

Pay for Performance in Primary Care: The Use of Administrative Data by Health Economists



Rita Santos, Sara Barsanti, and Chiara Seghieri

1 Introduction

Health economists have evaluated pay for performance (P4P) schemes to assess if they are efficient and effective and if they provide positive incentives regarding patient health and inequality reduction.

Researchers and policymakers are concerned with efficiency, effectiveness, value and behaviour in the production and consumption of health and health care. As will be demonstrated in this chapter, P4P schemes and performance measurements of health care systems more generally are usually evaluated on those criteria.

Furthermore, studies on P4P mainly use various administrative datasets that are linkable. The fact that it is possible to link different datasets is essential, since health care, well-being and economics are parts of a multifaceted society.

National statistics institutes produce publicly available socio-economic and demographic data for each country at different geographical scales. This is fundamental to understanding each country as a whole, but also the wide variation within a country and even within a city. It is essential to take these characteristics into account when analysing the performance of primary care providers, because those characteristics not only shape the health needs of their patients but are also a proxy for their responsiveness to continuity of care.

Since P4P is based on performance indicators, health authorities collect information on these for each primary care provider, which is publicly available in most

R. Santos (✉)

Centre for Health Economics, University of York, York, UK

e-mail: rita.santos@york.ac.uk

S. Barsanti · C. Seghieri

Laboratorio Management e Sanità, Institute of Management, Pisa, Italy

© The Author(s) 2019

N. Crato, P. Paruolo (eds.), *Data-Driven Policy Impact Evaluation*,

https://doi.org/10.1007/978-3-319-78461-8_19

countries. The information relates not only to the performance indicators but also to the population the provider serves and the main characteristics of the provider (e.g. number of doctors).

Once a country or a health region starts a P4P scheme in primary care, it is important to understand the benefits it brings to the health care system in the short and long term. This chapter will describe primary care P4P evaluation and incentives schemes in Italy and England.

The chapter describes the experience of the application of two management tools based on microdata to assist decision makers in monitoring and improving health care system performance, with a particular focus on primary care. In particular, the first case study describes the experience of P4P in general practice in England. The introduction of P4P in England is reported, along with its impact in improving chronic care conditions, measured through indicators of preventable emergency admissions (i.e. ambulatory care-sensitive conditions (ACSCs)), since these emergency admissions are likely to be reduced by improvement in the quality of primary care. The definition of ACSCs is internationally debated, and various definitions and measuring methods have been developed. However, it remains one of the common indicators derived from administrative data to compare primary care quality across providers.

The second case study refers to the development of a performance evaluation system (PES) for general practitioners in Tuscany Region (Italy). This section describes in detail how performance indicators from administrative and ad hoc data sources are measured, shared and made available to practitioners and policymakers to support performance improvement and alignment with the strategic goals of the health care system. Indeed, transparent, systematic benchmarking of performance and feedback mechanisms for health professionals are of strategic significance in designing effective P4P systems, based on reliability, recognition and trust of health care providers.

Considering these two case studies, the reader can be the judge of the importance of exploring and combining different administrative datasets used in the studies. Moreover, the England case study shows some results in terms of counterfactual impact assessment, considering the reduction of some sets of ACSC emergency admissions in England; the Italian case study focuses mainly on how micro-administrative data can support policymakers and public management decision-making, considering the performance reporting system for general practitioners and the first results in terms of improvements in the quality and appropriateness of care.

The plan for the remainder of the chapter is as follows. Section 2 discusses P4P schemes at international level, focusing on performance indicators, incentive schemes and data sources. Section 3 provides evidence on the English P4P scheme and the use of preventable emergency admissions to evaluate the impact of the scheme. Section 4 is dedicated to the Tuscany Region PES at primary care level, as an example of a reporting system for primary care using administrative data sources. Section 5 provides conclusions.

2 Pay for Performance in Primary Care

Within health care systems, countries use a combination of payment methods for doctors, including capitation formulae, salaries, fees for services and P4P schemes, to compensate for weakness associated with single payment approaches and suboptimal health care delivery (Eijkenaar et al. 2013).

Considering the principal–agent theory, when the interests of a principal (i.e. the policymakers) and its agents (i.e. family doctors, also known as general practitioners (GPs)) are not perfectly aligned, P4P may align agents' incentives with the principal's preferences. By making agents' reimbursement depend on their performance with respect to certain performance indicators, the hope is that agents respond by increasing effort on tasks valued by the principal (Prendergast 1999; Anell et al. 2015). Although P4P systems and PESs are used internationally, systematic review of P4P schemes report that the effects of the schemes remain largely uncertain (Houle et al. 2012).

In considering the P4P scheme, there is a need to emphasise that measuring primary care performance is a complex issue, since this level of care encompasses a myriad of activities and its functions differ considerably across countries and also in terms of provider payment and contractual schemes (Roland 2004). Furthermore, different primary care systems also have important effects on data collected and measured.

In P4P, care providers (i.e. family doctors or GPs) receive explicit financial incentives based on their scores on specific measures, which may pertain to certain performance expectations and standards in key domains such as clinical quality, resource use, access to services and patient-reported outcomes. In addition to the United States (USA), where P4P has become widespread, P4P programmes are being implemented in many other countries, including the United Kingdom, Canada, Australia, New Zealand, Taiwan, Israel, France, Italy and Germany, and in both primary care and inpatient care (Milstein and Schreyoegg 2016). Of all P4P physician's programmes in the US 10 years ago, more than 41% were related to primary care; the rest (59%) targeted both primary care physicians and specialists, or specialists only (Eijkenaar et al. 2013).

As reported by Sutherland et al. (2012), 'there is no accepted international definition of pay-for-performance', which may explain the degree of heterogeneity seen among P4P programmes. Moreover, P4P in primary care is context dependent and comparison can be difficult, and the way in which P4P schemes are designed and implemented can affect both the incentives provided and how physicians respond to them (Mehrotra et al. 2010; Eijkenaar et al. 2013).

2.1 Performance Domains and Indicators

Primary care P4P programmes usually select measures relating to conditions that are widespread (such as cardiovascular disease) and contribute significantly to the

overall burden of disease (such as coverage of vaccinations). In general, clinical indicators refer to compliance in following guidelines for common chronic care conditions, such as diabetes and cardiovascular disease. Considering efficiency, primary care providers may be incentivised for patients requiring below average levels of specialist services, inpatient hospital admissions, drugs prescription in terms of generic drugs consumption (e.g. statins and proton pump inhibitors (PPIs) dispensed) and pharmaceutical expenditure. France and New Zealand have specific targets for reducing pharmaceutical expenditure. Influenza vaccination rates for the elderly (or other target groups, such as children) and cancer screening participation are the main performance measures for preventative care. In the case of nonclinical indicators, performance measurement generally reflects the use of information and communication technology (ICT) (i.e. for registers, appointments and other facilities). Finally, in some countries, such as New Zealand, some indicators are measured separately for high-need sections of the population, and reduction in health inequality and disparities in access to services is also considered a general aim of the P4P scheme (Cashin et al. 2014).

2.2 Data Source

The role of data, information and reporting systems is crucial. Administrative health care data are collected by private health maintenance organisations or governmental institutions, for both managerial and epidemiological reasons (Gini et al. 2013). The use of administrative data to develop performance indicators related to primary care has been increasing over the years: case finding and algorithms are developed to estimate the prevalence of chronic care conditions (CCCs) and the quality of care for the same conditions. The content of databases varies from country to country: they may contain records collected at hospital discharge or during visits to GPs or specialists, or they may relate to drugs prescriptions or diagnostic procedures (Gini et al. 2013). In Canada, Sweden and the USA, administrative databases contain diagnosis codes from inpatient and outpatient care, enabling the estimation of CCC prevalence. In France, where only drugs prescriptions are available, the prevalence of diabetes, for example, is calculated using records for the prescription of anti-diabetic drugs. However, some authors (Green et al. 2012; Gini et al. 2013) are concerned that indicators estimated using administrative databases might not reflect the actual compliance with standard of care for chronic care patients.

In general, as P4P programmes evolve, data sources move from claim data to information directly derived from general practice. In the United Kingdom (UK), New Zealand and other countries, significant investments are being made in infrastructure for data collection (Eijkenaar 2012).

2.3 Incentive Payments

P4P programmes tend to have small rewards as a share of GPs' income; in terms of bonus payment, the percentages are generally 5% or less, with the exceptions of programmes in the UK (about 25%), Turkey (20%) and France (10%). Relatively low payments seem to be preferred because they are aligned with professionals' norms and values (Eijkenaar et al. 2013). As Cashin et al. (2014) emphasise, the incentive is more powerful if it increases as performance improves; almost all P4P in primary care uses higher payment rates for higher achievement levels. There are three main approaches to measuring achievement against performance (Cashin et al. 2014): (1) the absolute level of a measure towards a certain standard; (2) the change in terms of improvement of a measure; and (3) the relative ranking of a measure among the providers.

In general, participation in P4P is on a voluntary base (Eijkenaar 2012). For physicians participating in a P4P scheme, the rate varies from 99% in the UK, with the Quality and Outcomes Framework (QOF), to 80% in Poland.

3 The Case Study: Pay for Performance in Primary Care in England

The National Health Service (NHS) is a tax-financed system and free at point of demand (apart from a small charge for dispensed medicines). NHS primary care is provided by family doctors, known as GPs, who are organised in small surgeries known as general practices. All residents in England are entitled to register with a general practice, and have incentives to do so, as the practices provide primary care and act as the gatekeeper for elective (nonemergency) hospital care.

Most general practices are partnerships owned by GPs and they have, on average, five GPs (four full-time equivalents (FTEs)). They employ other medical staff, including nurses (an average head count (HC) of three, or two FTEs), direct patient care staff (average HC of two, or 1.3 FTEs) and administrative staff (average HC of 12, or eight FTEs) and have around 7500 patients (NHS Digital 2016). The NHS contracts with the practice rather than with the individual GPs. Practices are paid through a combination of lump sum payments, capitation, quality incentive payments and items of service payments. Quality incentives from the P4P scheme, the QOF (Roland 2004), generate a further 15% of practice revenue. Practices are reimbursed for the costs of their premises but have to fund all other expenses, such as hiring nurses and clerical staff, from their revenue.

3.1 The English Pay for Performance System: The Quality and Outcomes Framework

The NHS introduced the P4P contract for general practices—the QOF—in 2004/5.¹ This contract was intended to increase GPs' pay by up to 25%, depending on their performance with respect to 146 quality indicators relating to clinical care for ten chronic diseases, organisation of care and patient experience (Roland 2004).

Most of the organisation of care and patient experience indicators rely on a simple characteristic of the practice. For example, organisation of care Record 15 indicator in 2004/5 stated 'The practice has up-to-date clinical summaries in at least 60% of patient records', and patient experience indicator PE2 specifies 'The practice will have undertaken an approved patient survey each year'. While the clinical indicators can take the same form (e.g. Coronary Heart Disease indicator one states that 'The practice can produce a register of patients with coronary heart disease'), most of them are set to vary the attribution of points according to the proportion of patients for whom they achieve each target. On the latest indicators, points are awarded on a descending scale within the payment range. The payment range for each clinical indicator is between the minimum and maximum threshold. A practice whose achievement on a clinical indicator is below 25% does not earn a single point, while one that reaches the maximum threshold achievement (set between 55% and 90% in 2004/5) earns the maximum number of points.

Since some patients might not be eligible for specific indicators, practices are allowed to exclude them given certain circumstances. The NHS publishes the reasons for patient exclusion (in an 'exception report'), along with the P4P indicators.

The P4P indicators and disease groups have changed over the years. In 2004/5, there were four domains: clinical, organisational, patient experience and additional services. The clinical domain had 76 indicators in 11 areas (Coronary Heart Disease, Left Ventricular Dysfunction, Stroke and Transient Ischaemic Attack, Hypertension, Diabetes Mellitus, Chronic Obstructive Pulmonary Disease, Epilepsy, Hypothyroidism, Cancer, Mental Health and Asthma), while the organisational domain had 56 indicators in five areas (Records and Information, Patient Communication, Education and Training, Medicines Management and Clinical and Practice Management), the patient experience had four indicators in two areas (Patient Survey and Consultation Length), and the additional services domain had ten indicators in four areas (Cervical Screening, Child Health Surveillance, Maternity Services and Contraceptive Services). The organisational domain was retired in 2013/14. In 2012/13, the organisational domain still had 42 indicators and the average points per practice was 247.2 out of a maximum of 254 (HSCIC 2013). The patient experience domain was reduced to one indicator in 2013/14 (the average points per practice was 99.7 out of 100) and retired in 2014/15. The additional services domain was renamed

¹The UK fiscal year starts on 1 April and ends on 31 March of the following year.

‘public health—additional services’ in 2013/2014 and had nine indicators in the same four areas. However, from 2014/15, the public health—additional services domain includes only five indicators in two areas (Contraception (age below 55) and Cervical Screening (age 25–64)), which means that Child Health Surveillance and Maternity Services are no longer incentivised by the QOF. The clinical indicators also underwent changes in terms of new clinical areas. In 2015/16, the clinical domain has 65 indicators in 19 areas (Asthma, Atrial Fibrillation, Cancer, Coronary Heart Disease, Chronic Kidney Disease, Chronic Obstructive Pulmonary Disease, Dementia, Depression, Diabetes Mellitus, Epilepsy, Heart Failure, Hypertension, Learning Disabilities, Mental Health, Osteoporosis, Palliative Care, Peripheral Arterial Disease, Rheumatoid Arthritis, and Stroke and Transient Ischaemic Attack). A new public health domain was added to the QOF in 2013/14, with nine indicators in four areas (Blood Pressure (age above 40), Cardiovascular Disease—Primary Prevention, Obesity (age above 16) and Smoking (age above 15)).

3.2 The Impact of the Quality and Outcomes Framework on Preventable Emergency Admissions

The QOF was expected to raise clinical quality and, in particular, to prevent chronic condition events. However, Roland (2004) highlights that in addition to the expected benefits, some negative unintended consequences might also be expected, e.g. a focus on financial reward that is tied to specific tasks and a reduction in the quality of care for conditions not included in the incentive system.

The first studies on the impact of the QOF showed that English general practices had high levels of achievement in the first year of the scheme (Doran et al. 2006), but this did not imply that patients had better continuity of care (Campbell et al. 2009); on the contrary, once the targets were reached, the improvement in some conditions slowed.

Preventable emergency admissions, also known as ACSC emergency admissions, are defined as hospital emergency admissions that could be prevented or reduced through management of the acute episode in the community, or by preventive care (Purdy et al. 2009). Therefore, the ACSCs are a set of disease groups or, more precisely, of diagnosis codes using the tenth revision of the medical classification named International Statistical Classification of Diseases and Related Health Problems (*ICD-10*).

The records are usually retrieved from the patient’s hospital admissions administrative data, which include several ICD-10 codes (including the principal diagnosis for which the patient was admitted), treatment codes, admission and discharge data, the secondary care provider code, the diagnosis-related group, the age and gender of the patient, the geographical area of his/her address and, depending on the country, his/her primary care provider. ACSCs may reflect the suboptimal capacity of health services delivery to effectively prevent, diagnose, treat and/or manage

these conditions in primary care settings. ACSC rates are, therefore, inversely correlated with primary care performance, and studies have shown an inverse correlation between treatment guidelines adherence and ACSC inpatient rate (WHO 2016).

To understand the impact over time of the QOF on hospital emergency admissions, Harrison et al. (2014) analysed the time trends, between 2004 and 2011, of ACSC emergency admissions that were incentivised² by the QOF and those that were not.

Harrison et al. (2014) report a clear increase in non-incentivised preventable emergency admissions (and non-preventable emergency admissions) and a contrasting decrease in incentivised preventable emergency admissions. Analysis of the period extending from 5 years before the introduction of the QOF scheme to 7 years after its introduction revealed that the emergency admissions rates for all conditions increased by 34% between 1998/99 and 2010/11, while non-incentivised ACSC emergency admissions increased by 39% and non-ACSC emergency admissions rose by 41%. In contrast, incentivised ACSCs decreased by 10%. This decrease is even more important given that the rate of emergency admissions for incentivised ACSCs had been increasing by 1.7% per year before the introduction of the QOF scheme. The fact that the trends in incentivised ACSCs fell by 2.7–8% compared with non-incentivised ACSCs, and by 2.8–10.9% compared with non-ACSCs, between the first year of the QOF (2004/5) and (2010/11), shows that targeting chronic disease groups in primary care can reduce the emergency admission burden on resources and health care costs. The difference between the trends shows that the primary care P4P scheme had a significant impact on hospital emergency admissions.

To assess the impact of general practice characteristics on preventable emergency admissions, Gravelle et al. (2015) linked several administrative datasets between 2006/7 and 2011/12, including datasets on the QOF, ACSCs emergency admissions (from Hospital Episode Statistics), general practice workforce and list characteristics, patient satisfaction and catchment area characteristics. The authors show that the number of GPs and the different practice quality indicators have a significant negative effect on the number of ACSC admissions. The proportion of female GPs has an unexpectedly positive effect on admissions. The effect of GP average age is nonlinear, with the effect of average age declining and then reversing around 55 years of age. Neither the proportion of non-UK qualified GPs nor the proportion of salaried GPs has any significant effect. The patient-reported measures of ability to obtain urgent or advance appointments have negative coefficients, suggesting that this type of access reduces ACSC admissions. Being able to see a preferred GP also reduces admissions, hinting at the beneficial effects of continuity or good

²Harrison et al. (2014) included in the incentivised ACSCs group disease groups that had clearly been continuously incentivised under the QOF since the introduction of the scheme in 2004, and, as non-incentivised ACSCs, the remaining disease groups that were not targeted under the QOF at any time between 2004/5 and 2010/11.

interpersonal relations between patients and GPs. The more objective measure of clinical quality derived from the QOF also has a small negative effect on ACSCs. A practice that has a better QOF clinical quality achievement has significantly fewer ACSC emergency admissions, but the impact is small.

Given the opportunity to link administrative datasets on patient hospital admissions to general practice information, and general practice to census information, Kasteridis et al. (2016) and Goddard et al. (2016) examined the impact of the introduction of the 2006 QOF indicator for dementia on discharge destination and length of stay for patients admitted for dementia and on ACSC emergency admissions from 2006/7 to 2010/11, respectively. More precisely, Kasteridis et al. (2016) analysed if patients registered with GP practices that have better QOF indicator scores for the annual dementia review have a smaller likelihood of a care home placement following an acute hospital emergency admission. The major predisposing factors for institutionalisation in a care home were older age, female gender and the need factors of incontinence, fall, hip fracture, cerebrovascular disease, senility and total number of additional comorbidities. Over and above those factors, the dementia QOF review had no significant impact on the likelihood of care home placement for patients whose emergency admission primary diagnosis was dementia, but there was a small negative effect if the emergency admission was for an ACSC, with an odds ratio of 0.998. On the other hand, Goddard et al. (2016) found a significant and negative effect of the Dementia QOF indicator on length of stay among urgent admissions for dementia. Patients discharged to the community had significantly shorter hospital stays if they were cared for by practices that reviewed a higher percentage of their patients with dementia. However, this effect is not significant for patients discharged to care homes or who died in hospital. The authors also report that longer length of stay is associated with a range of comorbidities, markers of low availability of social care and intensive provision of informal care. Dusheiko et al. (2011) investigated another link between the QOF and a specific ACSC. The authors explored the association between general practices' quality of diabetic management, given by QOF indicators, and emergency admissions for short-term complications of diabetes between 2001/2 and 2006/07, i.e. before and after the introduction of the QOF. They reported that practices with better quality of diabetes care had fewer emergency admissions for short-term complications of diabetes. However, they did not find an association with hypoglycaemic admissions.

Some studies have also used cross-sectional analysis to assess the impact of the QOF in specific ACSCs. For example, Calderón-Larrañaga et al. (2011) analysed the association between the specific QOF Chronic Obstructive Pulmonary Disease (COPD) indicators and COPD hospital admissions in 2008/9 (the year before the influenza pandemic). The authors reported that smoking prevalence and deprivation were risk factors for admission, while the QOF indicator for patients with COPD who had received an influenza immunisation (an increase in the QOF indicator of 1% would be expected to decrease COPD admissions by a factor of 0.825), patient satisfaction with ability to book a GP appointment within 2 days and the number of GPs per 1000 patients in the practice were protective factors for COPD admissions.

The early impact of the QOF on angina and myocardial infarction 2006/7 hospital admissions was analysed by Purdy et al. (2011). While a higher overall clinical QOF score was associated with lower rates of admissions for angina and myocardial infarction, the four specific coronary heart disease indicators were not.

Overall, the NHS England primary care P4P scheme—the QOF—had a positive impact on the reduction of ACSC emergency admissions, overall and for specific conditions. The reductions were clear for the conditions incentivised by the QOF, especially during the period soon after the introduction of the scheme.

4 The Case Study: Pay for Performance in Primary Care in Italy

4.1 Health Care System and GPs in Italy

The Italian public health care system is inspired by the Beveridge model, and it is characterised by public taxation funding, free access at the point of delivery (with some copayments for specific services), and political control over providers.

In the Italian national health system, GPs are the first contact for most common health problems and act as gatekeepers for drug prescription and for access to secondary and hospital care. Their activities and responsibilities have three levels of governance (Barsanti et al. 2014): (1) the national level (through the ‘National Agreement’ between the central government and the national GPs’ trade unions (TUs)); (2) the regional level (through the ‘Regional Agreement’ between the regional government and the regional TUs); and (3) the local level (through the ‘Local Health Authority Agreement’, between the local health authority (LHA) managers and the local TUs). Primary care physicians are paid through a combination of methods, and regional and local health authorities have some degree of autonomy in defining additional payment. Each region may introduce economic incentives to complement the national current payment structure. These economic incentives can relate to performance, appropriateness of care or the adoption of patient referral.

The recent health planning legislation (Balduzzi Law No. 189/2012 and the *Patto per la Salute* (‘Agreement for Health’) 2014–2016) introduces strategies for the organisation of primary care according to operational forms that include a single unit of professional organisation, the *Aggregazioni Funzionali Territoriali* (‘Territorial Functional Aggregations’) (AFTs). AFTs represent the highest level of general practice organisation, and each serves a population of 30,000 patients, assisted by 25 GPs. Each AFT has a coordinator elected by the GPs. AFTs are expected to apply the philosophy of ‘clinical governance’ (Sccally 1998), whereby GPs have responsibility for continuously improving the quality of their services and safeguarding high standards of care.

Tuscany was the first Italian region to adhere to the new collaborative AFT model. The Tuscany Region has about 4 million inhabitants, and in 2015 its health care system comprised 12 LHAs (merged into three LHAs in 2016) and four teaching hospitals. In 2012, 115 AFTs were created through the Tuscany Regional Agreement and subsequent local agreements (Barsanti et al. 2016). On average, each AFT has 28,000 inhabitants with an average age of 52 years. In Tuscany, there are about 2700 GPs with on average 1100 patients each. The average age of GPs is about 60 years, with men in the majority.

4.2 The Making of the Tuscan Performance Evaluation System (PES) for Primary Care

In 2004, Tuscany Region commissioned the Scuola Superiore Sant'Anna of Pisa to design and implement a multidimensional PES to monitor the results of LHAs in terms of clinical quality and appropriateness, both for the hospital setting and for the district setting (Nuti et al. 2012, 2016). Starting from 2007, the performance indicators within the PES were presented in terms of a benchmark conducted across the health care providers and made available on a web platform for managers and professionals. The PES has now become the core of the P4P scheme of the CEOs of the LHAs. Starting in 2013, selected performance indicators within the PES were also calculated at AFT level to monitor and compare GPs' performance with respect to primary care activities and responsibilities, including (1) management of chronic disease; (2) prevention of avoidable hospital admission and inappropriate diagnostic tests; (3) preventive care and home care for the elderly; (4) drug prescriptions; (5) practice organisation; and (6) patient experience (Barsanti and Nuti 2016).

The Tuscan PES encompasses a large set of indicators grouped into about 25 indexes and classified in five dimensions (Table 1). Indicators are defined in regular meetings between the regional administration and representatives of general practices, including the perspectives of both managers and clinicians. The main source of data for clinical indicators are local health administrative data, which are centrally collected by the regional administration. Clinical indicators are measured using regional administrative data, which comprise electronic records of all inpatient and outpatient activity, as well of pharmaceutical consumption among all residents from 2009 to 2016. Patient experience and GP organisation measures at AFT level are collected through sample surveys (De Rosis and Barsanti 2016). Research and evaluators utilise various sources of microdata through data record linkage, including:

1. Linkage between patients, their usual GP and the GP's AFT.
2. Data linkage at individual (patient) level across different administrative databases (i.e. inpatient, outpatient and drug consumption data), to measure performance indicators along the different care pathways for chronic patients (e.g. process indicators for the care of diabetic patients).

Table 1 Domains, indicators and data sources of the PES for GPs in Tuscany

Domains	Example of indicators	Data source
<i>Regional primary care strategy compliance:</i> to guarantee that strategic regional goals are pursued in the appropriate time and manner, focusing on primary care setting	<ul style="list-style-type: none"> – Influenza vaccination rate for older people – Percentage of older people with home care assistance 	Ad hoc data sourced from GPs and LHAs, and administrative health data flow
<i>Management of chronic care patients and preventable inpatients:</i> focusing on quality, appropriateness and continuity of care for patients with chronic care disease	<ul style="list-style-type: none"> – Ambulatory care-sensitive condition inpatient rate (standardised by sex and age) – Access to emergency department for minor diseases (standardised by sex and age) – Percentage of patients with heart failure receiving angiotensin-converting enzyme (ACE) inhibitor therapy – Percentage of patients with heart failure receiving beta blocker therapy – Percentage of patients with diabetes receiving haemoglobin testing – Percentage of patients with diabetes receiving eye checks by an oculist 	Administrative health data flow
<i>Patient satisfaction:</i> patient experience and level of satisfaction with GPs	<ul style="list-style-type: none"> – Percentage of patients satisfied with their GP in terms of communication, general assistance, time for visiting, collaboration 	Statistical sample survey
<i>GPs satisfaction:</i> results of surveys on the satisfaction level and experience of GPs with their general practice organisation	<ul style="list-style-type: none"> – Meeting to share clinical guidelines – Frequency of clinical audit 	Survey of all GPs
<i>Diagnostic and pharmaceutical care:</i> focusing on appropriateness and compliance of care in drugs prescription and diagnostic visits	<ul style="list-style-type: none"> – Nuclear Magnetic Resonance (NMR) rate for older people – Consumption of specific drugs (statins, PPI) – Drug prescriptions for generic drugs 	Administrative health data flow

Note: *GP* General Practitioner, *LHA* Local Health Authority, *ACE* Angiotensin-Converting Enzyme, *NMR* Nuclear Magnetic Resonance, *PPI* Proton-pump inhibitors

Indicators were selected on the basis of the following criteria (Vainieri et al. 2016): (1) scientific soundness of the indicators in terms of their validity and reliability; (2) feasibility of obtaining regionally comparable data for all 115 AFTs; (3) ability to provide a comprehensive overview of primary care.

Each indicator considers regional, national or literature-based standards in order to assess performance. Where standards are absent or difficult to set, the regional variation of the practice is taken into account, benchmarking performances with

the quintile distribution. Moreover, each performance indicator is measured at a different level of governance: regional level, LHA level, AFT level and individual GP level.

PES indicators that are considered as evaluation measures are assigned performance assessment ratings for benchmarking reporting across AFTs. For each evaluation measure, five performance levels are derived for defining the performance of each AFT, from worst to best. These five evaluation tiers are associated with different colours, from dark green (excellent performance) to red (poor performance).

The PES considers three main perspectives of performance evaluation for each indicator:

- Performance assessment with respect to a standard derived either from the literature or from agreements at local and/or regional level. Data are displayed as histograms reporting the performance value of all 115 AFTs.
- Variability among AFTs and LHAs (Fig. 1) represented by a box plot with value of AFTs grouped by LHAs and by a boxplot with individual GPs grouped by AFT (Fig. 2).

Considering all the performance indicators, the summarising reporting system is visually represented by a 'target' diagram, which is divided into five coloured evaluation bands. Every year each AFT receives its own target and the more the AFT is able to reach the yearly defined objectives, the closer the performance indicators are to the centre (the dark-green area). Scarce and critical performance results are represented by indicators, which are positioned far from the centre, in the red area (Fig. 3).

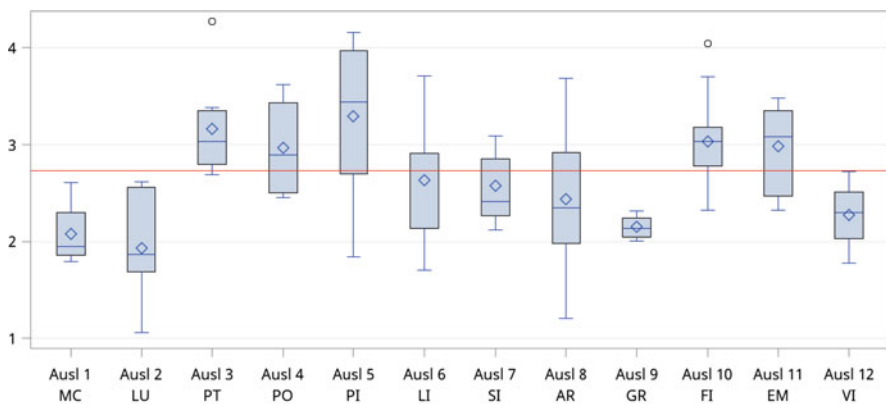


Fig. 1 Box plot of ACSC hospital admission between local health authorities in 2015. Source: Barsanti and Nuti (2016)

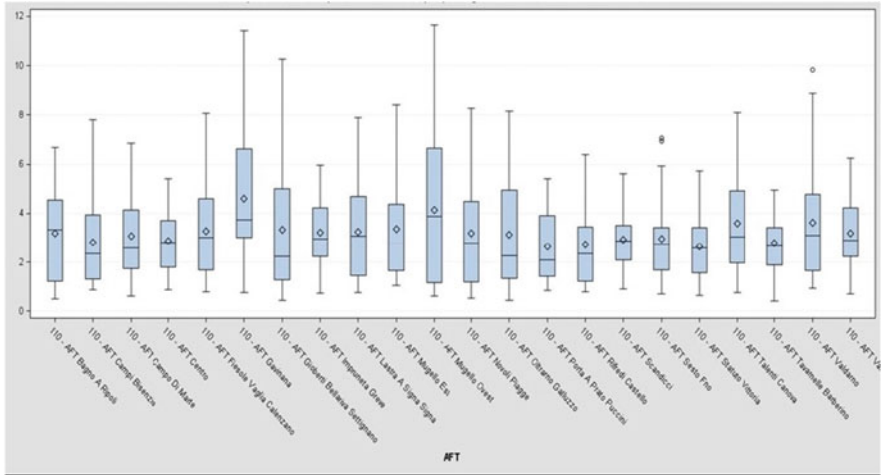


Fig. 2 Box plot of ACSC hospital admissions between AFTs of local health authorities (LHAs) in Florence in 2015 (in 2015 there were 22 AFTs grouped in the LHA of Florence). Source: Barsanti and Nuti (2016)

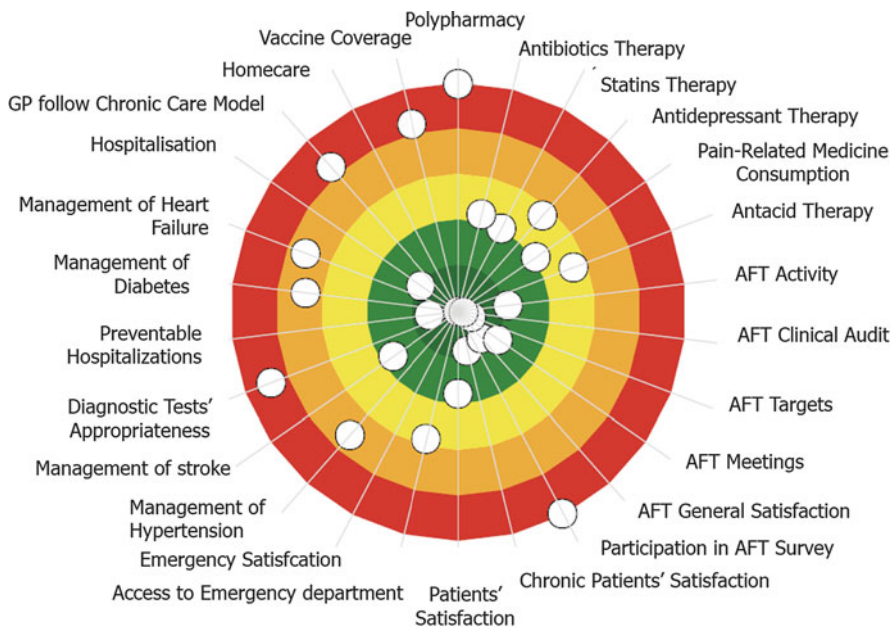


Fig. 3 Example of the system of reporting performance indicators AT AFT LEVEL in 2014'. Source: Barsanti and Nuti (2016)

4.3 The PES and the P4P Programme in Tuscany Region: First Results

A recent paper (Nutti et al. 2016) examines the various governance models in the health care sector that the Italian regions have adopted, and investigates the PESs associated with them, focusing on the experience of a network of ten regional governments that use the same health care PES as Tuscany Region. Considering 14 indicators measured in 2007 and in 2012 for all the regions, the study shows how different performance evaluation models are associated with differences in health care performance and whether or not the use of a PES has made any difference to the results achieved by the regions involved. In particular, Tuscany Region registered a high performance in 2007 and was still offering good general assistance in 2012, even improving both hospital and primary care processes (Nutti et al. 2016). In this sense, the authors conclude: systematic benchmarking and public disclosure of data based on the Tuscany PES are powerful tools that can guarantee sustained improvement of health care systems, if they are integrated with the regional governance mechanisms.

With regard to primary care, the P4P scheme for general practice in Tuscany Region has recently been developed following the P4P scheme for the chief executive officers of LHAs (Nutti et al. 2012). At this stage in the reform of primary care (see the constitution of the AFT), policymakers and professionals have some degree of autonomy in the design of the P4P programme, in terms of selection of performance indicators and incentives (Barsanti et al. 2014). Therefore, different groups of GPs might be incentivised on different sets of PES indicators. An analysis of the 12 Local Health Agreements for Tuscany Region in 2015 shows that almost 50% of the PES indicators were used also in the local P4P programmes for primary care. In particular, all indicators measuring pharmaceutical care, preventable hospital admission (ACSC inpatient rate) and management of chronic care disease are used to incentivise GPs. Each AFT has its own standard to achieve, based on the performance of the previous year. Usually, incentives are set both at AFT and at individual GP level.

Although the use of PESs in primary care and the linkage to GP incentives have only recently been introduced, and although the data do not allow for measuring any causal effect of P4P on performance, preliminary results show a positive impact of P4P on the quality and appropriateness of primary care in Tuscany (Barsanti et al. 2014), as measured by significant improvement over the years of performance of selected indicators included in the P4P system. About 50% of indicators improved from 2014 to 2015 at regional level (Barsanti and Nutti 2016).

In this sense, three selected primary care performance indicators used in the Tuscany GP PES are compared, in order to assess improvements over the years 2013–2015, considering:

1. The percentage of elderly people who received home care.
2. The rate of hospital admissions for ACSC per 1000 inhabitants (standardised by age and sex).

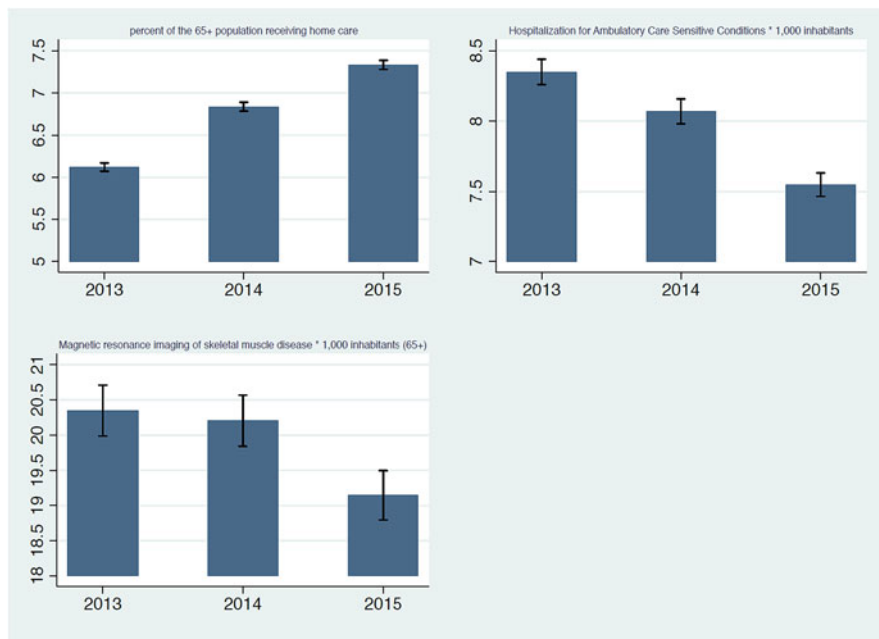


Fig. 4 Percentage receiving home care, ACSC inpatient rate standardised by age and sex and magnetic resonance imaging of skeletal muscle among the population aged over 65 years: trends 2013–2015 and confidence intervals

3. The rate of magnetic resonance imaging for musculoskeletal disease per 1000 elderly people (standardised by age and sex).

All the selected indicators show significant improvements, considering trend and confidence intervals (Fig. 4).

In Tuscany, ACSC inpatient rates decreased by 10% between 2013 and 2015 (from 8.35 per 1000 residents in 2013 to 7.54 per 1000 residents in 2015), whereas, in contrast, the number of elderly people (aged 65+) receiving home care increased by 20% (from 6.12% in 2013 to 7.33% in 2015). With regard to indicators of appropriateness, although the rate of magnetic resonance imaging for musculoskeletal disease among elderly people (those aged 65+) did not change appreciably from 2013 to 2014, it decreased by 6% from 2013 to 2015 (from 19.14 per 1000 elderly people in 2013 to 20.35 per 1000 in 2015) (Fig. 4).

Finally, considering the use of the PES, GPs use the PES results during clinical audits to formulate evidence-based improvement strategies for the care of their patients.

The Tuscan PES for primary care is playing an increasing role, providing valuable information to GPs and local decision-makers to support quality improvements, to define priorities and to set appropriate targets.

5 Conclusion

As we enter the ‘big data’ era, it is important to rethink the usage of administrative data to allow for timely evidence-informed clinical and policy decision-making. Although P4P has high face validity, the evidence for the effectiveness of such schemes in improving quality remains mixed (Rosenthal and Frank 2006; Mehrotra et al. 2010; Eijkenaar et al. 2013). A recent literature review on the effects of P4P programmes found 58 studies related to primary care P4P scheme and results (Mendelson et al. 2017). The authors found low-strength evidence that P4P programmes may improve process-of-care outcomes over the short term (2–3 years). Many of the studies reporting positive findings were conducted in the United Kingdom, where incentives are much larger than any P4P programmes in the United States. The largest improvements were seen in areas where baseline performance was poor.

This chapter describes two experiences of using performance measurement and P4P, whose efficacy and impact can be analysed only by linking data from regional population and administrative databases (i.e. from the P4P to the population census). The P4P schemes had a positive impact on the quality of primary care in both countries. However, more research is needed to understand the expected value of an indicator in terms of its impact on quality and its lifespan.

In England, the primary care P4P scheme—the QOF—had a positive impact on quality. However, practices achieved, on average, 958.7 points, representing 91.3% of the total 1050 points available, in the first year of the scheme. After changes to the P4P domains, clinical areas and indicators over the following 11 years, practices still had a high achievement, with an average of 532.9 points out of 559 in the clinical domain in 2015/16. In addition to the achievement of the P4P indicators, there was also a reduction in ACSC emergency admissions, especially for incentivised disease groups. Whether or not the long-term effects will compensate for the costs of P4P is a topic for further research.

The PES of Tuscany Region is presented as an example of how the combination of performance measurement and explicit incentives can be effectively used to promote accountability and to improve the quality of care in a regional health system. Moreover, the mix of systematic benchmarking between primary care providers over the years and the identification of best practices offer an overall strategic framework of evidence-based information to be used by professionals in everyday practice, and by decision-makers to define priorities and set performance targets.

It is essential to understand the contribution of the measurement of performance and use of incentives in primary care to the improvement of health care quality, to the reduction of unwarranted variation and to the ultimate decrease of secondary care burden. General practitioners and family practices are on the front line of prevention and treatment, and are key in any national health system to achieve its goal: a healthy population.

Acknowledgements Rita Santos was funded by a National Institute for Health Research Doctoral Fellowship (DRF 2014-07-055). The PES of the Tuscan health care system was founded by Tuscany Region through an agreement with Scuola Superiore Sant'Anna di Pisa. The views expressed are those of the authors and not necessarily those of the funders.

References

- Anell AD, Jens D, Ellegård LM (2015) Can pay-for-performance to primary care providers stimulate appropriate use of antibiotics? School of Economics and Management Working Paper 2015:36, Lund University, Lund
- Barsanti S, Nuti S (2016) Il sistema di valutazione della performance delle AFT toscane. Edizioni del Gallo, Perugia
- Barsanti S, Bonciani M, Pirisi L et al (2014) Trade union or trait d'union? Setting targets for general practitioners: a regional case study. Paper presented at the International Research Society for Public Management Conference 2015, University of Birmingham, 30 March–1 April 2015. <http://irspm2014.com/index.php/irspm/IRSPM2015/paper/viewFile/1460/585>
- Barsanti S, Bonciani M, Vola F (2016) Innovatori, indecisi, bisognosi o autonomi: I medici di medicina generale tra integrazione e accountability 2016. *Mecosan* 98:9–39. <https://doi.org/10.3280/MESA2016-098002>
- Calderón-Larrañaga A, Carney L, Soljak M et al (2011) Association of population and primary healthcare factors with hospital admission rates for chronic obstructive pulmonary disease in England: national cross-sectional study. *Thorax* 66(3):191–196
- Campbell SM, Reeves D, Kontopantelis E (2009) Effects of pay for performance on the quality of primary care in England. *N Eng J Med* 361(4):368–378. <https://doi.org/10.1056/NEJMsa0807651>
- Cashin C, Chi Y-L, Smith PC et al (2014) Paying for performance in health care: implications for health system performance and accountability. Open University Press and McGraw-Hill Education, Maidenhead
- De Rosis S, Barsanti S (2016) Patient satisfaction, e-health and the evolution of the patient–general practitioner relationship: evidence from an Italian survey. *Health Policy* 120(11):1279–1292
- Doran T, Fullwood C, Gravelle H et al (2006) Pay-for-performance programs in family practices in the United Kingdom. *N Engl J Med* 355(4):375–384. <https://doi.org/10.1056/NEJMsa055505>
- Dusheiko M, Doran T, Gravelle H et al (2011) Does higher quality of diabetes management in family practice reduce unplanned hospital admissions? *Health Serv Res* 46(1p1):27–46. <https://doi.org/10.1111/j.1475-6773.2010.01184.x>
- Eijkenaar F (2012) Pay for performance in health care: an international overview of initiatives. *Med Care Res Rev* 69(3):251–276
- Eijkenaar F, Emmert M, Scheppach M et al (2013) Effects of pay for performance in health care: a systematic review of systematic reviews. *Health Policy* 110(2–3):115–130. <https://doi.org/10.1016/j.healthpol.2013.01.008>
- Gini R, Francesconi P, Mazzaglia G et al (2013) Chronic disease prevalence from Italian administrative databases in the VALORE project: a validation through comparison of population estimates with general practice databases and national survey. *BMC Public Health* 13:15. <https://doi.org/10.1186/1471-2458-13-15>
- Goddard MK, Kasteridis P, Jacobs R (2016) Bridging the gap: the impact of quality of primary care on duration of hospital stay for people with dementia. *J Integr Care* 24(1):15–25
- Gravelle HSE, Martin S, Santos R (2015) GPs, practice quality and emergency hospital admissions for ambulatory care sensitive conditions. Paper presented at the Health Economists' Study Group Meeting, Lancaster University, 7–9 January 2015

- Green ME, Hogg W, Savage C et al (2012) Assessing methods for measurement of clinical outcomes and quality of care in primary care practices. *BMC Health Serv Res* 12:214. <https://doi.org/10.1186/1472-6963-12-214>
- Harrison MJ, Dusheiko M, Sutton M et al (2014) Effect of a national primary care pay for performance scheme on emergency hospital admissions for ambulatory care sensitive conditions: controlled longitudinal study. *BMJ* 349:g6423. <https://doi.org/10.1136/bmj.g6423>
- Houle SD, McAlister FA, Jackevicius CA et al (2012) Does performance-based remuneration for individual health care practitioners affect patient care? A systematic review. *Ann Intern Med* 157(12):889–899. <https://doi.org/10.7326/0003-4819-157-12-201212180-00009>
- HSCIC (2013) Quality and outcomes framework achievement, prevalence and exceptions data, 2012/13. Health and Social Care Information Centre, Leeds. <http://content.digital.nhs.uk/catalogue/PUB12262/qual-outc-fram-12-13-rep.pdf>
- Kasteridis P, Mason A, Goddard M (2016) Risk of care home placement following acute hospital admission: effects of a pay-for-performance scheme for dementia. *PLoS One* 11(5):e0155850. <https://doi.org/10.1371/journal.pone.0155850>
- Mehrotra A, Sorbero ME, Damberg CL (2010) Using the lessons of behavioral economics to design more effective pay-for-performance programs. *Am J Manag Care* 16(7):497–503
- Mendelson A, Kondo K, Damberg C et al (2017) The effects of pay-for-performance programs on health, health care use, and processes of care: a systematic review. *Ann Intern Med* 166(5):341–353. <https://doi.org/10.7326/M16-1881>
- Milstein R, Schreyoegg J (2016) Pay for performance in the inpatient sector: a review of 34 P4P programs in 14 OECD countries. *Health Policy* 120(10):1125–1140. <https://doi.org/10.1016/j.healthpol.2016.08.009>
- NHS Digital (2016) General and personal medical services, England, September 2015–March 2016, Provisional experimental statistics
- Nuti S, Seghieri C, Vainieri M (2012) Assessment and improvement of the Italian healthcare system: first evidence from a pilot national performance evaluation system. *J Healthc Manag* 57(3):182–199
- Nuti S, Vola F, Bonini A et al (2016) Making governance work in the health care sector: evidence from a ‘natural experiment’ in Italy. *Health Econ Policy Law* 11(1):17–38. <https://doi.org/10.1017/S1744133115000067>
- Prendergast C (1999) The provision of incentives in firms. *J Econ Lit* 37(1):7–63
- Purdy S, Griffin T, Salisbury C et al (2009) Ambulatory care sensitive conditions: terminology and disease coding need to be more specific to aid policy makers and clinicians. *Public Health* 123(2):169–173. <https://doi.org/10.1016/j.puhe.2008.11.001>
- Purdy S, Griffin T, Salisbury C et al (2011) Emergency admissions for coronary heart disease: a cross-sectional study of general practice, population and hospital factors in England. *Public Health* 125(1):46–54. <https://doi.org/10.1016/j.puhe.2010.07.006>
- Roland M (2004) Linking physicians’ pay to the quality of care: a major experiment in the United Kingdom. *N Engl J Med* 351(14):1448–1454. <https://doi.org/10.1056/NEJMp041294>
- Rosenthal MB, Frank RG (2006) What is the empirical basis for paying for quality in health care? *Med Care Res Rev* 63(2):135–157
- Scally G (1998) Clinical governance and the drive for quality improvement in the new NHS in England. *BMJ* 317:61
- Sutherland JM, Repin N, Crump RT (2012) Reviewing the potential roles of financial incentives for funding healthcare in Canada. Canadian Foundation for Healthcare Improvement, Ottawa. <http://www.cfhi-fcass.ca/Libraries/Reports/Reviewing-Financial-Incentives-Sutherland-E.sflb.ashx>. Accessed 9 Feb 2015
- Vainieri M, Vola F, Soriano Gomez G (2016) How to set challenging goals and conduct fair evaluation in regional public health systems: insights from Valencia and Tuscany regions 2016. *Health Policy* 120(11):1270–1278

WHO (2016) Assessing health services delivery performance with hospitalization for ambulatory care sensitive conditions. World Health Organization Regional Office for Europe, Copenhagen. http://www.euro.who.int/__data/assets/pdf_file/0010/305875/Assessing-HSD-performance-with-ACSH.pdf?ua=1

Rita Santos is a research fellow at Centre for Health Economics of the University of York. She holds an NIHR doctoral research fellowship on “Measuring and explaining primary care quality variation”. In the past years, she worked on GP, gender pay, differences, GP practice choice, hospital competition and GP practices competition. Her main interests are developing new applications of geographical information system for health economics and applying spatial econometric methods to economic theory.

Sara Barsanti is assistant professor of Management at the Laboratorio Management e Sanità of Scuola Superiore Sant’Anna of Pisa (Italy). She has a PhD in Health Care Management. In the past years, she has worked on a variety of health research projects and collaborated with several international institutions. Her research interests focus on organizational model of primary care and performance evaluation system in health care.

Chiara Seghieri is assistant professor of Social Statistics at the Laboratorio Management e Sanità of Scuola Superiore Sant’Anna of Pisa (Italy). In the past years, she has worked on a variety of health research projects and collaborated with several international institution and academic/research centers. Her current research interests include the development and use of statistical methods for measuring and comparing care quality and equity using observational data from both sample surveys and administrative records.

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

