

Revising Estonia's Quality Bonus Scheme in Primary Care

A report by the World Bank to the Estonian Health Insurance Fund

October 2018

Contents

Abbreviations.....	4
Acknowledgements.....	5
Executive summary and policy recommendations	6
1. Introduction	8
2. Health care needs and primary care in Estonia	9
The most common causes of death and disability in Estonia are not well reflected in QBS..	9
Socioeconomic inequalities in health are pronounced in Estonia.....	9
Rates of hospital admission for some chronic conditions are high in Estonia	10
There is scope to improve both patient-centredness and care coordination in Estonia.....	13
EHIF reimburses primary care through a complex array of mechanisms	14
QBS is blended with a complementary scheme for quality management.....	14
3. QBS compared to other national P4P schemes in primary care: what is incentivized?.....	16
QBS focuses on chronic disease, including complex bundles of care.....	16
QBS addresses few dimensions of primary care performance	21
Limited revision of QBS indicators has been undertaken since creation of the scheme	24
QBS indicators are constructed from routine administrative data	25
4. QBS compared to other national P4P schemes in primary care: who is incentivized?.....	27
QBS reimbursement is directed to the doctor, rather than the wider primary care team...	27
Group-level incentives may be more effective than incentives targeting individuals	27
5. QBS compared to other national P4P schemes in primary care: how is performance incentivized?	30
QBS is a small fraction of primary care’s overall income	30
The financial incentive in other schemes is larger than that offered by QBS.....	30

QBS reimbursement is annual and doctors receive QBS reports with a significant delay	31
Performance feedback in other schemes is quicker and more frequent than in QBS.....	32
QBS results are publicly reported by name for each Family Doctor	32
QBS generally rewards improvement on past performance, rather attainment of a specified threshold	33
Most other P4P schemes reward attainment of a threshold, rather than improvement against baseline	33
There is little opportunity for doctors to exempt patients from certain indicators	35
6. Family Doctors' current performance in QBS	36
QBS has been adopted, in principle, by all Family Doctors.....	36
One in three Family Doctors perform poorly on QBS	36
Preventive care (domain 1): good performance with little regional variation	40
Chronic disease management (Domain 2): moderately good performance with substantial regional variation	42
Additional professional competence (Domain 3): uniformly poor performance	48
Size of the patient population may influence QBS performance	48
7. Recommendations for modernizing QBS	50
What is incentivized.....	50
Who is incentivized.....	52
How is performance incentivized.....	53
8. Recommendations for further analysis and research.....	55
References.....	56
Annex 1: Overview of QBS indicators, including changes	58
Annex 2: Detailed description of QBS indicators in domain 2	62

Abbreviations

AMI	acute myocardial infarction
CPD	continuous professional development
DALY	disability-adjusted life year
ECM	enhanced care management
ED	emergency department
EHIF	Estonian Health Insurance Fund (<i>Haigekassa</i>)
ESFD	Estonian Society of Family Doctors
FFS	fee for service
HNDP	Hospital Network Development Plan
OECD	Organisation for Economic Cooperation and Development
OOH	out-of-office hours
P4P	pay for performance
QBS	Quality Bonus Scheme
QOF	Quality and Outcomes Framework (a primary care P4P scheme in the United Kingdom)
ROSP	Rémunération sur Objectifs de Santé Publique (a primary care P4P scheme in France)
TSH	thyroid stimulating hormone

Acknowledgements

The authors of this report are Ian Forde, Marvin Ploetz, Adrien Dozol and Iryna Postolovska. The support of Christoph Kurowski, Enis Baris and other managers at the World Bank Group is gratefully acknowledged.

Dionne Sofia Kringos - Pereira Martins (Amsterdam University Medical Centre - location AMC, and Head of the WHO Collaborating Centre for Quality and Equity of Primary Healthcare Systems) contributed substantial portions of the text. We also thank the patients, clinicians and managers of Dr Elle-Mall Keevallik's and Dr Pille Heinamets's practice for welcoming us on working visits to see Estonian primary care in action and to discuss the challenges they face, and the Estonian Society of Family Doctors for feedback on the proposals presented in this report.

Most importantly, we thank the board and staff of the Estonian Health Insurance Fund, and the Steering Committee overseeing the World Bank Group's engagement with EHIF. Without their guidance and support, particularly during complex working visits to Estonia in March, June and September 2018, this report would not have been possible.

Executive summary and policy recommendations

Effective management of chronic conditions is Estonia's most pressing health system challenge, alongside tackling socioeconomic and geographic inequities in health. Both challenges demand high-performing primary care. Although many measures of primary care performance in Estonia are good, there are signs of weakness. In particular, patient-centredness and care coordination could be improved to prevent unnecessary admission to hospital or attendance at hospital Emergency Departments. In order to address these challenges, the Estonian Health Insurance Fund and Estonian Society of Family Doctors launched a pay-for-performance scheme in 2006, the Quality Bonus Scheme. Specific aims of the QBS were to incentivize preventive care and management of chronic conditions. Nearly all Family Doctors in Estonia now participate in QBS. One in three, however, obtain low scores. Doctors generally perform well in the preventive care domain, but performance in the domain linked to care for chronic conditions is highly variable - with poor performance typically concentrated in Estonia's southern counties. Performance in domain of additional professional competencies is uniformly low.

Estonia's QBS differs in important respects from primary care P4P schemes in other countries. It has a more complex design than many schemes, rewarding relative improvement rather than absolute achievement, and is linked to a separate bonus scheme for Quality Management. QBS pays less (as a fraction of overall income) than several other schemes and feedback of results takes several months, slower than that seen in other countries. It also offers less opportunities to exempt patients from inclusion in the scheme, which may penalize Family Doctors who care for patients with unusually complex needs. The scheme also relies on claims data submitted by Family Doctors to EHIF for reimbursement. While this minimises administrative burden, it means that QBS is limited to measuring processes and activities. In contrast to primary care P4P schemes in other countries, key dimensions performance such as clinical outcomes or patient experience are not currently captured by QBS. Furthermore, most QBS incentives are directed to the individual Family Doctor. Primary care is increasingly team-based, however, meaning that there may be scope to consider wider use of group-level incentives within QBS.

Ten recommendations are made for modernizing QBS in light of Estonia's priority health care needs; international experience and best practice in using pay-for-performance in primary care; and, concurrent reforms to Estonia's primary care financing and service delivery model. Taking each of them forward will require close cooperation between EHIF and ESFD:

- 1. On-going revision of QBS indicators is recommended, dropping those which have high achievement or low disease burden*
- 2. The dimensions of primary care performance captured by QBS should be expanded*
- 3. Local elements should be developed*
- 4. QBS should reward both improvement and absolute level target achievement*
- 5. Consider wider application of practice-level incentives, to reflect team-base care*
- 6. Use QBS to encourage group practice*
- 7. Expand the criteria by which patients can be exempted from inclusion in QBS*
- 8. Strengthen the incentives within QBS, both financial and non-financial*
- 9. Shorten the feedback loop by which Family Doctors receive results*
- 10. Finally, ensure that QBS is part of an overall strategy to strengthen primary care quality.*

Of the ten, recommendations 1, 2, 3, 7 and 8 should receive the highest priority. The final recommendation, however, is perhaps the most important – namely, to ensure that QBS articulates effectively with other activities to monitor and improve primary care quality and is part of an overall strategy to strengthen the sector. Although EHIF and ESFD have developed several ambitious initiatives to improve the quality of primary care over many years, there is a risk that different activities do not synergise as effectively as they could. In particular, EHIF and ESFD should discuss how the new programme of Enhanced Care Management can best build upon and synergise with QBS to take Estonia’s primary care onto the next level of performance.

1. Introduction

For many years, the Estonian Health Insurance Fund (EHIF), in partnership with the Estonian Society of Family Doctors (ESFD), has developed an ambitious range of initiatives to strengthen primary care. Monitoring and improving the quality of care has been a central to this and in 2006 the Quality Bonus Scheme (QBS), one of the first national pay-for-performance (P4P) schemes in the world, was launched.

With more than a decade's experience of implementation, it is timely to review the design and operation of QBS. Several other OECD countries have developed P4P schemes in primary and secondary care (Table 1), meaning that substantial international experience is also available to inform a review of QBS. A detailed review of QBS was also recommended by the World Bank's recent review of EHIF's reimbursement systems, *Toward greater integration of care and improved efficiency: a critical review of EHIF's payment systems*, which recommended refining primary care payment methods to improve the scope and quality of primary care (particularly for the management of patients with chronic conditions) and strengthening the sector's coordination with other settings of care.

Table 1: Pay for performance schemes have emerged in many OECD, particularly for primary care

Country	Primary care	Specialist care	Hospitals	Country	Primary care	Specialist care	Hospitals
Australia	X		X	Korea, Rep. of	X	X	X
Austria				Luxembourg	X		
Belgium	X			Mexico	X		
Canada				Netherlands	X	X	X
Chile	X	X		New Zealand	X		
Czechia	X			Norway			
Denmark				Poland	X		
Estonia	X			Portugal	X		X
Finland				Slovakia			
France	X	X	X	Slovenia			
Germany	X			Spain	X	X	X
Greece				Sweden	X		X
Iceland				Switzerland			
Ireland				Turkey	X		X
Israel				U'td Kingdom	X	X	X
Italy				United States	X	X	X
Japan				United States	X	X	X

Source: *Better Ways to Pay for Health Care*, OECD Publishing, Paris 2016

Following this Introduction, Section 2 of the report considers the health and health care context within which QBS operates. Sections 3 - 5 considers the design of QBS, comparing it with primary care P4P schemes in other OECD countries. The framework developed by Eijkenaar et al. (2013), which considers *what* is incentivized, *who* is incentivized and *how* performance is incentivized, is adapted as a basis for sections 3, 4 and 5 respectively. Moving on from design, Section 6 describes trends, latest data and regional variation in Family Doctors' current performance within QBS. Section 7 offers a set of recommendations for how QBS could be improved, deploying the *what, how, who* framework. The report closes by identifying areas for further analysis and research.

2. Health care needs and primary care in Estonia

This Section 2 considers the health care needs and health care system within which QBS operates. Effective management of chronic conditions is Estonia's most pressing health system challenge, alongside tackling socioeconomic and geographic inequities in health. Both challenges demand high-performing primary care. Although many measures of primary care performance in Estonia are good, there are signs of weakness. In particular, patient-centredness and care coordination could be improved to prevent unnecessary admission to hospital or attendance at hospital Emergency Departments.

The most common causes of death and disability in Estonia are not well reflected in QBS

The leading causes of death in Estonia are cardiovascular diseases, including ischaemic heart disease, stroke and other heart diseases. These account for just over a third of deaths. Cancers (principally lung, colorectal, stomach, prostate and breast) are the next largest cause of mortality, accounting for around a sixth of deaths. In terms of morbidity (measured as disability-adjusted life years), ischaemic heart disease and stroke again top the list, together accounting for around a quarter of all DALYs. Neurological and sensory organ conditions (such as difficulty in seeing or hearing) account for the next largest group of DALYs, followed by lung cancer; falls; back and neck pain; alcohol use disorders; and, depression.

The prevalence of chronic conditions is due to ageing of the population, but it is likely also related to behavioural risk factors (1). It has been estimated that 37% of the overall burden of disease (in terms of Disability Adjusted Life Years) can be attributed to behavioural risks such as alcohol consumption, smoking, diet and low physical activity. Despite some improvements, men still smoke and drink heavily (in both cases more than the EU average), and a rapid increase has been observed in obesity rates. One in four Estonian children start primary care school already overweight or obese. Such risk factors are more prevalent among people with low levels of education or income, contributing to differences in health status between socioeconomic groups (2).

It is worth noting that QBS only covers few of these leading causes of mortality and morbidity. Of the leading causes of death, only hypertension, diabetes and colon cancer are part of QBS. Likewise, several leading causes of DALY burden, such as mental health and substance abuse related conditions, or age-related conditions (such as falls), are not included in QBS. In terms of health promotion and risk factors, QBS includes nurse counselling for people with some chronic conditions (such as diabetes), but it is not known how effective these activities are. Conversely, some conditions with a very low disease burden (such as hypothyroidism or vaccine-preventable infectious diseases) are included.

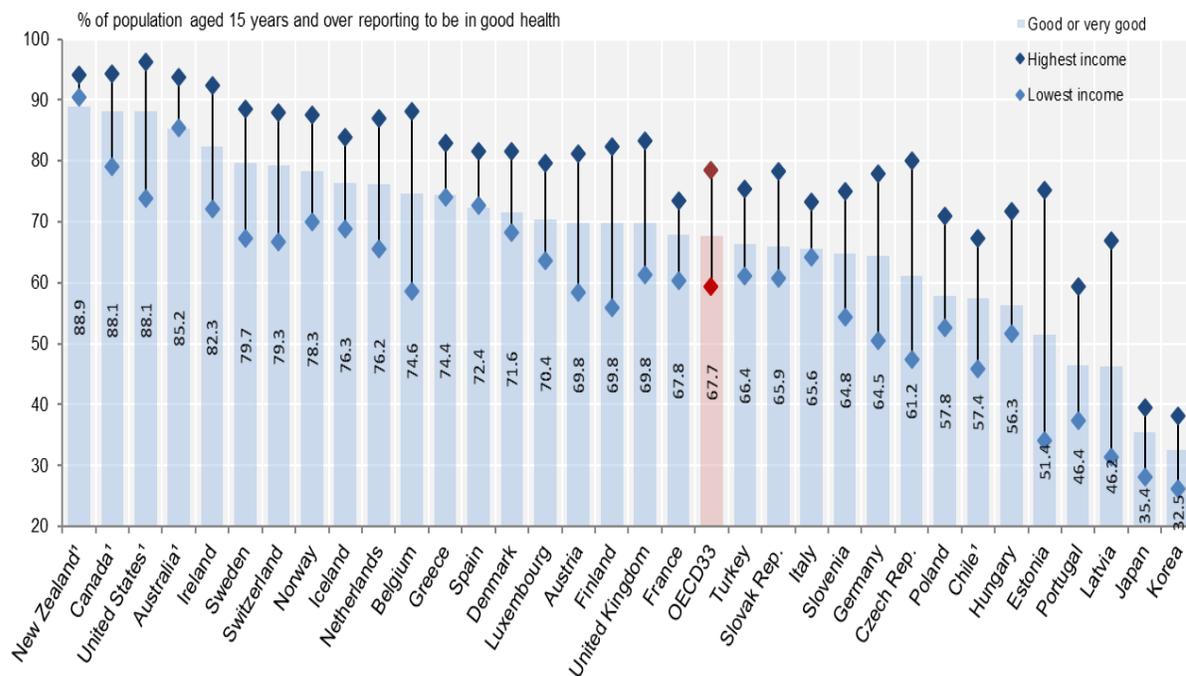
The selection of indicators to include within a P4P scheme depends upon many factors, including data availability and the evidence base for effective intervention. Arguably, however, disease burden should be a principal design criterion, if the final aim of the payment system or performance management mechanisms more generally is to improve population health.

Socioeconomic inequalities in health are pronounced in Estonia

Twice as many people in the highest income quintile assess their health status as good or very good in Estonia compared to people in the lowest income quintile. This is the largest gap among all EU countries (2), as shown in Figure 1. QBS is relevant here because primary care has the potential to reduce socio-

economic inequalities in self-perceived health (3). A strong primary care system can reduce the negative impact of income inequality on health (4).

Figure 1: Perceived health status in OECD countries by income level, 2015 (or nearest year)



Source: OECD Health Statistics, 2018

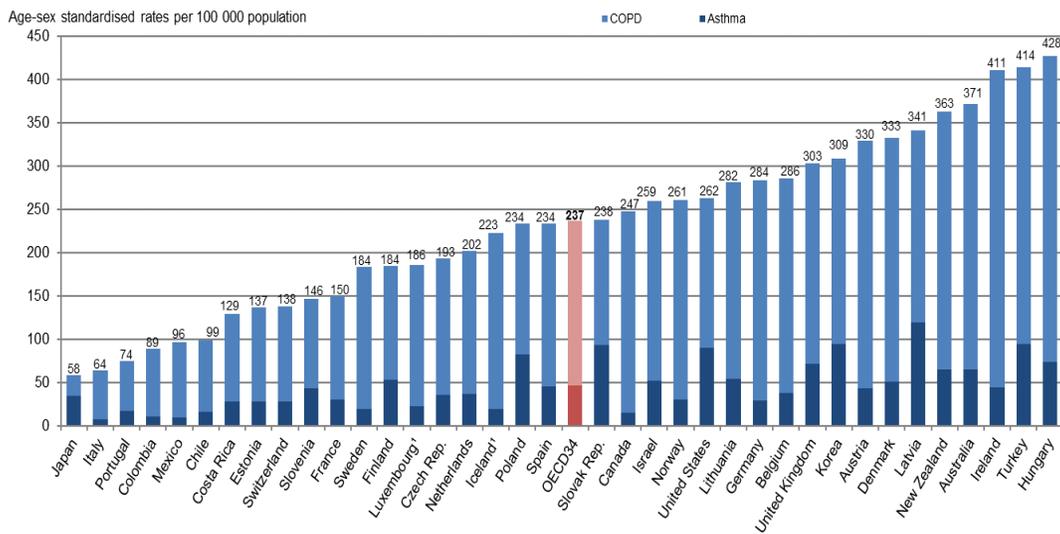
Disparities in health status are largely driven by the high prevalence of chronic conditions among people with lower educational levels and lower income levels. People with less education in Estonia, for example, are 1.5 times more likely to suffer from asthma or other chronic respiratory diseases, and 1.4 times more likely to have hypertension than those with the highest level of education. Disparities in health outcomes have also been observed when comparing Estonians with Russians, Ukrainians and Belarusians living in Estonia (1). Between 2006 and 2016, there was no improvement in self-assessed health for people in the lowest income quintile, while the rest of the population reported improvement. More promisingly, however, improvements were reported for all educational levels.

Rates of hospital admission for some chronic conditions are high in Estonia

Avoidable hospitalizations for ambulatory care sensitive conditions provide an indication of the effectiveness of the primary care system. The Estonian primary care system seems to be effective in helping to prevent hospital admission for patients with asthma and COPD, about average for congestive heart failure and diabetes, but performs relatively poorly for hypertension compared to OECD peers (Figures 2-4).

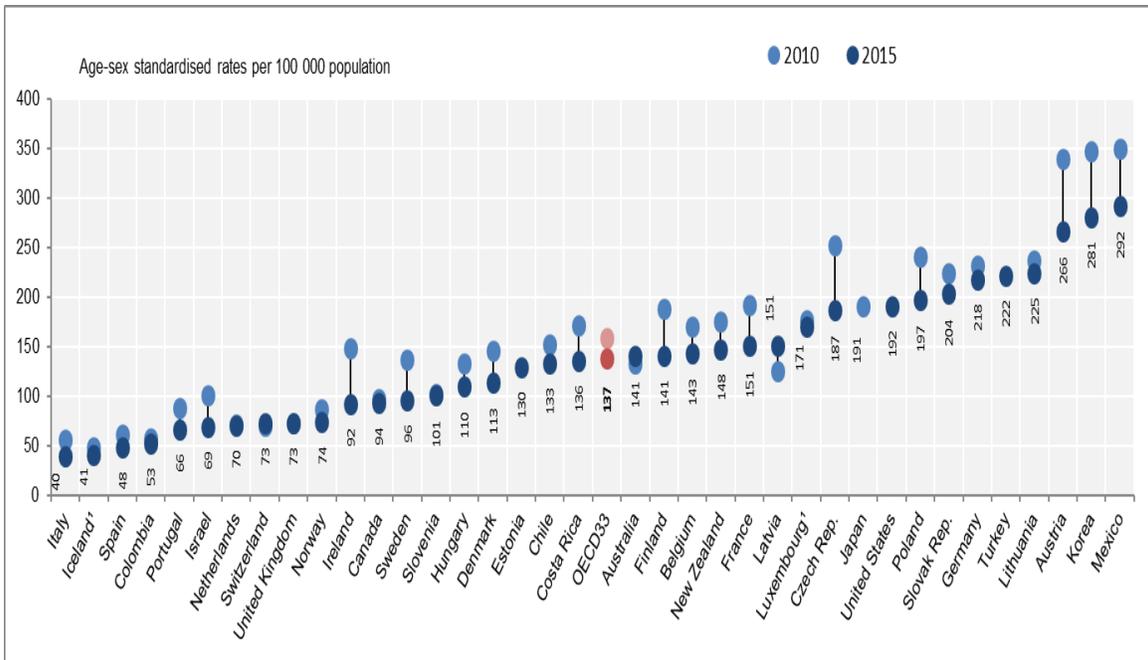
As shown later in this report, QBS includes a large set of process indicators for hypertensive care. Despite this, clinical outcomes for hypertension (as measured by hospital admission) appear poor. Crude admission rates can be misleading, however, since are also likely to be influenced by the local prevalence of disease, by the availability of hospital beds, and by financial or other (dis)incentives to admit. It is also worth noting, however, that Eijkenaar et al 2013, in a review of systematic reviews of primary care P4P schemes, concluded that there was no evidence of effects of pay-for-performance on hospital admissions or mortality.

Figure 2: Asthma and COPD hospital admission in adults, 2015 (or nearest year)



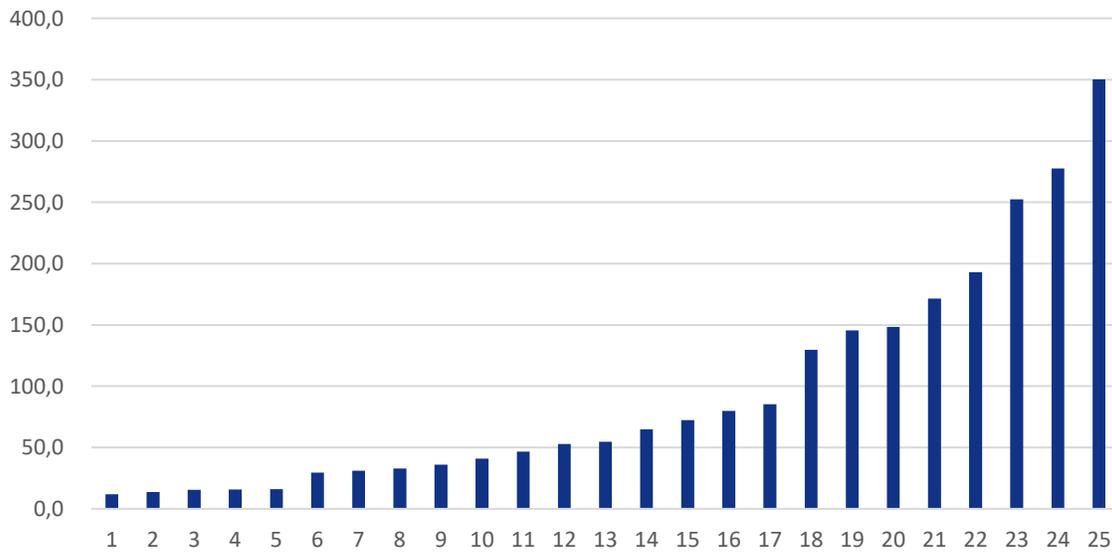
Source: OECD Health Statistics, 2018

Figure 3: Diabetes hospital admission in adults, 2015 (or nearest year)



Source: OECD Health Statistics, 2018

Figure 4: Hypertension hospital admission in adults, 2015 (or nearest year)

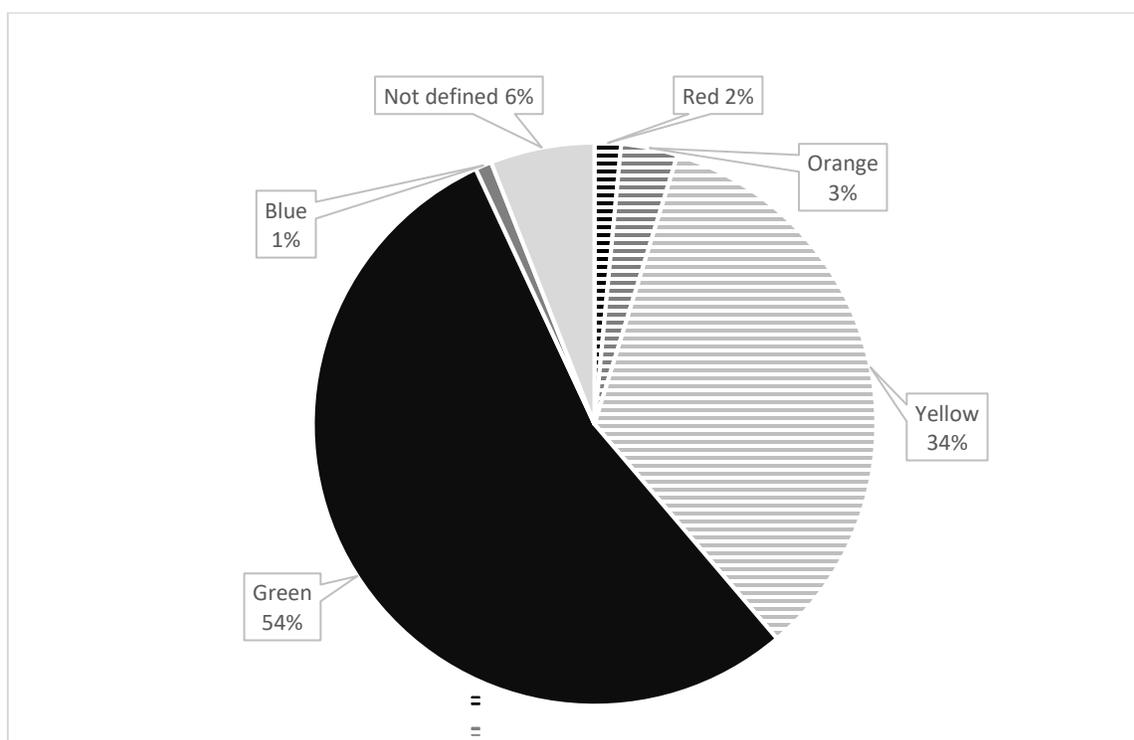


Source: OECD Health Statistics, 2018

There is scope to improve both patient-centredness and care coordination in Estonia

Primary care plays a central role in coordinating care for patients across the health care systems. Data on health outcomes suggest substantial room for improving the coordination between levels of care in Estonia. The 30-day fatality rates for acute myocardial infarction and stroke, for example, are among the worst in the EU (2). Survival after a heart attack or stroke is not first and foremost dependent on primary care, but a lack of clarity concerning roles and responsibilities for the management of people with chronic conditions across primary and secondary care, and between health and social care sectors may contribute to worse outcomes (1). There is also substantial evidence that many Estonians use hospital Emergency Departments (ED) as a substitute for primary care. Most ED attendees, for example, are triaged into the lowest severity categories, blue or green (Figure 5). Further evidence on substitution, as well as the lack of co-ordination between primary care and other health care sectors, is presented in the companion report *“Reforming reimbursement of hospital-based emergency care in Estonia”*. Of note, care coordination is not currently part of QBS.

Figure 5. ED attendees by triage category, 2017



Source: NAO, based on data of HNDP hospitals

Estonian Family Doctors score well on some, but not all, measures of patient-centredness. OECD data shows that the share of patients reporting that their doctor spent enough time with them during the consultation was 86.7%, well above the OECD average of 81.3%. In contrast, the share of patients reporting that their doctor provided easy-to-understand explanations was 85.9%, just below the OECD average of 88.9%, and the share of patients reporting that their doctor involved them in decisions about care and treatment was

79.5%, below the OECD average of 83.1%. Of note, measures of patient satisfaction and patient experience are not currently part of QBS.

EHIF reimburses primary care through a complex array of mechanisms

EHIF's reimbursement of primary care comprises a blend of four mechanisms: capitation, fee-for-service (FFS), the quality bonus scheme, and add-on payments. Capitation is the dominant mechanism, accounting for around two-thirds of a typical Family Doctor's income. FFS accounts for just over a fifth of total income, and is derived from three sources: the diagnostic fund (which reimburses investigations such as blood tests or X-rays); the therapeutic fund (which reimburses clinical psychology, speech therapy and physiotherapy services); and, the procedure fund (which reimburses minor surgery and other primary care procedures). Add-on payments account for around a sixth of total income, and are intended to reimburse the cost of premises, IT systems, medical devices and equipment, transportation and training. Other add-on payments are available to cover the salary, premises and equipment costs for a second family nurse; transport costs for doctors working in rural areas; and, work outside normal office hours. Patient co-payments for home visits represent another potential source of income. These, however, are discretionary (and cannot be charged for visits to infants or pregnant women).

QBS represents a small component of primary care reimbursement. Doctors qualifying for the full bonus amount can derive at most around 4% of total income through QBS; average reimbursement, however, represents just under 2% of total income. Total reimbursement through QBS was 2.95m EUR in 2017.

QBS is blended with a complementary scheme for quality management

Alongside QBS, a quality management bonus scheme (QMBS) also operates, focused on monitoring and encouraging Family Doctors to follow the guidelines set by ESFD. Participation is voluntary, even though compliance with the guidelines is compulsory and is specified in EHIF's contract with Family Doctors.

Participating doctors self-complete an online questionnaire concerning 20 criteria. One point is awarded for meeting each criterion, adding up to an accreditation level for each doctor (Level A, 19-20 points; Level B, 16-18 points; Level C, 0-15 points). A and B practices receive a bonus payment at the practice level. QBS results is one of the criteria evaluated. To receive a point for this criterion, at least 80% of QBS must be filled. Additional point thresholds apply if doctors share patients lists.

To validate the data, claims data and other routine administrative data are checked. In addition, a sample of practices is visited, based on the self-assigned levels. The final accreditation score (and thus accreditation level) is based on the practice visits, or on the check based on available data (for those doctors that were not visited). All A and B level doctors receive a financial incentive in the form of a quality management bonus.

Table 2 shows that almost half of the Family Doctors participated in 2017/18 in the Quality Management Bonus System, which was higher in 2016, but much lower in 2009. The share of practices that reached level A accreditation scores was similar in 2009 and in 2016.

Table 2: Quality Management Bonus System results

Year	% of Family Doctors participating	% reaching Level A	% reaching Level B	% reaching Level C
2009	16.3% (n=79)	22.8% (n=18)	NA	NA
2016	68.8% (n=315)	21.0% (n=66)	16.5% (n=52)	62.5% (n=197)
2017	48.8% (n=225)	NA	NA	NA

An important strength of the practice visits is not just the validation of the data, but particularly that the discussions with the reviewers could result in a better understanding of quality guidelines by Family Doctors. In addition, it can be stimulating for doctors, that the list of A and B-level doctors is reported on the website of EHIF.

3. QBS compared to other national P4P schemes in primary care: what is incentivized?

The QBS was launched in 2006, with the aim of incentivizing preventive care and management of chronic conditions. It relies on claims data submitted by Family Doctors to EHIF for reimbursement. While this minimises administrative burden, it means that QBS is limited to measuring processes and activities. In contrast to primary care P4P schemes in other countries, key dimensions performance such as clinical outcomes or patient experience are not captured by QBS.

QBS focuses on chronic disease, including complex bundles of care

The overall aim of QBS is to incentivize preventive care and management of chronic conditions in primary care. More specific objectives include:

- encouraging Family Doctors to actively engage in the prevention of illnesses, thereby avoiding subsequent high costs in relation to the treatment of those illnesses or people's premature incapacitation for work, invalidity or death;
- achieving and maintaining a critical level of vaccination, thereby helping prevent the spread of certain infectious diseases;
- assuring more effective monitoring of certain chronic conditions, thereby helping prevent the development of complications;
- incentivizing Family Doctors to provide a more broad-based health service.

Further detail on the current design and operation of QBS is given in Box 1, and Annexes 1 and 2.

Box 1: Current design of QBS

The QBS includes 19 indicators agreed upon by EHIF and the Family Doctor's Association. These indicators mostly capture care processes and fall under three domains: Domain 1 targets children (vaccinations and health checks), Domain 2 targets people with chronic conditions (chronic disease management such as monitoring and medication prescriptions), and Domain 3 targets pregnant women (monitoring of pregnancy) and monitors other miscellaneous activities (performing gynecological examinations, performing gynecological and minor surgical procedures paid for through the therapeutic fund, ensuring Family Doctor's and family nurse's professional competence, and participation in the Estonian Family Doctor Association's quality management audit).

The QBS uses a points-based system to determine performance. The point system gives greater weight to indicators that have been considered more important at the time of the design of QBS. These weights have not been revised since. For indicators in Domains 1 and 2, Family Doctors must meet a coverage threshold for the relevant target group to obtain points. In 2016, the threshold for indicators in Domain 1 was 90% of the target. For indicators in Domain 2, thresholds are revised annually based on the average performance of all practices participating in the

scheme in the previous year, ensuring a stepwise increase in coverage. In 2016, the threshold consisted of the previous year's average coverage among all participating Family Doctor practices for each indicator + 10% but no more than 90%. In both domains, if an indicator threshold is missed, no points are awarded for that indicator. No points are awarded for the indicators in Domain 3. Instead, a coefficient is determined (up to a maximum of 1.0) to calculate the financial bonus. Further detail on Domains 1 and 2 can be found in the Annexes to this report.

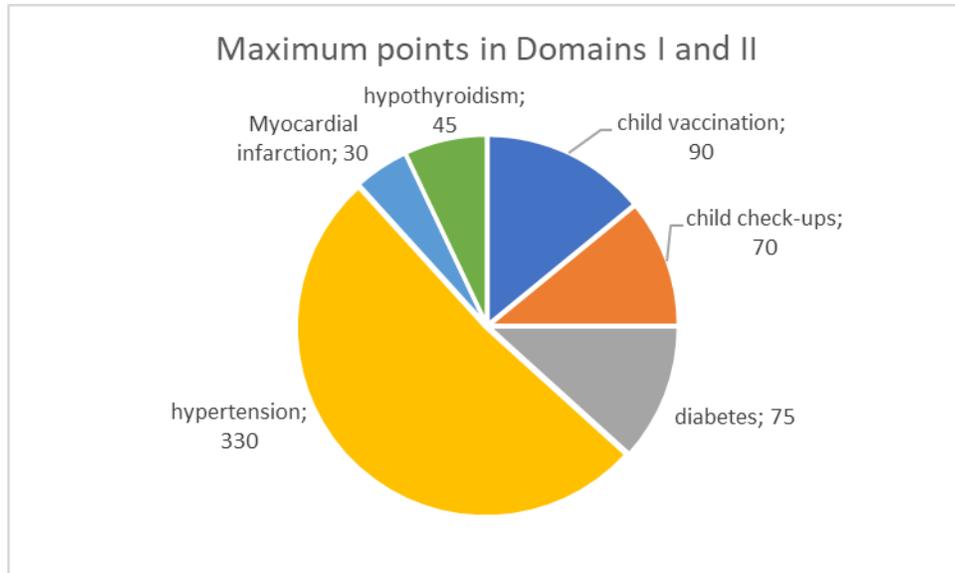
For all indicators, service bills must be submitted as evidence to support claims of achieving indicator thresholds. At the end of each year, EHIF awards bonuses to Family Doctors based on the total number of points achieved. To receive a bonus, Family Doctors must achieve at least 512 out of 640 points. Family Doctors meeting this threshold, but not attaining 90% of the maximum point yield (or 576 points) receive a bonus of 0.8 percent of the total value. If Family Doctors achieve at least 90% or 576 points, they are awarded the full bonus of EUR 4300 (2016). For those Family Doctors achieving the full bonus and employing a second family nurse, EHIF pays an additional EUR 1000 (2016). In 2016, 471 Family Doctors (59% of all Family Doctors) were paid the maximum bonus for performance on Domains 1 and 2, while 244 (30%) Family Doctors received bonuses for Domain 3.

Family doctors receive electronic performance feedback three times a year: mid-year results in the third quarter, full-year preliminary results in April and final results in June. Performance results for each Family Doctor are published on the EHIF's website.

Source: "Toward greater integration of care and improved efficiency: a critical review of EHIF's payment system", WBG report to EHIF, October 2017.

Figure 6 and Table 3 illustrate how QBS points are currently distributed. Looking at the distribution of maximum points that can be achieved, Figure 6 shows that more than half of the points in Domains 1 and 2 are awarded to hypertension. Although hypertension is an important condition (in itself, and also as an exacerbating factor in several other chronic cardiovascular and metabolic conditions), the distribution of points illustrated in Figure 6 is arguably skewed, with insufficient emphasis on other chronic conditions such as diabetes. Some chronic conditions that can be managed fully or in part by primary care, such as asthma or mild/moderate depression, are not covered at all (the QBS working group deciding that treatment options in primary care were not yet sufficiently developed to justify their incorporation into QBS).

Figure 6: Clinical areas linked to QBS points in Domains I (prevention) and II (chronic conditions)



Within the domain focussed on management of chronic disease, the focus of QBS has shifted from reimbursing single procedures or services towards complex targets with bundled (grouped) activities, using an 'all or nothing' approach for rewarding points. These complex targets cover an array of activities that must be achieved simultaneously for a given patient group. If one activity is not achieved, all other activities are not rewarded. For example, points are only awarded if all necessary vaccinations for children up to 3 years old are administered as opposed to performing a single vaccination, or if a complete cycle of services for hypertensive patients is performed, as opposed to single services.

This is an innovative aspect of QBS, as it forms an important incentive for comprehensive, proactive primary care. However, doctors have indicated that it has made it more complex to reach targets, and might reduce the motivation levels to improve performance.

Table 3: Summary overview of QBS points distribution by domain and indicator set

QBS Domain	Indicator topics	Indicators	Indicator sets	Single indicators	Maximum points for indicator(s)	Total points for domain	Minimum points required for bonus
Domain I: Disease prevention	Child vaccination	39	10	1	90	160	90 (100 per cent of max)
	Children’s preventive check-ups aged 1-, 3-, 12-months, 2 years, 6-8 years	5	5	2	70 (two indicators, points 60 and 10)		At least one of the indicators
	Colon cancer screening	1	1	1 (no points assigned)			
Domain II: Chronic disease management	Diabetes, type II	6	2	2	75 (two indicators, points 65 and 10)	480	At least one of the indicators
	Hypertension	16	5	5	330 (five indicators, points; 90; 175; 40; 5; 20)		At least one of the indicators
	Myocardial infarction	6	3	3	30 (three indicators, points: 20; 5; 5)		At least one of the indicators
	Hypothyroidism	1	1	1	45		45 (100 per cent of max)

Domain III: Additional professional competence	Recertification for Family Doctor and nurse	2	1	2 (no points assigned)	0.2 (coeff)	
	Maternity care: pregnancy monitoring	3	1	1 (no points assigned)	0.3 (coeff)	
	Gynaecological activities	3	1	1 (no points assigned)	0.2 (coeff)	
	Minor surgical activities	14	1	1 (no points assigned)	0.3 (coeff)	
	Albumin/creatinine ration in urine in diabetic patients	1	1	1 (no points assigned)	0 (tracking indicator)	
	Albumin/creatinine ration in urine in hypertension patients	1	1	1 (no points assigned)	0 (tracking indicator)	
	E-consultation	1	1	1 (no points assigned)	0 (tracking indicator)	
TOTAL						640
						512 (80 per cent of max)

QBS addresses few dimensions of primary care performance

The current design of QBS is restricted to three dimensions of primary care performance: preventive care; chronic disease management; and, enhanced services (including competence and some aspects of prescribing safety). This leaves several dimensions of performance untouched, such as equity, effectiveness, continuity of care, accessibility, acceptability and patient-centredness. Furthermore, within this limited number of dimensions, QBS uniquely addresses processes of care (such as undertaking a blood pressure check). QBS does not currently address clinical outcomes, namely, achievement of target blood pressure. As a result, it cannot be evaluated if primary care is achieving beneficial health outcomes for patients.

There is no right number for dimensions of performance to be addressed by a pay for performance schemes. On the one hand, fewer dimensions (and indicators) may make the programme simpler to administer and provide clarity. Few indicators, however, may risk distorting overall patterns of care, by overemphasising those services or dimensions of performance that are incentivised. Conversely, a scheme with many dimensions may give a more balanced set of incentives, but will add complexity in administering the programme and may dilute the value of individual incentives to providers (5). Whatever number is chosen, it is important that indicators reflect the objectives of the pay for performance programme, and that those objectives are aligned with and reinforce other strategies to monitor and improve the quality of primary care. Similarly, when providers are performing well, or no (further) improvement can be reached, an indicator should be removed from the incentive scheme.

Tables 4 and 5 show the various domains that national primary care pay for performance schemes cover. Most often, indicators relate to processes, as with QBS. Other commonly found dimensions of performance, however, are efficiency (such as pharmaceutical expenditure per user), patient experience or satisfaction, and improved equity or the reduction of health disparities.

The tables also illustrate how some national programmes (including Estonia's QBS) use a limited total number of indicators to emphasise clinical activities that are linked to conditions with a high local prevalence, emphasise selected dimensions of clinical quality, or emphasise particular problems such as low vaccination coverage. In contrast, other schemes use a bigger number of indicators to capture a wider care continuum (5). One example of the latter is the United Kingdom's QOF, which uses 77 indicators (in 2016/17) aimed at capturing primary care quality from prevention to clinical outcomes. It is worth noting, however, that the QOF initially comprised 150 indicators. The number was reduced after feedback from primary care practitioners advised that fewer indicators would be identify clinical priorities more clearly and be easier to manage (6).

Table 4: Primary care pay for performance domains in OECD countries (2012)

Participation of health providers	Country	Pay for performance typically relates to					Other
		Preventive care	Management of chronic diseases	Uptake of IT services	Patient satisfaction	Efficiency	
Mandatory participation	Estonia	X	X				X
	Chile	X	X		X	X	X
	Korea		X			X	
	Spain	X	X			X	
	Sweden	X	X	X	X	X	
	Turkey	X				X	X
Voluntary participation with conditions	Australia	X	X	X			X
	France	X	X	X		X	
	Mexico	X	X		X	X	X
	New Zealand	X	X				
	Portugal	X	X		X	X	
	United States	X	X	X	X	X	
Voluntary open participation	Czech Republic	X					X
	Hungary	X	X			X	
	United Kingdom	X	X		X	X	

Source: *Better Ways to Pay for Health Care*, OECD Publishing, Paris 2016

Table 5: focus areas by national pay for performance scheme

Canada	Australia	New Zealand	United Kingdom	United States	Estonia
Health status	Health status and outcomes	Health			Preventive care
Acceptability	Responsive	People-centred	Patient / carer experience	Patient-centred	
Accessibility	Accessible	Access	Fair access	Timeliness	
Appropriateness	Appropriate		Effective delivery of appropriate health care		Preventive care Chronic disease management
Competence	Capable		Effective delivery of appropriate health care		Enhanced services
Continuity	Continuous		Patient / carer experience		
Effectiveness	Effective	Effectiveness / Health	Health improvement and health outcomes	Staying healthy Getting better Living with illness End of life care	
Efficiency	Efficient	Efficiency	Efficiency		
Equity	Equity	Equity	Equity	Equity	
Safety	Safe	Safety		Safety	

Within the ‘effectiveness’ dimension shown in Table 5, it should be noted that these indicators refer to intermediate health outcomes, such as blood pressure. No schemes deploy indicators related to final health status, such as mortality, or self-reported health. This is due to the difficulty in reliably linking the activities of an individual care provider to patient outcomes. Examples of the clinical outcomes included in national primary care P4P schemes are given in Box 2.

Box 2: Inclusion of clinical outcomes in primary care performance schemes

Examples of clinical outcomes, for people with diabetes, included in other primary care P4P schemes include:

- **UK, QOF:** The percentage of patients with diabetes in whom the last blood pressure reading (measured in the preceding 12 months) is i) 150/90 mmHg or less; ii) 140/80 mmHg or less
- **UK, QOF:** The percentage of patients with diabetes in whom the last IFCC-HbA1c reading is i) 64 mmol/mol or less; ii) 75 mmol/mol or less
- **France, ROSP:** The percentage of patients with diabetes in whom the last IFCC-HbA1c reading is 8.5% or less; 7.5% or less
- **France, ROSP:** The percentage of patients with diabetes in whom the last cholesterol reading is i) 1.3g/L or less; ii) 1.5g/L or less
- **Portugal:** The percentage of patients with diabetes in whom the last blood pressure reading (measured in the preceding 12 months) is 140/80 mmHg or less
- **Portugal:** The percentage of patients with diabetes in whom the last IFCC-HbA1c reading is 8.0% or less

In New Zealand’s PHO Performance Programme, some indicators are measured separately for high-need-populations, which are rewarded at a higher rate. In Portugal for the Family Healthcare Units (FHU), there is a mix national set of indicators based on national health objectives, population characteristics, good practices, and historical data (covering the domains of access, clinical performance, efficiency and perceived quality); and additional indicators that are selected regionally (weight 15%), and indicators that are proposed by each FHU according to their own improvement quality plan (weight 15%). The clinical performance indicators are a mix of process indicators, and intermediate outcome indicators. Table 6 shows the indicators that are applied in the pay for performance schemes in Portugal and Canada (7).

Table 6: Indicators used in primary care-based quality programmes in Canada and Portugal

Indicator	Area	Type	Weight
Proportion of patients with at least one medical appointment during the last three years	Horizontal	Access	4.50%
Rate of nursing home visits per 1 000 patients	Horizontal	Access	3.00%
Proportion of pregnant women with adequate follow-up	Women Health	Clinical Performance (process)	4.50%
Proportion of women in reproductive age with appropriate monitoring in family planning	Women Health/ Family planning	Clinical Performance (process)	5.00%
Proportion of Infants within the first year of life with adequate follow-up	New-born, child and adolescent care	Clinical Performance (process)	6.00%
Proportion of seniors without prescription anxiolytics, sedatives and hypnotics	Mental Health	Clinical Performance (intermediate outcome)	2.00%
Proportion of patients of more than 13 years old characterised with smoking habits in the last three years	Horizontal	Clinical Performance (process)	2.50%
Proportion of hypertensive patients younger than 65 years old with controlled blood pressure	Chronic diseases - High blood pressure	Clinical Performance (intermediate outcome)	3.00%
Proportion of controlled diabetics (HgbA1c <= 8.0 %)	Chronic diseases - Diabetes	Clinical Performance (intermediate outcome)	3.00%
Pharmaceuticals expenditure per user	Horizontal	Efficiency	16.00%
Ancillary exams expenditure per user	Horizontal	Efficiency	8.00%
Proportion of patients satisfied and very satisfied	Horizontal	Perceived quality	5.00%

Source: (7)

Patient-centredness is another domain of performance that is not currently included in QBS. Explaining treatment choices in an understandable manner, responding to patient concerns and treating them as partners in health care decisions are essential in strengthening relationships with patients as well as supporting self-management and enabling adherence to treatment. Few primary care performance schemes internationally include patient-reported indicators. Nevertheless, OECD health systems are increasingly showing interest in collecting patient-reported experiences and outcomes, as a means to monitor and encourage more patient-centred care.

Limited revision of QBS indicators has been undertaken since creation of the scheme

QBS was one of the world's first national pay-for-performance schemes to be introduced in primary care, signalling significant innovation and ambition at that time. The principles applied when designing QBS were that the indicators should: 1) relate to an important health care issue; 2) be easy to measure and/or be part of routinely collected data; 3) not relate to a clinical area where performance is difficult to assess; and, 4) not cover too many aspects of care initially. QBS has, however, undergone limited systematic updating or revision since, besides creation of the complex indicator sets described earlier, and minor additions/removals of some indicators. Within Domain 1 (prevention), recent changes include the addition of the rotavirus vaccination in 2015, and the consolidation of nine individual childhood vaccinations into a single, complex indicator in the same year. Now, all these vaccinations must be completed for points to be awarded. Similarly, the four routine childhood examinations (at ages 1 month, 3 months, 12 months and 2 years) were combined into a single, complex indicator in 2016. Now, all four examinations must be done to complete this indicator.

Within Domain 2 (chronic conditions), monitoring of levels of albumin in the urine of diabetic patients was dropped in 2017 and monitoring of urinary albumin/creatinine ratio was introduced as a pilot; nurse-led patient counselling was added in 2018. Furthermore, the separate elements of diabetes care were combined into a single, complex indicator in 2016. Appropriate prescribing of medication was added in 2016. Monitoring of urinary albumin/creatinine ratio was also piloted for patients with hypertension, as was nurse-led patient counselling. The requirement for an ECG in low-risk hypertensive patients was dropped in 2013. For patients with an MI, measurement of cholesterol fractions was added in 2016; appropriate prescribing of beta-blockers and statins was added the same year. Within Domain 3 (additional competencies), the requirement for a midwifery appointment during pregnancy monitoring was added in 2018; e-consultation was introduced as pilot indicator in the same year.

It is timely to ask whether each indicator still meets the original design principles, and whether the design principles themselves are as relevant now as previously. The criterion that indicators be easy to measure, for example, would have favoured inclusion of hypothyroidism in QBS. This is a condition, however, with a low disease burden, and one that does not require high-performing primary care to manage it. Furthermore, there is no national quality guideline for managing hypothyroidism in primary care in Estonia. This both signals that hypothyroidism is not a priority condition and represents a disconnect between QBS and strategies to improve primary care quality more broadly.

Some QBS indicators and targets are now more up-to-date than the corresponding national clinical guidelines. The national clinical guideline for acute myocardial infarction (AMI) was last updated in 2004, for example, and diabetes in 2008. Although this means that QBS encourages more evidence-based care than the guidelines, the disconnect between QBS and guidelines means that Estonia lacks a coherent and strategic approach to primary care quality. In short, it is important that QBS indicators be regularly reviewed against inclusion criteria that are in widespread use internationally, as described later.

QBS indicators are constructed from routine administrative data

A key criterion in designing QBS was that the indicators could be drawn from EHIF's electronic billing system, the drug prescription database and National Birth Register (1). The electronic billing system covers all primary care doctors in the country and encompasses patient-level information such as diagnosis, activities included in the capitation payment, and other activities qualifying for fee-for-service or add-on payments.

It is positive feature in the design of QBS that no additional reporting is needed, beyond Family Doctors' recertification status (which is extracted by EHIF from the Health Board), and the audit visits that are performed regarding the quality management indicators. It should be noted, however, that claims data are not designed to measure performance and could thus provide an incomplete picture of provider performance. To measure the quality of primary care, particularly outcomes, data would from electronic medical records and the administrative data of hospitals is needed. Hospital admissions or emergency department attendance data for ambulatory sensitive conditions, for example, could provide valuable insight in the effectiveness of primary care. This is technically possible in Estonia, since hospitals record data on individuals' primary diagnoses, reason for attendance/admission, and Family Doctor.

Internationally, most pay for performance programmes to rely on claims data sent by care providers to health insurers, as with QBS. This not surprising given the detail of claims data, their general good quality and ready availability. As mentioned above, however, claims data are not designed for quality measurement and improvement: critical dimensions of quality, such as the effectiveness or patient-centredness of care, are not typically captured. Over time, therefore, pay for performance programmes have tended to move away from claims data or supplement it with other data sources (8), with the aim of providing a more complete picture of provider performance. A particularly sophisticated example of this concerns the QOF, which extracts data (including clinical outcomes and measurements) from patients' electronic medical records (5).

Whatever data sources are used, P4P schemes typically have a mechanism to validate the performance metrics that are calculated for or supplied by each provider. This is necessary to avoid overpayment due to inflated reporting or other forms of gaming. Data validation also offers a valuable opportunity for insurers or quality agencies to discuss current performance, barriers to improvement, and steps toward further improvement directly with care providers (5).

4. QBS compared to other national P4P schemes in primary care: who is incentivized?

Most QBS incentives are directed to the individual Family Doctor. Primary care is increasingly team-based, however, suggesting that there may be scope to consider wider use of group-level incentives within QBS.

QBS reimbursement is directed to the doctor, rather than the wider primary care team

Within QBS, all incentives are directed towards the Family Doctors as individuals. The exception to this arrangement concerns the QMBS, where the incentive is paid to the practice. One of the risks of incentivizing individuals as opposed to team or group level is that estimation of performance may not be reliable due to the small sample size (in this case, patient panels). If this is the case, it could lead to incorrect scoring and allocation of payments (9). In Estonia around 300 out of around 800 Family Doctors work as single practitioners. In these cases, there would be no gain in sample size from moving to practice-based payment. Nevertheless, it is important to note that sample sizes may affect the reliability of performance estimates of individual doctors.

More relevant is to consider the role of the wider primary care team, including within a single Family Doctor practice. Typically, achievement of QBS targets will involve the wider practice team, including nurses, receptionists and so on. Given that it is the Family Doctor who typically decides how the payment is used, there is a risk that the wider primary care team may not see the benefits of QBS achievement. This might reduce the impact of QBS. The same is true for the quality management bonus that is paid to the practice: there are no rules on how this bonus should be spent.

Incentives are only likely to work if the person(s) doing the work sees the benefit(s), such as additional income distributed to group members, or employing additional staff. For each indicator that is included in QBS, therefore, it is worth considering whether all the primary care professionals responsible for achieving the incentivised target are able to directly see the benefit from achieving good performance, and if not, considering how the financial incentive could be more fairly distributed to reward individuals' contribution. In most cases, this would mean that the income derived from QBS Domains I and II should be treated in the same way as that from Domain III, namely, paid to the practice rather than the individual Family Doctor.

Group-level incentives may be more effective than incentives targeting individuals

The impact of incentives on provider behaviour may be influenced by whether the payment is made to individual providers or made at group/organisation level. Given the increasingly collaborative forms of care provision, group effort is generally needed to improve primary care performance. It becomes less valid, therefore, to hold individual providers accountable for performance, particularly for team-based care (10). At the same time, however, if incentives are paid to at group-level, it is essential that they also reach individual care providers, otherwise they may not influence behaviour. In OECD countries, most national pay for performance schemes for primary care pay bonus payments to individual providers, as illustrated in Table 7

Table 7: Individual vs group-level targeting in P4P schemes in primary care

Participation of health providers	Country	Bonus payment is made to	
		Individual	Organisation
Mandatory participation	Estonia	X	
	Chile	X	
	Korea	X	
	Spain	X	
	Sweden		X
	Turkey	X	
Voluntary participation with conditions	Australia	X	X
	France	X	
	Mexico	X	
	New Zealand		X
	Portugal	X	
	United States	X	X
Voluntary open participation	Czech Republic	X	
	Hungary		X
	United Kingdom		X

Source: (7)

To date, only one randomized trial has compared individual versus clinical group incentives. It showed that individually incentivized doctors were more likely to achieve blood pressure control or change medication appropriately compared to doctors for whom performance incentives were paid to the group (11). In contrast, a recent study using administrative data from a Dutch insurer sought to ascribe performance to the profile of individual GPs. The study found that the correlation between GP profiles and performance was generally too low to justify relying on this mechanism for performance-related pay, or even any application of profiling (12). This finding may point, then, to a preference for team, group or even higher levels incentives (9, 13). Even in this case, however, it is important that performance payments are effectively distributed to its members. Even if some practices are still rather small (e.g. solo practices), it is important to include them (9).

Box 3 describes how the system of group level payments works in the Quality and Outcomes Framework in the United Kingdom (6).

Box 3: Practice-level payment in the United Kingdom's Quality and Outcomes Framework

Under the QOF, general practices are awarded points, each attracting a payment, for recording specific activities or outcomes described in a set of indicators.

The practice will agree in advance with the Primary Care Organisation the number of QOF points they are aiming for that year. A monthly payment equal to the value of 70% of these points is made to the practice. Remaining payments are made once the practice has actually achieved the points.

The number of points available varies by indicator. Some indicators reward a practice-level activity, for example, being able to identify a list of patients with a particular condition, and others reward the practice for the proportion of patients who have received a component of clinical care or who have achieved a particular outcome. For this latter type of indicator, the practice receives points on a sliding scale between a lower and upper threshold according to the proportion of relevant patients recorded as receiving the care or achieving the outcome. No points were given if the practice did not reach the lower threshold, and maximum points if it reached the upper threshold.

In 2016/17, a practice may earn a maximum of 559 points across 77 indicators, each point attracting a payment of £165.18. Payments are weighted by list size and measures of disease prevalence.

There may also be instances in which it could be more appropriate to reward higher organisational levels. This might apply, for example, to incentives for closer cooperation between primary care, hospital care and social care organisations for patients with multiple chronic conditions (13). Examples from other countries (e.g. bundled payments in the Netherlands) show to positive impact of network-level performance rewards (14). Another example for which higher level entity rewards could be appropriate is if indicators were developed to improve information flow between acute inpatient and ambulatory care settings in Estonia. Poor information flow is known to be a problem in Estonia, manifest as significant proportions of hospitalized patients that receive unnecessary preoperative diagnostic procedures, are discharged without appropriate medications, or do not receive appropriate follow-up care after discharge (1).

5. QBS compared to other national P4P schemes in primary care: how is performance incentivized?

Estonia's QBS differs in important respects from primary care P4P schemes in other countries. It has a more complex design than many schemes, rewarding relative improvement rather than absolute achievement, and is linked to a separate bonus scheme for Quality Management. QBS pays less (as a fraction of overall income) than several other schemes and feedback of results takes several months, slower than that seen in other countries. It also offers less opportunities to exempt patients from inclusion in the scheme, which may penalize Family Doctors who care for patients with unusually complex needs.

QBS is a small fraction of primary care's overall income

The size of the financial incentive is important for creating a meaningful incentive to which providers will find it worthwhile to respond, without distorting other activities. A too low incentive will not encourage a behaviour change, while a too high incentive may result in unintended consequences such as providers shifting excessive focus towards performance areas and services that are rewarded or risk selection (15, 16).

As explained earlier, QBS offers a relatively small financial contribution to the income of Family Doctors today (1), despite payment increases of up to 25% in 2016 compared to the previous year. In 2017, only 2.7% of Family Doctors' total revenue came from QBS (1). In addition, Family Doctors can earn up to a maximum of 41% of their total capitation payment if they perform well according to the quality bonus system standards. This additional payment is 39%, however, for Family Doctors who do not attain the quality goals (1). Though the differentiation in rate is intended to promote improvements in quality of care, the small difference (2%) is unlikely to have this effect.

It should be noted that Family Doctors with more patients (with diabetes, hypertension etc.) than the national average can apply a 1.5 coefficient to the bonus reward. Overall, however, QBS offers a marginal additional contribution to doctors' income, and may therefore have a small impact on behaviour.

The financial incentive in other schemes is larger than that offered by QBS

The impact of a P4P scheme on provider behaviour is not solely dependent on the size of the financial incentive. Several other factors will also determine providers' response, including other incentives in place, the characteristics of the patient population, and the flexibility and resources that providers have to make changes (17). Some studies have found that low size of incentive is the reason behind the modest impact of pay-for-performance programmes (18), other studies found no consistent relationship between incentive size and provider response (19, 20).

With this in mind, it is nevertheless worth noting that most pay-for performance schemes for primary care in OECD countries offer somewhat larger incentives than QBS: typically around 5-15% of total income. In the United Kingdom, up to 25% of GP practice income comes from QOF incentive payments; in Turkey, around 20% of primary care provider salaries comes from incentive payments, and in Brazil about 10% of

GP income comes from incentive payments (5). Furthermore, in several pay for performance schemes (including those in France, Germany, New Zealand, Turkey, United Kingdom) higher payment rates are given for higher achievement levels, typically after a minimum threshold is achieved. Increasing the incentive for increases in performance, has been shown to be powerful mechanism (5).

QBS reimbursement is annual and doctors receive QBS reports with a significant delay

Initially, QBS reimbursement was paid monthly. In 2008, however, this was changed into annual payments to reduce the administrative complexity. Doctors receive feedback on their results for all indicators electronically. There is, however, a 4 to 6-month delay in the feedback loop whereby Family Doctors see their final QBS performance scores and local/national comparisons. Table 8 shows the timing of the key steps in QBS process for a given year. Although doctors receive their initial results after 4 months, their final results after 6 months can be different, following verification.

Table 8: Milestones in the QBS data collection and feedback cycle

YEAR 1	
February	Lists of target groups made electronically available to Family Doctors
March - August	6-month review (includes new patients in the doctor’s list; activities done by that point)
September - November	9-month review (includes new patients in the doctor’s list; activities done by that point)
YEAR 2	
April	Preliminary results
May	Doctors present additional information if needed
June	Final results (EHIF management board confirms); Results communicated to doctors
July	Payments and results published on EHIF website

Six months is a long delay, reducing the chance and motivation to act on the data, and thereby possibly reducing the interest and commitment of Family Doctors to QBS. The relatively long feedback loop is relevant because QBS should, ideally, also drive performance through intrinsic motivation or reputational incentive. This is likely to be weakened if there is too long a delay in receiving feedback. EHIF and the ESFD should look, therefore, for opportunities to shorten the feedback loop.

Performance feedback in other schemes is quicker and more frequent than in QBS

In other countries that apply pay-for-performance schemes, there is at most, a four-month delay in feedback to providers, and a different timing for public reporting. Eijkenaar (2012) (9) identified 13 pay-for-performance programs in 9 countries. For the five programmes for which information was available, three provide provider feedback on a quarterly basis (including the regional Advancing Quality programme in the United Kingdom; the Performance Management Program in New Zealand; and, the Doctor Integrated Network in Manitoba, Canada). Two programmes provided continuous real-time feedback to providers (the regional Ergebnis Orientierte Vergütung scheme in Germany; and the QOF in the United Kingdom). These two programmes, however, used a lower frequency for public reporting of the results, either quarterly (in Germany) or annually (in the United Kingdom).

National P4P schemes in primary care either make annual payments, semi-annual payments, or quarterly payments. Most common are annual payments (seen, for example, in the United Kingdom's QOF, Israel's Clalit and Maccabi P4P schemes; the Netherlands; Canada's Primary Care Renewal Model in Ontario and Doctor Integrated Network in Manitoba; and Argentina's Program of Quality Improvement scheme) (9).

The evidence-base on the right frequency of performance payments is inconclusive (and is probably of marginal importance in the overall scheme design). Individuals tend to value immediate outcomes more than future outcomes of equivalent size. Increasing the frequency of lower-powered payment could potentially increase the effectiveness on provider behaviour, and increase indicator salience, as opposed to the current annual lump-sum payment (9). However, a US study from 2010 (21) in which primary care doctors were randomized into two study arms differing by the frequency of incentive payment, either four quarterly bonus checks or a single year-end bonus, showed no difference between the two arms in average quality measure score or in total bonus amount earned.

QBS results are publicly reported by name for each Family Doctor

QBS indicators, in the early years of the programme, were not publicly disclosed. This changed in 2015, however, when participation in QBS became obligatory for all Family Doctors. From that time, results for individual Family Doctors became publicly reported. Today, EHIF publishes the name of the Family Doctor, number of the patient list and total achieved points and coefficients for QBS on its website. Indicator-level results are not published as such.

Publishing the results of QBS is important feature of the programme, increasing accountability and transparency. Public interest in QBS results, however, is reported to be modest (5). This might very well be because the presented information is not easy to interpret. In addition, because results are not published at indicator level, it limits the actionability for patients (for example, to use the information to choose a Family Doctor with good performance on diabetes care).

EHIF also communicates QBS results to Family Doctors and the ESFD in yearly seminars. For these seminars, EHIF presents graphs and overviews showing comparisons of QBS results between counties, enabling benchmarking (see Section 6). These overviews are also published on EHIF's website and are intended to be a supportive approach in encouraging continuous quality improvement. The results of individual doctors are not made publicly available, however, nor made available within the closed community of QBS participants. This limits opportunities for Family Doctors to directly benchmark their performance against colleagues.

Internationally, almost all P4P schemes in primary care publish, at some level of aggregation, providers' results. The literature shows small, positive effects of public reporting on provider behaviour and patient choices (22). In addition, public reporting creates transparency of the performance of primary care providers and accountability of providers to the public and patients (5).

QBS generally rewards improvement on past performance, rather attainment of a specified threshold

Targets for the majority of QBS indicators in Domains I and II are based on doctors' previous years' achievement plus 10%, up to a ceiling of 90% attainment beyond which further gains are not rewarded. Doctors are therefore encouraged to improve their coverage each year by 10%, realizing that more than 90% is unlikely possible in all circumstances. In addition, each target has its own target frequency (e.g. once a year, or multiple times for a set period). In contrast, in Domain III (Professional Competence) doctors earn point for reaching absolute thresholds. Points are awarded on an 'all or nothing' basis: the doctor is awarded all points if the target is reached, and no points if the doctor fails to reach the target.

This is a sophisticated mechanism and, for low performing providers, can provide good motivation to participate in QBS given that a 10% improvement in performance is generally feasible. A possible disadvantage, however, is that could inhibit motivation amongst providers who are already high performers, given that an additional 10% improvement could be harder to achieve. In addition, even low levels of performance are rewarded, even if below a generally accepted minimum threshold.

Most other P4P schemes reward attainment of a threshold, rather than improvement against baseline

Looking internationally, there are three common designs used to determine level of incentive payments: i) absolute level (that is, whether a pre-specified target was achieved); ii) improvement (that is, whether a sufficient change in the performance measure was achieved); or, iii) relative ranking (that is, how the provider performs on a given measure compared to peers).

Absolute levels can be useful to clearly define performance standards and make pay-for-performance programmes transparent and objective. It is important to note, however, that absolute levels may require case-mix adjustment at local level to account for differences in patient populations. In addition, such targets may not be stimulating further improvements for providers who already achieved the threshold, and – if case-mix adjustment is not adequately done – targets may encourage providers to only focus on easy to reach patients. One solution is to periodically adjust the minimum threshold, and/or adjust them locally given national benchmarks. In the New Zealand PHO Performance Programme, provider-specific targets are adjusted each period as performance and priorities change. In France, the ROSP programme also takes provider baseline performance against national targets into account to compute achievement rates for bonus payments (5).

In contrast, basing provider reward on the change in a measure over time (improvement) has advantages and disadvantages. On the positive side, it has more intuitive appeal for providers, it can encourage continual progress, and it reduces the need for complex case-mix adjustments. On the downside, it adds

complexity to the system and is resource intensive, can favour poor performers, and could inhibit high performers who see further improvements as not worth pursuing.

Relative ranking has the advantage of filtering out the impact of large-scale events (such as an epidemic, or recession) that might affect providers absolute performance or relative improvement. There are, however, several concerns in using relative ranking as a basis of the reward in pay-for-performance programmes. Although it can encourage a greater effort among high performers for fear of losing a top-ranking position to be taken over, it may not incentivize low performance that are often in the greatest need for additional resources to improve performance. This could further increase inequalities in care for patients (5).

Of the three methods, the most commonly used is the absolute level, as illustrated in Table 9. It should be noted, however, that the ‘change over time’ method (that which is mostly used in Estonia’s QBS) is also common. In some schemes, incentives are paid on the basis of reaching a minimum performance threshold, with a sliding scale of additional payments up to a maximum value. In New Zealand’s Primary Health Organisation Performance Programme, for example, providers receive the full incentive payment if the targets are reached, or partial payments if the target is not reached but achievement was achieved. In the California IHA programme, doctor groups are scored on both attainment and improvement for each measure with the higher of the two used for each measure summed to the domain total, which is then weighted (5).

Table 9: Basis of rewards for primary care pay for performance schemes (2012)

Participation of health providers	Country	Performance measurement		
		Absolute	Change over time	Relative ranking
Mandatory participation	Estonia	X	X	
	Chile	X		
	Korea		X	X
	Spain	X	X	
	Sweden	X	X	X
	Turkey	X		
Voluntary participation with conditions	Australia	X		
	France		X	
	Mexico			X
	New Zealand	X	X	
	Portugal	X		
	United States	X	X	X
Voluntary open participation	Czech Republic			X
	Hungary	X		
	United Kingdom	X		

Source: (7)

There is little opportunity for doctors to exempt patients from certain indicators

It is important to allow Family Doctors to exempt particularly complex patients from a given QBS indicator. Otherwise, the lack of appropriate exemption criteria could result in unfairly low QBS scores for some Family Doctors with an unusually difficult case-load. Rules to exempt patients do exist in QBS but are currently rather limited, and focussed on Domain I. Patients can be exempted if: 1) there are contraindications to the management incentivised by QBS; 2) they are children who do not live in Estonia but have not yet been removed from population registry (this reportedly happens fairly often); 3) parents refuse their children to be vaccinated, or if they have a medical condition that does not allow vaccination.

In contrast, other primary care performance schemes allow exception reporting in a wider range of circumstances. The United Kingdom's QOF scheme, as illustrated for example in Box 1 for example, has a much wider range of exemption criteria that reflect the complexities of "real life" primary care criteria (23). These recognise that patients may refuse, be unresponsive to, or too frail for recommended treatment. The QOF's exemption criteria, developed in partnership with Family Doctors, are reportedly not overused or exploited. EHIF and ESFD may consider revising QBS exemption criteria such that they better reflect the complexities of chronic disease management in primary care. This may be particularly relevant for Family Doctors working in the southern counties of Estonia. These doctors do less well on QBS and have patient populations that are older and more rural, as explained in Section 4.

Box 4: Exception reporting in the United Kingdom's Quality and Outcomes Framework

The reasons for which Family Doctors in the United Kingdom's primary care P4P scheme can exclude patients include the following:

- Patients refusing to attend review who have been invited on at least three occasions during the preceding 12 months.
- Patients for whom it is not appropriate to review, e.g. terminal illness, extreme frailty.
- Patients newly diagnosed or who have recently registered
- Patients who are on maximum tolerated doses of medication whose levels remain sub-optimal.
- Patients for whom prescribing a medication is not clinically appropriate e.g. those who have an allergy, contra-indication or have experienced an adverse reaction
- Where a patient does not agree to investigation or treatment (informed dissent) and this has been recorded in their patient record following a discussion with the patient.
- Where an investigative service or secondary care service is unavailable.

6. Family Doctors' current performance in QBS

Nearly all Family Doctors in Estonia now participate in QBS. One in three, however, obtain low scores. Doctors generally perform well in the preventive care domain, but performance in the domain linked to care for chronic conditions is highly variable - with poor performance typically concentrated in Estonia's southern counties. Performance in the domain of additional professional competencies is uniformly low.

QBS has been adopted, in principle, by all Family Doctors

As described earlier, QBS was introduced in 2006 and the number of participating Family Doctors has risen from 50% in 2006 to 97% in 2014. The rapid growth in participation in the early years of the scheme points to an increasingly, broad acceptance over time of the pay-for-performance scheme, and the intrinsic motivation from Family Doctors to provide high quality care. Widespread participation is also important to avoid worsening socioeconomic inequalities in health.

Table 10: Participation in QBS before 2015

Year	Family Doctors	Participators (%)
2006	793	63
2007	800	57
2008	802	80
2009	802	85
2010	801	90
2011	801	94
2012	800	97
2013	798	96
2014	799	97

In 2015, participation became compulsory, with no possibility to opt-out of having the performance results calculated. Although the international literature is largely in favour of voluntary pay-for-performance schemes to prevent the risk of impairing provider's intrinsic motivation, it is acknowledged that efforts may be needed to yield high participation rates, such as making a scheme compulsory (9). Nevertheless, it is important that compulsory schemes keep providers motivated. One of the methods to do so is to keep providers closely involved in the planning, updating and revisions of the scheme to align indicators with their professional concerns, priorities and values. Studies have shown that providers are more comfortable with incentive programs when they are involved in planning and design (24). It is therefore a positive feature that the Estonian Family Doctors Association have always been closely involved in the design and operation of QBS.

One in three Family Doctors perform poorly on QBS

Looking first at trends in the number of doctors attaining the threshold for QBS payment (that is, at least 80% of total maximum points achieved), around 30% failed to meet this threshold in 2017, with more than

four-fold variation in attainment of the threshold across the country. Trends show steady, if gradual, improvement (Figure 7), nevertheless, many primary care doctors still struggle to respond to the incentives contained within QBS, particularly in southern Estonia (Figure 8).

Figure 7: Percentage of Family Doctors achieving at least 80% of maximum points, 2015–2017.

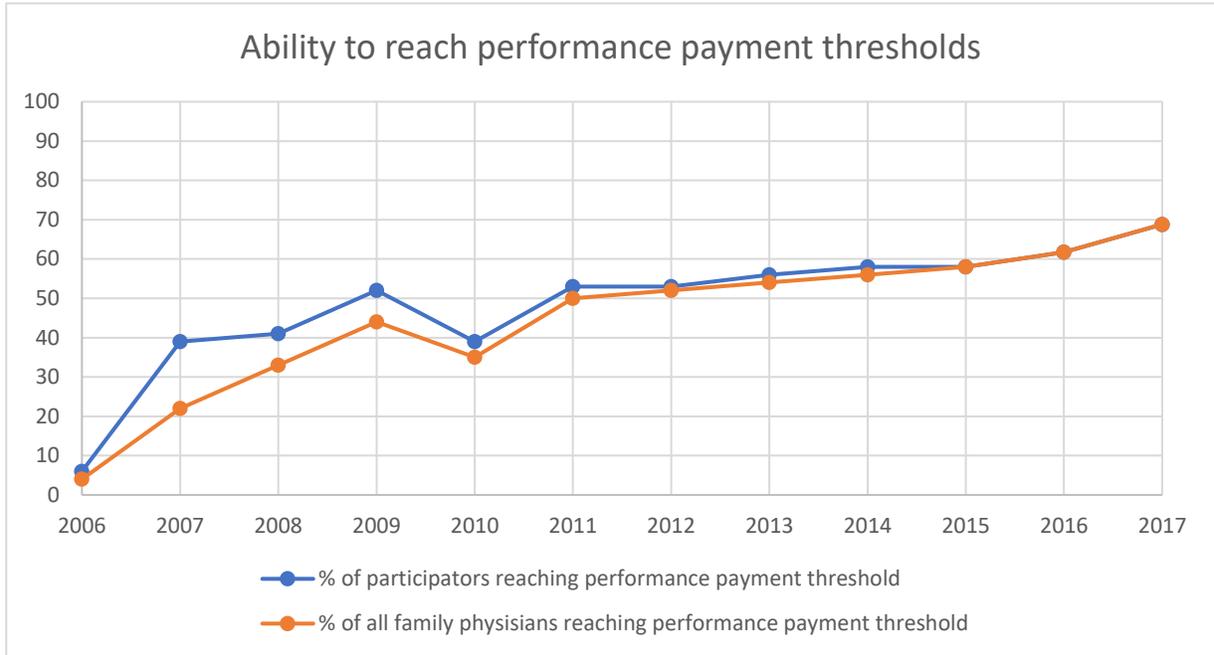
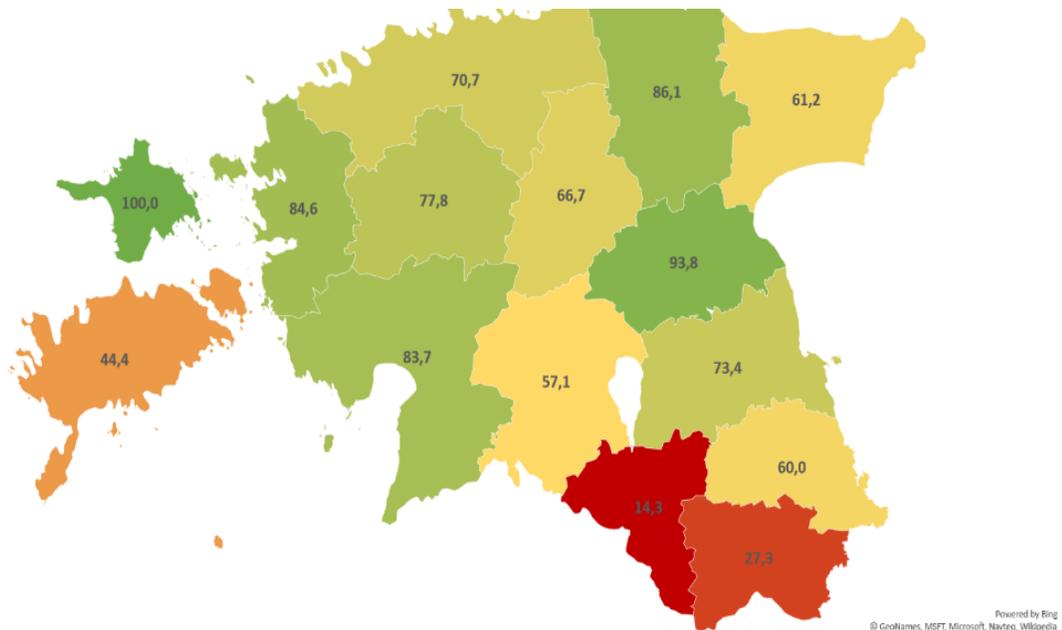


Figure 8: Percentage of Family Doctors achieving at least 80% of maximum points by county, 2017.



Looking next at trends in the number of doctors attaining the threshold for QBS payment within each domain, improvement is many, but not all, indicators. Within domain I, more than 80% of the doctors have reached all preventive care targets. Of some concern, however, attainment is falling for vaccinations and check-ups at age 3. Within domain II, QBS scores have improved for the medication management of low and moderate risk hypertensive patients, care of patients with an acute myocardial infarction and patients with hypothyroidism. Attainment is falling, however, for care of patients with diabetes, and the monitoring of patients with hypertension. Within domain III, very few doctors reach the targets for additional professional competencies, but scores are improving slowly in this domain (Table 11). Regional variation within each domain is discussed later in this section.

Table 11: National trends in goal achievement by indicator, 2014-2017

		National average (% of FPs reaching target coverage)				
Indicator		2014	2015	2016	2017	Difference
Prevention	Vaccinations of children up to 3 years old	91,6	85,4	86,8	88	-3.6
	Children's examination and general medical examination of children up to 3 years of age	89,6	91,2	82,6	85,2	-4.4
	Pre-school child examination and health check at 6, 7, 8 years of age	86,4	87	89,7	91,1	+4.7
Chronic disease management	Diabetes type II	74,7	78,4	66,2	72,6	-2.1
	diabetes medication	NA	NA	65	64,7	-0.3
	Hypertension I (low risk)	76,3	79,9	66,6	69,8	-6.5
	Hypertension II (moderate risk)	67,5	72,2	56,3	63,2	-4.3
	Hypertension III (high risk)	71	75,6	62,6	71,9	+0.9
	Hypertension medication 1	93,9	95,3	96,6	98,6	+4.7
	Hypertension medication 2	79,6	80,7	82,1	81,9	+2.3
	MI	77,7	82,5	79,4	81,6	+3.9
	MI medication 1	NA	NA	64,9	66,1	+1.2
	MI medication 2	NA	NA	58,1	59,9	+1.8
Hypothyroidism	78,8	82	84,1	85,3	+6.5	
Additional professional competence	Recertification (for FP and nurse)	16,6	22,4	22,2	20,3	+3.7
	Pregnancy monitoring	1,3	1,6	2,2	3,0	+1.7
	Gynaecological examinations	2,3	2,7	2,9	4,3	+2
	Small surgery	15,5	18,9	20,1	23,2	+7.7
		*Estimated averages				

Looking finally at trends in the actual QBS scores achieved, a striking bimodal distribution is apparent. Substantial numbers of doctors achieve high scores, but a peak is also seen at low scores (Figure 9). QBS is characterised, therefore, by two extremes in performance: Family Doctors meeting targets for most indicators and Family Doctors with almost no success in meeting targets. This metric is also characterised by large regional variation. There is more than a 5-fold difference in total QBS scores across Estonia, with southern countries scoring least (Figure 10).

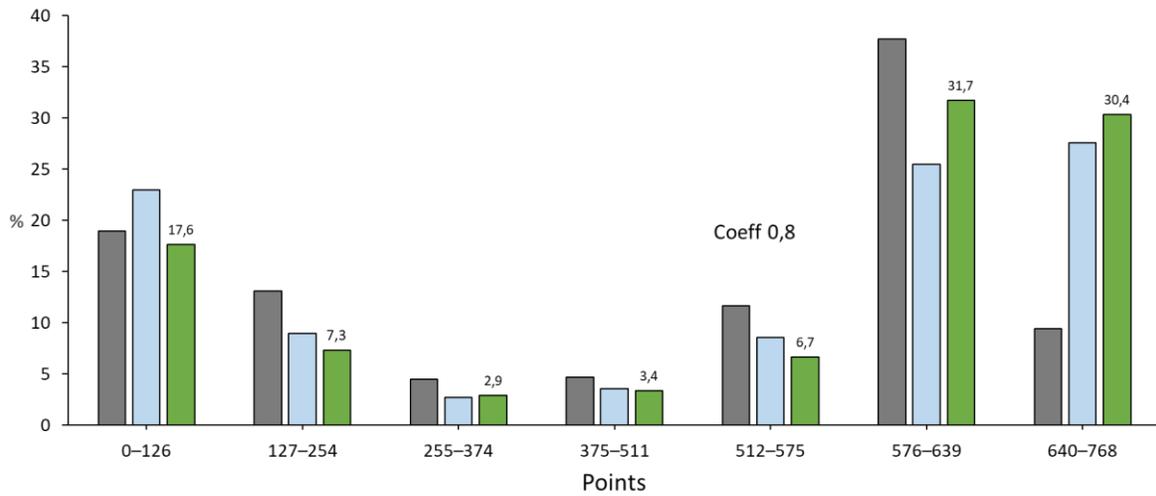


Figure 9: National distribution of QBS points, 2015–2017.

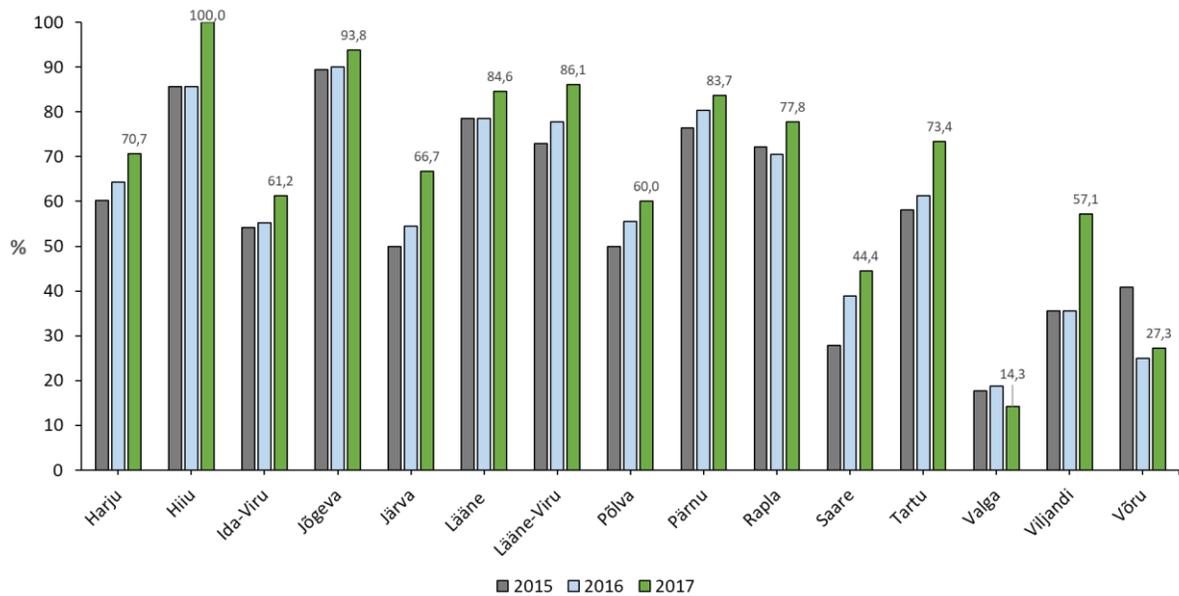


Figure 10: Average QBS points achieved by county, 2015-2017

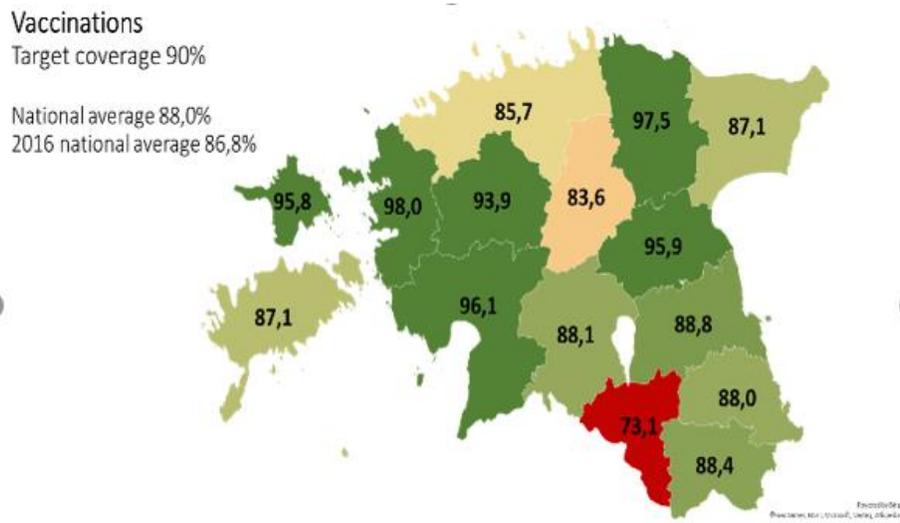
This pronounced regional variation is a reason to investigate persisting low achievement, particularly in southern countries. This would require practice visits and interviews with clinicians, managers and patients, in both Family Medicine as well as other health and social care sectors, to identify to causes of low achievement. One reason, for example, may be that QBS does not reflect the local health care needs

of the southern counties as well as northern neighbours. Estonia’s southern regions are more rural, with older populations and, presumably, a higher prevalence of chronic disease and multimorbidity. Currently, QBS does not allow for any flexibility or local variation, limiting opportunities to align indicators with local health care needs, and therefore their relevance for providers and patients. In addition, as explained earlier, the criteria by which patients can be exempted from QBS are currently rather limited, which may unfairly penalise doctors whose patients have more complex health care needs.

Preventive care (domain I): good performance with little regional variation

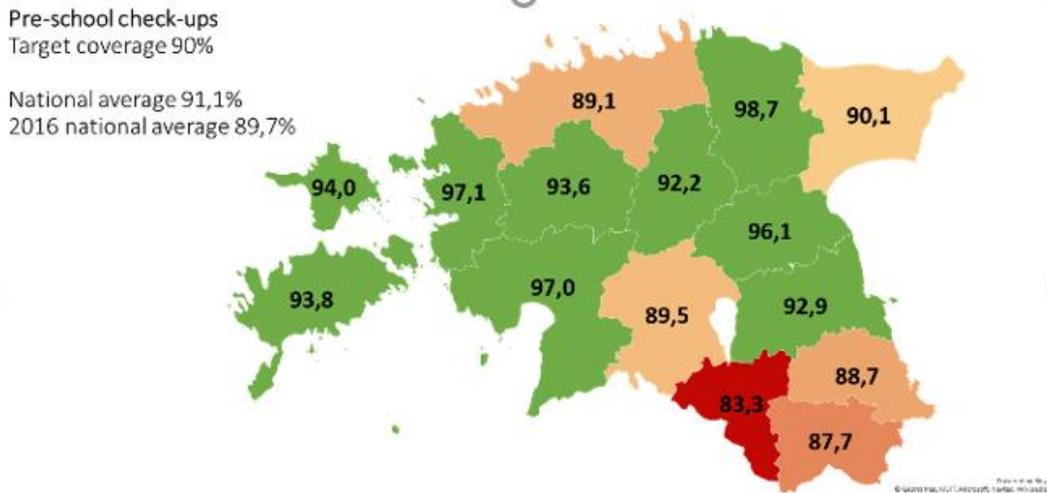
In 2017, most doctors were able to achieve QBS targets for childhood vaccinations, examination of children aged 0 to 2 years, and pre-school examinations for children aged 6-8 years, with relatively little regional variation compared to other QBS domains. Nevertheless, southern counties, particularly Valga, perform consistently worse in this domain.

Figure 11: Attainment of QBS threshold for childhood vaccination by county, 2017



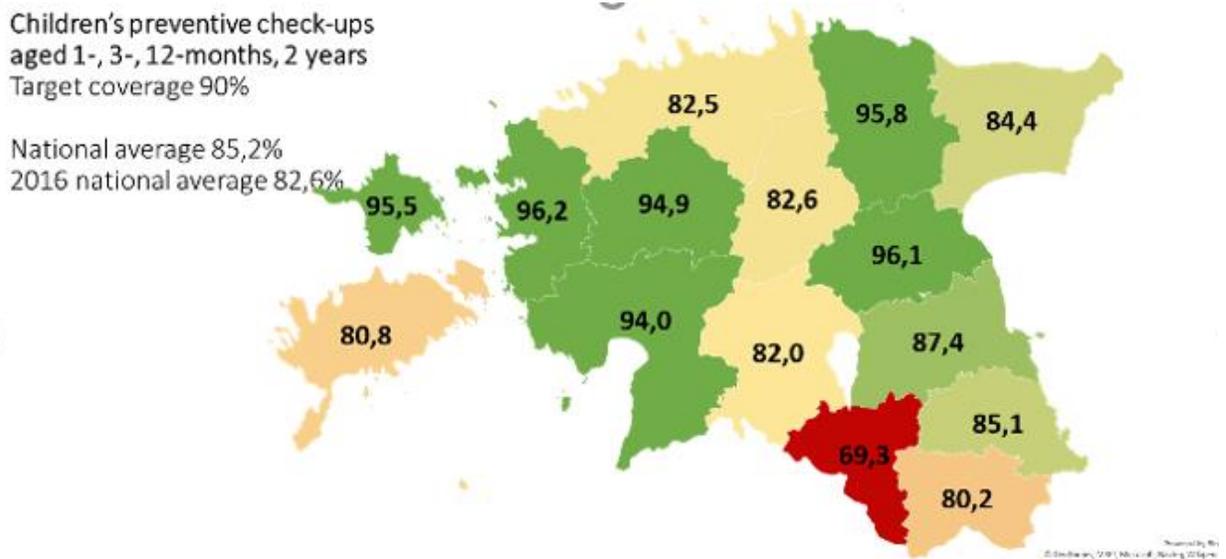
Source: EHIF data

Figure 12: Attainment of QBS threshold for pre-school check-ups by county, 2017



Source: EHIF data

Figure 13: Attainment of QBS threshold for children’s preventive check-ups by county, 2017



Source: EHIF data

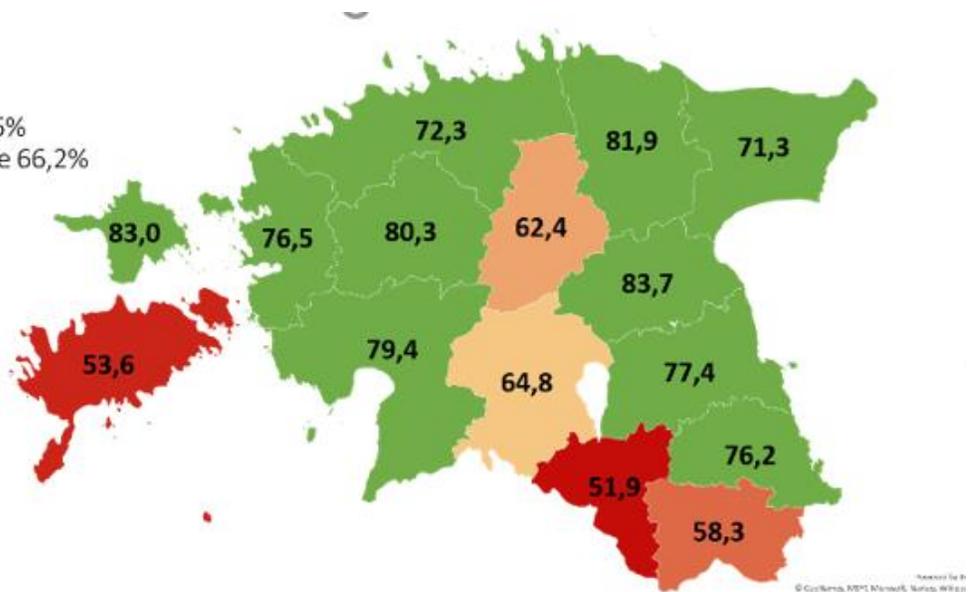
Chronic disease management (Domain 2): moderately good performance with substantial regional variation

Regarding care for people with Type II diabetes, substantial regional variation is seen in the share of doctors reaching the threshold for the package of care specified in QBS. In three counties (Võru, Valga and Saare) only around half of doctors reached the threshold (Figure 14). Regarding the share of doctors adequately prescribing metformin, a different regional pattern is seen. Võru and Valga in the south perform relatively well, but doctors in Saare and Ida-Viru perform relatively poorly (Figure 15).

Figure 14: Attainment of QBS threshold for diabetes care by county, 2017

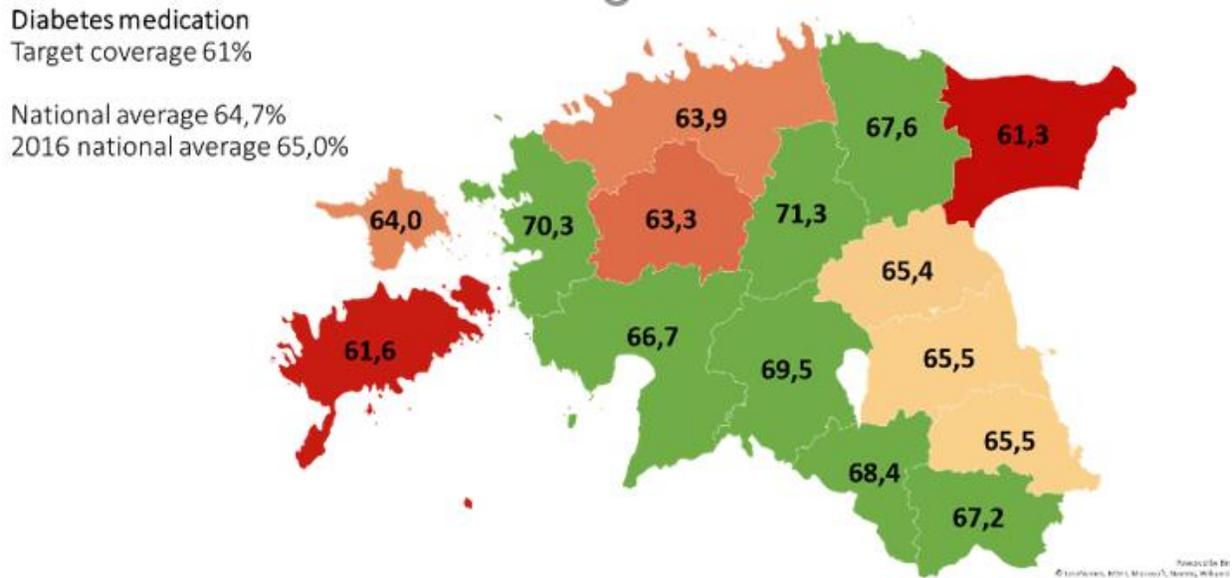
Diabetes, type II
Target coverage 70%

National average 72,6%
2016 national average 66,2%



Source: EHIF data

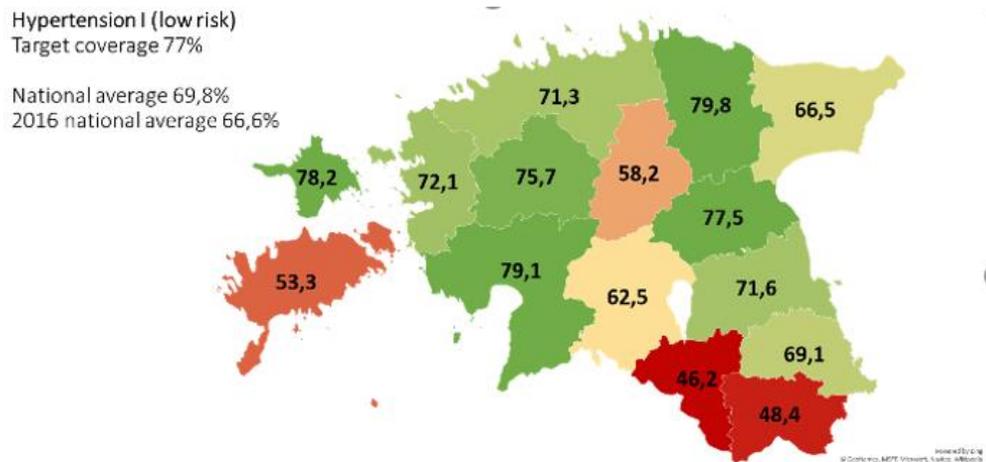
Figure 15: Attainment of QBS threshold for diabetes medication by county, 2017



Source: EHIF data

Regarding care for people with hypertension, substantial regional variation is also seen in the share of doctors reaching the threshold for the package of care specified in QBS, regardless of the risk group of patients. Southern counties and Saare county perform worse. Regarding the share of doctors adequately prescribing appropriate antihypertensives, a different regional pattern is seen. As with prescribing in diabetes, doctors in Ida-Viru county perform worst.

Figure 16: Attainment of QBS threshold for hypertension (low risk) care by county, 2017



Source: EHIF data

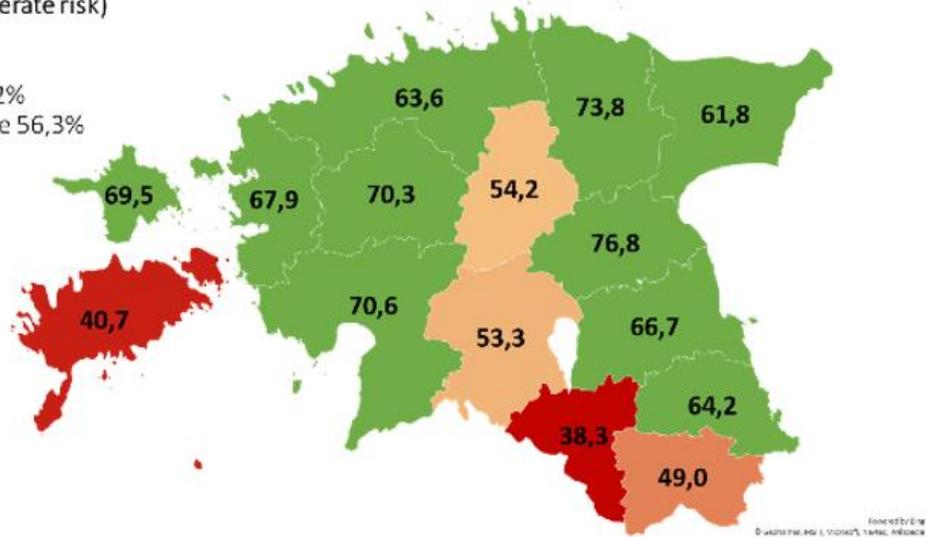
Figure 17: Attainment of QBS threshold for hypertension (moderate risk) care by county, 2017

Hypertension II (moderate risk)

Target coverage 58%

National average 63,2%

2016 national average 56,3%



Source: EHIF data

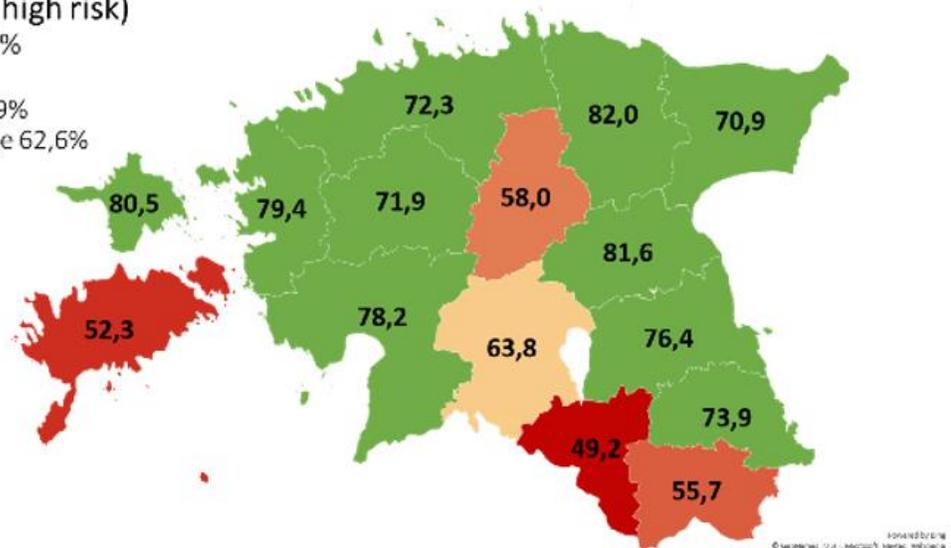
Figure 18: Attainment of QBS threshold for hypertension (high risk) care by county, 2017

Hypertension III (high risk)

Target coverage 67%

National average 71,9%

2016 national average 62,6%

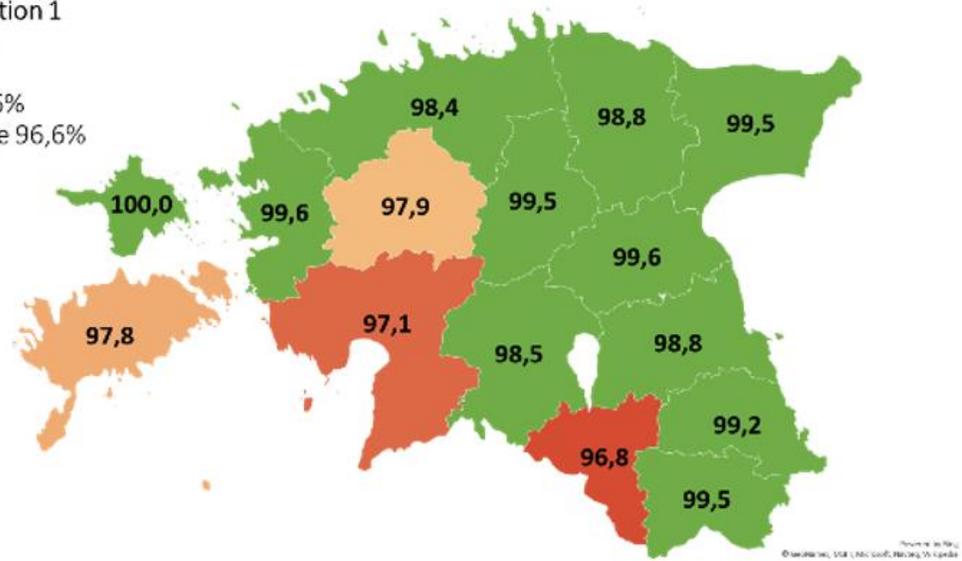


Source: EHIF data

Figure 19: Attainment of QBS threshold for hypertension medication (1) by county, 2017

Hypertension medication 1
Target coverage 90%

National average 98,6%
2016 national average 96,6%

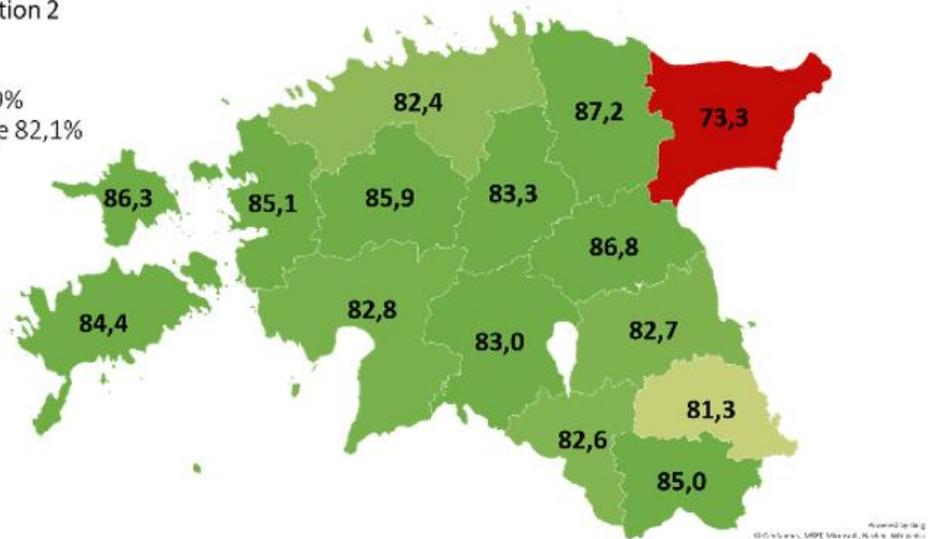


Source: EHIF data

Figure 20: Attainment of QBS threshold for hypertension medication (2) by county, 2017

Hypertension medication 2
Target coverage 83%

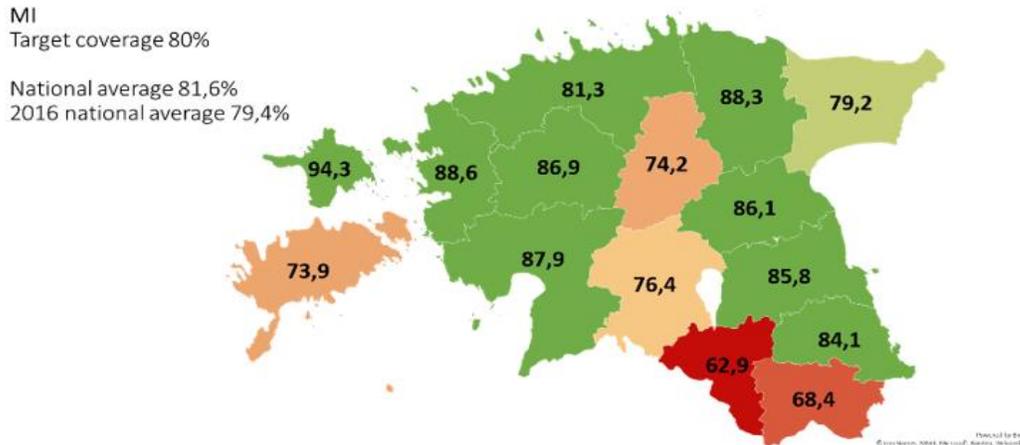
National average 81,9%
2016 national average 82,1%



Source: EHIF data

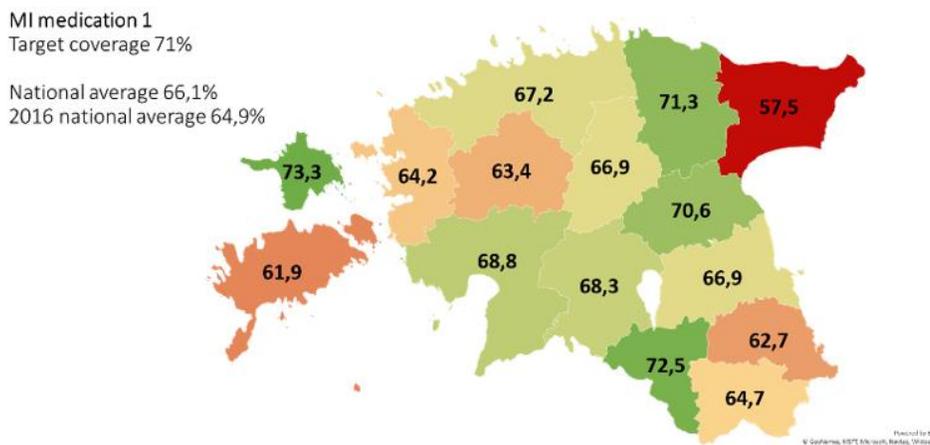
Regarding care for people who have suffered a heart attack, substantial regional variation is also seen in the share of doctors reaching the threshold for the package of care specified in QBS, with southern counties performing worse. Regarding the share of doctors adequately prescribing beta-blockers (*MI medication I* indicator), the same pattern as seen for prescribing in diabetes is apparent, where doctors in Ida-Viru county perform worst. Substantial regional variation is seen in the share of doctors adequately prescribing statins (*MI medication II* indicator), with doctors in many counties performing badly, but worse in Valga and Ida-Viru counties, pointing to ample room for improvement.

Figure 21: Attainment of QBS threshold for MI care by county, 2017



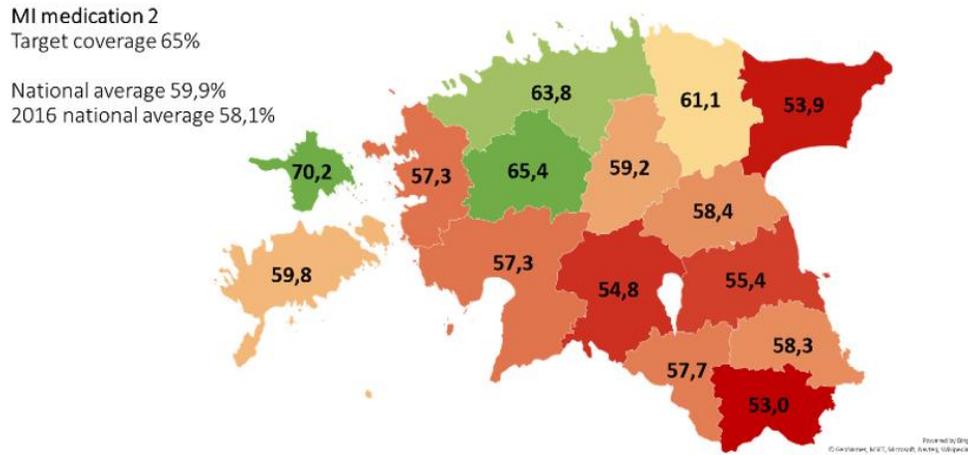
Source: EHIF data

Figure 22: Attainment of QBS threshold for MI medication (1) by county, 2017



Source: EHIF data

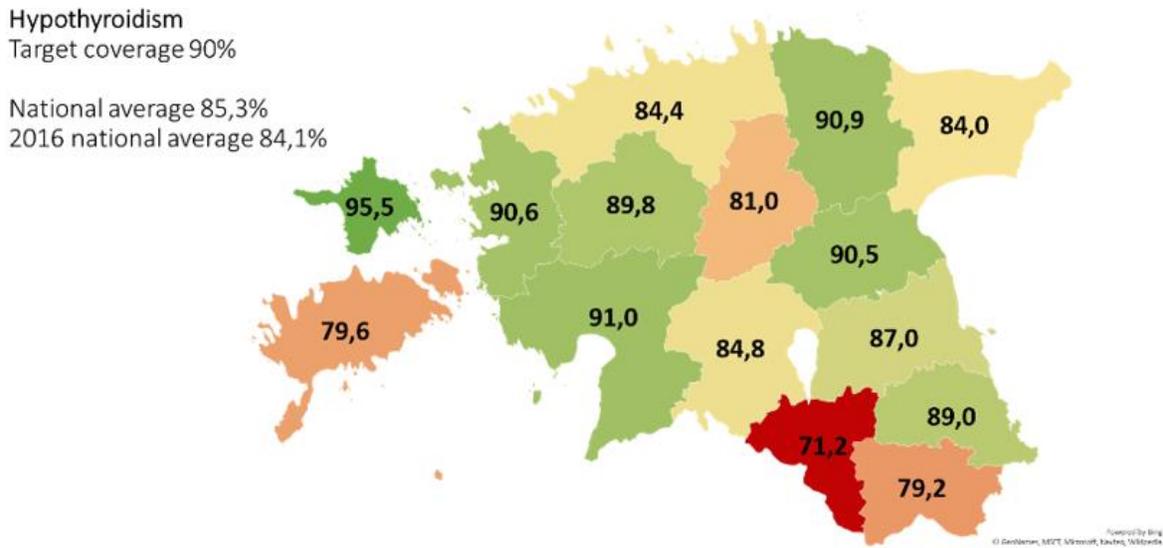
Figure 23: Attainment of QBS threshold for MI medication (2) by county, 2017



Source: EHIF data

Regarding care for people with hypothyroidism, 85% of the doctors reached the target for determining TSH level. Regional variation is relatively small, although southern counties again perform worse.

Figure 24: Attainment of QBS threshold for hypothyroidism monitoring by county, 2017

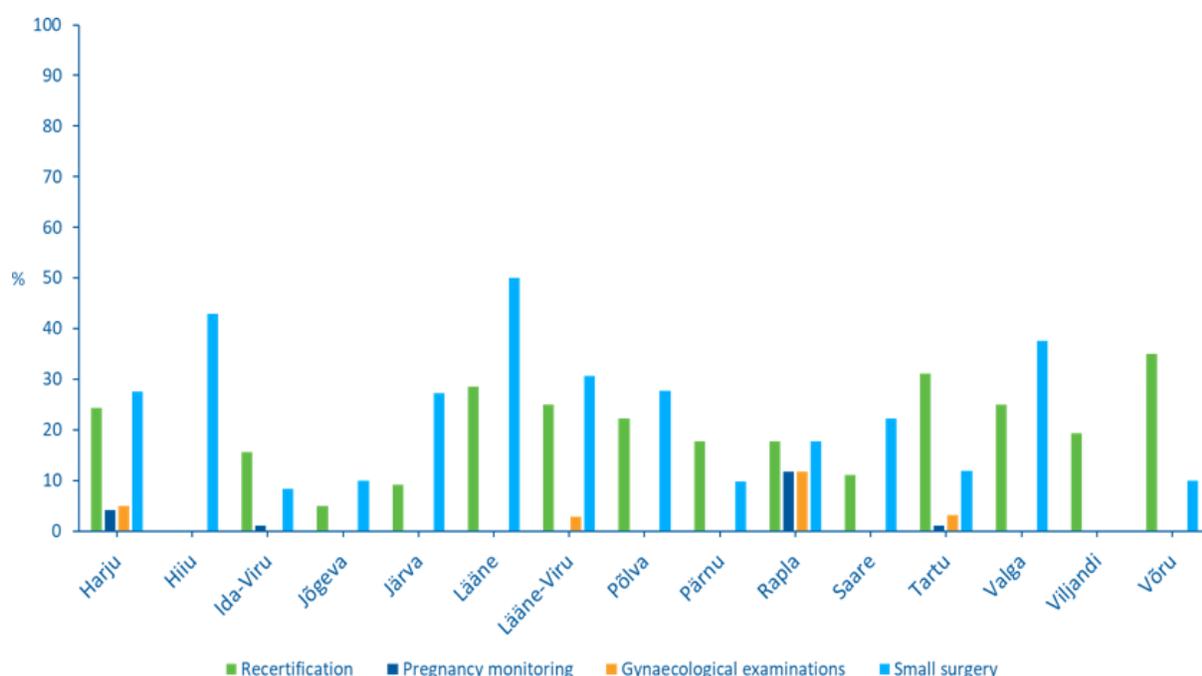


Source: EHIF data

Additional professional competence (Domain 3): uniformly poor performance

The below figure shows exceptional low performance levels in all counties for the additional professional competences in 2016. At best, 30% of doctors and nurses (both) have been recertified in Võru, while in other counties this was less or not the case at all. Only in four counties a few doctors (less than 10%) met the target coverage for pregnancy monitoring: i.e. providing pregnancy monitoring services at least 8 times a year. A similar result was found for gynaecological examination services, for which the target coverage was to provide such services at least 10 times a year. A more positive result was shown for small surgery, though with large regional variation. In Lääne half of the doctors provide small surgery services at least 40 times year. This was lower in the other counties. It would be interesting to get a better understanding why so few doctors meet the target coverage. Perhaps they do not agree that the indicators or their targeted frequency are part of their task profile or matches with the needs of their patient population. This would require practice visits, and discussions with clinical and managerial teams.

Figure 25: Attainment of QBS thresholds in Domain 3 (additional professional competence)



Source: EHIF data

Size of the patient population may influence QBS performance

The size of a Family Doctor's patient list on QBS attainment has been identified as a possible determinant of performance. In theory, a small list (with few patients with diabetes, for example) may support better performance, if a doctor is able to devote more time to her individual patients. Conversely, a small list may mean that it is harder to build up expertise or achieve economies of scale in delivering high quality

care. In other countries, it has been seen that larger practices have more resources available to achieve quality improvements than smaller practices.

Table 12 is a preliminary, unadjusted analysis showing the distribution of QBS results by average number of patients with diabetes, hypothyroidism, hypertension or MI in the patient list, 2015–2017. In 2015, with most of the conditions, poorer performing Family Doctors (that is, those with 0–79% score, or coefficient 0) had fewer patients with diabetes, hypertension (low and high risk), MI and hypothyroidism. Higher-performing doctors (that is, those with 100% score, or coefficient 1,0) had on average higher numbers of patients with those conditions in their list. In 2017, however, it seems that the opposite is happening. Family Doctors have indicated that a good QBS result is more difficult to achieve for those who have more patients in target group.

Table 12: Distribution of QBS results (coefficient) by average number of patients with diabetes, hypothyroidism, hypertension or MI in the patient list, 2015–2017.

Year	Average number of patients in list								
	2015			2016			2017		
Coefficient	0	0.8	1	0	0.8	1	0	0.8	1
Diabetes	82.1	69.5	86.4	85.4	83.2	82.3	91.0	85.8	83.3
Hypertension 1 (low risk)	95.4	118.6	131.0	94.8	111.8	132.4	104.3	125.9	133.0
Hypertension 1 (moderate risk)	247.4	178.6	216.8	254.6	224.9	201.4	258.8	220.2	203.4
Hypertension 1 (high risk)	55.0	47.7	58.4	51.8	47.6	51.6	57.5	48.6	52.9
MI	23.6	22.4	27.9	24.4	22.9	26.0	27.7	25.2	26.2
Hypothyroidism	50.3	46.3	58.2	56.6	58.5	60.0	59.2	61.0	60.4

This reversal of patterns between 2015 and 2017 is surprising. It could be due to the preliminary nature of the data, and may be attenuated or disappear once final data are available. Nevertheless, the findings provide another angle of enquiry into the determinants of QBS performance. Again, a mix of quantitative (e.g. differentiating results by total list size, location) and qualitative (e.g. by interviewing doctors) methods would be appropriate to understand the relationship between practice size and QBS attainment.

7. Recommendations for modernizing QBS

Ten recommendations are made for modernizing QBS in light of Estonia's priority health care needs; international experience and best practice in using pay-for-performance in primary care; and, concurrent reforms to Estonia's primary care financing and service delivery model.

Of the ten, recommendations 1, 2, 3, 7 and 8 should receive the highest priority. Taking these, as well as the remaining recommendations forward, will require close cooperation between Estonian Health Insurance Fund and the Estonian Society of Family Doctors.

What is incentivized

1. On-going revision of QBS indicators should continue, dropping those which have high achievement or low disease burden

The most important criterion for designing QBS indicators should be that they focus on common health care issues that cause significant morbidity or mortality, where the main responsibility for care is in family practice, and where there is evidence of health benefits arising from intervention. With this in mind, further revisions to QBS indicators would be beneficial in addition to those made in recent years, set out in Section 3. Specifically, the indicator linked to management of hypothyroidism should be dropped. It is understood that EHIF intend to do this in the next revision of the scheme.

There may also be scope, in the future, to drop indicators which Family Doctors are achieving easily and which no longer need to be incentivised. Childhood vaccination may be one such indicator, at some point, since many Family Doctors achieve high vaccination rates. It would be premature to drop it today, however, because some doctors (particularly in southern counties) still struggle to vaccinate all children appropriately. But if, in the future, doctors achieve uniformly high vaccination rates (demonstrated, for example, by all counties reaching target levels for three years in a row), consideration should be given to removing the indicator from QBS and offering other forms of targeted support to individual Family Doctors that do not achieve high vaccination rates.

More broadly, there may be scope to speed up and regularise the mechanisms by which QBS indicators are refreshed. Although a number of revisions occurred in the last year, they have generally happened infrequently. It will be important to keep providers closely involved to ensure the indicators reflect service priorities.

2. The dimensions of primary care performance captured by QBS should be expanded

QBS rewards predominantly process indicators, in contrast to other national P4P schemes in primary care which also include safety of care, equity, effectiveness, continuity, accessibility, and patient experiences of care. A particularly notable omission from QBS concerns clinical outcomes. As explained in Section 3, other national P4P schemes in primary care include these indicators, and thereby capture the effectiveness of primary care. QBS, now more than a decade old, should now be mature enough to join

the set of national P4P schemes that include outcome indicators. It is recommended, therefore, that EHIF and ESFD work together to pilot the introduction of some patient outcome indicators, guided by international experience. It is understood that EHIF will move forward in this area by using some of the clinical outcomes collected through the Enhanced Care Management programme. Over the longer term, reliable systems to routinely measure and record clinical outcomes will need to be established to allow QBS to embrace a wider set of patient outcome indicators.

The concern that clinical outcomes are determined by many factors, some of which cannot be influenced by primary care, should be addressed. Notably, the outcomes referred to are intermediate outcomes such as controlled blood pressure, blood sugar and cholesterol in people with diabetes and cardiovascular disease, rather than more final outcomes such as quality of life or survival. It is reasonable to hold primary care to account for these intermediate outcomes (if not, then who should be accountable?). Including outcomes will necessitate extension of the criteria that permit patients to be exempted, to make fair allowance for Family Doctors with unusually complex patients. This is discussed further below.

Primary care schemes in other countries offer substantial international experience for EHIF and ESFD to draw from, when considering how clinical outcomes could be included in QBS. Another source would be standardized set of outcome measures that are being developed by the International Consortium for Health Outcomes Measurement. This group of experts is working to identify outcomes for specific conditions covering about 45% of the disease burden in high-income countries.

In addition to outcomes, consideration should also be given to including patients' experience in QBS. These are novel indicators, that an increasing number of countries are experimenting with. In addition, indicators that reward the coordination of care between care settings should be considered. QBS has already made a step in this direction by including e-consultations as a pilot indicator; and this innovation should be continued. It is recognized that including new indicator types (whether clinical outcomes or patient experiences) will require new methods to record and extract primary care data, beyond the routine reimbursement data currently used.

3. Local elements should be developed

There is no facility within QBS to set local targets or facility-level targets, meaning that local health care priorities or health care system challenges cannot be addressed. It is therefore recommended to develop local variants of QBS to reflect local health care priorities. This would help build interest and strengthen commitment and may also help moderate the wide regional variation currently seen in QBS scores. Local variants may mean different thresholds (for example based on local disease prevalence) or new indicators for distinct patient groups. This could be further expanded by allowing Family Doctors themselves to select from a menu of QBS indicators, each year, to reflect their interests or local health care priorities. Again, primary care schemes in other countries offer substantial international experience, particularly in the United Kingdom, for EHIF and ESFD to draw from in designing local variants to the national scheme. It is understood that EHIF intend to do this in the next revision of the scheme.

4. QBS should reward both improvement and absolute level target achievement

QBS has an unusually complex mechanism for determining performance, awarding points for relative improvements against baseline. This stands in contrast to many other national P4P schemes in primary care, which simply reward attainment of an absolute threshold. The QBS design has the strong advantage of favouring low performers, but may disincentivise good performers, for whom it may be hard to achieve additional gains from a high baseline. Furthermore, the current design does not establish an optimal (or minimally acceptable) level of care.

One solution would be to establish a blended incentive system, that i) offers a reward for attaining the threshold that represents optimal care; and ii) offers a reward for doctors that fail to reach this threshold but have shown progress toward it. The incentive under scenario i) would be higher than that under scenario ii). EHIF could follow the example of New Zealand in the Primary Health Organisation Performance Programme, providers receive the full incentive payment if the targets are reached, or partial payments if the target is not reached but progress was achieved. The disadvantage of such a blended incentive scheme, however, is that it would further complicate an already complex incentive scheme.

Who is incentivized

5. Consider wider application of practice-level incentives, to reflect team-base care

With the exception of the QMBS, QBS incentivizes individual Family Doctors, as opposed to the wider primary care team. This may not reflect modern primary care practice, however, which increasingly depends upon a team effort, particularly in the management of patients with complex, chronic conditions. QBS should be developed as a lever, therefore, to encourage team-based care.

This could be pursued in three ways. First, wider participation in QMBS should be encouraged. This will require preliminary work to understand why current participation rates are so low. Second, advice and support should be given to Family Doctors on how to invest QBS payments in tools or activities that support team-based care, such as shared continuing professional development (CPD). Finally, increasing the proportion of QBS incentives paid to the practice could also be explored. This would be most feasible as new indicators are added to QBS.

The new program of Enhanced Care Management (ECM) also offers a route to better supporting team-based primary care. ECM relies on nurse-led care and close collaboration between the primary doctor and nurse in the management of patients with complex, chronic disease. The reimbursement mechanism for ECM has not yet been finalised, but it will inevitably interact substantially with QBS.

6. Use QBS to encourage group practice

In a similar vein, there may also be scope to offer enhanced QBS benefits to group practices. Domain 3, for example, could be modified to emphasise activities typical of group practices, such as peer review and learning. This would need to link to broader changes in the primary care reimbursement system. Within

FFS, for example, the range of services included in the diagnostic, therapeutic and procedure funds should be expanded, and their respective caps raised to allow group practices to extend their scope of care. Capitation and allowances may also need to be adjusted to reflect the services, case mix, and cost structures of group practices.

How is performance incentivized

7. Expand the criteria by which patients can be exempted from inclusion in QBS

As mentioned earlier, an essential step toward including clinical outcomes within QBS will be to expand the list of criteria by which patients can be exempted. This is needed as a form of case-mix adjustment, so as not to unfairly penalise Family Doctors with particularly complex patients. A wider set of exemption criteria would also be helpful for the new, composite indicators, whereby a package of care has to be delivered for it to be recognized within QBS. The exemption criteria used by QOF in the United Kingdom could be considered as a template to adapt for application in Estonia.

A complementary means of adjusting to case-mix differences would be to relax the ‘all or nothing’ for composite indicators and allow the possibility to break down them into separate elements. Composite indicators were introduced very recently, and with a strong rationale – namely, that primary care should offer coherent packages of care rather than lots of individual services. Nevertheless, some Family Doctors have reported that they are too difficult to attain. More time is needed to see how composite indicators play out in practice, but if it proves that they are too complex to achieve, and do not encourage quality improvement, revision should be considered.

8. Strengthen the incentives within QBS, both financial and non-financial

There are several means by which the incentive-effect within QBS, both financial and non-financial, could be strengthened. On the financial side, the most obvious way would be to pay more. The monetary impact under the current design of QBS on Family Doctors’ income is relatively small, and the difference in additional fee-for-services payments between high and low performing providers is also relatively small. Noting the caveats surrounding the evidence on P4P in primary care discussed earlier, it is nevertheless worth considering whether further quality improvement gains could be achieved if a larger financial reward were available, as per the “pay enough or not at all” logic. Resources are constrained in Estonia as in all countries. It is unlikely, therefore, that additional funds would be available for QBS; instead, reallocation of funding from capitation, FFS elements or allowances within the current resource envelope would have to be negotiated.

On the non-financial side, the QBS’s incentive effect could be improved by strengthening intrinsic motivation. Motivation to deliver high quality care is complex, and many factors beyond financial rewards will drive Family Doctors’ performance. It is important, therefore, to optimise other ways of motivating health professionals to deliver high quality care. There is scope, for example, to improve the feedback and support that Family Doctors receive regarding their performance on individual indicators, benchmarked against peers at county and national level. Including such benchmarks in the electronic personal feedback of QBS results, as well as support to use the data in plan-do-study-act (PDSA) cycles of continuous quality

improvement would be a powerful tool. The Institute for Healthcare Improvement (IHI) in Boston has collected multiple examples of national and local health systems applying rapid PDSA cycles in practice.

Patients should also be empowered to use QBS data more effectively. EHIF already publishes QBS results on its website, but it appears that this information is very little used by the public. Ways to increase the public accessibility, impact and relevance of QBS data should be explored therefore, especially for defined patient groups (such as those with diabetes, cardiovascular disease and other chronic conditions). National patient associations should be closely involved in these discussions, to ensure that any resulting publications or digital platforms align with their expectations and needs.

9. Shorten the feedback loop by which Family Doctors receive results

There are technical constraints which determine the speed by which Family Doctors receive their QBS scores and local/national comparisons. Nevertheless, opportunities to shorten the feedback loop should be explored, in order to optimise the relevance, impact and actionability of the data. Ideally, Family Doctors would be able to see real-time performance comparisons (as happens in the United Kingdom's QOF); failing this, quarterly report would still be an improvement on the current situation. Public reports on an annual basis, as currently happens, are reasonable.

10. Ensure that QBS articulates effectively with other primary care quality monitoring and improvement activities

The final recommendation, perhaps the most important, is to ensure that QBS articulates effectively with other activities to monitor and improve primary care quality and is part of an overall strategy to strengthen the sector. Both EHIF and ESFD have prioritised the quality of care over many years, developing several ambitious initiatives to do so. There is a risk, though, that different activities do not synergise as effectively as they could. The disconnect between QBS and national clinical guidelines has already been discussed; likewise, the audit visits that form part of QMBS have not been effective in identifying lessons on what determines good or substandard performance in QBS more generally.

Thought should be given, therefore, to ensure that QBS should form part of a cohesive quality monitoring and improvement strategy for primary care. EHIF and ESFD should ensure, for example, that guidelines are available for all diseases and conditions included in the QBS and that indicators and guidelines are regularly updated in line with international good practice. In particular, EHIF and ESFD should discuss how the new programme of Enhanced Care Management can best build upon and synergise with QBS to take Estonia's primary care onto the next level of performance.

8. Recommendations for further analysis and research

In addition to the reforms to QBS recommended above, further research and analytics are also needed to better understand primary care quality and how QBS may best support quality gains.

Three potential topics for further research and analysis are:

1) Reasons underlying persistent low QBS scores

Noting the striking bi-modal distribution of QBS scores presented earlier in the report, a better understanding is needed of why some Family Doctors consistently achieve low QBS scores, and why low scores are concentrated in specific regions. As well as quantitative analysis of the observable determinants of high/low scores, interviews with clinicians, managers and patients of in both high- and low-scoring practices will help elucidate explanatory factors. This should build upon the audit visits that are already undertaken as part of QMBS.

2) The effect of the size of patient list on QBS performance

As discussed earlier in the report, there is a theoretical risk that QBS scores may be a less reliable measure of performance in practices with small patient lists. The uncertainty or confidence intervals, given sample size, should be estimated for each indicator.

As a separate issue, preliminary data was also presented showing that list size may determine QBS scores. This effect should be confirmed, or otherwise, through more detailed quantitative analysis. If confirmed, this finding will add to the understanding of the determinants of high- or low scores.

3) Reasons underlying low participation in QMBS

A better understanding is needed of why Family Doctors almost universally achieve low QMBS scores. As in the first research topic, interviews with clinicians, managers and patients of in both high- and low-scoring practices will help elucidate explanatory factors. By interviewing both participating and non-participating doctors, information to improve QMBS design and implementation could emerge, thereby encouraging future participation.

References

1. Habicht T, Reinap M, Kasekamp K, Sikkut R, Laura Aaben L, Van Ginneken E. Estonia: Health system review. *Health Systems in Transition*. 2018;20(1):1-193.
2. OECD/European Observatory on Health Systems and Policies. Estonia: Country Health Profile 2017, State of Health in the EU. Paris/Brussels: OECD/European Observatory on Health Systems and Policies; 2017.
3. Kringos DS, Boerma W, van der Zee J, Groenewegen P. Europe's strong primary care systems are linked to better population health but also to higher health spending. *Health affairs (Project Hope)*. 2013;32(4):686-94.
4. Detollenaere J, Desmarest AS, Boeckxstaens P, Willems S. The link between income inequality and health in Europe, adding strength dimensions of primary care to the equation. *Social science & medicine (1982)*. 2018;201:103-10.
5. European Observatory on Health Systems and Policies. Paying for performance in health care. Implications for health system performance and accountability. Cashin C, Chi Y, Smith PC, Borowitz M, Thomson S, editors. London: Open University Press; 2014.
6. Forbes L, Marchand C, Peckham S. Review of the Quality and Outcomes Framework in England. London: PRUComm; 2016.
7. OECD. Better Ways to Pay for Health Care. Paris: OECD; 2016.
8. Eijkenaar F. Key issues in the design of pay for performance programs. *The European journal of health economics : HEPAC : health economics in prevention and care*. 2013;14(1):117-31.
9. Eijkenaar F. Pay for performance in health care: an international overview of initiatives. *Medical care research and review : MCRR*. 2012;69(3):251-76.
10. Berenson B, Pronovost P, Krumholz H. Achieving the potential of health care performance measures. Washington D.C.: Urban Institute; 2013.
11. Petersen LA, Simpson K, Pietz K, Urech TH, Hysong SJ, Profit J, et al. Effects of individual physician-level and practice-level financial incentives on hypertension care: a randomized trial. *Jama*. 2013;310(10):1042-50.
12. Eijkenaar F, van Vliet RC. Profiling individual physicians using administrative data from a single insurer: variance components, reliability, and implications for performance improvement efforts. *Medical care*. 2013;51(8):731-9.
13. Roland M, Dudley RA. How Financial and Reputational Incentives Can Be Used to Improve Medical Care. *Health services research*. 2015;50 Suppl 2:2090-115.
14. Struijs JN, Baan CA. Integrating care through bundled payments--lessons from The Netherlands. *The New England journal of medicine*. 2011;364(11):990-1.
15. Deci EL, Koestner R, Ryan RM. A meta-analytic review of experiments examining the effects of extrinsic rewards on intrinsic motivation. *Psychological bulletin*. 1999;125(6):627-68; discussion 92-700.
16. Frey B. On the relationship between intrinsic and extrinsic work motivation. *International Journal of Industrial Organization*. 1997;15(4):427-39.
17. Dudley RA, Rosenthal M. Pay for performance: a decision guide for purchasers. Washington, D.C.: Agency for Healthcare Research and Quality; 2006.
18. Damberg CL, Raube K, Teleki SS, Dela Cruz E. Taking stock of pay-for-performance: a candid assessment from the front lines. *Health affairs (Project Hope)*. 2009;28(2):517-25.
19. Van Herck P, De Smedt D, Annemans L, Remmen R, Rosenthal MB, Sermeus W. Systematic review: Effects, design choices, and context of pay-for-performance in health care. *BMC health services research*. 2010;10:247.
20. Frolich A, Talavera JA, Broadhead P, Dudley RA. A behavioral model of clinician responses to incentives to improve quality. *Health policy (Amsterdam, Netherlands)*. 2007;80(1):179-93.
21. Chung S, Palaniappan L, Wong E, Rubin H, Luft H. Does the frequency of pay-for-performance payment matter?--Experience from a randomized trial. *Health services research*. 2010;45(2):553-64.
22. Robinowitz DL, Dudley RA. Public reporting of provider performance: can its impact be made greater? *Annual review of public health*. 2006;27:517-36.
23. NHS Employers. Exception reporting Leeds: NHS Employers; 2018 [Available from: <https://www.nhsemployers.org/-/media/Employers/Documents/Primary-care-contracts/QOF/FAQS/Exception-reporting-FAQs.pdf?la=en&hash=B94AE91AD26B2B3FB4CA423549B934C79BF5E129>].

24. Begum R, Smith Ryan M, Winther CH, Wang JJ, Bardach NS, Parsons AH, et al. Small practices' experience with EHR, quality measurement, and incentives. *The American journal of managed care*. 2013;19(10 Spec No):eSP12-8.

Annex 1: Overview of QBS indicators, including changes

Domain I: Prevention

Vaccinations of children up to 3 years old

- In QBS since: 2006.
- Includes: vaccinations against whooping cough, diphtheria, tetanus, poliomyelitis, measles, mumps, rubella, hepatitis B, haemophilus influenza type b, rotavirus (since 2015).
- Other major changes: since 2015, instead of 9 separate indicators all vaccinations were combined so all vaccinations must be done to complete this indicator.

Children's examination and general medical examination of children up to 3 years of age

- In QBS since: 2006.
- Includes: 4 examinations in different age (1 month, 3 months, 12 months and 2 years).
- Other major changes: since 2016, in QBS as a combined indicator - all examinations must be done to complete this indicator.

Pre-school child examination and health check at 6, 7, 8 years of age

- In QBS since: 2006.
- Includes: pre-school examination of a 6, 7 or 8 years old child.
- Other major changes: no.

Part II Monitoring chronically ill patients

Diabetes (type II DM)

- In QBS since: 2007
- Includes: ECG (dropped in 2007), examination of the fundus of the eye (dropped in 2007), levels of albumin in urine (dropped in 2017), glycosylated haemoglobin, creatinine values, cholesterol values, cholesterol fraction values, appointment by family nurse, counselling for chronic patient (added in 2018).
- Other major changes: a combined indicator since 2016.

Diabetes medication

- In QBS since: 2016.
- Includes: metformin or combination of metformin prescribed for patients with diabetes II.
- Other major changes: no.

Hypertension I (low risk)

- In QBS since: 2007

- Includes: Glucose or glycosylated haemoglobin, cholesterol, counselling for chronic patient (added in 2018), appointment by family nurse, examination of the fundus of the eye (dropped in 2007), ECG (dropped in 2013).
- Other major changes: no, just general simplification (dropping some components).

Hypertension II (moderate risk)

- In QBS since: 2007
- Includes: cholesterol determined for patients under 80yo, cholesterol fractions determined for patients under 80yo, glucose or glycosylated haemoglobin, creatinine, ECG, counselling for chronic patient (added in 2018), appointment by family nurse (added in 2013).
- Other major changes: no.

Hypertension III (high risk)

- In QBS since: 2007
- Includes: cholesterol determined for patients under 80yo, cholesterol fractions determined for patients under 80yo, glucose or glycosylated haemoglobin, creatinine, counselling for chronic patient (added in 2018), appointment by family nurse (added in 2013).
- Other major changes: no.

Hypertension medication 1

- In QBS since: 2013
- Includes: proportion of active substance-based prescriptions out of prescriptions issued for hypertensive patients of all risk levels.
- Other major changes: no.

Hypertension medication 2

- In QBS since: 2013
- Includes: Prescription for angiotensin-converting enzyme inhibitors, calcium channel blockers, beta-blockers or angiotensin antagonist (II) treatment group (incl combination drugs).
- Other major changes: no.

MI

- In QBS since: 2008
- Includes: cholesterol (added in 2009), glucose (added in 2009), cholesterol fractions (added in 2016).
- Other major changes: no.

MI medication 1

- In QBS since: 2016

- Includes: prescription of beta-blockers treatment group (incl combination drugs).
- Other major changes: no.

MI medication 2

- In QBS since: 2016
- Includes: prescription of statins treatment group (incl combination drugs).
- Other major changes: no.

Hypothyroidism

- In QBS since: 2008
- Includes: TSH (thyroid stimulating hormone) determined
- Other major changes: no.

Part III Additional professional competence

Recertification (for FP and nurse)

- In QBS since:
- Includes: Both, FP and nurse are recertified
- Other major changes:

Pregnancy monitoring

- In QBS since: 2007
- Includes: detection of pregnancy, setting up pregnancy monitoring plan, monitoring pregnancy, midwife's appointment (since 2018),
- Other major changes: since 2008, introduced as combination indicator.

Gynaecological examinations

- In QBS since: 2008
- Includes: gynaecological examination, simple and complex gynaeco-cytological tests, insertion of intrauterine device.
- Other major changes: no.

Small surgery

- In QBS since: 2006
- Includes: several procedures (wound dressing, catheterisation, casting fractures, surgical suturing etc). The list of components is updated almost every year.
- Other major changes: no.

Newest indicators

Albumin/creatinine ration in urine in diabetic patients

- In QBS since: 2018
- Includes: albumin/creatinine ration in urine. Expected coverage 50%. No points assigned.
- Other major changes: no.

Albumin/creatinine ration in urine in hypertension patients

- In QBS since: 2018
- Includes: albumin/creatinine ration in urine. Expected coverage 50%. No points assigned.
- Other major changes: no.

E-consultation

- In QBS since: 2018.
- Includes: e-consultation using health information system. Introduced as tracking indicator. No points assigned.
- Other major changes: no.

Indicators dropped from QBS

CVD

- In QBS since: 2007
- Includes: ECG (dropped in 2008), blood glucose, cholesterol, cholesterol fractions, appointment by family nurse for target group patients aged 40–60 years.
- Other major changes: indicator dropped by 2016.

Breast cancer screening

- In QBS since: 2007
- Includes: mammography
- Other major changes: indicator dropped by 2008.

Cervical cancer screening

- In QBS since: 2007
- Includes: gynaeco-cytological test.
- Other major changes: indicator dropped by 2008. Introduced as part of Gynaecological examinations indicator in 2008.

Continuous professional development

- In QBS since: 2007.
- Includes: list of courses attended (minimum of 60h required for both, doctor and nurse).
- Other major changes: dropped by 2008.

Annex 2: Detailed description of QBS indicators in domain 2

Indicator	Code	Description	Target coverage	Points	Comments
Diabetes (type II DM)	66118 or 9118	Glycosylated haemoglobin determined for type II DM patients	Previous year's coverage+10% but not more than 90%	65	1 x in a year
	66102 or 9102	Creatinine values determined for type II DM patients			
	66104 or 9104	Cholesterol values determined for type II DM patients			1 x in 3 years
	66105 or 9105	Cholesterol fraction values determined for type II DM patients			
	9044	Counselling for chronic patient			
	9061	Appointment by family nurse			
Diabetes medication	Prescription for metformin treatment group	Prescribed for all type II DM patients	Previous year's coverage+10% but not more than 90%	10	6 prescriptions in 14 months
Hypertension I (low risk)	66101 or 9101 or 66118 or 9118	Glucose or glycosylated haemoglobin	Previous year's coverage+10% but not more than 90%	90	1 x in 3 years
	66104 or 9104	Cholesterol			
	9044	Counselling for chronic patient			1 x in a year
	9061	Appointment by family nurse			
Hypertension II (moderate risk)	66104 or 9104	Cholesterol determined for patients under 80yo	Previous year's coverage+10% but not more than 90%	175	1 x in a year
	66105 or 9105	Cholesterol fractions determined for patients under 80yo			
	66101 or 9101 or 66118 or 9118	Glucose or glycosylated haemoglobin			
	66102 or 9102	Creatinine			1 x in 3 years
	6320 or 6322 or 6323 or 9320	ECG			
	9044	Counselling for chronic patient			
	9061	Appointment by family nurse			

Indicator	Code	Description	Target coverage	Points	Comments
Hypertension III (high risk)	66104 or 9104	Cholesterol determined for patients under 80yo	Previous year's coverage+10% but not more than 90%	40	1 x in a year
	66105 or 9105	Cholesterol fractions determined for patients under 80yo			
	66101 or 9101 or 66118 or 9118	Glucose or glycosylated haemoglobin			
	66102 or 9102	Creatinine			
	9044	Counselling for chronic patient			
	9061	Appointment by family nurse			
Hypertension medication 1	Prescription containing diagnosis I10-I15	Percentage of active ingredients based prescriptions for hypertension patients (all risk levels)	90%	5	1 x in a year
Hypertension medication 2	Prescription for angiotensin-converting enzyme inhibitors, calcium channel blockers, beta-blockers or angiotensin antagonist (II) treatment group (<i>incl</i> combination drugs)	Prescriptions for moderate or high-risk hypertension patients	83%	20	6 prescriptions in 14 months
MI	66104 or 9104	Cholesterol	Previous year's coverage+10% but not more than 90%	20	1 x in a year
	66101 or 9101 or 66118 or 9118	Glucose or glycosylated haemoglobin			
	66105 or 9105	Cholesterol fractions			
	9044	Counselling for chronic patient			1 x in a year
MI medication 1	Prescription of beta-blockers treatment group (<i>incl</i> combination drugs)	Medication for patients with MI	Previous year's coverage+10% but not more than 90%	5	6 prescriptions in 14 months
MI medication 2	Prescription of statins treatment group (<i>incl</i> combination drugs)			5	6 prescriptions in 14 months
Hypothyroidism	66706 or 9706	TSH (thyroid stimulating hormone) determined	Previous year's coverage+10% but not more than 90%	45	1 x in a year
Points total				480	

Note: EHIF is responsible for making lists of patients eligible for specific QBS indicators for each Family Doctor. A patient is categorized as a eligible for a given indicator if he/she has had at least one claim sent

to EHIF by the Family Doctor in the last three years with a relevant diagnosis. The doctors are presented with this list and can confirm its accuracy. To validate the data, regular electronic controls are applied to all claims. In addition, regular medical records' monitoring takes place to ensure data quality.