

Appendixes with Chapter 4 – Evidence base for P4Q

APPENDIX 15 DESCRIPTION OF STUDIES

Author(s)	Publication year	Country	Setting		Study period	Study design	Sample size		
			Pri- mary	Hos- pital			Organizations/ practices	Providers	Patients
Ritchie et al	1992	UK	X		1990-1991	Historic, Multi	95	313	6600 vs. 6400
Kouides et al	1993	US	X	X	1990-1991	Concurrent + Historic	NR	53 vs. 82	12271 vs. 30387
Langham et al	1995	UK	X		1992-1993	Concurrent + Historic	78 vs. 85	NR	NR
Lynch	1995	UK	X		1991-1992	Cross section	208	NR	NR
Morrow et al	1995	US	X		1987-1990	Historic, Multi	1607, 418, 271	NR	50 per practice
Fairbrother et al	1997	US	X		1993-1996	Historic, before-after	NR	23	173 vs. 528
Grady et al	1997	US	X		3 years	Randomized	21 vs. 21 vs. 23	109	11716
Hillman et al	1998	US	X		1993-1995	Randomized	26 vs. 26	NR	NR
Kouides et al	1998	US	X		1990-1991	Randomized	54	NR	NR
Cameron et al	1999	Australia		X	1991-1997	Historic, before-after	21	NR	NR
Fairbrother et al	1999	US	X		1995-1996	Randomized	NR	60	50 per provider
Hillman et al	1999	US	X		1993-1995	Randomized	49	NR	15 per practice
Lebaron	1999	US	X		NR	Cross section	73-116 per state (4), 8-25 per city (2)	NR	4639-18000, 714-5276 per clinic
Safran et al	2000	US	X		1996	Cross section	NR	NR	6018
Cattaneo et al	2000	Italy		X	1998-1999	Historic, before-after	10	NR	9264
Fairbrother et al	2001	US	X		1997-1998	Randomized	NR	57	50 per provider
Shortell et al	2001	US	X		NR	Cross section	56	1797	NR
Bond et al	2002	Australia	X		1997-2000	Cross section	47	NR	1578 vs. 1793
Amundson et al	2003	US	X		1996-1997	Historic, Multi	20	NR	14489

Author(s)	Publication year	Country	Setting		Study period	Study design	Sample size		
			Primary	Hospital			Organizations/practices	Providers	Patients
Casalino et al	2003	US	X		2000-2001	Cross section	1587	NR	NR
Chung et al	2003	US	X		1997-2000	Historic, Multi	NR	1600	NR
Larsen et al	2003	US	X	X	1998-2002	Concurrent + Historic	NR	NR	5785-9463 per year
McMenamin et al	2003	US	X	X	2000-2001	Cross section	1104	NR	NR
Roski et al	2003	US	X		1999-2000	Randomized	15 vs. 15 vs. 10	NR	4813 vs. 4734
Armour et al	2004	US	X		2000-2001	Historic, before-after	NR	NR	3058 vs. 3691
Greene et al	2004	US	X	X	1999-2001	Historic, before-after	NR	900	96766
Hippisley-Cox et al	2004	UK	X		2004	Cross section	237	NR	53687
Li et al	2004	US	X	X	2000-2001	Cross section	1104	NR	NR
McMenamin et al	2004	US	X	X	2000-2001	Cross section	1104	NR	NR
Schmittdiel et al	2004	US	X	X	2000-2001	Cross section	1104	NR	NR
Ashworth et al	2005	UK	X		2001-2002	Cross section	151	NR	NR
Beaulieu & Horrigan	2005	US	X		2001-2002	Concurrent + Historic	NR	21	624 vs. 600
Mentari et al	2005	US		X	2003-2004	Historic, Multi	12	NR	1600
Pourat et al	2005	US	X		2002	Cross section	NR	948	NR
Rosenthal et al	2005	US	X		2003-2004	Concurrent + Historic	134 vs. 33	NR	NR
Shenkman et al	2005	US	X		1999-2001	Cross section	NR	NR	2333
Averill et al	2006	US		X	2000	Cross section	324	NR	49809
Doran et al	2006	UK	X		2004-2005	Cross section	8105	NR	NR
Ettner et al	2006	US	X		2000-2001	Cross section	NR	NR	6194
Grossbart	2006	US		X	2003-2004	Concurrent + Historic	4 vs. 6	NR	4964 and 6025 vs. 8641 and 9295
Jaiveer et al	2006	UK	X		2004-2005	Historic, before-after	13	NR	3453 vs. 3173
Levin-Scherz et al	2006	US	X	X	2001-2003	Concurrent + Historic	8	5100	NR
Reiter et al	2006	US		X	NR	Cross section	66	NR	NR
Reschovsky et al	2006	US	X	X	2000-2001	Cross section	NR	12406	NR
Rittenhouse &	2006	US	X		2003	Cross section	123	NR	NR

Author(s)	Publication year	Country	Setting		Study period	Study design	Sample size		
			Pri- mary	Hos- pital			Organizations/ practices	Providers	Patients
Robinson									
Sigfrid et al	2006	UK	X		2004-2005	Cross section	49	NR	7157
Simpson et al	2006	UK	X		2004-2005	Historic, before-after	310	NR	21901 vs. 32401
Srilangalingam et al	2006	UK	X		2003-2004	Historic, before-after	NR	NR	328 vs. 319
Strong et al	2006	UK	X		2004-2005	Cross section	38	NR	12920
Sutton & McLean	2006	UK	X		2004-2005	Cross section	60	NR	NR
Wang et al	2006	UK	X		2005	Cross section	638	NR	NR
Williams et al ²	2006	US	X		2003-2004	Historic, before-after	225	NR	More than 35000
Williams et al ³	2006	UK	X		2004	Cross section	2	NR	11109 and 9557
Wright et al	2006	UK	X		2004-2005	Cross section	NR	NR	8569
Ashworth et al (a)	2007	UK	X		2004-2005	Cross section	8430	NR	NR
Ashworth et al (b)	2007	UK	X		2004-2006	Historic, before-after	8515 vs. 8264	NR	NR
Campbell et al	2007	UK	X		1998, 2003, 2005	Concurrent + Historic	42	NR	2300, 1495, 1482
Casale et al	2007	US		X	2006-2007	Historic, Multi	NR	NR	137 vs. 117
Coleman et al ⁴	2007	US	X		2002-2004	Concurrent + Historic	NR	46	1166
Coleman et al ⁵	2007	UK	X		1990-2005	Historic, Multi	NR	NR	384259 per year
Downing et al	2007	UK	X		2004-2005	Cross section	94	NR	NR
Gene-Badia et al	2007	Spain	X		2002-2003	Historic, before-after	257	3439 and 3781	200 per PCT
Gilmore et al	2007	US	X		1998-2003	Concurrent + Historic	NR	NR	222213 per year
Glickman et al	2007	US		X	2003-2006	Concurrent + Historic	54 vs. 446	NR	105383
Gray et al	2007	UK	X		2005-2006	Cross section	32	NR	7605
Gulliford et al	2007	UK	X		2005	Cross section	8484	NR	NR
Lindenauer et al	2007	US		X	2003-2005	Concurrent + Historic	207 vs. 406	NR	NR

Author(s)	Publication year	Country	Setting		Study period	Study design	Sample size		
			Primary	Hospital			Organizations/practices	Providers	Patients
McLean	2007	UK	X		2003-2005	Cross section	8214 vs. 1023 vs. 362 vs. 459	NR	NR
McLean et al	2007	UK	X		2004	Cross section	912	NR	NR
Mehrotra	2007	US	X		2005	Cross section	100	NR	NR
Millett et al ⁶	2007	UK	X		2003-2005	Historic, Multi	32	NR	4284
Millett et al ⁷	2007	UK	X		2003-2006	Historic, Multi	32	NR	4284
Millett et al ⁸	2007	UK	X		NR (after 2004)	Cross section	8970	NR	1852762
O'Malley et al	2007	US	X	X	1996-2005	Historic, Multi	7057 vs. 8487	NR	NR
Saxena et al	2007	UK	X		2004-2005	Cross section	8970	NR	2039919, 8970, 6300476, 839758
Shohet et al	2007	UK	X		2004-2005	Cross section	291	NR	14224
Simon et al	2007	US	X	X	2000-2001	Cross section	1104	NR	NR
Simpson et al	2007	UK	X		2005	Cross section	310	NR	32401
Steel et al	2007	UK	X		2003-2005	Historic, before-after	18	NR	586 vs. 570
Tahrani et al	2007	UK	X		2004-2006	Historic, Multi	66	NR	16867
Twardella & Brenner	2007	Germany	X		NR	Randomized	82	94	577
Weber et al	2007	US	X		2006-2007	Historic, Multi	NR	124-136	18511-19494
Young et al	2007	US	X		1999-2004	Concurrent + Historic	NR	334	NR
An et al ⁹	2008	US	X		2005-2006	Randomized	25 vs. 24	NR	NR
Ashworth et al	2008	UK	X		2004-2007	Historic, Multi	8515, 8264, 8192	NR	NR
Bhattacharyya et al	2008	US		X	NR	Cross section	257	NR	NR
Cupples et al	2008	UK	X		2004-2006	Cross section	16 vs. 32	NR	350 vs. 648
Doran et al (a)	2008	UK	X		2004-2006	Cross section	7637	NR	NR
Doran et al (b)	2008	UK	X		2005-2006	Cross section	8105	NR	NR
Greenberg et al	2008	US		X	2001-2003	Cross section	1	63	150
Herrin et al	2008	US		X	2001-2005	Concurrent + Historic	5 vs. 200	NR	13673
Karve et al	2008	US		X	2004-2005	Cross section	3449	NR	NR
MacBride-Stewart et al	2008	UK	X		2002-2006	Historic, Multi	92	NR	NR

Author(s)	Publica-tion year	Country	Setting		Study period	Study design	Sample size		
			Pri- mary	Hos- pital			Organizations/ practices	Providers	Patients
McGovern et al ¹⁰	2008	UK	X		2004-2005	Historic, before-after	310	NR	37329 vs. 56561
McGovern et al ¹¹	2008	UK	X		2004-2005	Historic, before-after	310	NR	58406 vs. 75495
Millett et al ¹²	2008	UK	X		2000-2005	Historic, Multi	15	NR	1968
Millett et al ¹³	2008	UK	X		2003-2005	Historic, before-after	32	NR	2891 vs. 3101
Millett et al ¹⁴	2008	UK	X		2005-2006	Historic, before-after	16	NR	8876
Pearson et al	2008	US	X		2001-2003	Concurrent + Historic	154	5350	NR
Rosenthal et al	2008	US	X	X	2003-2006	Concurrent + Historic	NR	405 vs. 3916, 91 vs. 1204	NR
Steel et al	2008	UK	X		2004-2005	Concurrent	NR	NR	4417
Tahrani et al	2008	UK	X		2004-2006	Historic, Multi	66	NR	16858
Vaghela et al	2008	UK	X		2004-2008	Historic, Multi	8423, 8264, 8192, 8255	NR	NR
McLean et al	2006	UK	X		2005	Cross section	1024	NR	NR

APPENDIX 16A DETAILED DESCRIPTION

Preventive care results

Immunization and vaccination results

There exists strong evidence of the positive effect of Pay for Quality on the influenza immunization rate with an increase in immunization rate between 6.8% and 8.4%^{16, 17}. In a later study, however with a weak design, a positive relationship between Pay for Quality and influenza reminder use was shown with an odds ratio of 1.5¹⁸. There is conflicting evidence, based on studies with a strong design, concerning the effect of Pay for Quality with regard to the children immunization rate (measles, mumps, rubella, diphtheria, tetanus, polio, pertussis, haemophilus influenza type B), ranging from no significant effect to an increase in immunization of 25.3%¹⁹⁻²². These positive findings are confirmed by other studies with a weaker design, with an increase in immunization ranging from no significant effect to an increase of 24.3%²³⁻²⁹.

Children preventive screening

There is a strong indication of absence of effect concerning the effect of Pay for Quality on children preventive screening, like TBC screening, lead screening, sickle cell screening, bacteriuria screening²². In contrast, there was some positive evidence, however based on a study with a weak design, that Pay for Quality had a positive effect on some of these children's preventive screening indicators, with an increase in screening rate for TBC of 28.8% and an increase in screening rate of lead between 16.9% and 23.4%²⁶.

Cancer preventive screening

There exists strong evidence of no significant effect of Pay for Quality on the compliance with guidelines of mammography, mammography referral and colorectal cancer screening^{30, 31}. However a later study, with a weak design, shows a positive effect of Pay for Quality on colorectal cancer screening, with an increase of 3%³². There is conflicting strong evidence concerning the effect of Pay for Quality on breast cancer screening ranging from no significant effect to a increase of 2.2%, and on cervical cancer screening ranging from no significant effect to a 3.9% increase in screening rate^{30, 31, 33-35}. Concerning the mammography reminder use, there is weak evidence of no significant effect of Pay for Quality on the reminder use¹⁸.

Well child visits

There is strong conflicting evidence concerning the effect of Pay for Quality on well child visits⁷ ranging from a decrease in well child visits with 5% to an increase in well child visits with 5%^{22, 33}. A positive effect is shown in one other study, however with a weak design, with an increase in well child visits of 6.6%²⁶.

Sexually transmitted diseases

There is conflicting evidence concerning the effect of Pay for Quality on Chlamydia screening in women. One study with a strong design shows a negative effect on screening rate with a decrease in screening of 11%³³. In contrast, one other study with a weaker design shows no significant association between Pay for Quality and annual Chlamydia screening³⁶. Moreover, there is weak evidence of no association between Pay for Quality and obtaining the sexual history of the patient, between Pay for quality and providing drugs for the partners treatment, and between Pay for Quality and providing services to minors without parental notification/consent³⁶.

Cholesterol screening in adults

Concerning cholesterol screening in adults, there is weak evidence of a positive effect of Pay for Quality on the use of repeated profiling, diet therapy or medication prescription if the cholesterol level is above 6.21mmol/l, with an increase in use of 3%²⁸.

⁷ For your information: the equivalent of well child visits in Belgium are the preventive paediatric consultations organized by Kind & Gezin and ONE.

Acute care results

Emergency care

There is weak evidence for an increase of respectively 10 and 2% on the target less than 20 minutes waiting time for category two patients, and on the target less than 30 minutes waiting time for the category three patients³⁷. In addition, there is weak evidence of no significant effect on zero waiting time in receiving emergency care for category one triage patients, it must be noted that the performance on this indicator was already 100% beforehand³⁷. Finally, there is weak evidence of no significant effect of Pay for Quality on number of patients waiting longer than 12 hours before being admitted to the hospital from the emergency department³⁷.

There is weak evidence of a positive relationship between Pay for Quality and smoking cessation referral rate for heart disease, gastritis, pregnancy and respiratory illness³⁸.

Myocardial infarction/acute cardiac event

With regard to the effect of Pay for Quality on incentivized targets there is strong evidence of no effect on receiving thrombolytic agent within 30 minutes after arrival and there is strong evidence of a positive effect on PCI within 120 min after arrival with an increase of 5.4%³⁹. There is conflicting strong evidence of the effect of Pay for quality on aspirin at arrival, aspirin at discharge, beta blocker at arrival and beta blocker at discharge ranging between no significant to 3.3%, 8.5%, 2.8% and 2.8% respectively. In addition there is conflicting strong evidence for the effect on Angiotensin Converting Enzyme (ACE) inhibitor use for left ventricular systolic dysfunction (LVSD) ranging from no significant effect to an increase of 9.9%, and for the effect on smoking cessation advice, ranging from no significant effect to an increase of 5.2%^{40, 39, 41, 42}.

Concerning the effect of Pay for Quality on not incentivized targets, which have been monitored during the P4Q program to assess any spill over of neglecting effects, there is strong evidence of no effect on heparin use, on glycoprotein IIb/IIIa inhibitor use, on the use of clopidogrel at discharge, on dietary modification counselling, on cardiac rehabilitation referral, on cardiac catheterization within 48 hours, on electrocardiogram (ECG) within 10 min, on the use of thrombolytics within 30 min after arrival, on the use of percutaneous coronary intervention (PCI) within 120 min after arrival and on in-hospital mortality rate^{40, 41}. In addition, there is strong evidence of a positive effect of Pay for Quality on lipid lowering agent at discharge, with an increase of 4.3%⁴⁰.

Coronary Artery Bypass Grafting (CABG)

There is weak evidence of an absence of effect of Pay for Quality on readmission within 30 days, on the number of patients with complication, on the number of patients receiving blood products, on readmission to Intensive Care Unit, on pulmonary complications, on operative mortality, on atrial fibrillation, on deep sternal wound infection, on reintubation during hospital stay, on total ventilation hours, on neurologic complications⁴³. In addition, there is weak evidence for a positive 10% effect on percentage of patients discharged to home⁴³. This study targeted mainly on long term patient outcomes.

Heart failure (acute phase)

There is strong evidence of a large effect of Pay for Quality on provision of discharge instructions, with an increase of 25.5%³⁹. Furthermore there is strong evidence of having no effect on ACE inhibitor use for left ventricular systolic dysfunction (LVSD) and on smoking cessation advice^{39, 42}. In addition, there is conflicting strong evidence concerning the effect on left ventricular failure (LVF) assessment, ranging from a negative effect with a decrease of 2.4% to a positive effect with a 5.1% increase^{39, 41, 42}. There is strong evidence of no effect of P4Q on not incentivized targets⁴¹.

Community acquired pneumonia (acute phase)

There is strong evidence for a positive effect of Pay for Quality on pneumococcal screening and/or vaccination with an effect ranging from 9.5% to 44.7%^{39, 41, 42}. Furthermore, there is also strong evidence of a positive effect on blood cultures with an increase of 3.5% and a negative effect on smoking cessation advice with a decrease of 16.7%. In addition there is conflicting strong evidence for the effect on oxygenation assessment ranging from a negative effect with a 1.9% decrease to no significant effect, as well as for the effect on antibiotic use within 4 hours after arrival, ranging from a negative effect with a 3.2% decrease to a positive effect with a 4.3% increase^{39, 41, 42}.

Acute sinusitis (Primary and hospital care)

One study, however with a weak design, states a positive effect concerning inappropriate antibiotics prescription, with a decrease in inappropriate description of 29%⁴⁴. In addition, there is weak evidence of an effect on first line antibiotics prescription with a 14% increase, and on the use of sinus plain X ray films and sinus computed tomographic scans with a decrease of respectively 28 and 29%⁴⁴. In terms of numbers of consultations, there is weak evidence of a decrease of 31% in allergist consultations and weak evidence of absence of a significant effect on the number of otolaryngologist consultations⁴⁴.

Breastfeeding

Concerning breastfeeding at discharge there is weak evidence of the effect of Pay for Quality in one Italian study. There is an increase in exclusive breastfeeding and full breastfeeding of respectively 12 and 6%. There is a decrease in predominant and complementary breastfeeding with respectively 7 and 5%. Furthermore, there is weak evidence of no effect on the number of mothers giving no breastfeeding⁴⁵.

Chronic care results

Diabetes

CLINICAL EFFECTIVENESS

There is conflicting strong evidence of the effect of P4Q on HbA1c (glycated haemoglobin) testing as a process measure. In eight studies with a strong design this effect ranges from having no significant effect in two studies^{33, 46} to a maximal 25.5% effect size in one of six studies with a positive effect⁴⁷. A positive effect was confirmed in four studies with a weaker design.

HbA1c as intermediate outcome target levels also showed conflicting strong evidence. Results range from no significant effect³⁴ to a maximal 13.9% effect size in one of three studies with a positive effect⁴⁷. A positive effect was confirmed in nine HbA1c targets in studies with a weak design.

There is strong evidence on the absence of effect on hypoglycaemia symptoms recording in patients receiving sulfonurea⁴⁸. Positive weaker evidence is available on the oral hypoglycaemic agent prescription rate (14% effect size) and on the insulin prescription rate (8.1% effect size)⁶. One study showed weak evidence on the absence of an effect in terms of glycaemia control rate⁴⁹.

There is conflicting strong evidence of P4Q affecting the lipid and cholesterol testing rate as process measure. Four studies found a positive effect with a maximal effect size of 25.8%⁵⁰. Three other studies came to no significant effect.

A similar discrepancy was confirmed in studies with a weaker design, four finding a positive effect and three finding an absence of effect.

There is strong evidence of a positive effect on cholesterol and lipid levels as intermediate outcome measures, as supported by three studies with a maximal effect size of 23.5%⁴⁷. These findings are confirmed by four studies with a weaker design, although one study of this type did not find a significant effect and another found a negative difference of 1.3%^{10, 51}.

There is weak evidence for an increasing effect on the lipid lowering drug prescription rate with an effect size of 21.9%⁶.

There is strong evidence of P4Q having no effect on blood pressure recording⁴⁸, although three studies with a weaker design came to positive effects.

There is strong evidence of a positive effect on blood pressure as an intermediate outcome, supported by two studies with an effect size ranging from 1.6 to 6.3%^{47, 48}. This positive finding was confirmed in six studies with a weaker design.

There is weak evidence of a relationship between P4Q use and the advice rate to take aspirin (1.17 relative risk)⁴⁹ and of the effect on ACE inhibitor prescription rate (effect size ranging from 12.2 to 17.1%^{52, 6}).

There is conflicting strong evidence on the effect of P4Q on the nephropathy testing rate. While three studies found a positive effect up to a maximal effect size of 25.6%⁴⁷, three other studies did not find a significant effect. A positive effect has been confirmed in four studies with a weaker design.

With regard to weight recording rate there is strong evidence of a positive effect with an effect size of 1.82%⁴⁸. This is confirmed by the findings of one study with a weaker design.

There is conflicting strong evidence on P4Q effects on the retinal exam rate. Five studies reported a positive effect with a maximal effect size of 25.6%⁴⁷, while one study found no significant effect³³. The positive results are confirmed in two studies with a weaker design, the no effect result is confirmed in one.

There is strong evidence of a positive P4Q effect on the foot exam rate, with an effect size ranging from 2.69 to 45.4%^{47, 48}. However, one study with a weaker design came to no significant effect⁴⁹.

There is strong evidence of a positive P4Q effect on the peripheral pulse testing rate with an effect size of 4.85%⁴⁸. This is confirmed by two studies with weaker design.

There is weak evidence of a positive P4Q effect on the neuropathy testing rate, with an effect size ranging from 42.8 to 59%^{52, 53}.

There is strong evidence of a positive effect on smoking status recording rate, with an effect size of 1.78%⁴⁸. This has been confirmed by three studies with weaker design. One study of this category did not find a significant result⁵¹.

There is weak evidence of a positive P4Q effect on the smoking advice rate, with an effect size ranging from 12 to 35.5% in three studies.

There is strong evidence of an absence of effect of P4Q on the influenza vaccination rate⁴⁶. However, two studies with a weaker design did find a positive effect^{49, 51}.

There is conflicting weak evidence on the effects of P4Q on the use of care management processes. Whereas two studies found a positive relationship with a regression coefficient ranging from 0.20 to 0.41^{54, 55}, one study didn't find a significant association¹⁸.

CARE COORDINATION

With regard to care coordination and integration there is weak evidence of a positive effect of P4Q on the referral rate for poor glycaemia control (23% effect size) and a 0.9% decrease in glycaemia threshold for referral⁵⁶. This was not associated with a change in the global number of referrals or the number per reason for referral.

Heart failure (chronic phase)

There is weak evidence of a positive effect of Pay for Quality on heart failure treatment in primary care, with an 23.4% increase in Angiotensin Converting Enzyme (ACE) inhibitor or Angiotensin Receptor Blocking (ARB) use²⁵.

Coronary heart disease (CHD)

There is strong evidence of an absence of effect of the QOF system on recording of angina attacks, on recording of exercise capacity, on recording of referral to specialist for exercise stress testing or assessment, on blood pressure and serum control to a fixed target level, and on recording of weight advice for overweight patients⁴⁸. On the contrary, another study, however with a weak design, shows a positive effect of the QOF system on recording of exercise testing or assessment by specialist and on recording of controlled blood pressure with an effect size of respectively 2.3 and 0.7%¹¹. For the recording of controlled cholesterol level this study reports a negative effect, with a decrease of 10.8%¹¹. In addition, there is strong evidence of a positive effect of Pay for Quality on the recording of blood pressure and cholesterol level, with an effect size of respectively 1.3 and 1.71%⁴⁸. This is being confirmed by one other study, however with a weak design with an increase of respectively 21.5 and 41.7%¹¹. Furthermore, there is strong evidence of a positive effect on recording of dietary advice, on recording of prescription/advice to take aspirin and on recording of smoking status, with an increase of respectively 1.9%, 2.04% and 2.39%⁴⁸. One other study, with a weak design confirms a positive effect for the last indicator with an increase of 26.2%¹¹. This study shows also a positive effect on smoking advice, on prescription of anti platelet/anticoagulant therapy, on beta blocker and ACE prescription and on influenza vaccination, with an increase of respectively 15.2%, 24.5%, 27.4%, 11.5% and 28.1%¹¹.

Concerning coronary heart diseases, there is also weak evidence for a positive association between higher scores on the additional service domain within QOF and emergency admission rate with an odds ratio of 1.03 in one primary care trust, whereas there is no evidence for this association in another primary care trust. Moreover, there is also weak evidence of absence of a relationship between scores on the clinical domain or on the organisation domain within QOF and emergency admission rate⁵⁷.

Stroke (including CVA and TIA in the chronic phase)

There is some weak evidence of P4Q having no effect on the achievement of the blood pressure target and cholesterol target, although there is a positive effect on blood pressure and cholesterol recording, with an increase of respectively 33.4% and 52.1%⁵⁸. Moreover there is an increase of 43.3% in MRI/CT scan recording, of 49.4% in recording of smoking status, of 17% in smoking cessation advise, of 32.2% in anti platelet or anticoagulant therapy, of 34.2% in flu vaccination and 28.8% in body mass index (BMI) recording⁵⁸. In addition, there is weak evidence of no relationship between total QOF score for stroke and adherence to evidence based clinical practice guidelines for stroke care, and between higher scores on clinical domains or organisation domains and emergency admission of stroke patients^{3, 57}. Furthermore, there is weak evidence of a positive relationship between higher scores on the additional services domain and the emergency admission rate of stroke patients with an odds ratio ranging from 1.02 and 1.05⁵⁷.

Asthma

There is strong evidence of an absence of effect of Pay for Quality on asthma controller use and on recording of peak expiratory flow or difficulty using this instrument⁵⁹, but there is strong evidence of a positive effect on recording of daily, nocturnal of activity-limited symptoms with an increase of 2.48%, there is also a positive effect on recording of smoking status with an increase of 1.80% and on recording of inhaler technique with an increase of 2.03%⁴⁸. There is weak evidence of a positive effect on the percentage of quality indicators achieved, with an increase of 14%⁶⁰. Furthermore, there is weak evidence of a positive relationship between Pay for quality and use of different organizational indicators, like use of registries, reminders, medical flow sheets etc. with a correlation index of 0.41⁵⁵. There is some conflicting evidence in a study with a weak design on the association between higher scores on the additional services domain and the emergency admission rate of asthma patients, ranging from no relationship in one primary care trust, to a positive association with an odds ratio of 1.04 in one other primary care trust⁵⁷.

Finally, there is some weak evidence of no relationship between higher scores on the clinical domain and the emergency admission rate of asthma patients, between higher scores on the clinical domain and the emergency admission rate, and between higher scores on the organisation domain and the emergency admission rate⁵⁷.

Hypertension

Concerning hypertension, there is some weak evidence of the effect of P4Q on the percentage of quality indicators achieved, with an increase of 12%⁶⁰.

Smoking cessation

There is strong evidence of an absence of effect of Pay for Quality on the percentage of smokers receiving advice or assistance to quit and on smoking abstinence^{61, 62}. In addition, there is strong evidence of a positive effect of Pay for Quality on the percentage of patients whose tobacco use status has been identified, with an increase of 7.9% and on the referral rate of smoking patients, with an increase of 6.2%. Some weak evidence can be found for a positive effect of P4Q on the recording of the smoking status with an increase of 24% and a positive rate ratio of 1.88 and on the advice to quit-rates with an increase of 21% and a positive rate ratio of 3.03^{5, 63}.

Finally, there is weak evidence that receiving better contracts based on quality performance is not related to the perceived need and provision of smoking cessation interventions. When quality was aligned with additional income, then the evidence became mixed, with a positive relationship with some perceptions and provisions and no relationship with others.

When financial incentives were closely linked to providing target interventions themselves, then both types of relationships became positive⁶⁴.

Depression/mental illness

There is weak evidence of an absence of effect of Pay for Quality on the percentage of quality indicators achieved concerning depression/mental illness⁶⁰.

Epilepsy

There is weak evidence of within the QOF contract of a positive relationship between the proportion of seizure free epilepsy patients and the proportion of epilepsy related emergency hospitalisations over the number of epilepsy-treated patients with a regression coefficient of 0.30⁶⁵.

Kidney disease/Chronic kidney disease

There is weak evidence that there was no significant P4Q effect on some targets (Kt/V > 1.2 target, shortening of treatments > 10% per patient month, hospital admission rate). However, there was an increase in the following targets: number of visits, Kt/V, and ultrafiltration volume.⁸ Furthermore, the number of patient months with zero visits, albumin rate, albumin > 3.8 g/dl, haemoglobin rate, haemoglobin > 11 g/dl, the phosphorus rate, the calcium rate, the number of patients with catheters and the number of treatments skipped all decreased significantly⁶⁶.

Osteoarthritis

There is weak evidence of an absence of effect of pay for quality on the percentage of quality measures achieved concerning osteoarthritis⁶⁰.

⁸ Kt/V is a way of measuring dialysis adequacy.

Generic findings

Chronic disease management for multiple patient groups

There is strong evidence of a positive association between overall achievement on prevention and chronic care indicators and visiting only P4Q participating physicians with an odds ratio ranging from 1.06 to 1.27²¹. In addition there is weak evidence that having a quality improvement initiative targeting an indicator is positively associated with having a P4Q incentive tied to the measure with an odds ratio of 1.6⁶⁷. Furthermore, there is strong evidence of no significant relationship between P4Q quality performance and mortality rate, a not incentivized quality measure⁶⁸. This has been confirmed by one other study, however with a weak design⁵⁷. Additionally, there is weak evidence of no significant relationship between performance on incentivized targets and the performance on not incentivized targets in the same target condition. However, concerning the not incentivized indicators, there is some weak evidence of a positive relationship between meeting the cholesterol target for diabetes and CHD and the statin prescribing volume with a regression coefficient of respectively 0.26 and 0.08⁶⁷.

Quality and outcomes framework effect

There is weak evidence that the percentage of indicator achievement is 16.9% less for indicators excluded from the QOF, compared to the indicators included in the QOF⁶⁹. Furthermore there is weak evidence of a positive effect on the overall QOF achievement with an increase of 4.2%, and a positive effect on drug prescription, with an increase between 0.69 and 1.09%^{70,71}.

Use of care management processes

There is weak evidence of a positive relationship between P4Q programs and the use of IT related care management processes, with an increase ranging from 9 to 27%². Likewise, there is weak evidence of a positive relationship between the reported effect of clinical practice guidelines on clinical practice and P4Q with a regression coefficient of 0.09⁷². Furthermore, there is conflicting evidence (however based on studies with a weak design) concerning the relationship between P4Q and a wide array of care management processes, ranging from no significant relationship to a positive relationship with a regression coefficient up to 0.74. It can be noted that the combination of several care management processes gives a higher score^{55, 73-75}. Additionally, there is weak evidence of no association between income for quality and the use of health care promotion programs and weak evidence of a positive association between the external incentives index and adoption of guidelines in order entry systems with decision support, with an odds ratio of 1.12 and a regression coefficient of 0.119^{76,77}.

Patient and provider satisfaction

There is weak evidence of absence of effect of P4Q on overall patient satisfaction in general as well as on the sub dimensions of patient satisfaction. Furthermore, there is absence of an effect of P4Q on physicians' and nurses' perceived overall quality of life, and on nurses' perceived demands⁷⁸. Additionally, there is some weak evidence (however based on studies with a weak design) of absence of effect of P4Q on physicians' perceived intrinsic motivation, on nurses' perceived management structure and on nurses' perceived intrinsic motivation. There is weak evidence of a positive effect of P4Q on physicians' perceived management structure support and on physicians' perceived demands⁷⁸. In addition, there seems to be no relationship between financial incentives as a function of patient satisfaction and continuity, integration of care, clinical interaction, interpersonal treatment and trust⁷⁹. However, there seems to be weak evidence of a positive relationship between these financial incentives and perceived access to care, patient knowledge and use of preventive counselling with a regression coefficient of respectively 2.57, 2 and 3.5⁷⁹. Furthermore, there is weak evidence of a positive relationship between provider perceived incentives to increase services and the provider perceived ability to provide quality of care. Finally, there is weak evidence a negative relationship between provider perceived incentives to reduce services and the provider perceived ability to provide quality of care⁸⁰.

APPENDIX I6B EVIDENCE TABLES

Legend:
 * < 0.05
 ** < 0.01
 *** < 0.001
 **** < 0.0001
 NS not significant

∞ Practice
 ▪ Physician
 △ Organization

I. Immunization and vaccination results

a. Effectiveness

i. Influenza immunization

I. Randomized

		Immunisation rate (Δ)
Kouides et al, 1998	US, ▪	6.8% * Variation C: -5.5% - 8.6% Variation I: 0% - 19.0%

ii. Concurrent + historical comparison studies

		Immunisation rate patients (Δ)	>70% achievement rate practice(Δ)	> 85% achievement rate practice(Δ)
Kouides et al, 1993	US, ▪	8.7% *	43%***	14%*

iii. Cross-sectional studies

			Influenza vaccine reminder use (OR)
Schmittziel et al, 2004	US, ∞△	Physician	OR 1.5 *

Three studies in the US focused on influenza immunization and vaccination, one randomized, two concurrent + historical, and one cross sectional study ¹⁶⁻¹⁸. The difference in influenza immunisation rates between the intervention group and the control group ranges from 6.8% in the randomized study to 8.7% in one concurrent + historical study (p<0.05) ^{16, 17}.

The difference in attaining the immunization target level in one concurrent + historical study ranges from 14% (p<0.05) to 43% (p<0.001) depending on the target percentage level (70 and 85% thresholds) ¹⁶.

As indicated by one cross sectional study, income for quality is positively related to influenza vaccine reminder use with an odds ratio of 1.5 (p<0.05) ¹⁸.

2. Children immunization (Measles, mumps, rubella, diphtheria, tetanus, polio, pertussis, haemophilus influenza type B)

a. Randomized

		Up to date Immunisation rate patients(Δ)
Fairbrother et al, 1999	US, \square	25.3% **
Hillman et al, 1999	US, ∞	NS
Fairbrother et al, 2001	US, ∞	5.9% *

i. Concurrent + historical comparison studies

		Immunisation rate patients (Δ):
Gilmore et al, 2007	US, \square	NS (1999) 10% * (2000) NS (2001) NS (2002) NS (2003)

ii. Historical comparison studies, multiple time points

		Up to date immunisation rate (Δ)	$\geq 90\%$ achievement rate (Δ)	$\geq 95\%$ achievement rate (Δ)	booster immunisation 90% achievement rate (Δ)	booster immunisation 95% achievement rate (Δ)
Ritchie et al, 1992	UK, Scotland, ∞	-	50% (1st quarter ****, 2nd quarter *, following NS)	20% (1st quarter ****, 2nd quarter *, following NS)	41% (1st and 2nd quarter NS, 3th and 4th quarter ***, following small deterioration ****)	42% (1st and 2nd quarter NS, 3th and 4th quarter ***, following small deterioration ****)
Morrow et al, 1995	US, \square	18% * Variation C: 73.96%-82.14% Variation I: 93.47%-97.65%	45% * Variation C: 38.63%-48.37% Variation I: 84.77%-91.23%	-	-	-
Chung et al, 2003	US, \square	NS	-	-	-	-

iii. Historical comparison studies, before-after time point

		Up to date immunisation rate (Δ)
Fairbrother et al, 1997	US, \square ∞	24.3% *

iv. cross-sectional studies

		Up to date immunisation rate (Δ)
Lebaron et al, 1999	US, ∞	5% *
Bond et al, 2002	Australia, ∞	9% *** Variation C: 82.4%-86.3% Variation I: 92.1%-94.9%
Ashworth et al, 2005	UK, ∞	NS

Two studies in the UK focused on child disease immunization, one historical comparison study, multiple time points and one cross-sectional study^{23, 29}. The difference in up to date children immunisation rates between intervention group and control group is not significant in the cross-sectional study²³. The difference between the intervention group en control group in attaining the immunization target level in one historical comparison study, multiple time points ranges from 20% for a 95% achievement rate to 50% for a 90% achievement rate. For both rates the effect was highly significant at the onset ($p < 0.0001$), significant in the second period ($p < 0.05$) and not significant in the following periods²⁹. For booster immunizations another evolution was found with a 41 and 42% effect. There was no significance in the first period, a positive effect in the second and third period ($p < 0.001$) and a small deterioration effect in the following periods ($p < 0.0001$).

Eight studies in the US focused on children immunization: three randomised studies, one concurrent + historical study, two historical comparison studies, multiple points, one historical comparison study, before-after point and one cross sectional study^{19-22, 25-28}. The difference in up to date children immunisation rates between the intervention group and control group ranges from no significant effect in one randomized, one concurrent + historical comparison study and one historical comparison study, multiple time points to 5.9% ($p < 0.05$) and 25.3% ($p < 0.01$) in two randomized studies, a one out of five period significant 10% ($p < 0.05$) in one concurrent + historical comparison study, 18% ($p < 0.05$) in one historical comparison study with multiple time points, 24.3% ($p < 0.05$) in one historical comparison study, before-after time point, and an average 5% difference ($p < 0.05$) in one cross-sectional study^{19-22, 25-28}. One historical comparison study indicated a difference between the intervention group en control group in attaining a 90% immunization target level of 45% ($p < 0.05$)²⁸.

One cross-sectional study in Australia focused on children immunization. The difference in up to date children immunization rates between the intervention group and control group amounted to 9% ($p < 0.001$)²⁴.

3. Children preventive screening

a. Effectiveness

i. Randomized

		Screening rate TBC (Δ)	Screening rate lead (Δ)	Screening rate sickle cell (Δ)	Screening rate bacteriuria (Δ)
Hillman et al, 1999	US, ∞	NS	NS	NS	NS

ii. Historical comparison studies, before-after time point

		Screening rate TBC (Δ)	Screening rate lead (Δ)
Fairbrother et al, 1997	US, ∞	28.8% *	16.9% * (high risk group) 23.4% * (low risk group)

Two studies in the US focused on children screening, one randomized and one historical comparison study, before-after time point ^{22, 26}. The difference in children screening rate for TBC between the intervention group and the control group ranges from no significant difference in one randomized study to 28.8% ($p < 0.05$) in one historical comparison study, before-after time point ^{22, 26}. The difference in children screening rate for lead between the intervention group and the control group ranges from no significant difference, in one randomized study to 23.4% ($p < 0.05$) for a low risk group and 16.9% ($p < 0.05$) for a high risk group in one historical comparison study, before-after time point ^{22, 26}. The difference in children screening rate for sickle cell and the difference in children screening rate for bacteriuria between the intervention group and the control group are not significant in one randomised study ²².

4. Cancer preventive screening

a. Effectiveness

i. Randomized

		Compliance with guidelines of mammography (Δ)	Mammography referral (Δ)	Screening rate		
				Breast cancer screening (Δ)	Colorectal cancer screening (Δ)	Cervical cancer screening (Δ)
Grady et al, 1997	US, ∞	NS	NS	NS	-	-
Hillman et al, 1998	US, ∞	-	-	NS	NS	NS

ii. Concurrent + historical comparison studies

5.		Screening rate	
		Breast cancer screening (Δ)	Cervical cancer screening (Δ)
Rosenthal et al, 2005	US, ∞/Δ	NS	3.6% *
Rosenthal et al, 2008	US, $\square \infty$	2.2% * Variation C: 87.2%-89.1% Variation I: 85.5%-86.4%	3.9% * Variation C: 84.6%-85.3% Variation I: 88.1%-89.7%
Pearson et al, 2008	US, $\square \infty$	NS	NS

i. Historical comparison studies, before-after time point

		Screening rate
		Colorectal cancer screening (Δ)
Armour et al, 2004	US, \square	3% **

ii. cross-sectional studies

		Mammography reminder use (OR)
Schmittziel et al, 2004	US, $\infty \Delta$	NS

Seven studies in the US focused on cancer: two randomized studies, three concurrent + historical comparison studies, one historical comparison study, before-after time point and one cross-sectional study ^{18, 21, 30-35}. Because one study only provides generic results over different patient groups, its results are discussed elsewhere (see section X) ²¹. The difference in compliance with guidelines for mammography and the difference in mammography referral between the intervention group and the control group is not significant in one randomised study ³⁰. The difference in screening rate for breast cancer between the intervention group and the control group ranges from not significant in two randomized and two concurrent+ historical comparison study to 2.2% ($p < 0.05$) in one concurrent+historical comparison studies ^{30, 31, 33-35}.

The difference in screening rate for colorectal cancer between the intervention group and the control group ranges from no significant difference in one randomized study to 3% ($p < 0.01$) in one historical comparison study, multiple time points ^{31, 32}.

The difference in screening rate for cervical cancer between the intervention group and the control group ranges from no significant difference in one randomized and one concurrent + historical study to 3.9% ($p < 0.05$) in two concurrent+historical studies^{31, 33-35}.

As indicated by one cross sectional study, income for quality is not significantly related to mammography reminder use¹⁸.

I. Well child visits

a. Effectiveness

i. Randomized

		Well child visits (Δ)
Hillman et al, 1999	US, ∞	NS

ii. Concurrent + Historical comparison studies

		Well child visits, ages 3-6 (Δ)	Well child visits, adolescents (Δ)
Pearson et al, 2008	US, ∞	5% *	-5% **

iii. Historical comparison studies, before-after time point

		Well child visits (Δ)
Fairbrother et al, 1997	US, ∞	6.6% *

Three studies in the US focused on well child visits: one randomized study, one concurrent + historical study and one historical comparison study, before-after time point^{22, 26, 33}. The difference in well child visits between the intervention group and the control group ranges from a negative effect in one concurrent + historical comparison study to 6.6% ($p < 0.05$) in one historical comparison study, before-after time point. One randomized study found no significant effect^{22, 26, 33}.

2. Sexually transmitted diseases

a. Effectiveness

i. historical comparison study, multiple time points

		Chlamydia screening in women ages (16-20)(Δ)
Pearson et al, 2008	US, ∞	-11% **

ii. cross-sectional studies

		Association with annually screening for Chlamydia (15-19 year) (OR)		Association with annually screening for Chlamydia (20-25 year) (OR)		Association with obtaining sexual history (OR)		Association with providing Chlamydia drugs for partner's treatment (OR)		Association with providing services to minors without parental notification/consent (OR)	
		capitation and quality of care	salary and quality of care	capitation and quality of care	salary and quality of care	capitation and quality of care	salary and quality of care	capitation and quality of care	salary and quality of care	capitation and quality of care	salary and quality of care
Pourat et al, 2005	US, ∞	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS

One cross sectional study and one historical comparison study, multiple time points in the US focused on sexually transmitted diseases^{33, 36}. The difference in Chlamydia screening rate in women aged 16-20 between the intervention group and the control group amounts to minus -11% ($p < 0.01$) in one historical comparison study³³. As indicated in one cross sectional study, income for quality is not significantly related to annually screening for Chlamydia³⁶. According to the same study income for quality is also not significantly related to obtaining the sexual history of the patient, with providing Chlamydia drugs for the partner's treatment and to providing services to minors without parental notification or consent³⁶.

3. Adult screening

a. Effectiveness

4. Cholesterol screening

i. Historical comparison studies, multiple time points

		Use of repeated profile, diet therapy or medication if >6,21 mmol/l, >240mg/dl (Δ)	> 90% of patients achievement rate (Δ)
Morrow et al, 1995	US, \square	3% (first period *, second period NS) Variation C: 87.77%-94.4% Variation I: 91.74%-97.39%	10% (first period *, second period NS) Variation C: 73.53%-83.67% Variation I: 85.44%-93.16%

One historical comparison study with multiple time points in the US focused on general adult cholesterol screening²⁸. Pay for quality led to a 3% increase in the use of repeated profiling, diet therapy or medication if the cholesterol level was above 6.21 mmol/l, and a 10% increase in performance on the 90% achievement rate. Both findings were significant in the first post implementation period ($p < 0.05$), but became not significant during a second period.

5. Emergency care

a. Continuity and integration

6. Smoking cessation referral at the emergency department

a. Cross-sectional studies

		Positive relationship with smoking cessation referral rate for heart disease, gastritis, pregnancy, respiratory illness
Greenberg et al, 2008	US, \triangle	**

One US cross sectional study focused on continuity and integration in terms of smoking cessation referral at the emergency department for patients with heart disease, gastritis, pregnancy or respiratory illness³⁸. A positive significant relationship was found with the use of pay for quality incentives ($p < 0.01$). No regression coefficient or odds ratio was reported.

7. Myocardial infarction/acute cardiac event

a. Effectiveness

i. Incentivized targets

I. Concurrent + Historical comparison studies

		Aspirin at arrival (Δ)	Aspirin at discharge (Δ)	ACE inhibitor for LVSD (Δ)	Smoking cessation advice (Δ)	Beta blocker at arrival (Δ)	Beta blocker at discharge (Δ)	Thrombolytic agent within 30 min after arrival (Δ)	PCI within 120 min after arrival (Δ)
Grossbart, 2006	US, \triangle	NS	0.7% **	NS	NS	NS	NS	NS	5.4% **
Glickman et al, 2007	US, \triangle	NS	2.3% *	NS	5.2% *	NS	NS		
Lindenauer et al, 2007	US, \triangle	3.3% **	NS	9.9% ***		2.8% **	2.8% *		
Herrin et al, 2008	US, \square	NS	8.5% ***			NS	NS		

i. Not incentivized targets

2. Concurrent + Historical comparison studies

		Heparin, any (Δ)	Glycoprotein IIb/IIIa inhibitor (Δ)	Clopidogrel at discharge (Δ)	Lipid lowering agent at discharge (Δ)	Dietary modification counseling (Δ)	Cardiac rehabilitation referral (Δ)	Cardiac catheterization within 48h (Δ)	ECG within 10 min (Δ)	Thrombo-lytics within 30 min after arrival (Δ)	PCI within 120 min after arrival (Δ)	In hospital mortality rate (Δ)
Glickman et al, 2007	US, △	NS	NS	NS	4.3%*	NS	NS	NS	NS			NS
Herrin et al, 2008	US, ▪									NS	NS	NS

Acute cardiac events, and myocardial infarction more in particular, have been focused upon by four concurrent plus historical comparison studies in the US as a target condition for pay for quality^{40, 39, 41, 42, 81}.

With regard to incentivized targets, the effects on PCI within 120 minutes after arrival and receiving a thrombolytic agent within 30 minutes after arrival have been investigated by only one study³⁹. Respectively a 5.4% ($p < 0.01$) and a not significant effect was found. Receiving smoking cessation advice was included in two studies, with a 5.2% ($p < 0.05$) and no significant difference^{40, 39}. Lindenauer et al (2007) reported a 9.9% ($p < 0.001$) difference in ACE inhibitor use for LVSD. This wasn't confirmed by two other studies that came to not significant effects^{40, 39}. The same is true for the aspirin at arrival target, the beta blocker at arrival target and the beta blocker at discharge target. Although Lindenauer et al (2007) came to an effect of respectively 3.3% ($p < 0.01$), 2.8% ($p < 0.01$) and 2.8% ($p < 0.05$), all three targets showed no significant result in three other studies^{40, 39, 41}. Finally, at the aspirin at discharge target, where Lindenauer et al (2007) found no significant effect, the other three studied did. They report an 8.5% ($p < 0.001$), a 2.3% ($p < 0.05$) and a 0.7% ($p < 0.01$) effect^{40, 39, 41}.

Two of these studies also focused on not incentivized targets which had been monitored during the P4Q program to assess any spillover or neglecting effects^{40, 41}. On the eleven measures included in one or both of these studies ten showed no significant change. Only the target of receiving a lipid lowering agent at discharge showed a 4.3% ($p < 0.05$) difference in one study⁴⁰.

8. CABG

a. Effectiveness

i. Historical comparison studies, multiple time points

		Readmission within 30 days (Δ)	Discharged home (Δ)	Patients with any complication (Δ)	Patients receiving blood products (Δ)	Readmitted to ICU (Δ)	Pulmonary complications (Δ)
Casale et al, 2007	US, ■	NS	10% *	NS	NS	NS	NS
		Operative mortality (Δ)	Atrial fibrillation (Δ)	Deep sternal wound infection (Δ)	Reintubated during hospital stay (Δ)	Total ventilation hours (Δ)	Neurologic complication (Δ)
		NS	NS	NS	NS	NS	NS

One historical comparison study with multiple time points in the US focused on CABG patients as P4Q target group⁴³. Out of the twelve targets only one reached statistical significance: The percentage of patients discharged to home showed a 10% effect ($p < 0.05$).

9. Heart failure (in hospital care)

a. Effectiveness

i. Incentivized targets

1. Concurrent + Historical comparison studies

		Provision of discharge instructions (Δ)	LVF assessment (Δ)	ACE inhibitor for LVSD (Δ)	Smoking cessation advice (Δ)
Grossbart, 2006	US, Δ	25.5% ***	-2.4% ** (neg.)	NS	NS
Lindenauer et al, 2007	US, Δ		5.1% ***	NS	
Herrin et al, 2008	US, ■		NS		

ii. Not incentivized targets

I. Concurrent + Historical comparison studies

		Discharge instruction
Herrin et al, 2008	US, ▫	NS

Three US concurrent plus historical comparison studies focused on heart failure within the hospital setting as a target condition^{39, 41, 42}. Smoking cessation advice showed no significant effect in one study³⁹. The same is found for ACE inhibitor use for LVSD in two studies^{39, 42}. However, the provision of discharge instructions demonstrated a large effect of 25.5% ($p < 0.001$) in one study³⁹. Finally, the LVF assessment target showed mixed findings, with one not significant result⁴¹, one negative result of minus 2.4% ($p < 0.01$)³⁹ and one positive result of 5.1% ($p < 0.001$)⁴². Herrin et al, included a not incentivized target in their study. Implementing a P4Q program had no significant effect on the not incentivized target⁴¹.

10. Community acquired pneumonia (in hospital care)

a. Effectiveness

i. Concurrent + Historical comparison studies

		Oxygenation assessment (Δ)	Pneumococcal screening and/or vaccination (Δ)	Blood cultures (Δ)	Smoking cessation advice (Δ)	Antibiotic within 4 hours after arrival (Δ)
Grossbart, 2006	US, \triangle	-1.9% *** (neg.)	9.5% ***	3.5% ***	-16.7% *** (neg.)	-3.2% *** (neg.)
Lindenauer et al, 2007	US, \triangle	NS	10.9% ***			4.3% ***
Herrin et al, 2008	US, ▫		44.7% **			NS

Community acquired pneumonia as an in hospital P4Q target condition has been studied in three concurrent plus historical evaluations in the US^{39, 41, 42}. The use of blood cultures and smoking cessation advice was investigated in one study with respectively a positive effect of 3.5% ($p < 0.001$) and a negative effect of minus 16.7% ($p < 0.001$)³⁹. Oxygenation assessment was included in two studies, with a not significant result⁴² and a negative result of minus 1.9% ($p < 0.001$)³⁹. The target of pneumococcal screening and/or vaccination showed a positive effect in all three studies with an effect size of 9.5% ($p < 0.001$), 10.9% ($p < 0.001$) and 44.7% ($p < 0.01$). Finally there were mixed findings on the target of receiving antibiotics within 4 hours after arrival. One study found no significant effect⁴¹, one study found a negative effect of minus 3.2% ($p < 0.001$)³⁹ and one study found a positive effect of 4.3% ($p < 0.001$)⁴². It should be noted that the three negative findings for the community acquired pneumonia target condition occurred in the same study³⁹.

11. Heart failure (in primary care)

a. Effectiveness

i. Incentivized targets

I. Historical comparison studies multiple time points

		ACE inhibitor or ARB (angiotensine receptor blockers) use (Δ)
Chung et al, 2003	US, ▫	23.4 % ***

One historical comparison study, multiple time points in the US focused on P4Q effects with heart failure as target condition²⁵.

As indicated by this study, there was a 23.4% ($p < 0.001$) difference between intervention and control group in receiving an Angiotensin converting enzyme inhibitor or an Angiotensin receptor blocker²⁵.

12. Acute sinusitis (primary + hospital care)

a. Effectiveness

i. Historical comparison studies, before-after time point

		Firstline antibiotics prescriptions (Δ)	Appropriate secondline antibiotics prescriptions (Δ)	Inappropriate antibiotics prescription (Δ)	Sinus plain x ray films (Δ)	Sinus computed tomographic scans (Δ)	Otolaryngologist consultations (Δ)	Allergist consultations (Δ)
Greene et al, 2004	US, $\infty\triangle$	14% **	NS	-29% **	-28% **	-39% **	NS	-31% **

One study focuses on acute sinusitis as a P4Q target condition. This study was performed in the US with a before after design ⁴⁴. P4Q implementation led to a 14% ($p<0.01$) increase in firstline antibiotics prescription, no significant effect on appropriate secondline antibiotics prescription and a 29% ($p<0.01$) decrease in inappropriate antibiotics prescription. In addition, the use of sinus plain x ray films and the use of sinus computed tomographic scans decreased with 28% ($p<0.01$) and with 39% ($p<0.01$). In terms of number of consultations, there was a decrease of 31% ($p<0.01$) in allergist consultations and no significant effect on the number of otolaryngologist consultations.

13. Breastfeeding

a. Effectiveness

i. Historical comparison studies, before-after time point

		Exclusive breastfeeding at discharge (Δ)	Predominant breastfeeding at discharge (Δ)	Full breastfeeding at discharge (Δ)	Complementary breastfeeding at discharge (Δ)	No breastfeeding at discharge (Δ)
Cattaneo et al, 2001	Italy, \triangle	12% ***	-7% ***	6%***	-5% ***	NS

One historical comparison study, before-after time point in Italy focused on breastfeeding at discharge ⁴⁵. The difference in exclusive breastfeeding between the intervention group and the control group amounted to 12% (p<0.001). The difference in full breastfeeding between the intervention group and the control group amounted to 6% (p<0.001). The difference in predominant breastfeeding between the intervention group and the control group equals minus 7% (p<0.001) and the difference in complementary breastfeeding between the intervention group and the control group amounted to minus 5% (p<0.001). There was no significant difference in no breastfeeding between the intervention group and the control group. ⁴⁵.

14. Diabetes

a. Effectiveness

i. Concurrent + Historical comparison studies

1. Incentivized

		hbA1c testing rate (Δ)	LDL cholesterol screening (Δ)	Second hbA1c testing rate (Δ)	hbA1c < 9.5% (Δ)	hbA1c < 7.4% (Δ)	Lipid testing rate (Δ)	LDL < 100 mg/dl (Δ)	LDL < 130 mg/dl (Δ)	Blood pressure recorded (Δ)	Blood pressure < 140/80 mmHg (Δ)	Nephro-pathy testing rate (Δ)	Foot exam rate (Δ)	Retinal exam rate (Δ)	Peri-pheral pulses or vibration sense (Δ)	BMI recorded (Δ)	
Larsen et al, 2003	US, ∞△	12% ***	25.8% ***														
Beaulieu & Horrigan, 2005	US, *			25.5% ***	13.9% ***		18.3% ***	10.5% ***	23.5% ***		6.3% *	37.0% ***	45.4% ***	25.6% ***			
Rosenthal et al, 2005	US, ∞					NS											
Levin-Scherz et al, 2006	US, ∞△	2.2% *	5.8% *									10.4% *		18.7% *			
Campbell et al, 2007	UK, ∞	0.54% (transformed) ***	NS			0.40% (transformed) *				NS	0.49% (transformed) *	NS	0.99% (transformed) ***	0.58% (transformed) *	1.58% (transformed) ***	0.60% (transformed) **	
Coleman et al, 2007 ⁴	US, *			16.2% ****													
Young et al, 2007	US, *			NS			NS					NS		5.18% ***			
Pearson et al, 2008	US, *∞	NS					NS					NS		NS			
Rosenthal et al, 2008	US, *	5.5% *					5.0% * Variation C: 81.3%-83.0% Variation I: 85.2%-89.1%					17.7% * Variation C: 57.6%-59.7% Variation I: 72.6%-79.8%		0.9% * Variation C: 98.1%-98.4% Variation I: 98.6%-99.7%			

		Smoking status recording (Δ)	Total serum cholesterol < 190 mg/dl (Δ)	Patient education documentation if diagnosed < 5 years (Δ)	Hypoglycemia symptoms recording in patients receiving sulfonylurea (Δ)	Flu vaccination (Δ)
Campbell et al, 2007	UK, ∞	0.58% (transformed) **	0.42% (transformed) *	NS	NS	
Young et al, 2007	US, \square					NS

2. Not incentivized

		hbA1c < 7.0% (Δ)	Average hbA1c (Δ)	Most recent hbA1c > 9.5% (Δ)	Most recent LDL cholesterol < 130 mg/dl (Δ)	Annual eye exam (Δ)
Larsen et al, 2003	US, $\infty\Delta$	19.3% ***	-0.8 ***	-13.2% ***	29.9% ***	10% **

15. Historical comparison studies, multiple time points

		hbA1c testing rate (Δ)	hbA1c < 7.0% (Δ)	hbA1c < 7.4% (Δ)	hbA1c < 10% (Δ)	Oral hypoglycaemic agent prescription (Δ)	Insulin prescription (Δ)	Cholesterol < 5 mmol/l (Δ)	Lipid lowering drug prescription (Δ)	Blood pressure < 140/80 mmHg (Δ)	ACE inhibitor prescription (Δ)
Chung et al, 2003	US, \square	28.1% ****									
Millett et al, 2007 ⁶	UK, ∞		2.3% **			14% ***	8.1% ***	12.9% ***	21.9% ***	10.9% ***	12.2% ***
Tahrani et al, 2007	UK, ∞	19% ***		21% ***	19% ***			24% ***		18% ***	
Weber et al, 2007	US, \square		2.6% ***							4.2% ****	
Vaghela et al, 2008	UK, ∞			7.6% ***				11% ***		9.3% ***	

		Smoking status recording ever (Δ)	Smoking status recording during previous period (Δ)	Smoking cessation advice recording (Δ)	Pneumococcal vaccination (Δ)	Influenza vaccination (Δ)
Millett et al, 2007 ⁷	UK, ∞	8.8% ***	19.1% ***	35.5% ***		
Tahrani et al, 2007	UK, ∞		51% ***	12% ***		
Weber et al, 2007	US, \square		NS		24.3 ****	15.9% ****

		Lipid testing rate (Δ)	LDL < 100 mg/dl (Δ)	Blood pressure recording (Δ)	Microalbuminuria recording (Δ)	Serum creatinine testing (Δ)	Peripheral pulse testing (Δ)	Neuropathy testing (Δ)	Retinal exam rate (Δ)
Tahrani et al, 2007	UK, ∞	15% ***		10% ***	70% ***	13% ***	59% ***	59% ***	37% ***
Weber et al, 2007	US, \square	NS	NS						

a. Not incentivized

		Prevalence of smoking (Δ)
Millett et al, 2007 ⁷	UK, ∞	-3.8% ***

i. Historical comparison studies, before after time points

		BMI recording (Δ)	Smoking status recording (Δ)	Offering of smoking cessation advice (Δ)	Influenza immunization (Δ)	Blood pressure recording (Δ)	Blood pressure < 145/85 mmHg (Δ)	Total cholesterol recording (Δ)	Total cholesterol < 5 (Δ)	ACE inhibitor or all blocker use when proteinuria or microalbuminuria (Δ)
Jaiveer et al, 2006	UK, ∞	16.1% *** Variation C: 4.5%-85.3% Variation I: 53%-97.2%	12% *** Variation C: 22.3%-97% Variation I: 91%-100%	24.6% **** Variation C: 25%-100% Variation I: 77%-100	12.9% ** Variation C: 56.3%-93.6% Variation I: 67.5%-98%	16.2% ** Variation C: 11.7%-98.7% Variation I: 91%-99%	15.7% *** Variation C: 15.5%-92.8% Variation I: 46%-92.7%	14.6% *** Variation C: 57.7%-92.8% Variation I: 78.2%-97%	21% **** Variation C: 36.2%-58.2% Variation I: 56%-91.6%	17.1% *** Variation C: 42.1%-80% Variation I: 0%-100%
McGovern et al, 2008 ¹⁰	UK, ∞					22.7% *	6.3% *		-1.3% * (neg.)	
		Microalbuminuria recording (Δ)	Serum creatinine testing (Δ)	hbA1c recording (Δ)	hbA1c < 7.4% (Δ)	hbA1c < 10% (Δ)	Retinal screening (Δ)	Peripheral pulses testing (Δ)	Neuropathy testing (Δ)	Lipid testing (Δ)
Jaiveer et al, 2006	UK, ∞	21.5% **** Variation C: 1.1%-59% Variation I: 1.9%-81.4%	17.7% **** Variation C: 43.3%-95.2% Variation I: 75.4%-95.6%	6.6% * Variation C: 57.3%-97% Variation I: 85.4%-97%	21.9% **** Variation C: 27.2%-70% Variation I: 57.6%-88.5%	11.1% ** Variation C: 49.5%-95.3% Variation I: 81.5%-96.8%	14.4% ** Variation C: 30.4%-74% Variation I: 43.3%-89.5%	16.6% *** Variation C: 33.6%-86% Variation I: 45.2%-92.3%	42.8% **** Variation C: 7.2%-51% Variation I: 52.4%-91.5%	
McGovern et al, 2008 ¹⁰	UK, ∞		39.8% *	34.1% *	7.7% *	10.3% *				38.5% *

b. Cross-sectional studies

		Eye exam reminder use	
Schmittziel et al, 2004	US, ∞△	NS	
		Use of diabetic patient registries, clinical practice guidelines with reminder systems, case management, physician feedback	
		Receiving income for quality (β)	Receiving better contracts for quality (β)
Li et al, 2004	US, ∞□	0.20 *	0.23 *

		Use of registries, reminders, medical record flowsheets, physician feedback, formal training on guidelines, on site health promotion, patient education classes, support groups for parents or adolescents, self management support programs, case management (β)
Rittenhouse & Robinson, 2006	US, ∞ △	0.65 *

		Glycemic control assessed (RR)	Lipid profile assessed (RR)	Proteinuria assessed (RR)	Dilated eye exam (RR)	Foot exam (RR)	Advised to take aspirin (RR)	Influenza immunization (RR)
Ettner et al, 2006	US, ▪	NS	NS	1.25 *	NS	NS	1.17 *	1.26 *

		Association between higher scores on additional services domain and emergency admission (OR)	Association between higher scores on clinical domain and emergency admission (OR)	Association between higher scores on organisation domain and emergency admission (OR)
Downing et al, 2007	UK, ∞	NS	NS	NS

Diabetes as a P4Q target condition has received most interest in evaluation studies. The effectiveness results are described by country and grouped further according to clinical targets.

In the UK eight studies have focused on clinical P4Q effects on diabetes targets. These include one concurrent plus historical comparison study ⁴⁸, four historical comparison studies with multiple time points ^{6, 7, 53, 82}, two before after studies ^{52, 10} and one cross sectional study ⁵⁷. Results reported by Campbell et al (2007) are recalculated towards a comparable effect size, using an e^x formula to take the logit transformation into account.

HbA1c testing as a process measure, in terms of receiving one and/or two tests per period, shows positive effects in most UK studies. One concurrent plus historical comparison study found a 1.71% (p<0.001) effect ⁴⁸. This is confirmed by multiple time point findings, with a result of 19% (p<0.001) ⁵³, and by before after findings, with results of 6.6% (p<0.05) ⁵² and 34.1% (p<0.05) ¹⁰.

HbA1c level as an intermediate outcome measure can be further subdivided into target levels of <7.0%, <7.4%, and <10% in UK studies. At the <7.0% target level one historical comparison study with multiple time points reports a 2.3% ($p<0.01$) effect⁶. At the <7.4% level results include a 1.49% ($p<0.05$) effect in one concurrent plus historical comparison study⁴⁸, 21% ($p<0.001$) and 7.6% ($p<0.001$) in multiple time point comparisons^{53, 82}, and 21.9% ($p<0.0001$) and 7.7% ($p<0.05$) in two before after studies^{52, 10}. At the <10% level one multiple time point study reports a 19% ($p<0.001$) difference⁵³, which is in line with two before after results: 11.1% ($p<0.01$) and 10.3% ($p<0.05$)^{52, 10}.

Next to the HbA1c targets a few other measures also focus on glycemia control. One concurrent plus historical comparison study found no significant effect in terms of hypoglycemia symptoms recording in patients receiving sulfonurea⁴⁸. Another multiple time point comparison study found a 14% ($p<0.001$) effect on the oral hypoglycemic agent prescription rate, and an 8.1% ($p<0.001$) effect on the insulin prescription rate⁶.

Lipid and cholesterol testing as a process measure shows a difference of 15% ($p<0.001$) in one multiple time point study⁵³ and a 14.6% ($p<0.001$) and a 38.5% ($p<0.05$) difference in two before after studies^{52, 10}. However, Campbell et al (2007) came to a not significant result in one concurrent plus historical comparison study⁴⁸.

As an intermediate outcome a 1.52% ($p<0.05$) difference is found on the total serum cholesterol < 190mg/dl target by one concurrent plus historical comparison study⁴⁸. In terms of cholesterol < 5 mmol/l three multiple time point studies report a 12.9% ($p<0.001$), a 24% ($p<0.001$), and an 11% ($p<0.001$) difference^{6, 53, 82}. Before after studies found conflicting results on this measure with a 21% ($p<0.0001$) and minus 1.3% ($p<0.05$) effect^{52, 10}. Finally, one multiple time point study reported a 21.9% ($p<0.001$) effect on the lipid lowering drug prescription rate⁶.

Blood pressure monitoring as a process measure shows a not significant effect in one concurrent plus historical comparison study⁴⁸. This is in contrast with one multiple time point study reporting a 10% ($p<0.001$) effect⁵³ and two before after studies reporting a 16.2% ($p<0.01$) and a 22.7% ($p<0.05$) effect^{52, 10}.

Effects on the target blood pressure level as intermediate outcome include a 1.63% ($p<0.05$) difference in a concurrent plus historical comparison study⁴⁸, a 18% ($p<0.001$), 4.2% ($p<0.0001$) and 9.3% ($p<0.001$) difference in three multiple time point studies^{6, 53, 82} and a 15.7% ($p<0.001$) and 6.3% ($p<0.05$) difference in two before after studies^{52, 10}. ACE inhibitor prescription changed with 12.2% ($p<0.001$) in one multiple time point study⁶ and with 17.1% ($p<0.001$) in one before after study⁵².

Nephropathy testing rate showed no significant result in one concurrent plus historical comparison study⁴⁸. In terms of microalbuminuria testing one multiple time point study found a 70% ($p<0.001$) effect⁵³ and one before after study found a 21.5% ($p<0.0001$) effect⁵². In terms of serum creatinine testing one multiple time point study reported a 13% ($p<0.001$) effect⁵³. Similar positive results are found in two before after studies, with an effect size of 17.7% ($p<0.0001$) and 39.8% ($p<0.05$)^{52, 10}.

Concerning weight control in diabetes patients a 1.82% ($p<0.01$) effect on BMI recording was reported by one concurrent plus historical comparison study⁴⁸. One before after study found a 16.1% ($p<0.001$) effect on this measure⁵².

According to one concurrent plus historical comparison study there was an effect of 1.78% ($p<0.05$) on the retinal exam rate⁴⁸. Higher effect sizes were reported in one multiple time point study: 37% ($p<0.001$)⁵³ and in one before after study: 14.4% ($p<0.01$)⁵².

With regard to foot exam rate there was a difference of 2.69% ($p<0.001$) in one concurrent plus historical study⁴⁸. Peripheral pulses testing rate was affected with an effect size of 4.85% ($p<0.001$) in one concurrent plus historical study⁴⁸. One multiple time point study found a result of 59% ($p<0.001$) on this measure⁵³. One before after study found a 16.6% ($p<0.001$) effect⁵². Neuropathy testing rate was included in one multiple time point study and in one before after study with a result of 59% ($p<0.001$) and 42.8% ($p<0.0001$)^{52, 53}.

Smoking status recording rate effects were evaluated in one concurrent plus historical comparison study with an effect of 1.78% ($p < 0.01$)⁴⁸, in two multiple time point comparison study with an effect of 8.8% ($p < 0.001$), 19.1% ($p < 0.001$) and 51% ($p < 0.001$)^{7, 53}. Finally, one before after study found a 12% ($p < 0.001$) effect⁵².

An effect on smoking advice rate was found in two multiple time point comparison studies, with a result of 35.5% ($p < 0.001$) and 12% ($p < 0.001$)^{7, 53}, and in one before after study with an effect of 24.6% ($p < 0.0001$)⁵². One before after study in the UK focused on the change in prevalence of smoking in diabetes patients as a concurrent not directly incentivized measure⁷. They found a reduction of 3.8% ($p < 0.001$).

Vaccination targets are not specifically included in UK based P4Q evaluation studies.

One concurrent plus historical comparison study found no significant effect on the rate of documented patient education provided to patients first diagnosed with diabetes during the last five years, although this was incentivized⁴⁸.

In the diabetes domain no significant relationship was found in a cross sectional study between the performance on the QOF clinical domain, the QOF organization domain and/or the QOF additional services domain on the one hand and the not incentivized emergency admission rate on the other hand⁵⁷.

In the US 14 studies have focused on clinical P4Q effects on diabetes targets. These include eight concurrent plus historical comparison studies^{33-35, 46, 47, 50, 59, 4}, two historical comparison studies with multiple time points^{25, 51}, and four cross sectional studies^{18, 49, 54, 55}.

HbA1c testing as a process measure, in terms of receiving one and/or two tests per period, shows positive effects in most US studies. In the concurrent plus historical comparison studies the results include a not significant effect in two studies^{33, 46}, a 2.2% ($p < 0.05$)⁵⁹, a 5.5% ($p < 0.05$) effect³⁴, a 12% ($p < 0.001$) effect⁵⁰, a 16.2% ($p < 0.0001$) effect⁴, and a 25.5% ($p < 0.001$) effect⁴⁷. Similar results were found in a multiple time point comparison study with an effect of 28.1% ($p < 0.0001$)²⁵.

HbA1c level as an intermediate outcome measure can be further subdivided into target levels of <7.0%, <7.4%, and <9.5% in US studies. At the <7.0% target level one historical comparison study with multiple time points reports a 7.6% ($p < 0.001$) effect⁵¹. In addition, one concurrent plus historical comparison study included this measure as a not incentivized concurrent target⁵⁰. These authors found a 19.3% ($p < 0.001$) difference. On average HbA1c levels were 0.8% ($p < 0.001$) lower.

At the <7.4% target level one concurrent plus historical comparison study found no significant results³⁴. And at the <9.5% target level another concurrent plus historical comparison study came to a 13.9% ($p < 0.001$) effect⁴⁷. A third study of this category included this measure as a not incentivized concurrent target. The result was 13.2% in effect size ($p < 0.05$).

One cross sectional study found no significant result on the glycemic control rate as an incentivized measure⁴⁹.

Lipid and cholesterol testing as a process measure shows no significant difference in two concurrent plus historical comparison studies^{33, 46}. The same is true in one multiple time point comparison study⁵¹. However, four concurrent plus historical studies found an effect of, respectively, 25.8% ($p < 0.001$)⁵⁰, 18.3% ($p < 0.001$)⁴⁷, 5.8% ($p < 0.05$)⁵⁹, and 5.0% ($p < 0.05$)³⁵. One cross sectional study found no significant difference⁴⁹.

With regard to cholesterol and lipid intermediate outcome measures one concurrent plus historical comparison study found a 10.5% ($p < 0.05$) effect on LDL < 100mg/dl and a 23.5% ($p < 0.001$) effect on LDL < 130mg/dl⁴⁷. On this last measure one concurrent plus historical study found a concurrent not incentivized effect of 29.9% ($p < 0.001$)⁵⁰. One multiple time point comparison study found no significant effect on the LDL < 100mg/dl target⁵¹.

Blood pressure monitoring as a process measure wasn't included in any US based P4Q evaluation study. The effect on the target blood pressure level as intermediate outcome includes a 6.3% ($p<0.001$) difference in one concurrent plus historical comparison study⁴⁷ and a 4.2% ($p<0.0001$) difference in one multiple time point comparison study⁵¹. One cross sectional study found a positive 1.17 ($p<0.05$) relative risk of advising to take aspirin to diabetes patients in relationship with P4Q use⁴⁹.

There were no diabetes specific reports on the use of ACE inhibitors as a target in US studies.

Nephropathy testing showed a not significant result in two concurrent plus historical comparison studies^{33,46}. Three other studies of this design type found an effect of 37.0% ($p<0.001$)⁴⁷, 10.4% ($p<0.05$)⁵⁹ and 17.7% ($p<0.05$)³⁵. One cross sectional study found a 1.25 ($p<0.05$) positive relative risk of the use of proteinuria assessment in relationship with P4Q use⁴⁹.

Weight control is not further addressed in the US based studies.

At the concurrent plus historical comparison study design level one study found no significant effect on the retinal exam rate³³. However, four other studies at the same level came to an effect of 25.6% ($p<0.001$)⁴⁷, 18.7% ($p<0.05$)⁵⁹, 5.18% ($p<0.001$)⁴⁶ and 0.9% ($p<0.05$)³⁵. One study of this design type included this measure as a not incentivized concurrent monitoring target, with an effect of 10% ($p<0.01$)⁵⁰. One cross sectional study found no significant difference⁴⁹.

One concurrent plus historical comparison study reports an effect of 45.4% ($p<0.001$) on the foot exam rate⁴⁷. One cross sectional study found no significant difference⁴⁹. This measure was not included in any other US based P4Q evaluation study.

Peripheral pulses testing and neuropathy testing are not included in US evaluation studies.

Smoking status recording rate was only included in one study with a not significant result⁵¹. Smoking cessation advice rate is not reported upon.

Influenza vaccination is included in one concurrent plus historical comparison study with a not significant effect⁴⁶. One multiple time point comparison study found a 15.9% ($p<0.001$) effect⁵¹. The same study included pneumococcal vaccination with a 24.3% ($p<0.0001$) effect. One cross sectional study found a positive 1.26 ($p<0.05$) relative risk of the use of influenza immunization in relationship with P4Q use⁴⁹.

Three cross sectional studies in the US focus on the relationship of the use of care management processes with the use of P4Q. One study came to a not significant relationship with the use of eye exam reminders¹⁸. However, reminder use was positively associated with P4Q use in both other studies^{54,55}. Together with other care management processes such as the use of patient registries, guidelines, IT integrated flowsheets, case management, physician feedback, training and education significantly positive regression coefficients of 0.20 ($p<0.05$) to 0.41 ($p<0.05$) were reported by these authors.

16. Continuity and integration

a. Historical comparison studies, before after time points

		Global number of referrals (Δ)	Referral rate for poor glycemic control (Δ)	Glycemic threshold for referral (Δ)	Number per reason for referral (newly diagnosed, type I diabetes, poor glycemic control, renal disease, foot problems, lost to follow-up) (Δ)
Srirangalingam et al, 2006	UK, ∞	NS	23% ***	-0.9% **	NS

With regard to continuity and integration one before after study in the UK focused specifically on a number of referral targets in the treatment of diabetes ⁵⁶. There was no significant effect in terms of the global number of referrals or in terms of the number per reason for referral (newly diagnosed, type I diabetes, poor glycemic control, renal disease, foot problems, lost to follow-up). However, the referral rate for poor glycemic control increased with 23% ($p < 0.001$). In addition, the glycemic threshold for referral decreased with 0.9% ($p < 0.01$).

17. Coronary heart disease

a. effectiveness

i. Concurrent + historical comparison studies

		Frequency or pattern of angina attacks recorded during the previous 15 mo	Blood pressure recorded during the previous 15 mo	Exercise capacity recorded during the previous 15 mo	Cholesterol level recorded during the previous 5 yr	Dietary advice recorded during the previous 5 yr	Smoking status recorded during the previous 5 yr
Campbell et al, 2007	UK, ∞	NS	0.26% (transformed) *	NS	0.54% (transformed) ***	0.64% (transformed) *	0.87% (transformed) ***
		Referral to specialist for exercise stress testing or assessment (ECG) ever recorded	Prescription or advice to take aspirin recorded unless record of contraindication or intolerance	Blood pressure controlled to $\leq 150/90$ mm Hg	Serum cholesterol controlled to $190 \mu\text{g/dl}$	Weight advice for overweight patients recorded during the previous 5 yr	
Campbell et al, 2007	UK, ∞	NS	0.71% (transformed) **	NS	NS	NS	

ii. Historical comparison studies, before after

		Angina patient exercise test/specialist assessment	Smoking status recorded	Blood pressure recorded	Blood pressure controlled	Smokers given advice	Cholesterol Recorded
Mc Govern et al, 2008b	UK, ∞	2.3% *	26.2% *	21.5% *	0.7% *	15.2% *	41.7% *
		Cholesterol with measurement ≤ 5 mmol/l	Antiplatelet or anticoagulant therapy prescription	b-Blocker therapy prescription	ACE inhibitor prescription	Influenza vaccination	
Mc Govern et al, 2008b	UK, ∞	-10.8% *	24.5% *	27.4% *	11.5% *	28.1% *	

iii. Cross sectional studies

		Association between higher scores on additional services domain and emergency admission (OR)	Association between higher scores on clinical domain and emergency admission (OR)	Association between higher scores on organisation domain and emergency admission (OR)
Downing et al, 2007	UK, ∞	1.03 * NS	NS NS	NS NS

Three studies in the UK focus on P4Q effects with coronary heart diseases as target condition. One concurrent + historical comparison study⁴⁸, one historical comparison study before after¹¹ and one cross-sectional study⁵⁷. Results reported by Campbell et al (2007) are recalculated towards a comparable effect size, using an e^x formula to take the logit transformation into account. The difference in recording of blood pressure between the intervention and control group ranged from 1.3% ($p < 0.05$) in one concurrent + historical comparison study to 21.5% ($p < 0.05$) in one historical comparison study, before after⁴⁸,¹¹. The difference in recording the cholesterol level between intervention and control group ranges from 1.71% ($p < 0.001$) in one concurrent + historical comparison study to 41.7% ($p < 0.05$) in one historical comparison study, before after⁴⁸,¹¹. The difference in recording of smoking status between intervention and control group ranges from 2.39% ($p < 0.001$) to 26.2% ($p < 0.05$) in these two studies⁴⁸,¹¹. For controlled blood pressure the difference between the intervention and control group ranges from no significant difference in the concurrent + historical comparison study, to 0.7% ($p < 0.05$) in the historical comparison study, before after. The difference in controlled cholesterol (with a target level of \leq mmol/l) between intervention and control group, ranges from no significant difference to -10.8 ($p < 0.05$) in the two studies. Concerning the prescription of anti-platelet or anti-coagulant therapy prescription, the study from Campbell et al (2007) reported a 2.04% increase between intervention and control group for recording of the prescription or advice to take aspirin, The study from McGovern et al showed a 24.5% for the more general prescription of anti-platelet or anti-coagulant therapy prescription.¹¹ Furthermore, Campbell et al,(2007) stated there is no difference between the intervention and control group for recording of frequency patterns of angina attacks, recording of exercise capacity, recording of referral to specialists for exercise stress testing or assessment, and recording of weight advice for overweight patients, this study shows a difference between intervention and control group for recording of dietary advice of 1.9% ($p < 0.05$). In addition, the study of McGovern et al shows for β -blocker therapy prescription, for the ACE inhibitor prescription and for influenza vaccination a difference between intervention and control group of respectively 27.4% ($p < 0.05$), 11.5% ($p < 0.05$), 28.1% ($p < 0.05$).¹¹

As indicated in one cross-sectional study, there was no significant association between higher scores on clinical domain and emergency admission and between higher scores on organizational domain and emergency admission. According to the same study, the association between higher scores on additional services domain and emergency admission ranged from no significant association to a positive association with an odds ratio of 1.03 ($p < 0.05$), depending on the primary care trust.

18. Stroke (CVA, TIA)

a. Effectiveness

i. Historical comparison studies, before after time points

		MRI/CT scan recording (Δ)	Smoking status recording (Δ)	Smoking cessation advice (Δ)	Cholesterol recording (Δ)	Cholesterol < 5 mmol/l (Δ)
Simpson et al, 2006	UK, ∞	43.3% *	49.4% *	17% *	52.1% *	NS
		Blood pressure recording (Δ)	Blood pressure < 145/85 mmHg (Δ)	Antiplatelet or anticoagulant therapy (Δ)	Flu vaccination (Δ)	BMI recording (Δ)
		33.4% *	NS	32.3% *	34.2% *	28.8% *

ii. Cross-sectional studies

		Total QOF score for stroke
Williams et al, 2006 ³	UK, ∞	Relationship with adherence to RCP guidelines NS

		Association between higher scores on additional services domain and emergency admission (OR)	Association between higher scores on clinical domain and emergency admission (OR)	Association between higher scores on organisation domain and emergency admission (OR)
Downing et al, 2007	UK, ∞	1.05 * 1.02 *	NS NS	NS NS

Three studies in the UK focused on P4Q effects with stroke as target condition. One before after study found no significant difference in the performance on the cholesterol < 5 mmol/l target and on blood pressure < 145/85 mmHg. There was a significant improvement of 34.3% (p<0.05) in MRI/CT scan recording, 49.4% (p<0.05) in smoking status recording, 17% (p<0.05) in giving smoking cessation advice, 52.1% (p<0.05) in cholesterol recording, 33.4% (p<0.05) in blood pressure recording, 32.3% (p<0.05) in prescription of antiplatelet or anticoagulant therapy, 34.2% (p<0.05) in influenza vaccination and 28.8% (p<0.05) in BMI recording⁵⁸.

One cross sectional study in the UK found no significant relationship between the total QOF score for stroke and the level of adherence to an evidence based clinical practice guideline for stroke care².

Finally, one cross sectional study found no significant relationship between higher scores at the clinical and/or organization domain of the QOF on the one hand and the rate of emergency admissions of stroke patients on the other hand. There was a positive significant relationship of scoring higher on the additional services domain with being more likely to experience a hospital emergency admission (OR 1.05 and 1.02, p<0.05)⁵⁷. The direction of this relationship cannot be specified within this study design.

19. Asthma

a. Effectiveness

i. Concurrent + historical comparison studies

		Improvement Asthma controller use
Levin-scherz et al, 2006	US, ∞, △	NS

		Normal or predicted peak expiratory flow or record of difficulty using meter recorded during the previous 5 yr	Daily, nocturnal, or activity-limiting symptoms recorded during the previous 15 mo	Smoking status recorded during the previous 5 yr	Inhaler technique recorded during the previous 5 yr
Campbell et al, 2007	UK, ∞	NS	0.91 ** (transformed)	0.59 ** (transformed)	0.71 * (transformed)

ii. Historical comparison studies, before-after time point

		% of quality indicators achieved (Δ) ‡
Steel et al, 2007	UK, ∞	14% *** Variation C: 59 ±24 Variation I: 73 ±23

‡ Indicators:

- The percentage of patients aged eight and over diagnosed as having asthma from 1st April 2003 where the diagnosis has been confirmed by spirometry or peak flow measurement.
- The percentage of patients with asthma who have had an asthma review in the last 15 months.
- Patients with asthma, if on current medication, should have their predicted peak flow calculated on at least one occasion.
- Patients presenting with asthma in the last 5 years but not on current medication, should have their predicted peak flow calculated on at least one occasion.
- Patients on current medication or presenting with asthma should have their inhaler technique checked at least once every 5 years.

- For patients on current medication or presenting with asthma, patients should be asked at every asthma consultation in the last year about:
 - any difficulty sleeping due to asthma.
 - any asthma symptoms during the day (eg cough, wheeze).
 - whether asthma has interfered with usual daily activities.

cross-sectional studies

		Association between higher scores on additional services domain and emergency admission (OR)	Association between higher scores on clinical domain and emergency admission (OR)	Association between higher scores on organisation domain and emergency admission (OR)	Use of registries, reminders, medical record flowsheets, physician feedback, formal training on guidelines, on site health promotion, patient education classes, support groups for parents or adolescents, self management support programs, case management (β)
Downing et al, 2007	UK, ∞	1.04 * NS	NS NS	NS NS	
Rittenhouse & Robinson, 2006	US, ∞ \triangle				0.41 *

Four studies in the UK focused on asthma: one concurrent + historical comparison study, one historical comparison study, before-after time point and two cross sectional studies ^{48, 55, 57, 60}. Results reported by Campbell et al (2007) are recalculated towards a comparable effect size, using an e^x formula to take the logit transformation into account.

One concurrent + historical comparison study reports no significant difference between intervention and control group for normal or predicted peak expiratory flow or record of difficulty using meter recorded during the previous 5 year ⁴⁸. In addition this study indicates a significant difference for recording of Daily, nocturnal, or activity-limiting symptoms which amounts to 2.48 ($p < 0.01$), for recording of Smoking status which amounts to 1.80 ($p < 0.01$) and for recording of inhaler technique, which amounts to 2.03 ($p < 0.05$) ⁴⁸.

The difference in percentage quality indicators achieved, based on nine indicators, amounts to 14% ($p < 0.001$) in one historical comparison study, before-after time point ⁶⁰.

As indicated by one cross-sectional study, specifically for asthma the relation between higher scores on the additional services domain of the QOF and emergency admission rate from no significant relation to a positive relation with an odds ratio of 1.04 ($p < 0.05$), depending on the primary care trust. This study also reports no significant relation between higher scores on clinical domain and emergency admission, and between higher scores on organisation domain and emergency admission ⁵⁷. As indicated by one other cross-sectional study, the use of registries, reminders, medical record flowsheets, physician feedback, formal training on guidelines, on site health promotion, patient education classes, support groups for parents or adolescents, self management support programs, case management is positively associated with the external incentive index with a correlation index of 0.41 ($p < 0.05$)

Two concurrent + historical comparison studies in the US focused on asthma ^{21, 59}. The generic results provided by Glimore and colleagues are described elsewhere (see section X). There was no significant difference in improved asthma controller use between the intervention group and the control group in the first study ⁵⁹.

20. Hypertension

a. Effectiveness

i. Historical comparison studies, before-after time point

		% of quality indicators achieved (Δ) \downarrow
Steel et al, 2007	UK, ∞	12% *** Variation C: 58 \pm 17 Variation I: 70 \pm 16

\downarrow indicators

- The percentage of patients with hypertension whose notes record smoking status at least once.
- The percentage of patients with hypertension who smoke, whose notes contain a record that smoking cessation advice has been offered at least once.
- The percentage of patients with hypertension in which there is a record of the blood pressure in the past 9 months.
- The percentage of patients with hypertension in whom the last blood pressure (measured in last 9 months) is 150/90 or less.
- An individual in whom hypertension is identified or for whom hypertension is treated is offered lifestyle advice at the following times:
 - a. initially.
 - b. periodically.
- If a person aged 65 or older is diagnosed with hypertension, THEN nonpharmacological therapy with lifestyle modification for treatment of hypertension should be recommended.
- When an individual is identified as having hypertension, a formal cardiovascular risk assessment including the following is carried out: a) medical history. Ascertain patients' alcohol consumption and encourage a reduced intake if patients drink excessively.
- Initial history should document assessment of the following within 3 months of diagnosis (list as specified in left hand column).
- When an individual is identified as having hypertension, a formal cardiovascular risk assessment including the following is carried out:
 - urine strip test for blood and protein.
 - blood electrolytes and creatinine.
 - blood glucose.
 - serum total and HDL cholesterol.
 - 12-lead electrocardiogram.
- Initial laboratory investigations should include the following tests within 3 months of diagnosis (list as 1-5 specified in left hand column).

One historical comparison study, before-after time point in the UK focused on hypertension, based on 10 indicators (see section X) ⁶⁰. The difference in percentage quality indicators achieved between the intervention group and the control group amounts to 12% ($p < 0.001$) ⁶⁰.

21. Smoking cessation

a. Effectiveness

i. Randomized

		Percentage of patients, tobacco use status identified at last visit (Δ)	Percentage of smokers who received advice to quit at last visit (Δ)	Percentage of smokers who were offered assistance to quit at last visit (Δ)	Referral rate (Δ)	Effect of intervention on smoking abstinence (OR)	Prevalence of smoking abstinence 12 months after recruitment (Δ)
Roski et al, 2003	US, Δ	7.9% **	NS	NS	-		
An et al, 2008 ⁹	US, Δ	-	-	-	6.2% *** Variation C: 8.0%-14.9% Variation I: 1.5%-6.9%		
Twardella & Brenner, 2007	Germany, Δ	-	-	-		NS	NS

ii. Historical comparison studies, multiple time points

		Recording of smoke status (Δ)	Advice to quit- rates (Δ)
Amundson et al, 2003	US, ∞	24% *** Variation C: 24%-78% Variation I: 52%-97%	21% ** Variation C: 0%-71% Variation I: 22%-93%
Coleman et al, 2007 ⁵	UK, ∞	1.88 * (Rate Ratio)	3.03 * (Rate Ratio)

iii. cross-sectional studies

		Relationship between receiving financial incentives from HMO to promote smoking cessation interventions and (OR)					Relationship between receiving additional income from health plan for scoring well on quality measures and (OR)					Relationship between receiving better contracts with health plans for scoring well on quality measures (OR)				
		Perceived specific interventions need	Perceived intervention evaluation need	Offer Health Promotion Program	Provide NRT Starter Kit	Provide Written Materials	Perceived specific interventions need	Perceived intervention evaluation need	Offer Health Promotion Program	Provide NRT Starter Kit	Provide Written Materials on:	Perceived specific interventions need	Perceived intervention evaluation need	Offer Health Promotion Program	Provide NRT Starter Kit	Provide Written Materials on:
McMenamin et al, 2003	US, ∞	3.67 ***	14.46 ****	3.63 ***	2.75 **	(a) 2.13 * (b) 3.11 ** (c) NS	1.9 **	NS	NS	NS	(a) NS (b) NS (c) 1.49 *	NS	NS	NS	NS	(a) NS (b) NS (c) NS

- (a) Pharmaco-therapy
- (b) Counseling
- (c) Self-help

Four studies in the US have focused on smoking cessation: two randomized studies, one historical comparison study, multiple time points and one cross sectional study ^{61, 63, 64, 9}. The difference in recording of smoking status, between intervention group and comparison group is 7.9% (p<0.01) in one randomized study and 24% (p<0.001), in one historical comparison study, multiple time points ^{61, 63}. The difference in advice to quit rates between the intervention group and the control group, ranges from no significant difference in one randomized study to 21% (p<0.01) in one historical comparison study, multiple time points ^{61, 63}.

There was no significant difference in the percentage of smokers who were offered assistance to quit at the last visit in one randomized study ⁶¹. The difference in referral rate between intervention group and control group amounts to 6.2% (p<0.001) for another randomized study ⁹.

One cross sectional study has found a relationship between financial incentives to promote smoking cessation interventions and the perception that specific interventions are required, perception that intervention evaluation is needed, offering health promotion programs, providing NRT Starter Kit, providing written materials on pharmacotherapy, providing written materials on counselling and providing written materials on self-help with an odds ratio of respectively 3.67 (p<0.001), 14.46 (p<0.0001), 3.63 (p<0.001), 2.75 (p<0.01), 2.13 (p<0.05), 3.11 (p<0.01) and no significant odds ⁶⁴.

The same study came to not significant results on the same indicators for receiving better contracts for scoring well on quality measures. Receiving additional income for scoring well on quality measures showed only a positive significant relationship with the perceived specific interventions need (OR 1.9, p<0.01) and with providing written materials on self-help (OR 1.49, p<0.05).

One historical comparison study, multiple time points in the UK has focused on smoking cessation ⁵. This study indicates a relationship between income for quality and recording of smoke status with a positive rate ratio of 1.88 (p<0.05)⁵. This study also indicates a relationship between income for quality and advice to quit smoking rates with a positive rate ratio of 3.03 (p<0.05) ⁵.

One randomized study in Germany has focused on smoking cessation ⁶². The difference in the effect of the intervention on smoking abstinence and on the prevalence of smoking abstinence 12 months after recruitment, between intervention group and control group were both not significant ⁶².

22. Depression/mental illness

a. Effectiveness

i. Historical comparison studies, before-after time point

		% of quality indicators achieved (Δ) ‡‡
Steel et al, 2007	UK, ∞	NS

‡‡ indicators

- Healthcare professionals should always ask patients with depression directly about suicidal ideas and intent.
- IF a person aged 65 or older receives a diagnosis of a new depression episode, THEN the diagnosing physician should ask on the day of diagnosis whether the person aged 65 or older had any thoughts about suicide.
- The presence or absence of suicidal thoughts should be sought out routinely in all patients found to be depressed.
- In the assessment of depression, enquiry should be made about:
 - alcohol use.
 - substance misuse.
 - current medication.
- Patients started on antidepressants who are not considered to be at increased risk of suicide should normally be seen after 2 weeks.
- IF a person aged 65 or older receives a diagnosis of a new depression episode, THEN they should be offered a follow-up appointment within 4 weeks.
- Patients with depression prescribed antidepressant drug treatment should be invited for review by a health care professional within 4 weeks of initiating of initiating antidepressant drug treatment.

One historical comparison study, before-after time points in the UK focused on depression as a mental illness in a primary care setting, based on seven indicators ⁶⁰. There is no significant difference in the percentages of quality indicators achieved between the intervention group and the control group in this study ⁶⁰.

23. Epilepsy

a. Effectiveness

i. cross-sectional studies

		relationship between the proportion of seizure-free epilepsy patients and the percentage of epilepsy patients who had at least one epilepsy-related emergency hospitalisation (β)	relationship between the proportion of seizure-free epilepsy patients and the proportion of epilepsy-related emergency hospitalisations over the number of epilepsy-treated patients (β)
Shohet et al, 2007	UK, ∞	0.43 *	0.30 *

One cross sectional study in the UK focused on P4Q effectiveness in epilepsy ⁶⁵. As this study indicates, the proportion of seizure free epilepsy patients under the QOF contract is related to the percentage of epilepsy patients who had at least one epilepsy related emergency hospitalisation with a regression coefficient of 0.43 ($p < 0.05$). It is also related to the percentage rate of epilepsy-related emergency hospitalisations over the number of epilepsy-treated patients with a regression coefficient of 0.30 ($p < 0.05$) ⁶⁵.

24. Kidney dialysis/chronic kidney disease

a. Effectiveness

i. Historical comparison studies, multiple time points

		Number of visits per patient month (Δ)	Patient-months with 0 visit (%) (Δ)	Kt/V (Δ)	Kt/V ≥ 1.2 (%)	Albumin (g/dL) (Δ)
Mentari et al, 2005	US, ■	1.62 ***	-11.8 ***	0.02 **	NS	-0.07 ***
		Albumin ≥ 3.8 g/dL (%) (Δ)	Hemoglobin (g/dL) (Δ)	Hemoglobin $11 \geq$ g/dL (%) (Δ)	Phosphorus (mg/dL) (Δ)	Calcium (mg/dL) (Δ)
Mentari et al, 2005	US, ■	-5.4 ***	-0.05 **	-0.8 *	-0.15 ***	-0.08 ***
		Patients with catheter (%) (Δ)	Ultrafiltration volume (L) (Δ)	Treatments shortened $\geq 10\%$ per patient-month (Δ)	Skipped treatments per patient-month (Δ)	Hospital admissions per patient-month (Δ)
Mentari et al, 2005	US, ■	-2.1 **	0.11 *	NS	-0.05 ***	NS

One historical comparison study, multiple time points in the US focused on chronic kidney disease ⁶⁶. This study indicates no significant differences in percentage of patients meeting the Kt/V target (dialyzer clearance of urea* dialysis time/ patient's total body water $\geq 1.2\%$), in number of shortened treatments per patient-month, and in number of hospital admission per patient-month. As indicated by this study there is a difference between intervention group and control group in number of visits per patient month: 1.62 ($p < 0.001$), percentage of patient months with no visits: minus 11.8% ($p < 0.001$), mean Kt/V: 0.02 ($p < 0.01$), mean albumin values: 0.07 ($p < 0.001$), percentage of patients meeting the albumin target (albumin ≥ 3.8 g/dl): minus 5.4% ($p < 0.001$), mean haemoglobin values: minus 0.05 ($p < 0.01$), percentage of patients meeting the haemoglobin target (haemoglobin ≥ 11 g/dl): minus 0.8% ($p < 0.05$), mean phosphorus value: minus 0.15 ($p < 0.001$), mean calcium value: minus 0.08 ($p < 0.001$), percentage of

patient with a catheter: minus 2.1% ($p<0.01$), mean ultrafiltration volume: 0.1 l ($p<0.05$), and the number of skipped treatments per patient month: minus 0.05 ($p<0.001$)⁶⁶.

25. Osteoarthritis

a. Effectiveness

i. Historical comparison studies, before-after time point

		% of quality indicators achieved (Δ) \Downarrow
Steel et al, 2007	UK, ∞	NS

\Downarrow indicators

- IF oral pharmacological therapy is initiated to treat osteoarthritis among people aged 65 or older, THEN paracetamol should be the first drug used, unless there is a contraindication to use.
- Patients with a new diagnosis of osteoarthritis who wish to take medication for joint symptoms should be offered a trial of paracetamol if not already tried.
- IF oral pharmacological therapy for osteoarthritis is changed from paracetamol to a different oral agent among people aged 65 or older, THEN the patient should have had a trial of maximum dose paracetamol (suitable for age/co-morbidities).
- If NSAIDS are considered, ibuprofen should be considered for first line treatment unless contraindicated or intolerant.
- Cox II selective inhibitors should be used, in preference to standard NSAIDs, when clearly indicated as part of the management of RA or OA only in patients who may be at 'high risk' of developing serious gastrointestinal adverse effects.
- IF a person aged 65 or older is treated for symptomatic osteoarthritis, THEN functional status and degree of pain should be assessed at least annually.
- IF an ambulatory person aged 65 or older has a diagnosis of symptomatic osteoarthritis, THEN education regarding the natural history, treatment and self-management of the disease should be offered at least once.
- IF a person aged 65 or older is treated with a non-selective NSAID, or IF a person aged 65 or older is treated with a COX-2 selective NSAID THEN the patient should be advised of the gastrointestinal and renal risks associated with this drug.
- IF a person aged 65 or over is treated with an NSAID (selective or non-selective), THEN they should be asked about gastro-intestinal symptoms at least annually.
- IF a person aged 65 or older with severe symptomatic osteoarthritis of the knee or hip has failed to respond to non- pharmacological and pharmacological therapy, THEN the patient should be offered referral to an orthopaedic surgeon to be evaluated for total joint replacement within 6 months unless surgery is contraindicated.
- Patients with severe symptomatic osteoarthritis of knee or hip who have failed to respond to conservative therapy should be offered referral to an orthopaedic surgeon for consideration of joint replacement.

One historical comparison study, before-after time point in the UK focused on osteoarthritis, based on eleven indicators (see section X)⁶⁰. There is no significant difference in percentage of quality indicators achieved between the intervention group and the control group in this study⁶⁰.

Generic findings

I. Chronic disease management for multiple patient groups

a. Effectiveness

i. Concurrent + Historical comparison studies

1. Incentivized

		Overall achievement ^{]]}
Gilmore et al, 2007	US, *	Visiting only participating physicians: OR 1.06-1.27 *

^{]]} Included indicators:

- Breast cancer screening: Receiving at least one screening mammogram
- Cervical cancer screening: Receiving at least one Papanicolaou smear
- Colorectal cancer screening: Receiving at least one fecal occult blood test, barium enema, sigmoidoscopy or colonoscopy
- Use of ACE inhibitor in CHF: Receiving at least one prescription for an ACE inhibitor, angiotensin receptor blocker, or nitrates and hydralazine
- Use of long term asthma control drugs: Receiving at least one prescription for a long-term asthma control drug
- Diabetic retinal exam: have an ophthalmoscopic examination performed by an eye care professional
- hbA1c for diabetics: Receiving at least two hbA1c tests
- Antihypertensive drug compliance: Receiving antihypertensive prescription coverage for at least 80% of eligible days
- Lipid lowering drug compliance: Receiving lipid-lowering drug prescription coverage for at least 80% of eligible days
- Childhood immunizations VZV: Receiving vaccination between the first and second birthdays or with a history of varicella disease
- Childhood immunizations MMR: Receiving vaccination between the first and second birthdays of disease diagnosis for measles, mumps, and rubella

2. Not incentivized

		P4Q payment amount as a function of health promotion
Langham et al, 1995	UK, ∞	Standardized mortality ratio NS

ii. Cross sectional studies

1. Incentivized

		P4Q incentive tied to the measure ^{]]}
Mehrota et al, 2007	US, ∞	Having a quality improvement initiative targeting an included indicator: OR 1.6 *

^{]]} Indicators:

- Receiving mammography
- Receiving chlamydia screening
- Appropriate well child visits
- Diabetes hbA1c testing
- Asthma controller medication
- Hyperlipidemia screening and appropriate control after an acute cardiac event
- Patients with hypertension with blood pressure < 14/9 mmHg

2. Not incentivized

		Cholesterol < 5 mmol/l for diabetes and CVD
Ashworth et al, 2007a	UK, ∞	Statin prescribing volume: $\beta = 0.26$ *** and $\beta = 0.08$ ***
		Incentivized CHD, asthma and diabetes 2 indicators
Campbell et al, 2007	UK, ∞	Not incentivized indicators NS
		QOF additional services, clinical, organizational domain
Downing et al, 2007	UK, ∞	Not incentivized all cause mortality NS

2. Quality and Outcomes Framework effect

a. Effectiveness

i. Concurrent comparison studies

		% indicators achieved on indicators included vs. excluded in QOF (Δ)
Steel et al, 2008	UK, ∞	16.9% less for excluded * Variation included indicators: 73.2%-75.9% Variation excluded indicators: 56.3%-59.2%

ii. Historical comparison studies, multiple time points

		Overall achievement (Δ)
Doran et al, 2008a	UK, ∞	4.2% **** Variation C: 79.0%-89.1% Variation I: 86.0%-91.5%

		Increase in prescribing defined daily doses in QOF vs. non QOF (Δ) (after implementation series)
MacBride-Stewart et al, 2008	UK, ∞	1.09% *** (period 1), 0.69% *** (period 2)

b. Use of care management processes

i. Historical comparison studies, before after time points

		Use of registries, actionable reports, HEDIS results, electronic checks for interactions, retrieval of lab results, access to clinical notes (Δ)
Williams et al, 2006 ²	US, Δ	Between 9 and 27% increase associated with P4Q program *

		The reported effect of clinical practice guidelines on clinical practice
O'Malley et al, 2007	US, Δ ∞	Quality measures began to affect compensation: $\beta = 0.09$ *

ii. Cross-sectional studies

		% of conditions with protocols or guidelines	% of quality of care data elements reported to the board	% of quality of care data elements for which benchmarks exist	CMP's combined use
Shortell et al, 2001	US, Δ	Cost containment, productivity and quality combined incentives use NS	Cost containment, productivity and quality combined incentives use NS	Cost containment, productivity and quality combined incentives use NS	Cost containment, productivity and quality combined incentives use: $\beta = 0.05$ *
				Use of case management, physician feedback, disease registries, clinical guidelines, self management skills	
Casalino et al, 2003	US, Δ			Better contracts for quality: $\beta = 0.74$ ***, bonus from health plans NS	
				Use of health promotion programs	
McMenamin et al, 2004	US, Δ			Income for quality NS	
		Reporting of making structural changes	Reporting of making process changes	Use of clinical algorithms, guidelines, case management, education, computerized order entry, electronic medical records	
Reiter et al, 2006	US, Δ	'MD incentives' NS, 'MD's incentives aligned with board' NS,	'MD incentives' NS, 'Incentive provides leverage with MDs': $z = 2.748$ **, 'earn program incentive payment': $z = 4.765$ **	NS except case management in relationship to structure changes: $z = 2.349$ *	
				Use of registries, reminders, medical record flowsheets, performance feedback to physicians, formal training for physicians on established clinical guidelines, on site health promotion, patient education classes, support groups for parents or for adolescents, self management support programs, case managers	
Rittenhouse & Robinson, 2006	US, $\infty\Delta$			External incentives index positively related: $\beta = 0.41$ to 0.65 **	
				Adoption of guidelines in order entry systems with decision support	
Simon et al, 2007	US, $\Delta\infty$			External incentives index: $\square = 0.119$ ***, OR 1.12 *	

The use of care management processes as an outcome measure of a P4Q intervention is studied in eight US based studies: two before after studies^{3,72} and six cross sectional studies^{55,73-77}. One before after study noted a between 9 and 27% increase ($p<0.05$) in the use of IT related care management processes such as electronic checks for interactions, retrieval of lab results and access to clinical notes³. The second before after study found a positive relationship between P4Q and the reported effect of clinical practice guidelines on clinical practice (0.09 regression coefficient, $p<0.05$)⁷².

Of the six cross sectional studies three investigated the relationship of P4Q use with a wide array of care management processes, including guideline use, case management, reminder use, feedback use, education, etc. One of these studies found a clear positive relationship of CMP use with a combined external incentive index (0.41 to 0.65 regression coefficient, $p<0.01$)⁵⁵. This was partially confirmed by Casalino et al (2003) who found a positive relationship of CMP use with receiving better contracts for quality (0.74 regression coefficient, $p<0.001$), but not with receiving a bonus⁷³. Finally, one cross sectional study found on the one hand no significant relationship of P4Q with CMP use, and no relationship with the reporting of making structural changes. On the other hand earning program incentive payments as a factor and incentives providing leverage with physicians as a factor were positively related to the reporting of making process changes ($z=4.765$, $p<0.01$ and $z=2.748$, $p<0.05$ respectively)⁷⁴.

One cross sectional study focused on the relationship between the score on an external incentives index and the adoption of guidelines in order entry systems with decision support. They came to a positive finding with an OR of 1.12 ($p<0.05$) and a 0.119 ($p<0.001$) regression coefficient⁷⁷. One cross sectional study found no significant relationship of receiving income for quality and the use of health promotion programs⁷⁶. Finally, one cross sectional study reported no significant relationship of the combined use of quality incentives together with cost containment and productivity incentives on the one hand and the percentage of conditions with protocol or guideline use, the percentage of quality of care data elements reported to the board and the percentage of quality of care data elements for which benchmarks exist on the other hand⁷⁵. They did find a positive relationship with the combined use of these three CMP indicators (0.05 regression coefficient, $p<0.05$).

c. Patient and provider satisfaction

i. Historical comparison studies, before after time points

		Physicians Perceived management structure support	Physicians Perceived demands	Physicians Perceived intrinsic motivation	Physicians Perceived overall quality of life
Gene badia et al, 2007	Spain, ■	Significant pos effect (Signif. positive effect on 8 statements ***, 4 NS) Variation C: ± 0.764 Variation I: ± 0.654	Significant pos effect (Signif. positive effect on 10 statements *, ***, 3 NS (higher workload)) Variation C: ± 0.704 Variation I: ± 0.594	NS (Signif. positive effect on 1 statement (family support) ***, 9 statements NS) Variation C: ± 0.498 Variation I: ± 0.448	NS
		Nurses Perceived management structure support	Nurses Perceived demands	Nurses Perceived intrinsic motivation	Nurses Perceived overall quality of life
		Significant pos effect (Signif. positive effect on 8 statements ***, 4 NS) Variation C: ± 0.869 Variation I: ± 0.716	NS	NS (Signif. positive effect on 3 statements (job satisfaction, motivation, job proudness) *, **, 7 statements NS) Variation C: ± 0.493 Variation I: ± 0.476	NS

		Patient satisfaction on organization	Patient satisfaction on physicians	Patient satisfaction on nurses	Patient satisfaction on support personnel
		NS	NS	NS	NS
		Patient satisfaction on healthcare	Patient satisfaction on premises and facilities	Overall patient satisfaction	
		NS	NS	NS	

ii. . Cross-sectional studies

		Financial incentives as a function of patient satisfaction
Safran et al, 2000	US, ▫	$\beta = 2.57^{**}$ for patient perceived access to care, $\beta = 2^{*}$ for patient knowledge, $\beta = 3.5^{*}$ for preventive counseling use, continuity NS, integration of care NS, clinical interaction NS, interpersonal treatment NS, trust NS

		Provider perceived financial incentive to reduce services	Provider perceived financial incentive to increase services
Reschovsky et al, 2006	US, △	Provider perceived ability to provide quality care negatively related ****	Provider perceived ability to provide quality care positive related ****

The influence of P4Q on patient and provider satisfaction is only described here based on quantitative evaluation studies. A large group of qualitative studies exist focusing on the influence of P4Q on the different aspects of the provider and patient experience, going beyond the scope of this study. Focusing on quantitative evaluation studies only three studies can be identified, of which one before after study and two cross sectional studies.

The before after study was conducted in Spain ⁷⁸. According to this study P4Q showed no significant relationship with patient satisfaction in general or with sub dimensions of patient satisfaction such as satisfaction on organization, on physicians, on nurses, on support personnel, on healthcare, and on premises and facilities. However, there was a significant relationship with both physicians and nurses perceived management structure support ($p < 0.001$). The perceived demands (higher workload) were significant for physicians ($p < 0.001$), but not for nurses. The majority of statements concerning perceived intrinsic motivation showed no significant relationship, both for physicians as for nurses.

The statements that showed a significant relationship were of a positive nature in direction, i.e. higher job satisfaction, motivation and pride for nurses, and higher family support for physicians. There was no significant relationship with the physicians and nurses perceived overall quality of life.

Both cross sectional studies were conducted in the US. One reported a negative relationship of provider perceived financial incentives to reduce services with the perceived ability to provide quality care ($p < 0.0001$) and, vice versa, a positive relationship of provider perceived financial incentives to increase services with the perceived ability to provide quality care ($p < 0.0001$) ⁸⁰.

The second study reported on the use of patient satisfaction as a P4Q target ⁷⁹. P4Q performance on this measure wasn't significantly related to patient perceived continuity, integration of care, clinical interaction, interpersonal treatment or trust. But there was a positive relationship with perceived access to care (2.57 regression coefficient, $p < 0.01$), with patient knowledge (2 regression coefficient, $p < 0.05$) and with the use of preventive counseling (3.5 regression coefficient, $p < 0.05$).

APPENDIX 16C SYSTEMATIC REVIEW RESULTS PERIOD JANUARY – JULY 2009

Because the literature on P4Q is quickly expanding, with new studies being published every month, this Appendix presents most recent evidence for the period January – July 2009. The methods used are identical to the systematic review methods described in Chapter Two. This Appendix provides an update of the results presented in Chapter Four and Five. Major findings adding new insights are described, whereas findings identical to previously reported study results are not focused upon.

A total of 18 evaluation studies are included for the first half of 2009.⁸³⁻¹⁰⁰ Nine studies originate from the USA, six from the UK, one from Argentina, one from Germany, and one from Spain. Thirteen studies focus upon primary care, four upon acute hospital care and one study upon both. In terms of study design these studies do not include randomization. Three studies make use of a combined concurrent plus historical comparison design. Three studies apply interrupted time series analysis. One study addresses economic modeling. Three studies perform a before after P4P effect comparison without concurrent comparison group, and finally, eight studies make use of a cross sectional study design.

Study results confirm the previous findings with regard to P4Q effects on clinical effectiveness. With regard to sustainability of change, there is evidence that a plateau of performance might be reached, with attenuation of the initial improvement rate.⁸⁵ The effect of P4Q on non-incentivized quality measures varied from none to positive. However, one study reported a declining trend in improvement rate for non-incentivized measures of asthma and CHD after a performance plateau was reached.⁸⁵ One study found positive effects for P4Q targets when applied to non-incentivized medical conditions (10.9% effect size), suggesting a spillover effect.⁹⁸ Literature on equity effects confirmed an absent or positive effect⁸⁴ and literature on cost effectiveness confirmed previous positive findings⁹⁶. One time-series study reported no effect on non-incentivized access and communication measures.⁸⁵ This study, however, did observe a patient self-reported decrease in timely access to patients' regular doctors, which might be a negative spillover effect. With regard to patient-centeredness, one Spanish study found positive P4Q effects on patient experience.⁹³ Another before-and-after study, this one from Argentina, reported that P4Q had no significant effect on patient satisfaction, due to a ceiling effect.⁹⁴ The occurrence and relevance of ceiling effects is emphasized also in other studies from the update set.^{83, 84}

One study confirmed that gaming is currently kept to a minimum for USA settings, as was previously reported for UK settings.⁹⁵ One study found a strong relationship between the program adoption rate by physicians and incentive size.⁸⁷ In this instance, the reward level, which was also determined by the number of eligible patients per provider, explained 89 to 95% of the variation in participation.

Regarding programs in which the provider is either a team or organization, one study found no relationship between the role of leadership and P4Q performance.⁹⁹ These authors also reported a positive relationship between P4Q performance effects and an organizational culture that supports the coordination of care, the perceived pace of change in the organization, the willingness to try new projects, and a focus on identifying system errors rather than blaming individuals. Furthermore, they found a positive relationship between P4Q performance and the multidisciplinary team approach, the use of clinical pathways, and having adequate human resources for quality improvement projects. A study that examined the reduction of overuse found a negative relationship between the number of providers in a practice and P4Q performance.⁹¹

REFERENCES

1. Coleman T, Wynn AT, Stevenson K, Cheater F. Qualitative study of pilot payment aimed at increasing general practitioners' antismoking advice to smokers. *British Medical Journal*. 2001;323(7310):432-5.
2. Williams PH, De Lusignan S. Does a higher 'quality points' score mean better care in stroke? An audit of general practice medical records. *Inform.Prim.Care*. 2006;14(1):29-40.
3. Williams TR, Raube K, Damberg CL, Mardon RE. Pay for performance: its influence on the use of IT in physician organizations. *J Med.Pract.Manage*. 2006;21(5):301-6.
4. Coleman K, Reiter KL, Fulwiler D. The impact of pay-for-performance on diabetes care in a large network of community health centers. *Journal of Health Care for the Poor and Underserved*. 2007;18(4):966-83.
5. Coleman T, Lewis S, Hubbard R, Smith C. Impact of contractual financial incentives on the ascertainment and management of smoking in primary care. *Addiction*. 2007;102(5):803-8.
6. Millett C, Gray J, Saxena S, Netuveli G, Khunti K, Majeed A. Ethnic disparities in diabetes management and pay-for-performance in the UK: The Wandsworth prospective diabetes study. *Plos Medicine*. 2007;4(6):1087-93.
7. Millett C, Gray J, Saxena S, Netuveli G, Majeed A. Impact of a pay-for-performance incentive on support for smoking cessation and on smoking prevalence among people with diabetes. *Canadian Medical Association Journal*. 2007;176(12):1705-10.
8. Millett C, Car J, Eldred D, Khunti K, Mainous AG, Majeed A. Diabetes prevalence, process of care and outcomes in relation to practice size, caseload and deprivation: national cross-sectional study in primary care. *Journal of the Royal Society of Medicine*. 2007;100(6):275-83.
9. An LC, Bluhm JH, Foldes SS, Alesci NL, Klatt CM, Center BA, et al. A randomized trial of a pay-for-performance program targeting clinician referral to a state tobacco quitline. *Archives of Internal Medicine*. 2008;168(18):1993-9.
10. McGovern MP, Williams DJ, Hannaford PC, Taylor MW, Lefevre KE, Boroujerdi MA, et al. Introduction of a new incentive and target-based contract for family physicians in the UK: good for older patients with diabetes but less good for women? *Diabetic Medicine*. 2008;25(9):1083-9.
11. McGovern MP, Boroujerdi MA, Taylor MW, Williams DJ, Hannaford PC, Lefevre KE, et al. The effect of the UK incentive-based contract on the management of patients with coronary heart disease in primary care. *Family Practice*. 2008;25(1):33-9.
12. Millett C, Gray J, Bottle A, Majeed A. Ethnic disparities in blood pressure management in patients with hypertension after the introduction of pay for performance. *Ann.Fam.Med*. 2008;6(6):490-6.
13. Millett C, Gray J, Wall M, Majeed A. Ethnic Disparities in Coronary Heart Disease Management and Pay for Performance in the UK. *J Gen.Intern.Med*. 2008.
14. Millett C, Netuveli G, Saxena S, Majeed A. Impact of pay for performance on ethnic disparities in intermediate outcomes for diabetes: longitudinal study. *Diabetes Care*. 2008.
15. Doran T, Fullwood C, Gravelle H, Reeves D, Kontopantelis E, Hiroeh U, et al. Pay-for-performance programs in family practices in the United Kingdom. *New England Journal of Medicine*. 2006;355(4):375-84.
16. Kouides RW, Lewis B, Bennett NM, Bell KM, Barker WH, Black ER, et al. A Performance-Based Incentive Program for Influenza Immunization in the Elderly. *American Journal of Preventive Medicine*. 1993;9(4):250-5.
17. Kouides RW, Bennett NM, Lewis B, Cappuccio JD, Barker WH, LaForce FM. Performance-based physician reimbursement and influenza immunization rates in the elderly. *American Journal of Preventive Medicine*. 1998;14(2):89-95.
18. Schmittiel J, McMenamin SB, Halpin HA, Gillies RR, Bodenheimer T, Shortell SM, et al. The use of patient and physician reminders for preventive services: results from a National Study of Physician Organizations. *Preventive Medicine*. 2004;39(5):1000-6.
19. Fairbrother G, Hanson KL, Friedman S, Butts GC. The impact of physician bonuses, enhanced fees, and feedback on childhood immunization coverage rates. *American Journal of Public Health*. 1999;89(2):171-5.

20. Fairbrother G, Hanson KL, Butts GC, Friedman S. Comparison of preventive care in Medicaid managed care and medicaid fee for service in institutions and private practices. *Ambulatory Pediatrics*. 2001;1(6):294-301.
21. Gilmore AS, Zhao YX, Kang N, Ryskina KL, Legorreta AP, Taira DA, et al. Patient outcomes and evidence-based medicine in a preferred provider organization setting: A six-year evaluation of a physician pay-for-performance program. *Health Services Research*. 2007;42(6):2140-59.
22. Hillman AL, Ripley K, Goldfarb N, Weiner J, Nuamah I, Lusk E. The use of physician financial incentives and feedback to improve pediatric preventive care in Medicaid managed care. *Pediatrics*. 1999;104(4):931-5.
23. Ashworth M, Armstrong D, de Freitas J, Boullier G, Garforth J, Virji A. The relationship between income and performance indicators in general practice: a cross-sectional study. *Health Serv.Manage.Res*. 2005;18(4):258-64.
24. Bond L, Davie G, Carlin JB, Lester R, Nolan T. Increases in vaccination coverage for children in child care, 1997 to 2000: an evaluation of the impact of government incentives and initiatives. *Australian and New Zealand Journal of Public Health*. 2002;26(1):58-64.
25. Chung RS, Chernicoff HO, Nakao KA, Nickel RC, Legorreta AP. A quality-driven physician compensation model: four-year follow-up study. *J Healthc.Qual*. 2003;25(6):31-7.
26. Fairbrother G, Friedman S, Hanson KL, Butts GC. Effect of the vaccines for children program on inner-city neighborhood physicians. *Archives of Pediatrics & Adolescent Medicine*. 1997;151(12):1229-35.
27. LeBaron CW, Mercer JT, Massoudi MS, Dini E, Stevenson J, Fischer WM, et al. Changes in clinic vaccination coverage after institution of measurement and feedback in 4 states and 2 cities. *Archives of Pediatrics & Adolescent Medicine*. 1999;153(8):879-86.
28. Morrow RW, Gooding AD, Clark C. Improving physicians' preventive health care behavior through peer review and financial incentives. *Arch.Fam.Med*. 1995;4(2):165-9.
29. Ritchie LD, Bisset AF, Russell D, Leslie V, Thomson I. Progress and the 1990 Contract. *British Medical Journal*. 1992;304(6830):816-9.
30. Grady KE, Lemkau JP, Lee NR, Caddell C. Enhancing mammography referral in primary care. *Preventive Medicine*. 1997;26(6):791-800.
31. Hillman AL, Ripley K, Goldfarb N, Nuamah I, Weiner J, Lusk E. Physician financial incentives and feedback: Failure to increase cancer screening in Medicaid managed care. *American Journal of Public Health*. 1998;88(11):1699-701.
32. Armour BS, Friedman C, Pitts MM, Wike J, Alley L, Etchason J. The influence of year-end bonuses on colorectal cancer screening. *American Journal of Managed Care*. 2004;10(9):617-24.
33. Pearson SD, Schneider EC, Kleinman KP, Coltin KL, Singer JA. The impact of pay-for-performance on health care quality in Massachusetts, 2001-2003. *Health Affairs*. 2008;27(4):1167-76.
34. Rosenthal MB, Frank RG, Li ZH, Epstein AM. From concept to practice. *Jama-Journal of the American Medical Association*. 2005;294(14):1788-93.
35. Rosenthal MB, de Brantes FS, Sinaiko AD, Frankel M, Robbins RD, Young S. Recognizing High-Quality Care: Analysis of Physician Quality and Resource Use. *American Journal of Managed Care*. 2008;14(10):670-7.
36. Pourat N, Rice T, Tai-Seale M, Bolan G, Nihalani J. Association between physician compensation methods and delivery of guideline-concordant STD care: Is there a link? *American Journal of Managed Care*. 2005;11(7):426-32.
37. Cameron PA, Kennedy MP, McNeil JJ. The effects of bonus payments on emergency service performance in Victoria. *Medical Journal of Australia*. 1999;171(5):243-6.
38. Greenberg MR, Weinstock M, Fenimore DG, Sierzega GM. Emergency department tobacco cessation program: staff participation and intervention success among patients. *J Am.Osteopath.Assoc*. 2008;108(8):391-6.
39. Grossbart SR. What's the return? Assessing the effect of "pay-for-performance" initiatives on the quality of care delivery. *Medical Care Research and Review*. 2006;63(1):29S-48S.
40. Glickman SW, Ou FS, DeLong ER, Roe MT, Lytle BL, Mulgund J, et al. Pay for performance, quality of care, and outcomes in acute myocardial infarction. *Jama-Journal of the American Medical Association*. 2007;297(21):2373-80.

This page is left intentionally blank.

Legal depot : D/2009/10.273/53

KCE reports

- 33 Effects and costs of pneumococcal conjugate vaccination of Belgian children. D/2006/10.273/54.
- 34 Trastuzumab in Early Stage Breast Cancer. D/2006/10.273/25.
- 36 Pharmacological and surgical treatment of obesity. Residential care for severely obese children in Belgium. D/2006/10.273/30.
- 37 Magnetic Resonance Imaging. D/2006/10.273/34.
- 38 Cervical Cancer Screening and Human Papillomavirus (HPV) Testing D/2006/10.273/37.
- 40 Functional status of the patient: a potential tool for the reimbursement of physiotherapy in Belgium? D/2006/10.273/53.
- 47 Medication use in rest and nursing homes in Belgium. D/2006/10.273/70.
- 48 Chronic low back pain. D/2006/10.273.71.
- 49 Antiviral agents in seasonal and pandemic influenza. Literature study and development of practice guidelines. D/2006/10.273/67.
- 54 Cost-effectiveness analysis of rotavirus vaccination of Belgian infants D/2007/10.273/11.
- 59 Laboratory tests in general practice D/2007/10.273/26.
- 60 Pulmonary Function Tests in Adults D/2007/10.273/29.
- 64 HPV Vaccination for the Prevention of Cervical Cancer in Belgium: Health Technology Assessment. D/2007/10.273/43.
- 65 Organisation and financing of genetic testing in Belgium. D/2007/10.273/46.
66. Health Technology Assessment: Drug-Eluting Stents in Belgium. D/2007/10.273/49.
70. Comparative study of hospital accreditation programs in Europe. D/2008/10.273/03
71. Guidance for the use of ophthalmic tests in clinical practice. D/2008/10.273/06.
72. Physician workforce supply in Belgium. Current situation and challenges. D/2008/10.273/09.
- 74 Hyperbaric Oxygen Therapy: a Rapid Assessment. D/2008/10.273/15.
76. Quality improvement in general practice in Belgium: status quo or quo vadis? D/2008/10.273/20
82. 64-Slice computed tomography imaging of coronary arteries in patients suspected for coronary artery disease. D/2008/10.273/42
83. International comparison of reimbursement principles and legal aspects of plastic surgery. D/2008/10.273/45
87. Consumption of physiotherapy and physical and rehabilitation medicine in Belgium. D/2008/10.273/56
90. Making general practice attractive: encouraging GP attraction and retention D/2008/10.273/66.
- 91 Hearing aids in Belgium: health technology assessment. D/2008/10.273/69.
92. Nosocomial Infections in Belgium, part I: national prevalence study. D/2008/10.273/72.
93. Detection of adverse events in administrative databases. D/2008/10.273/75.
95. Percutaneous heart valve implantation in congenital and degenerative valve disease. A rapid Health Technology Assessment. D/2008/10.273/81
100. Threshold values for cost-effectiveness in health care. D/2008/10.273/96
102. Nosocomial Infections in Belgium: Part II, Impact on Mortality and Costs. D/2009/10.273/03
- 103 Mental health care reforms: evaluation research of 'therapeutic projects' - first intermediate report. D/2009/10.273/06.
104. Robot-assisted surgery: health technology assessment. D/2009/10.273/09
108. Tiotropium in the Treatment of Chronic Obstructive Pulmonary Disease: Health Technology Assessment. D/2009/10.273/20
109. The value of EEG and evoked potentials in clinical practice. D/2009/10.273/23
111. Pharmaceutical and non-pharmaceutical interventions for Alzheimer's Disease, a rapid assessment. D/2009/10.273/29
112. Policies for Orphan Diseases and Orphan Drugs. D/2009/10.273/32.
113. The volume of surgical interventions and its impact on the outcome: feasibility study based on Belgian data
114. Endobronchial valves in the treatment of severe pulmonary emphysema. A rapid Health Technology Assessment. D/2009/10.273/39
115. Organisation of palliative care in Belgium. D/2009/10.273/42
116. Interspinous implants and pedicle screws for dynamic stabilization of lumbar spine: Rapid assessment. D/2009/10.273/46
117. Use of point-of care devices in patients with oral anticoagulation: a Health Technology Assessment. D/2009/10.273/49.

118. Advantages, disadvantages and feasibility of the introduction of 'Pay for Quality' programmes in Belgium. D/2009/10.273/52.

This list only includes those KCE reports for which a full English version is available. However, all KCE reports are available with a French or Dutch executive summary and often contain a scientific summary in English.

