Bundled Payment for Ischemic Acute Stroke

An Assessment of the Estonian Health Insurance Fund Pilot and Recommendations

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

June 2020

Contents

LIST OF ACRONYMS

BPCI	Bundled Payment of Care Improvement Initiative
DRG	Diagnosis-Related Group
EOC	Episode of care
EHIF	Estonian Health Insurance Fund
FFS	Fee-for-service
ICD-10	International Statistical Classification of Diseases and Related Health Problems, 10th revision
ICHOM	International Consortium for Health Outcomes Measurement
IT	Information technology
ITS	Interrupted time series
MoSA	Ministry of Social Affairs
MS-DRG	Medicare Severity-Diagnosis Related Group
PROMIS	Patient-Reported Outcomes Measurement Information System
OECD	Organisation for Economic Co-operation and Development
PHC	Primary health care
PROMETHEUS	Provider Payment Reform for Outcomes, Margins, Evidence, Transparency, Hassle Reduction, Excellence, Understandability, and Sustainability
USA	United States of America

ACKNOWLEDGMENTS

The authors of this report are Adrien Dozol, Dean Farley, Ian Forde, and Iryna Postolovska. The support from Christoph Kurowski at the World Bank Group is gratefully acknowledged.

We are grateful to the peer reviewer, Jack Langenbrunner, for providing excellent comments.

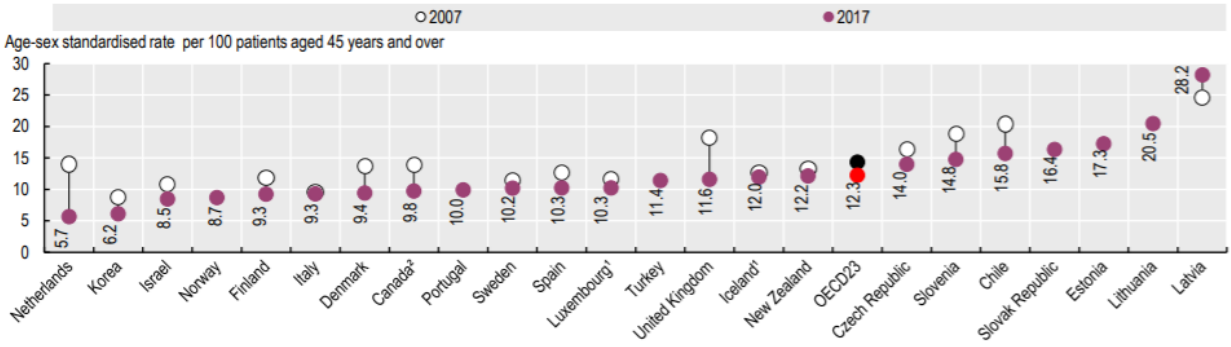
Most importantly, we thank the board and staff of the Estonian Health Insurance Fund and the Steering Committee overseeing the World Bank Group's engagement with the Fund. Without their guidance and support, this report would not have been possible.

1. INTRODUCTION

1.1 Background

Strokes represent a large share of the disease burden in Estonia, with almost 8 % of deaths in the country attributable to the condition (IHME, 2019). According to the National Institute for Health Development, on average approximately 4,000 patients per year experience a stroke, representing a rate of 393.9 cases per 100,000 population (National Institute for Health Development, 2019). Compared to other Organisation of Economic Co-operation and Development (OECD) countries, Estonia has one of the highest 30-day mortality rates after admission to hospital for ischemic stroke, indicating gaps in the quality of stroke care (Figure 1).[1]

Figure 1: Thirty-day Mortality After Admission to Hospital for Ischemic Stroke Based On Linked Data, 2007 and 2017 (or nearest years)



Source: OECD (2019

Notes: [1]Three-year average. [2] Results for Canada do not include deaths outside acute care hospitals.

The Estonian Specialist Association of Neurologists approached the Estonian Health Insurance Fund (EHIF) to address some of the endemic problems in the delivery of stroke care. To better understand the root causes and identify bottlenecks to delivering high quality stroke care, EHIF conducted a series of workshops with the main stakeholders involved in stroke care, including patients. The discussions highlighted several key problems across different levels of care. These included insufficient attention to secondary prevention, problems with data collection and information flow across the pathway, poor collaboration across levels of care, challenges in the delivery of rehabilitation services, and limited engagement with patients and lack of services to patient's needs. In addition, stakeholders identified both demand and supply constraints related to underutilization of inpatient and outpatient rehabilitation care (e.g., poor advanced planning, unavailability of services on weekends, limited volume of services, and lack of transportation).

Given the size and severity of the problem, EHIF is undertaking a pilot program designed to improve care coordination and outcomes for ischemic stroke patients in Estonia. Its goal is to improve patient outcomes by developing and implementing (1) integrated clinical pathways; (2) a process for measuring and

[1] Some improvements in the quality of care were observed after the introduction of guidelines for stroke management in Tallinn Emergency Medical Services in 2008, including a reduction in the door to needle time (Gross-Paju, 2017).

reporting on treatment decisions, patient outcomes and costs of care; and (3) a reimbursement system – bundled payment – that will support the clinical and operational reforms that emerge from the pilot program. This pilot is also accompanied by an initiative led by the Ministry of Social Affairs (MoSA) to restructure the provision of stroke care to ensure timely access to quality care. Since September 2019, every patient with clinical symptoms of stroke must be transported to a hospital with a stroke unit or a stroke center. This report evaluates EHIF's approach for piloting bundled payment and provides recommendations based on international experience.

BPCI Stroke Episodes

Bundled Payment of Care Improvement Initiative (BPCI) is the only bundled payment program in the United States of America (USA) to include acute stroke. The Model 2 stroke bundle included the cost of an initial hospital stay assigned to one of the 6 stroke Medicare Severity-Diagnosis Related Groups (MS-DRGs), most physician and post-acute services for up to 90 days following discharge, and any readmissions occurring during that time period. Eighty-four hospitals (20% of the hospitals in Model 2) participated in the stroke episode. The results were modest. The average cost of a stroke bundle for 90-day episodes was US$30,050 during the intervention period, an estimated savings of US$247 (not statistically significant). None of the estimated quantitative effects were significant for stroke, but they were broadly consistent with other BPCI results: reductions in readmissions and the use of institutional post-acute care combined with increases in home health services. The experience in the USA underscores the importance of making sure delivery systems have resources available to undertake desired changes in clinical practice if payment reform is to succeed in its objectives.

While bundled payment is a relatively new reimbursement methodology, there is international experience that may help guide the EHIF. The following discussion brings that experience to bear on current EHIF plans, and is organized as follows. Section 1 provides a definition of bundled payment and an overview of global experience with the payment mechanism. Section 2 presents a summary of the proposed EHIF pilot based on discussions between World Bank staff and EHIF as well as written materials and analyses provided by EHIF. Section 3 provides an overview of the components that comprise an episode, while Section 4 outlines the operational issues that need to be addressed in implementing a bundled payment program. Section 5 assesses the methodology used to measure patient reported outcomes. Section 6 presents an overview of the proposed pilot, together with the evidence on best practices in bundled payment, to identify the strengths of the current pilot design and to suggest areas where the design might be strengthened. The final section addresses the issue of evaluation and describes options available to EHIF to assess the impact of the pilot in Estonia.

1.2 Definition of bundled payment

Payment is an important lever to incentivize behavior and improve health system performance. Traditional payment methods, such as fee-for-service (FFS), global budgets, or diagnosis-related groups (DRGs), however, are not well adapted to meet the growing challenges of rising chronic conditions and aging populations. Such payment methods often have undesirable consequences, including overprovision of services, and do not directly incentivize quality of care (Table 1).[2] New ways of paying providers are needed to improve coordination of care and align provider behavior with health systems' objectives.

Bundled payment - a form of prospective payment – is designed to improve the efficiency of service delivery while maintaining or enhancing the quality of care across multiple providers. Similar to DRGs, bundled payments are episode-based, covering the cost of all services delivered during a defined episode of care. Unlike DRGs, however, payments are bundled for services delivered by different providers. They also differ from global payments because the payment covers a specific condition or set of services.

[2] While the established payment methods themselves may not necessarily incentivize quality of care, the accompanying measures and policies may impact quality (e.g. addressing "quicker and sicker" or excluding readmissions within 30 days from DRG payment).

Bundled payments are based on the concept of an "episode of care" (EOC) – a defined set of services delivered by multiple providers (and levels of care).[3] Loosely speaking, an EOC is a set of services designed to treat an acute medical event, to manage a chronic condition, or to assess and maintain the health status of an individual. In its simplest form, a health plan or payer establishes a fixed fee or target price for the set of services associated with an EOC. This shifts the financial risk of variations in the cost of care from the payer to one or more providers. It creates an incentive to organize the delivery of services more efficiently and to manage cases in a way that minimizes complications and other undesirable outcomes. As such, it also provides a framework for associating services to episodes and measuring outcomes in a way that supports value-based reimbursement. A common characteristic of bundled payment models is the use of quality metrics for providers, such as inpatient mortality or readmission rate. Bundled payments for chronic conditions aim to incentivize continuity of care along a clinical pathway rather than paying for discrete episodes or interventions (OECD, 2016).

[3] Hussey and others (2009) provide an accessible overview of EOC methodology and its relationship to reimbursement and performance assessment. See also Barchi and others (2014), which discusses two of the most common episode methodologies used in the United States.

Table 1: Risk-Sharing and Provider Incentives Under Different Payment Mechanisms

Payment mechanism	Risk borne by		Provider incentive to					
	Purchaser	***Provider***	***Increase enrollment/ patients***	***Decrease services per episode***	***Increase reported illness severity***	***Select healthier people/ patients***	***Improve integration and coordination of care***	***Improve quality of care***
Line-item budget	All	None	No	Yes	No	Yes	No	No
Global budget	None	All	No	Yes	No	Yes	No	No
Fee-for-service	All	None	Yes	No	Yes	No	No	No
Per diem	Partly	Partly	Yes	Yes	No	No	No	No
Case-based	Partly	Partly	Yes	Yes	Yes	Yes	No	No
Bundled payment	Partly	Partly	Yes	Yes	Yes	Yes	Yes	Yes

1.3 Global experience with bundled payment

A number of high-income countries have introduced bundled payment for select conditions (see Table 2 for examples).

Table 2: Examples of Bundled Payment Initiatives

Country	Conditions
USA (PROMETHEUS)*	Chronic conditions, orthopedic surgeries
USA (BPCI)	Congestive heart failure, pneumonia, chronic obstructive pulmonary disease, sepsis, acute myocardial infarction, stroke
UK	Maternity care
Sweden	Episodes of care (hip replacement, spine surgery)
Portugal	Selected high cost chronic conditions (HIV/AIDS, multiple sclerosis, end-stage renal disease)
Netherlands	Type 2 diabetes, chronic obstructive pulmonary disease, vascular risk management
Taiwan	Breast cancer
Denmark	Diabetes mellitus
New Zealand	Maternity care

Source: OECD (2016) and Struijs et al. (2020)
*Provider Payment Reform for Outcomes, Margins, Evidence, Transparency, Hassle Reduction, Excellence, Understandability, and Sustainability

Perhaps the most documented among these is the BPCI under the U.S. Medicare program that was introduced in 2013.[4,5] Under this program, providers volunteered to accept financial risk for up to 48 different episodes of care. An episode started with a hospital admission that was consistent with 1 of the 48 definitions and extended for a period of time, up to 90 days, following discharge from the hospital.

Providers could choose one of four models, which differed in terms of bundle definition and payment methodology. Models 2 and 4 are most similar to the EHIF stroke pilot. The Model 2 bundle included all Medicare reimbursements that occurred during the initial hospital stay as well as the post-discharge period selected by the participant. All providers were paid under standard Medicare methods, but there

[4] Several states introduced initiatives shortly after the enactment of the Affordable Care Act in 2010: Arkansas (2012), Tennessee (2015), and Ohio (2015). Each program established financial incentives for providers based on the difference between average cost and a risk-adjusted target price. Bonuses were only paid to providers that met episode-specific quality standards, such as 30-day mortality rates for coronary artery bypass grafting patients and gestational diabetes screening for perinatal services. For all episodes, at-risk providers received provider-specific reports of average cost and quality performance relative to peer groups. In fact, the states generally introduced new episodes with an initial "reporting only" period, often a year long, in which providers received and reacted to performance reports without any change in payment.

[5] The BPCI demonstration officially ended in 2018, but the evaluation research continues, and the Centers for Medicare and Medicaid Services have launched a new program, BPCI-Advanced, that builds on the lessons from the original BPCI initiative.

was a subsequent reconciliation against a target price. Three different types of providers could join BPCI under Model 2: acute hospitals, physicians, and post-acute providers. The Centers for Medicare & Medicaid Services developed hierarchical logic to determine which provider (or provider group) was actually at-risk when more than one provider could "claim" a patient. Model 4 differed from Model 2 in that participation was only available to hospitals and only included services provided during the initial hospital stay and any subsequent readmissions. Participating hospitals received a single, fixed payment for each episode to cover the costs of hospital and physician services. The hospital was responsible for distributing payments to the individual physicians.

Model 2 was the largest of the models and the most important from an evaluation perspective. Interestingly, Model 4 never attracted much interest; 23 hospitals enrolled but only 5 hospitals remained in 2017. Anecdotally, the lack of interest in Model 4 was largely the result of difficulties hospitals experienced in administering the single payment that encompassed both hospital and physician services.

Several other countries have also implemented bundled payment for different clinical episodes and conditions. In England, the National Health Service has adopted a bundled payment approach for obstetrics and maternity-related services (Henderson, 2016). Similarly, in 2010 the Netherlands introduced a nationwide bundled payment program for diabetes care, chronic obstructive pulmonary disease, and cardiovascular risk management (de Bakker, 2012). In Sweden, the Stockholm County Council introduced a bundled payment program for primary elective hip and knee replacement at the same time that it increased patient choice of providers in 2009.

Evidence on the impact of bundled payment remains limited. In a recent review of 23 bundled payment initiatives, Struijs et al. (2020) found that less than half were empirically evaluated. The evidence from these evaluations suggests that bundled-payment models have the potential to reduce cost without a negative impact on quality. The gains, however, depend on the condition or episode (Agarwal et al., 2020; Struijs et al., 2020) (see Figure 2). This underlines the importance of risk-stratification and risk-adjustment to account for patient heterogeneity in designing bundled payment programs and in evaluating their effects (Agarwal et al, 2020).

Evidence from the BPCI shows generally favorable, but modest, results (The Lewin Group, 2018):[6]

> "Under the BPCI initiative, Medicare payments declined for most clinical episodes and over half of the relative payment reductions were statistically significant. The declines were primarily due to relative reductions in the use of PAC [post-acute care]. The Medicare payment reductions occurred under Model 2 and 3 and across participant types as well as a range of surgical, acute, and chronic clinical episodes. Quality of care, measured as emergency department visits, mortality, and readmissions, was not affected in the vast majority of clinical episodes. Changes in functional status did not differ between beneficiaries in BPCI episodes and comparison beneficiaries, based on survey results, although fewer BPCI beneficiary respondents reported the highest level of satisfaction with their care."

A recent evaluation of the program in Stockholm County Council, using a difference-in-difference methodology, reports a combined 14% reduction in costs per episode along with a substantial reduction in the frequency of repeat procedures (Wohlin, et al., 2017). The introduction of bundled payment, however, is usually part of a larger reform, making it difficult to disentangle the direct impact of bundled payment on service provision and patient outcomes (OECD, 2016).

[6] Savings estimates are based on a difference-in-difference analysis, using matched non-BPCI hospitals as controls.

Figure 2: Selected Evidence of Bundled Payment Initiatives on Quality of Care and Cost Outcomes

Bundled-Payment Initiative	Quality of Care Results	Medical Spending Results
Acute Care Episode (ACE) Payment model (cardiac)	Quality of care levels maintained Reduction in the use of internal mammary artery grafts in patients undergoing coronary artery bypass graft surgery	Savings of $319 per episode Total of $4 million in net savings for 12,501 episodes of care
UnitedHealthcare episode payment model (oncology)	No differences between the groups on multiple quality measures	34% reduction of predicted total medical cost (Study used two interventions — financial incentives and data-sharing — to change behavior; relative effect of each incentive could not be determined)
Horizon Health Care division (Blue Cross Blue Shield) (orthopedic)	Total knee arthroscopy (TKA) length of stay decreased from 3.3 days to 1.9 days Total hip arthroscopy (THA) length of stay decreased from 2.9 days to 1.8 days Discharge to inpatient rehabilitation significantly decreased from 66.3% in 2011 to 33.17% in 2013–2014. In-hospital complication rate increased from 6.4% to 8.67%, but a review of this data revealed a significant increase in hospital coding for clinically insignificant complications Transfusion rate decreased from 23.2% to 4.45% 30-day readmission rates decreased from 3.2% to 2.7%	Average device cost decreased from $6,301 per patient to $4,972 per patient with the last six months averaging $4,585 per patient Average episode budget was $25,365 for TKA and $23,580 for THA Under budget for 65 of 78 TKA episodes and under budget for 27 of 38 THA episodes Total savings relative to budget for all Horizon patients over this two-year period exceeded $524,000, resulting in a savings of $262,445 during this time or an average of $2,262 per patient
Bundled Payment for Care Improvement (BPCI) (total joint arthroplasty)	18% reduction in length of stay Shift from home health and skilled nursing facility discharge to home self-care (54.1% to 63.7%)	No significant differences in implant cost Improvements resulted in 6% reduction in average total allowed claims per case
Bundled Payment, Netherlands (diabetes)	Decrease in specialist care Increase in regular checkups Increase in foot exams Increase in kidney exams Decrease in eye exams	Increase in total medical spending of 388 euro compared to control group No increase in medical spending for diabetes specialist care

Source: Struijs et al. (2020)

Implementation of bundled payments is a complex process and requires time to ensure that all the elements work together. Global experience highlights several implementation challenges (Hussey et al., 2011; Ridgley et al., 2014; OECD, 2016). These include (a) agreement on the definition of the bundle (finding a balance between the objectives of the payer and those of the provider); (b) provider uptake; size of the financial incentive (including risk-adjustment); and (c) information systems. Stakeholder support and involvement in the design process is critical for successful implementation (OECD, 2016). In the U.S., many of the earlier initiatives generally introduced new episodes with an initial "reporting only" period, often a year long, in which providers received and reacted to performance reports without any change in payment. Although the PROMETHEUS[7] Initiative pilot – an earlier attempt to introduce bundled payment in the U.S. – started in 2008, 3 years into the pilot (by May 2011) bundled payment contracts

[7] Provider Payment Reform for Outcomes, Margins, Evidence, Transparency, Hassle Reduction, Excellence, Understandability, and Sustainability

still had not been agreed between payers and providers due to a number of implementation challenges (Figure 3).

Figure 3: Summary of Implementation Challenges in The PROMETHEUS Payment Pilot

Implementation challenge	Pilot sites' experiences
Defining "bundles"	Pilot sites have had difficulty applying case-rate definitions to their own data, particularly in real time Evidence base for "potentially avoidable services" questioned and has so far been exposed to relatively few frontline physicians
Defining the payment method	No contracts between payers and providers have been executed yet Payers hesitant to allocate shared savings payments, providers hesitant to accept withhold "Chicken or egg" problem of reengineering care before payments change
Implementing quality measurement	Focus on commonly used process quality measures Electronic health record is crucial, but implementation of electronic measures is time and resource intensive
Determining accountability	Providers have not yet resolved issue of how to allocate payments/accountability "Leakage" of patients outside of network recognized as a major issue, but no solutions Pilot has driven increased functional integration at all sites
Engaging providers	Physicians, management at pilot sites very supportive of the PROMETHEUS concepts, but frontline physicians not yet broadly engaged Perceived need to translate analytics into clinically actionable information
Delivery redesign	Changes to this point generally focus on increasing coordination; in early stages of planning for clinical redesign PROMETHEUS viewed as complementary to other initiatives, including medical homes and accountable care organizations Despite potential to save money by reducing cost of typical services (for example, implantable devices), pilot sites have not focused on this

Source: Hussey et al. (2011)

2. EHIF PILOT PROGRAM TO IMPROVE STROKE CARE

EHIF is undertaking a pilot program to improve care that stroke patients receive in Estonia. One key to reducing mortality and improving clinical outcomes is to ensure that stroke patients are identified quickly after the onset of a stroke and then admitted to facilities that have specialized resources to treat stroke patients and expertise to use those resources effectively. Prior to the establishment of a network of stroke units/centers and referral of all suspected cases to these facilities in September 2019, nearly 20 % of stroke patients in Estonia were admitted initially to general and local hospitals, where resources are more limited, rather than being taken immediately to regional and central hospitals for care.

EHIF plans to introduce the bundled payment pilot in 2020. All six regional and central hospitals that provide acute stroke treatment in Estonia applied to participate in the pilot. EHIF plans to define the scope of the bundle for reimbursement purposes as a subset of the care episodes associated with the clinical pathway. Specifically, the payment bundle will include the cost of:

- Hospital and physician services associated with an initial admission where the principal diagnosis, using the 10th revision of the International Statistical Classification of Diseases and Related Health Problems (ICD-10), is I63 "Cerebral infarction".
- Subsequent non-emergency hospital care and outpatient visits (including day treatment) with a primary diagnosis of stroke, more specifically cerebrovascular diseases (I60-69) or Hemiplegia and hemiparesis (G81).
- Subsequent emergency visits and outpatient visits (including day care) with a primary diagnosis of stroke (I60-69 or G81) within one year of the initial admission.

- Outpatient or inpatient rehabilitation with primary diagnosis codes of stroke (I60-69 or G81) that occur within 12 months of the initial hospital admission[8].
- Nursing care within 12 months of the initial admission billed on invoices with a primary diagnosis of stroke (I60-69 or G81).
- Treatment of complications within 30 days of the initial or subsequent hospital admissions.

The payment bundle will end 365 days after the initial emergency admission, with the death of the patient, or another emergency admission for acute stroke.

In addition to piloting the new reimbursement model, EHIF established a service delivery innovation fund using a two-part solicitation. In Round 1, the six regional and central hospitals applied for planning grants of EUR 15,000. These grants are helping four selected hospitals create an initial demonstration concept, assemble a team of providers with whom to work, and prepare a detailed plan for developing and testing their concept, including information technology (IT) solutions. The planning phase will last four months. Hospitals that receive a planning grant will then compete for larger awards (up to EUR 300,000) in a second round of bidding. Round 2 grants will support the implementation and refinement of the proposed solutions. EHIF expects that grant funding will be used to develop the infrastructure needed to support the solution across the entire project team and to both measure and evaluate the results of the demonstration. EHIF expects to end the pilot project no later than December 31, 2021, at which point it plans to deploy best practices among Round 2 awardees more broadly across its provider network.

EHIF recognizes that the success of the program will depend upon its ability to obtain clinical leadership and institutional support from the six target hospitals. It has been working with physician and institutional providers to prepare for this pilot by identifying clinically appropriate outcome measures, discussing the clinical pathways that will promote improvements in those measures, and exploring organizational changes that will be necessary to support the changes in clinical services that it envisions.

In general, EHIF is relying on the hospitals to define the interventions that they will pilot rather than prescribing the specific content of those interventions. At the same time, the projects must address the development needs identified in the stakeholder workshop on October 5, 2019 (Box 1). For example, EHIF will define a minimum set of clinician and patient-reported outcome measures that all Round 2 grantees must use, however the grantees will be free to supplement these measures as they think appropriate.

[8] The costing of the bundle took into account a maximum of 6 months of rehabilitation over the course of a year.

Box 1. Development Needs for Improving the Quality of Stroke Care

- Adopt a patient-centered approach. Patients and their relatives must be provided with timely, understandable and relevant information and be involved as equal partners in decision-making concerning them.
- Implement a single care plan. The care plan should be accessible to all parties involved in the care pathway, including the goals identified together with the patient, treatment plan, risk factors, important contacts, visits, and other necessary information. The care plan should follow the patient along the pathway, providing an overview of the current treatment and outcomes for providers and allowing providers to supplement or amend the plan.
- Ensure cooperation and division of roles. The parties in the care pathway must agree on the division of roles, clearly outlining the responsibilities and processes, and ensuring that the patient moves smoothly from one stage of the pathway to another.
- Create a coordinating role. The coordinator must be the first point of contact for the patient and their family members throughout the whole pathway and guide them from one stage to the next.
- Develop home and community services. The patient must be provided services according to their individual needs and preferences either at home or as close to their home as possible.
- Measure outcomes. A single system for measuring patient outcomes should be established to capture all relevant aspects and ensure comparability.

Source: EHIF

EHIF has also established a definition for acute stroke episodes that will underline the pilot. An episode, referred to as care pathway in the pilot documentation, will begin with an emergency admission to a participating hospital with a principal diagnosis of acute ischemic stroke.[9] Clinical pathways will then be defined for managing patients during episode starting with the acute phase of the treatment and extend into the post-discharge period to address the need for rehabilitation and support services as well as primary care and social services for at least a year after the episode begins.

An objective of the pilot is to test new outcome indicators and an IT solution to collect the required information. It was agreed with stakeholders that the indicators will be based on the International Consortium for Health Outcomes Measurement (ICHOM) "stroke standard set". The ICHOM stroke standard set includes measures that are based on administrative and clinical data, as well as patient-reported information (Salinas and others, 2016). After comparing alternative data collection instruments, EHIF and its stakeholders have agreed to rely on the Patient-reported Outcomes Measurement Information System (PROMIS-10) platform to collect patient-reported outcomes. All participating hospitals will be required to report on process and outcomes indicators. EHIF is working with the piloting hospitals to procure the IT services necessary to collect and process data to support outcome measurement, benchmarking and performance assessment across the continuum of care. These services will need to accommodate patient-reported outcome measures and provide a platform for integrating services across the entire scope of the clinical pathways that grantees will put in place. Reports will also

[9] The emergency admission will need to be preceded by a 12-month "clean period," i.e., a period of at least 12 months during which the patient did not require hospitalization for a previous stroke.

need to support bundled payment by focusing on the subset of services associated with the bundle. Because there will be no historical data available on the broad spectrum of measures included in the ICHOM standard set, EHIF will not adjust payment based on results during the pilot phase. However, collecting and disseminating process and outcome measures during the pilot will be important for the long-term success of the EHIF initiative, since the experience of the pilot will help to establish a baseline for the future and to provide important information to providers on opportunities for clinical improvement.

3. DEFINING EPISODES OF CARE

3.1 Triggers.

The first step in defining an episode of care, whether for payment or analytic purposes, is to specify a "trigger." A trigger is diagnosis or procedure code, or a set of diagnosis and/or procedure codes, along with contextual rules, that indicate an episode has begun. In the case of the EHIF pilot, for example, the trigger for a stroke episode will be an emergency admission with a principal diagnosis of I63 (Cerebral Infarction). A recurrent stroke will trigger a new episode.

3.2 Rules for gathering services

The next step in defining an episode of care is to specify rules that determine the services that are included in the episode. A program can either use inclusion or exclusion logic in defining these rules. Programs that use inclusion logic establish rules that identify services that are part of an episode. Any service that does not conform to the established rules is considered to be outside of the episode. Exclusion logic works in the exact opposite way. That is, all services delivered within a specified time period are considered to be part of the episode unless they are specifically excluded by rule.

EHIF is proposing to use inclusion logic to define its episodes. Specifically, episodes will consist of:[10]

- All services delivered as part of the triggering admission.
- Non-emergency inpatient stays, day treatment, day surgery or outpatient visits with a primary diagnosis of I60-I69 or G81 within 1 year of the trigger.
- Emergency department visits with a primary diagnosis of I60-69 or G81 within 1 year of the trigger.
- Non-emergency services with a primary diagnosis of I60-I69 or G81 and health services list code 2048 or main specialty V10 within 1 year of the trigger admission.
- Rehabilitation services within 6 months of the trigger with a principal diagnosis of I60-I69 or G81.
- Nursing care with a principal diagnosis of I60-69 or G81 within 1 year of the trigger.
- Services associated with complications of the stroke have been specified and are presented in Table 3.

[10] While stroke episodes will generally last one year from the triggering admission in the EHIF pilot, the time period may be shorter in the event that the patient dies or has another stroke within the one-year time period.

Table 3: Services Associated with Complications of Stroke

ICD-10 code main diagnosis	ICD-10 main diagnosis description	Period during which the main diagnosis is included in the treatment route
I26	Pulmonary embolism	0 - 30 days after hospital discharg (valid at each stage of treatment)
L23, L89;	Allergic contact dermatitis, pressure ulcer	
I80 –I82;	Phlebitis and thrombophlebitis, portal vein thrombosis, embolism and thrombosis of other veins	
J69	Pneumonitis due to solids and liquids	
J09-J18	Pneumonia	
O08, O07	Sepsis	
K22.8, K25, K26, K27, K28, K62.5, K92.2	Gastrointestinal bleeding	
N10, N12 –N13, N17, N28.0, N30; N39	Acute tubulointerstitial nephritis, Unspecified whether acute or chronic tubulointerstitial nephritis, Obstructive and reflux ophthalmic, Acute renal failure, Other conditions of kidney and ureter not elsewhere classified, Cystitis or other inflammation of the bladder, Urinary tract	

Source: EHIF

While the use of inclusion logic is more intuitive, it presents some challenges. Rules must be articulated precisely and without ambiguity. The rules also need to be updated and revised as the underlying coding systems change. Inclusion logic is also more conservative in the sense that it tends to exclude invoices where the relationship to the underlying episode of interest is ambiguous. Stroke patients often have one or more co-morbidities, such as atrial fibrillation, that can affect their treatment and the course of their condition. The inclusion logic specified by EHIF will generally result in the exclusion of cardiology services from the stroke episode definition. Such exclusions may be appropriate for reimbursement purposes, but EHIF may want to include the cost of treating co-morbidities for evaluation purposes to understand how improved stroke care affects the overall costs of treating stroke patients.

Exclusion logic provides a more expansive episode definition. It defines a time period, which can begin before the triggering event and ends sometime after the event, and then gathers all services into the episode unless they are specifically excluded by rules. For example, the BPCI-Advanced demonstration uses exclusion logic to construct episodes, but it excludes costs associated with blood and blood products, medical devices, and end-of-life care. For each type of episode, it also created a list of MS-DRGs that represent inpatient admissions that are unlikely to be related to the initial triggering event. Medicare then excludes the costs of readmissions for MS-DRGs that appear on this list.

Exclusion logic is easier to implement and maintain because it relies less on code-specific rules to determine the content of the bundle. It also creates more financial risk for providers, which in turn encourages greater emphasis on case-management over the duration of the episode. However, the purpose of the EHIF stroke pilot is to define and implement clinical pathways that will improve the care of stroke patients in Estonia. For that reason, the decision to rely on inclusion logic for the pilot seems appropriate, especially given the level of agreement between EHIF and the clinical community. Once the pilot is over, EHIF may want to re-visit how episodes are defined to simplify maintenance and to strengthen incentives for future cost-savings.

3.3 Length of Episodes. The simplest and most common way to determine the length of episodes is to specify the duration *a priori*. At the same time, there are circumstances that require flexibility in determining the length of an episode. Not all patients survive the full length of the episode; their episodes end upon their death. In other cases, there may be compelling clinical information that the episode is not complete at the time it is scheduled to end, as would be case with a hospital readmission for the same underlying condition that triggered the episode. In these situations, the episode definition can extend the episode, terminate the episode at the start of the readmission, or exclude the episode entirely from reporting and analysis.

As a general rule, short durations are appropriate for acute events where an episode begins, treatment ensues, and the clinical problems is resolved within a relatively short time period. Longer durations are appropriate for chronic conditions that are never completely resolved by definition, and for acute conditions that require long-term clinical management. Regardless of the approach, definitions of episodes usually do not exceed one year because of the operational need to collect, analyze and report on data.

Current EHIF plans call for episodes to last for one year after the start of the triggering hospital admission. The duration will be shortened if the patient dies or has another stroke before the year is complete.[11] The decision to use a one year episode definition seems to be aligned with EHIF's interests in encouraging the use of rehabilitation, physiotherapy and social services to support patient recovery, given the long-term health consequences of acute strokes. In the pilot, a recurrent stroke will trigger a new episode and will receive a new bundled payment.

3.4 At-risk providers

Bundled payment programs set a prospectively determined price for an episode of care and then hold one provider accountable for the costs incurred relative to the target price. Designation of at-risk providers is an important policy decision and an essential part of defining a bundled payment program. Bundled payment programs have taken a variety of different approaches to designating at-risk providers. Some place hospitals at risk; others place physicians at risk or allow the at-risk provider to vary by type of episode.[12]

EHIF has made a simple, and reasonable decision to place the hospital, where the initial emergency admission for an ischemic stroke occurs, at risk for the episode. In the context of the EHIF pilot, this decision will hold participating regional and central hospitals responsible for the cost, clinical content and outcomes of episodes where they initiated treatment for the acute stroke event.

3.5 Measuring clinical and financial outcomes

EOCs are generally defined for analytic purposes. As discussed earlier, episodes are created to understand the health status of an individual or to group services into discrete bundles of care to study treatment patterns and their relationship to patient outcomes. Those outcomes can either be clinical (e.g., complication rates, mortality, length of episode) or financial (e.g., cost per episode). The episode definition (trigger event combined with rules to gather services) determine the invoices or patient-physician interactions that are part of the episodes and analysts can then look at a variety of outcome measures.

[11] Note that a recurrent stroke will not be excluded from the pilot and will trigger a new bundled payment.

[12] In its two BPCI demonstrations, the U.S. Medicare program adopted an even more complicated strategy. Because hospitals and groups of physicians could join the program separately, more than one provider could potentially bear risk for individual episodes. Medicare had to develop logic that determines which provider is at risk for any situation where more than one provider could "claim" the patient.

The data used to measure outcomes can be drawn from a number of sources. The most common sources are insurance claims, medical records, and other types of administrative data. In some situations, providers may be asked to submit supplemental information through a medical registry or "no-pay" invoices, especially if routine administrative data do not provide enough detail to support outcome measures of interest. Examples of such supplemental information include patient satisfaction surveys, mobility assessments, and cognitive impairment scores.

Bundled payment requires the development of a target price for a bundle of services associated with an EOC. The target price can be derived from historical behavior, analyzed with the use of an EOC definition, or it can be set normatively based on a prescribed clinical protocol. An example of the behavioral approach would be to set the target price for stroke episodes at the median cost of episodes observed during a preceding two year period, stratifying the target by the DRG associated with the triggering hospital stay. The current EHIF proposal is an example of a more normative approach. Specifically, the current plan is to set the target price based on two years' historical fee-for-service utilization among stroke patients, but to supplement the historical record to adjust for missing and "insufficient" services based on clinical pathways developed for ischemic strokes. The target price will then be set as the average cost of episodes, pricing individual services at rates specified by the current fee schedule.

To calculate the bundled payment price, EHIF took into account all patients with a stroke in 2016 meeting the inclusion criteria for the bundle, and added all costs incurred within a year (including rehabilitation). To be considered the start of an episode, an admission needed to be preceded by a 360-day "clean period," meaning that there was no similar admission in the previous year. Since the current levels of rehabilitation are relatively low, EHIF is considering adjusting the price for rehabilitation based on new guidelines being developed. Currently the proposal is to calculate the price on the assumption that average inpatient rehabilitation should be 21 days and about 15% of patients will require outpatient rehabilitation. During the consultations, stakeholders identified several supply-side constraints related to rehabilitation, including low volume capacity. Successfully transitioning patients into post-acute rehabilitation and nursing services requires such services to be available. Some at-risk providers may have difficulty in accessing these services for their patients. At a minimum, EHIF may want to encourage applicants to describe the supply of post-acute services in their networks and to consider expanding training programs for rehabilitation and nursing services.

Determining the target price is a critical step in designing a bundled payment program. If the price is not properly aligned with the bundle definition, it can lead to provider resistance and discourage the delivery of clinically appropriate services. For this reason, it is important to benchmark target prices against historical data before providers are placed at risk. If historical data are considered incomplete or unreliable, then it is often reasonable to defer risk-bearing until additional data have been collected.

The episode definition itself may vary across outcome measures, even for the same clinical event. In the EHIF stroke pilot, for example, the clinical pathway may include primary care and social support services that are outside the scope of the services covered by the bundled payment *per se*. In effect, EHIF can use one episode definition to set the target price, and a more expansive definition to measure clinical outcomes, patient satisfaction, and, potentially, total cost of care. The written materials that EHIF has already shared with the World Bank identify process and clinical outcome measures that could be used in the stroke pilot, using indicators endorsed by the Council for Quality of Medicines Indicators, many already available in Estonia, and indicators of the ICHOM "stroke dataset". These include the percent of patients that:

- Die within 30 days of their acute onset;
- Undergo thrombolysis or thrombectomy;
- Receive a one-hour CT or MRI scan of the brain;
- Have carotid artery imaging performed with 96 hours of admission;
- Have a physiotherapist assess their mobility function within 48 hours of admission;
- Have their swallowing function studied within 24 hours of admission;
- Receive continuous oral anticoagulant therapy;
- Receive rehabilitation services;
- Die within 7 days, 30 days, 90 days or one year after discharge;
- Have an acute complication of care (Intracranial haemorrhage); and
- Report favourably on a variety of patient-reported outcome measures (PROMs).

These materials also indicate financial and process measures that could be useful for the pilot. These include length-of-stay and average cost measures per episode and by stage within episodes, as well as more focused utilization measures for the year following the trigger admission, such as the number of outpatient visits, the number of prescriptions, and the occurrence of repeat inpatient stays. The assessment of the choice of indicators for patients reported outcome is presented in Section 5.

3.6 Risk-adjusting outcomes

Health care analytics often involve comparisons between actual and expected performance, where the expectation reflects the demographic and clinical attributes of a patient as well as the environment or market in which services are delivered. "Risk-adjustment" is the term that is often used to characterize the way expected values are adjusted based on such factors.

Risk-adjustment is important in the context of the bundled payment program for several reasons.[13] Patients vary along many dimensions and risk-adjustment makes it possible to compare outcome measures across patients. The better the risk-adjustment model, the more likely clinicians and other stakeholders are to accept those comparisons as valid and reliable. Risk-adjustment also gives payer organizations more certainty about provider performance assessments and the impacts of their payment policies.

Bundled payment programs have used a variety of risk-adjustment strategies. One option is to sort episodes into categories and use group averages within those categories to set "expected outcomes," relying on the law of large numbers to control for difference across episodes within a category. This is the approach that EHIF is considering for its stroke pilot. Episodes would be grouped based on the age of the patient, and on how patients are treated during the triggering hospital stay (thrombolysis, thrombectomy, other treatment). EHIF would then set target prices based on group averages, adjusting for missing and insufficient services.

Another option is to develop a risk-adjustment model based on multivariate regressions estimated on historical utilization and claims data. The risk factors in these models can include patient demographics, as well as clinical information about pre-existing conditions, health-related behaviors, and prior use of medical services. The more complex approach strengthens analytic results but increases development and maintenance costs.[14]

[13] Courtney et al. (2018) and Ellimoottil et al. (2016) examine the potential consequences of not including risk-adjustments.

[14] EHIF also tested if co-morbidities, prior hospitalization, sex, age, or first service was associated with the episode cost. Among these, only sex and age were found to be statistically significant predictors of episode cost.

Regardless of the risk-adjustment method that is used, most operational bundled payment programs exclude certain episodes that appear to be so different from the typical case that they can be considered "outliers." Outliers can be defined financially or clinically. Financial outliers are defined in terms of the distribution of actual costs per episode. For example, extremely low-cost episodes can be excluded by setting a low-cost fixed-price low-cost threshold (e.g., EUR 400) or set based on percentiles or confidence intervals. High-cost thresholds are usually set based on statistical criteria. By contrast, clinical outliers involve cases where a patient presents unusual clinical challenges due to co-morbidities. Typically, operational programs define clinical outliers as episodes with specific co-morbidities. Common examples include HIV/AIDs, active cancer, and transplant recipients. Using a sample from 2016-17, EHIF examined potential outliers and did not identify any significant outliers.[15] However, stroke severity or daily physical activity prior to stroke onset were not available in EHIF's dataset and could not be included in this analysis. Similarly, it was not possible to differentiate a patient hospitalized for a complication requiring urgent care from a patient hospitalized for a recurring stroke. Therefore, both cases have been excluded using a "clean period" of 360 days from the price calculation as explained in Section 3.5. However, a recurring event will trigger a new episode. Patients treated for a recurring stroke might differ from other patients from a clinical and/or financial perspective and impact the performance of the providers. Therefore, EHIF should consider further analytics once the data on recurrence become available or treat these cases separately during the pilot if outliers are expected.

4. OPERATIONAL CONCEPTS IN IMPLEMENTATION

The previous section described the various components of an episode-based payment program, starting with the triggering event, the rules that govern the scope of the services that are grouped into the episode, and the associated financial and clinical outcome measures. While these components are important, the success of an episode-based payment program depends critically on decisions about how it is implemented. This section describes some of the key operational issues that EHIF will want to address as it moves forward with its stroke pilot program.

4.1 Clinical pathways

In most situations, episode-based payment programs are designed to create incentives that will drive changes in the way that target populations are treated. The goal is to improve efficiency and effectiveness in clinical decisions. For that reason, it is important for clinicians to be involved early in the planning process to ensure that episode definition is in line with the current state of the art and adjusted to local Estonian context as needed. It also helps providers to understand the goals of the payment program and to earn support for the effort. Finally, it helps to identify clinical leaders who can serve as allies and lines of communication during the implementation phase of the program.

Physician engagement often takes place around the development of a written clinical guidelines or pathway. Such documents summarize the current state of knowledge, describe how patients should be evaluated and treated, and potentially identify opportunities to deliver care more efficiently and effectively than current practices. The clinical pathway can become a "living document" to be modified and updated over time in response to changes in market conditions and technology. It also serves as the

[15] To identify clinical outliers, EHIF applied the Charlson co-morbidity index to test for the differences in prices. This index gives a high score for patients with AIDS, metastatic solid tumors, and other non-mild chronic conditions. EHIF did not observe statistically significant differences between patients with this index compared to other patients.

foundation for developing an episode definition and for determining process and outcome measures that will be needed to help change clinical practice and evaluate the impact of new payment policies.

As discussed in Section 2, EHIF has already been working with neurologists in Estonia on a variety of planning activities for the proposed stroke pilot program. Physician engagement has helped to create a shared understanding of what will be required to improve care for stroke patients. Specific steps include:

- Channeling more patients into specialized stroke units and centers;
- Adopting clinical guidelines for treating patients experiencing an acute ischemic stroke;
- Developing a care coordination function to coordinate services across providers during the episode of care; and
-
- Enhancing follow-up care for stroke patients by increasing the use of post-acute rehabilitation services.

Currently, Estonian providers are following European Stroke Organization guidelines, and most of the monitored clinical indicators are derived from these guidelines. As such, EHIF is able to use proxy measures for monitoring adherence to clinical guidelines.[16] Further work is needed to ensure effective implementation and adherence to clinical guidelines (including any needed modifications of European Stroke Organization guidelines based on the Estonian context).

4.2 Claims processing

One of the critical issues in developing a bundled payment program is how to manage the invoicing and claims management for services included in the bundle. There are two possible approaches. One approach requires an at-risk provider to submit a claim that covers all of the services in the episode; the payer reimburses at-risk providers, who in turn compensate other providers involved in the episode.[17] The alternative is to require all providers to submit claims for reimbursement exactly as they did prior to implementation of the bundled payment program. The payer then develops a retrospective analytic function in which they reconcile the actual costs of each episode against target prices and then resolve differences in a subsequent financial transaction.

While the first of these approaches may seem more compatible theoretically with the idea of "bundled payment", it tends to be difficult to implement on a practical level. It forces the at-risk provider to take on the role of an insurer or third-party payer, a role for which they may not be fully prepared. One simple problem is that the at-risk provider may not know whether a patient is part of a bundled payment program when the patient is admitted. Further, the insurance function involves many more activities than simply receiving invoices and sending money; other functions include network development, contracting, data management and reporting. These difficulties are one of the reasons that Model 4 in the BPCI program was not successful.

In practice, many bundled payment programs have found it easier to maintain historic claims management workflows and to build an analytic and reporting function on the back end of claims processing systems. This means that only one organization needs to make substantial changes to its IT. Providers continue to focus on treating patients while insurers continue to operate as usual. The only change is the creation of a periodic, retrospective reconciliation process. Note that this retrospective approach has the advantage of being compatible with a variety of episode-based payment models. This

[16] A patient record audit could also be conducted to measure adherence to European Stroke Organization guidelines.
[17] In some examples of bundled payment, providers who are not at risk are still required to submit "no-pay" insurance claims containing clinical information about the case and the services that they delivered.

enables payers to transition to bundled payment over time. It is our understanding that EHIF and its stakeholders have decided to use the retrospective approach for the stroke pilot, while potentially transitioning to a prospective model at some point in the future.

4.3 Reporting

Another key to success in bundled payment is to develop a set of reports that help at-risk providers identify opportunities to reduce costs and improve outcomes. Typically, such reports identify the patients for which the provider is responsible, the services they received, and the outcomes associated with their care. The reports compare the provider's individual performance to their peer group as well as to the standards articulated in clinical pathways. Such reports enable providers to understand the basis for their performance assessment, identify areas where their performance can improve, and engage in discussions with external partners with whom they work to deliver services.

The development of such reports does not depend upon the implementation of bundled payment. In fact, these types of reports are often developed well before the details of a bundled payment program are finalized. This strategy enables providers to understand how they are performing, in comparison to clinical standards and their peer group, during a baseline period so that they can prepare for – and take full advantage of – opportunities created by bundled payment programs.

4.4 Payment for improving coordination of care

There are several forms of payment that can improve the coordination of care; not all require payers to set a single fee in advance for the entire set of services needed to manage an episode of care. What different episode-based payment models have in common is the episode definition, a clinical pathway, outcome measurement, and reporting. In general, a payer should focus on designing a payment model that uses the episode framework to incentivize behaviors that it wants to encourage.

Episode-based pay-for-performance is one example of this approach. This model combines a basic fee-for-service payment system with incentive payments or adjustments in fee schedules based on how well providers perform against a set of episode-based performance objectives. In the case of stroke patients, for instance, each of the outcome measures identified in Section 3.5 could serve as the basis for a bonus payment or rate adjustment in the current EHIF fee schedule. Provider-specific reports could still be used to help guide at-risk providers to drive performance improvements in stroke care.[18]

A similar option is to reimburse at-risk providers for care coordination that spans different phases of treatment and provider settings. Care coordinators can be hired to oversee treatment throughout the episode, armed with information available from provider-specific reports. The cost of the care coordinators can be covered partially from a set, per-episode fee with additional monies available from a shared savings program or a bonus arrangement tied to achieving specified clinical objectives.[19]

Bundled payment, in which an at-risk provider receives a fixed payment in exchange for delivering all of the services associated with an episode of care is most appropriate in circumstances where the health plan or payer wants to reduce costs or achieve other efficiencies by rationalizing the delivery process (e.g., substituting home health services instead of admitting patients to a skilled nursing facility). Incentives to improve value can be enhanced by adjusting payments based on clinical performance, or even making bonus payments contingent on meeting clinical performance standards. It is our understanding that EHIF plans to implement the data collection and reporting necessary to measure clinical performance

[18] See McDonald and others (2012) for further discussion of integrating quality metrics into payment policy.

[19] Several participants in U.S. bundled payment programs have developed this case-manager role, labelled as a "nurse-navigator" or a "patient advocate."

(including patient-reported outcomes) during the stroke pilot, but it does not plan to adjust payment to providers based on clinical performance metrics. EHIF should recognize that this approach means that the objectives of the EHIF initiative and the incentives of the payment system may not be aligned properly. One way to mitigate this risk is to move aggressively on its PROMIS-10 and ICHOM standard set implementation and reporting for both internal monitoring and for provider benchmarking.[20]

4.5 Other operational issues

Several other aspects of episode-based payment should be mentioned in this discussion. One is documentation. Episode-based payment typically involves multiple providers operating in different environments, often independent of one another. Transparent, publicly available documentation ensures that all parties understand program goals, performance standards, and the processes that will be used in the program. Documentation should be sufficiently detailed that financial calculations and performance scores can be replicated from a common set of data.

Another issue that arises in any payment system that is based on clinical information has to do with updates and maintenance. It is helpful to agree on a maintenance schedule in advance of the implementation date, so that all parties understand the rules under which they will be operating, how long those rules will be in place, and what changes are likely to occur in the next update cycle.

Finally, complex payment systems benefit from an explicit governance and dispute resolution framework. Mistakes and misunderstandings occur, and it is helpful to agree on how disagreements will be resolved in advance. As parties gain experience with episode-based payments, a body of shared knowledge will develop, and the frequency of those disagreements will likely decline. However, the introduction of any new payment system often creates unforeseen issues that need to be addressed.

5. PATIENT REPORTED OUTCOMES METHODOLOGIES

To complement the suite of measures that will be used to assess the quality and outcomes of stroke care, a module of patient-reported outcomes will be included. Patient-reported outcomes are critical to fully understand health system performance and are increasingly finding their place in health sector performance frameworks internationally. This pilot offers a valuable opportunity, therefore, to build Estonia's experience in systematizing use of these important metrics.

There are, however, many tools available to measure patient-reported outcomes, whether relating to general health or specific conditions/procedures. In this pilot, EHIF has principally considered the choice between EQ-5D and PROMIS, both well-established and used internationally, opting in the end to use PROMIS. The EQ-5D tool comprises five questions relating to five dimensions of well-being: mobility, self-care, usual activities, pain/discomfort, and anxiety depression. The PROMIS comprises a core set of questions to assess common outcomes such as pain, fatigue, emotional distress, physical functioning, social role participation. Commonly used sets of questions within PROMIS include the Global Health Instrument (10 questions) and the PROMIS Adult Profile Instrument (29 questions).

PROMIS has several advantages. It was first developed as a research tool and, as such, has undergone rigorous testing of reliability and validity in diverse patient groups using sensitive statistical models, including item response theory. The PROMIS instrument yields a scaled numerical score, meaning that it can be used to detect meaningful change in a patient's health over time. This is important in the care of

[20] The following section contains a more detailed discussion of clinical performance measurement and data collection, especially in the context of patient-reported outcomes.

patients with stroke, where successful rehabilitation is a critical outcome (that health insurance funds may want to reward). Furthermore, PROMIS measures multiple domains of health and well-being (not just physical health). This is important for a complex, chronic condition such as stroke where mental health is substantially affected and requires specific care.

Two minor drawbacks to PROMIS should be mentioned. First, PROMIS is not as well-established internationally in routine clinical use as EQ-5D (perhaps because PROMIS was developed as a research instrument, rather than an instrument for patient or system management). This may limit EHIF's ability to benchmark Estonian scores with international comparators. However, this may not be a priority for EHIF. In any case, the situation is changing as increasing numbers of national and sub-national health systems opt to use PROMIS as a routine clinical measure; tools to convert between PROMIS and EQ-5D scores also exist, although they are only moderately successful (Revicki et al, 2009). Furthermore, ICHOM's adoption of PROMIS in its outcome set for stroke patients is likely to accelerate international use. Second, PROMIS yields a numerical score, in contrast to EQ-5D's visual-analogue score (like a thermometer, which can be easily understood by patients and caregivers). Straightforward algorithms, however, are available to convert PROMIS results into EQ-5D scores or equivalents. It should also be noted that comparative ease of use is *not* a drawback that disadvantages PROMIS, as is sometimes believed. Although the PROMIS Global Health instrument contains more questions than EQ-5D, the overall word-count (in English) is less, and both instruments take the same time (two minutes, on average) to complete.

EQ-5D, as already mentioned, has a good track-record of successful system-wide use in routine clinical care. It is a routinely and frequently used patient-reported outcome instrument in Canada, Sweden and England, for example – three dynamic and innovative health systems that are important benchmarks for Estonia. Disadvantages are, though, that EQ-5D predominantly measures physical health. It has also been found to be less valid, reliable and responsive to change in a patient's well-being over time than other instruments.

Overall, PROMIS is an appropriate patient reported outcome measurement tool for EHIF to choose for this pilot. Questions that EHIF should consider going forward, however, are:

- What further conditions will be included in future bundled-payment pilots, or other reforms that seek to measure and reward quality and outcomes?
- How important is international benchmarking to EHIF?
- What broader plans does EHIF have to integrate patient-reported experiences and outcomes into routinely collected health system statistics? How should this pilot fit into those plans?

EHIF's choice of instrument should be future proof and, ideally, internationally comparable. If hip/knee replacement, for example, are likely candidates for a future bundled-payment pilot, it is worth noting that EQ-5D is currently the more commonly used instrument in OECD health systems.

6. ASSESSMENT AND RECOMMENDATIONS

The EHIF and its partners in this effort have accomplished much good work in preparing to initiate a pilot program for stroke care in Estonia. There is a well-designed definition of a stroke episode pathway with clear rules for triggering the episode, gathering services around the initial hospital stay, terminating the episode, identifying at-risk providers, and calculating relevant outcome measures (both clinical and financial). The absence of a strong risk-adjustment model is understandable, given data limitations associated with missing and insufficient services, as well as the scope of the changes that EHIF wants to encourage. The risk-adjustment for clinical and financial performance should be refined as new data

collected during the pilot become available, for example on recurring stroke. Based on a review of the proposed pilot and understanding of other bundled payment programs, EHIF should consider the following suggestions for refining the current plan in Estonia.

6.1 ***Finalize the adoption of clinical guidelines***. EHIF has been working closely with the physician community throughout the development of the pilot program, and needs to continue that work to ensure that the guidelines covering the acute and post-acute care period that underlie the pilot are finalized and aligned with the bundle definition. It appears that the focus to date has been more on improving care coordination. Given the relatively high stroke mortality in Estonia, it is important to ensure that providers are delivering appropriate and coordinated care according to established guidelines and pathways.

6.2 ***Distinguish more clearly between the bundle used for payment purposes and the episode definition that will be used for performance measurement.*** EHIF has been properly cautious in defining the scope of the initial payment bundle, but there is no necessary reason to limit outcome measures or the evaluation of financial performance to services included in the payment bundle. EHIF can draw this distinction explicitly from the outset of the pilot.

6.3 ***Further develop and deploy provider-specific performance reports***. Publicizing performance data on hospital quality is a powerful tool and has been found to improve the quality of care by stimulating hospitals to improve their performance and introduce quality improvement activities (even in the absence of financial incentives) (Hibbard et al., 2003; Lindenauer et al., 2007; Fung et al., 2008). Performance reports enable providers to understand for which patients they are accountable, the services they receive and their collective performance in terms of clinical and financial outcomes. Performance should be presented in relation to standards associated with the clinical pathways and actual peer group outcomes, including patient reported outcomes. EHIF already publishes a set of quality indicators for neurology patients. These contain a limited number of outcomes, however, and no patient-reported outcomes. Continued development of this indicator set, in collaboration with the provider community, will enhance the impact of these reports.

6.4 ***Consider clinical as well as financial outliers in operationalizing the pilot.*** It is necessary to use historical data to develop performance standards and target prices, but data can be messy, and it is difficult to anticipate all situations that can create anomalous results. EHIF plans to exempt episodes that cost more than EUR 100,000 from the pilot, but it may also want to exclude certain patients entirely based on co-morbidities such as transplant status, HIV/AIDS, or active cancer.

6.5 ***Adopt a retrospective bundled payment approach.*** Based on international experience, paying at-risk providers and requiring them to distribute funds to other providers involved in stroke cases may create challenges for the pilot. There is also a risk that hospitals may not make all of the operational improvements needed for successes, given the temporary nature of the pilot. Leaving current payment processes in place, and simply adding a back-end reporting function onto the EHIF systems, will be simpler to implement. It is also consistent with a phased approach to implementation.

6.6 ***Consider phasing in the pilot program over a period of time.*** The proposed timeline is ambitious. EHIF could begin, for example, with a "reporting only" baseline period, followed by a period in which providers are incentivized financially for meeting certain clinical performance standards. At the same time, we note that EHIF is already supporting case-management through its innovation fund. The current plan appears to call for implementing a bundled-payment arrangement with two-sided financial risk without any adjustments for changes in outcomes or clinical processes. EHIF's intent with the proposed pilot is to pursue a multi-dimensional reform involving improvements in financial and clinical outcomes. The danger with the current plan is that incentives under bundled payment to pursue cost-savings may

discourage clinical improvements. A phased approach with an initial reporting-only period could help to mitigate the risks of a payment model that rewards savings without any adjustments for clinical performance. In the long run, it may be better to delay actual financial incentives tied to average costs per episode until EHIF is comfortable with provisions for clinical changes and reporting.

6.7 ***Develop a formal documentation system with version control that is available to EHIF and its at-risk providers.*** At a minimum, this system would include clinical pathways, episode definitions, specifications for calculating performance measures, and report templates.

6.8 ***Formalize agreements between EHIF and its at-risk providers regarding maintenance schedules and dispute resolution in the pilot program***. It is possible that such issues can be addressed under current arrangements. If not, it is a good idea to address them at the start of the pilot, when there is little urgency to the discussion.

7. POTENTIAL EVALUATION OPTIONS

To evaluate the pilot, EHIF could adopt a mixed methods approach, using qualitative and quantitative data. An interrupted time series (ITS), a quasi-experimental design, is particularly well suited for this type of intervention. In an ITS, a time series of a particular outcome of interest is used to establish a trend, which is interrupted by an intervention. The counterfactual, or the expected trend in the absence of an intervention, provides a comparison for the evaluation by estimating any changes after the introduction of the pilot. ITS has been frequently used to examine the effects of population-based interventions for which it is not possible or practical to conduct a randomized trial. In addition, administrative data can be used for an ITS evaluation.[21]

After controlling for the baseline trend and other covariates, the ITS models can be used to evaluate the impact of the bundled payment on pre-specified outcomes of interest. For the bundled payment pilot, the main outcomes of interest could include length of stay, 30 or 90-day readmissions, mortality, provision of thrombolysis or thrombectomy, number of outpatient visits, and whether the patient received rehabilitation care. The choice of outcomes will depend on the availability of historical claims data to establish a pre-intervention trend and the frequency of occurrence. ITS is most suitable for short-term outcomes that are expected to change relatively quickly after the introduction of the pilot or within an identified time period. For example, mortality may be a rare event and would not enable detecting statistically significant differences over a short period of time. Hospital claims data for select outcomes could be aggregated on a monthly basis to examine changes before and after the introduction of the bundled payment pilot. It is recommended that the pre-intervention period at a minimum include the two years prior to the pilot. ITS can be applied to the four hospitals that are participating in the bundled payment pilot and the innovation fund, however, due to small sample sizes it will be difficult to disentangle the effect of bundled payment separately from the innovation fund. Since hospitals will only begin to collect patient-reported outcome measures once the pilot begins, it will not be possible to include these measures in the ITS analysis.

As an alternative to the proposed ITS approach, EHIF could apply a difference-in-difference methodology to all six hospitals providing stroke care. To ensure validity, several assumptions must be met: intervention (i.e. enrollment in the pilot) is unrelated to outcome at baseline; treatment and control groups exhibit

[21] See Lagarde (2011) and Bernal et al. (2017) for a more detailed methodological explanation of ITS.

parallel trends in outcome over time (i.e., in the absence of the pilot, the difference in outcome between participating and non-participating hospitals is constant over time); the composition of the treatment and control groups does not change; and there are no spillover effects. To ensure that the parallel trends assumption holds, EHIF should inspect the trend in monthly outcomes during the past two years for both groups of hospitals. If the assumption holds, a difference-in-difference approach could be used to compare performance of the three selected hospitals versus the other hospitals not participating in the pilot. Randomization at the patient level (within each facility) is not suitable, as it may create confusion and dilute the incentives for providers. In addition, it may also raise ethical concerns for patients.

To evaluate the performance of hospitals selected for the service delivery innovation fund, EHIF could monitor monthly outcome performance indicators and benchmark the performance with other hospitals. Due to small sample sizes, it will not be possible to conduct a rigorous econometric evaluation of the service delivery innovation fund to identify the specific elements associated with observed changes. A process evaluation, supplemented by qualitative information, would be more appropriate and would help to understand the processes behind the piloted interventions.

In addition to the ITS analysis, EHIF could implement a patient experience survey to examine changes in patients' perceptions and experiences with stroke care as a result of the bundled payment pilot and organizational changes.[22] The survey should be implemented at two points in time: before the introduction (or at the very beginning) of the bundled payment pilot to establish a baseline, and at the end of the pilot. The gaps in patient experience identified through the survey could be complemented with qualitative data.

The qualitative component could include structured interviews with patients and providers to assess coordination of care, patient pathways, and implementation challenges. For hospitals receiving additional funds through the service delivery innovation fund, the qualitative interviews would assist EHIF in identifying the procedural elements associated with observed changes.

Importantly, the evaluation should also attempt to examine the unintended consequences of the pilot. In particular, given the limited capacities to deliver rehabilitation care, it would be important to monitor access to rehabilitation care for non-stroke patients.

The evaluation should be conducted by an independent third party. Clear guidelines and triggers for ending the pilot prematurely in case of declines in quality of care should also be established.

[22] It is important to distinguish between patient experience and patient satisfaction. While patient satisfaction surveys tend to measure whether patient's expectations were met and how the patient felt about their care, patient experience surveys attempt to measure whether something that should happen in a health care setting actually happened or how often it happened. As such, patient experience surveys are an important instrument to measure process quality. A number of instruments are available to measure patient experience (Beattie et al., 2015). For more information see: https://www.ahrq.gov/cahps/about-cahps/patient-experience/index.html

REFERENCES

Agarwal, R., Liao, J. M., Gupta, A., & Navathe, A. S. (2020). The Impact Of Bundled Payment On Health Care Spending, Utilization, And Quality: A Systematic Review: A systematic review of the impact on spending, utilization, and quality outcomes from three Centers for Medicare and Medicaid Services bundled payment programs. *Health Affairs*, *39*(1), 50-57.

Barchi D., de Brantes F., Hughes G., Nguyen T., Poulsen G., Redfearn D. (2014) Episode Analytics: Essential Tools for New Healthcare Models. Institute for Health Technology Transformation. Accessed on 16 October 2019 at https://www.sas.com/content/dam/SAS/en_us/doc/whitepaper2/iht2-episode-analytics-106886.pdf .

Beattie, M., Murphy, D. J., Atherton, I., & Lauder, W. (2015). Instruments to measure patient experience of healthcare quality in hospitals: a systematic review. Systematic reviews, 4(1), 97.

Berlinguet, M., Vertrees, J., Freedman, R., D'Andrea, R., Tinker, A. (2007). Case-mix analysis across patient populations and boundaries: a refined classification system designed specifically for international use. Accessed on 04 September 2019 at http://multimedia.3m.com/mws/media/425433O/white-paper-international-refined-ir-drgs-01-07.pdf .

Bernal, J. L., Cummins, S., & Gasparrini, A. (2017). Interrupted time series regression for the evaluation of public health interventions: a tutorial. International journal of epidemiology, 46(1), 348-355.

Centers for Medicare and Medicaid Services (2019). Design and development of the Diagnosis-Related Group (DRGs). ICD-10-CM/PCS MS-DRG v36.0 Definitions Manual. accessed on 04 September 2019 at https://www.cms.gov/ICD10Manual/version36-fullcode-cms/fullcode_cms/P0001.html.

Chernew, M. E., Conway, P. H., & Frakt, A. B. (2020). Transforming Medicare's Payment Systems: Progress Shaped By The ACA: A narrative review of Affordable Care Act payment reforms. Health Affairs, 39(3), 413-420.

Courtney, P. M., Bohl, D. D., Lau, E. C., Ong, K. L., Jacobs, J. J., & Della Valle, C. J. (2018). Risk adjustment is necessary in Medicare bundled payment models for total hip and knee arthroplasty. The Journal of arthroplasty, 33(8), 2368-2375.

de Bakker DH, Struijs JN, Baan CB, Raams J, de Wildt JE, Vrijhoef HJ, Schut FT (2012). Early Results from adoption of bundled payment for diabetes care in the Netherlands show improvement in care coordination. Health Aff (Millwood). 2012 Feb;31(2):426-33. doi: 10.1377/hlthaff.2011.0912.

Fetter, R.B., Shin, Y., Freeman, J.L., Averill, R.F., Thompson, J.D. (1980). Casemix definition by diagnosis-related groups. Medical Care, 18(2):1–53.

Fung, C. H., Lim, Y. W., Mattke, S., Damberg, C., & Shekelle, P. G. (2008). Systematic review: the evidence that publishing patient care performance data improves quality of care. Annals of internal medicine, 148(2), 111-123

Gross-Paju, K., Orav, K., Adlas, R., Thomson, U., Jaakmees, H., Kannel, K., . . . Ütt, S. (2017). Changes in Quality of Prehospital Care and Time Delays in Acute Stroke in Tallinn, Estonia from 2005 to 2016. Prehospital and Disaster Medicine, 32(S1), S94-S95. doi:10.1017/S1049023X17002436

Hibbard, J. H., Stockard, J., & Tusler, M. (2003). Does publicizing hospital performance stimulate quality improvement efforts? Health Affairs, 22(2), 84-94.

Hussein, Saeed & Moonesar Ph.D. R.D., Immanuel Azaad. (2018). Abu Dhabi Healthcare System-Diagnosis Related Group Perspective. International Journal of Medical, Pharmacy and Drug Research. 2. 10.22161/ijmpd.2.1.2.

Hussey P, Sorbero M, Mehrotra A, Liu H, Damberg C (2009). Using episodes of care as a basis for performance measurement and payment: moving from concept to practice. Health Aff (Millwood). 2009 Sep–Oct; 28(5): 1406–1417.

Lindenauer, P. K., Remus, D., Roman, S., Rothberg, M. B., Benjamin, E. M., Ma, A., & Bratzler, D. W. (2007). Public reporting and pay for performance in hospital quality improvement. New England Journal of Medicine, 356(5), 486-496.

McDonald R, Zaidi S, Todd S, Konteh F, Hussain K, Roe J, Allen T, Fischera E, Sutton M (2012). A qualitative and quantitative evaluation of the introduction of best practice tariffs. Department of Health (England).

Miika Linna and Martti Virtanen (2011). Chapter 16: NordDRG: The benefits of coordination in Diagnosis-Related Groups in Europe: moving towards transparency, efficiency and quality in hospitals. European observatory on health systems and policies. Accessed 04 September 2019 at http://eurodrg.projects.tu-berlin.de/publications/DRGbook/Ch16_Linna.pdf .

National Institute for Health Development. Health Statistics and Health Research Database. http: //pxweb.tai.ee /.

OECD (2016). Better Ways to Pay for Health Care, OECD Health Policy Studies, OECD Publishing, Paris.

Revicki, D.A., Kawata, A.K., Harnam, N., Chen, W.H., Hays, R.D. & Cella, D. (2009). Predicting EuroQol (EQ-5D) scores from the patient-reported outcomes measurement information system (PROMIS) global items and domain item banks in a United States sample. Quality of Life Research, 18(6), pp.783-791.

Ridgely, M. S., De Vries, D., Bozic, K. J., & Hussey, P. S. (2014). Bundled payment fails to gain a foothold in California: the experience of the IHA bundled payment demonstration. Health Affairs, 33(8), 1345-1352.

Salinas J, Sprinkhuizen SM, Ackerson T, et al. An International Standard Set of Patient-Centered Outcome Measures After Stroke. *Stroke*. 2016;47(1):180–186. doi:10.1161/STROKEAHA.115.010898.

Struijs, J. N., de Vries, E. F., Baan, C. A., van Gils, P. F., & Rosenthal, M. B. (2020). Bundled-Payment Models Around the World: How They Work and What Their Impact Has Been. https://www.commonwealthfund.org/sites/default/files/2020-04/Struijs_bundled_payment_models_around_world_ib.pdf

The Lewin Group (2018). CMS bundled payments for care improvement initiative, Models 2-4: Year 5 Evaluation and Monitoring Annual Report. Accessed 04 September 2019 at https://downloads.cms.gov/files/cmmi/bpci-models2-4-yr5evalrpt.pdf

Wohlin J, Stalberg H, Strom O, Rolfson O, Willers C, Brommels M (2017). Effects of introducing bundled payment and patients' choice of provider for elective hip and knee replacements in Stockholm county.

World Health Organization (2013). Bulletin of the World Health Organization (2013) 91:746-756A. doi:http://dx.doi.org/10.2471/BLT.12.115931. Accessed 04 September 2019.