World Health Organization
Centre for Health Development

Scoping review on alternative purchasing arrangements

Impact of purchasing, service delivery and institutional design on quality of chronic care

Cassandra Simmons, Mirjam Pot and Kai Leichsenring

World Health Organization
Centre for Health Development

Scoping review on alternative purchasing arrangements

Impact of purchasing, service delivery and institutional design on quality of chronic care

Cassandra Simmons, Mirjam Pot and Kai Leichsenring

Scoping review on alternative purchasing arrangements: impact of purchasing, service delivery and institutional design on quality of chronic care/Cassandra Simmons, Mirjam Pot, Kai Leichsenring

ISBN 978-92-4-008397-4 (electronic version)

ISBN 978-92-4-008398-1 (print version)

Suggested citation. Simmons C, Pot M, Leichsenring K. Scoping review on alternative purchasing arrangements: impact of purchasing, service delivery and institutional design on quality of chronic care. Geneva: World Health Organization; 2023. Licence: CC BY-NC-SA 3.0 IGO.

Cataloguing-in-Publication (CIP) data. CIP data are available at https://iris.who.int/.

Design and artwork by Howdy LLP.

Contents

Acknowledgements

The authors of this report are Cassandra Simmons, Mirjam Pot and Kai Leichsenring, of the European Centre for Social Welfare Policy and Research, Vienna, Austria.

Sarah L Barber, World Health Organization (WHO) Centre for Health Development (WHO Kobe Centre – WKC), Kobe, Japan, Inke Mathauer, WHO Department of Health Systems Governance and Financing, Geneva, Switzerland, and Megumi Rosenberg, WKC, contributed to the study design and led the technical review and revision.

This research was commissioned by WKC.

The Kobe Group, which includes Hyogo Prefecture, Kobe City, the Kobe Chamber of Commerce and Kobe Steel, in Japan, contributed financially to the development and production of this research report.

Abbreviations

ACO	accountable care organization
AMSTAR	A MeaSurement Tool to Assess systematic Reviews
AQC	Alternative Quality Contract
BPCI-A	Bundled Payments for Care Improvement – Advanced
CCO	coordinated care organization
CCFE	Care Chain Frail Elderly programme
ED	emergency department
FFS	fee for service
GP	general practitioner
GRADE	Grading of Recommendations Assessment, Development and Evaluation
HMO	health maintenance organization
JBI	Joanna Briggs Institute
MCO	managed care organization
NCDs	noncommunicable diseases
P4P	pay for performance
PCMH	patient-centred medical home
PCP	primary care provider
PICO	Population, Intervention, Comparison, Outcomes
POP	paediatric-only plan
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-analyses
PSN	provider service network
RCCO	regional collaborative care organization
RCT	randomized controlled trial
WHO	World Health Organization

Glossary

Accountable Care Organizations (ACOs): ACOs are provider-led networks of health care professionals in different settings (e.g. physicians, hospitals and others) that agree to be accountable for the costs and quality of care for the assigned population of patients under their coverage, thus directly taking on the financial risks and rewards of their care. As a result, ACOs are responsible for coordinating, managing and providing services to their patients. ACOs depend on their ability to incentivize hospitals, primary care providers (PCPs), postacute care facilities and other providers to form partnerships *(1)*. ACOs may be contracted either entirely publicly or by private payers (i.e. commercial insurers), or by a combination of both. Contracts and organizational structures can vary widely across ACOs.

Alternative purchasing arrangements: Alternative purchasing arrangements, similar to alternative payment models, the term used in the literature in the United States, refer to purchasing models that go beyond traditional fee-for-service (FFS) payment for individual services and instead reimburse providers based on the quality and coordination of services. Alternative purchasing arrangements may include capitation, global budgets, shared savings and shared risk, and more.

Blended capitation (also known as partial capitation): Blended capitation refers to a population-based model of purchasing in which reimbursement is provided per enrolled patient for a predetermined time for the provision of certain services included within a bundle. Services outside of the bundle are likely to be reimbursed on an FFS basis. In the scope of this review, the blended capitation model is found in Ontario, Canada. A related model, also found in Ontario, Canada, is the blended FFS model, which reimburses providers on a capitation basis for a more limited set of services, with FFS comprising the remaining and larger portion of reimbursement.

Bundled payments: Bundled payments refer to the reimbursement of providers for a defined bundle of services for a defined episode of care, based on expected costs. Bundled payments can cover a range of services offered by different types of providers and across different health care settings (e.g. hospital-based care, care by PCPs, care in postacute facilities). Bundled payments are typically defined for patients who have chronic diseases or conditions, but may also relate to care episodes stemming from surgical procedures and aftercare.

Bundled Payments for Care Improvement – Advanced (BPCI-A): The BPCI-A programme is the successor to the Bundled Payments for

Care Improvement initiative in the USA. Similar to the original programme, the BPCI-A initiative requires providers to take accountability for their care of patients; therefore, providers receive a single retrospective payment aimed at covering care provided within a certain time frame after the onset of the care episode. A clinical episode is defined either by inpatient admission to an acute care hospital or the start of outpatient procedures within a hospital. A number of types of care are included, ranging from cardiac care to gastrointestinal surgery and aftercare to critical care, spinal procedures and orthopaedics. In the BPCI-A programme, payment is tied to performance on certain quality measures. Further information about the programme can be found on the website of the Centers for Medicare and Medicaid Services (https://innovation.cms.gov/innovation-models/bpci-advanced#episodes).

Capitation: Capitation refers to reimbursement on the basis of per-patient fees designed to cover all services for a predetermined amount of time. Capitation models will often adjust for additional risks associated with patients who are likely to require more care than others (e.g. on the basis of gender, age, chronic disease). Variations of the capitation model may be based on the type of care being provided (i.e. primary care, secondary care) or on the extent or scope of services covered (see Blended capitation).

Comprehensive care: Comprehensive care refers to the coordinated delivery of services, provided either directly or indirectly, along a continuum based on all of the needs of a patient. These may include "health promotion, prevention, diagnosis and treatment of common conditions, referral to other providers, management of chronic conditions, rehabilitation, palliative care and, in some models, social services" *(2)*.

Global budget: Global budgets, more commonly found in hospitals, are population-based models that provide a lump sum for a defined set of services to be provided for a defined period of time. The method for determining reimbursement can be based on revenues from previous years; anticipated revenue, given population details; or even capitation, based on the needs of the population.

Health maintenance organizations (HMOs): HMOs are a type of managed care insurance plan that covers comprehensive care provided by doctors within the contracted HMO network or medical or physician group. Out-of-network care is typically not covered unless part of emergency treatment. HMOs take on the risks for their voluntarily enrolled patient population and coordinate and manage the delivery of their services. More information about different types of HMO models can be found on the website of the United States Centers for Disease Control and Prevention (https://www.cdc.gov/nchs/hus/sources-definitions/hmo.htm).

Managed care: Managed care refers to a health insurance approach that integrates the financing and delivery of services in a way that manages costs, utilization and the quality of care. Common features tend to include a network of providers working together, financial incentives for the network, management of utilization and quality of care, and financial incentives or risk-sharing by the provider *(3)*. The design of payment in managed care arrangements can vary substantially, from retrospective FFS to capitation based on a per-member per-month rate, with varying degrees of additional financial incentives.

Related to this, **managed care organizations** (MCOs) are contracted organizations that accept payment from US Medicaid agencies to provide services to Medicaid beneficiaries on Medicaid's behalf. More information is provided on the official Medicaid website (https://www.medicaid.gov/medicaid/managed-care/index.html).

Medicare (United States): Medicare is the largest federal health insurance programme in the United States, covering people aged 65 years and older, younger individuals with disabilities, and individuals with end-stage renal disease. Traditional Medicare consists of several parts. Part A provides hospital insurance (and covers hospital stays, care in a skilled nursing facility, hospice care and some home health care); Part B provides medical insurance (and covers doctors' services, outpatient care, medical supplies and preventive services); and Part D consists of a voluntary outpatient prescription medication coverage plan for some of the costs of medications (including recommended immunizations and vaccines). Beneficiaries may receive coverage through traditional Medicare or under a Medicare Advantage plan (Part C), under which beneficiaries enrol in a private HMO or MCO. Beneficiaries can also be covered by both Medicare and private insurance (i.e. through employment-related coverage). More information about Medicare can be found on the official Medicare website (https://www.medicare.gov/what-medicare-covers/your-medicare-coverage-choices/whats-medicare).

Medicaid (United States): Similar to Medicare, Medicaid is a federally mandated and state-administered health insurance plan in the United States, providing health coverage to low-income adults, children, pregnant women, older people and those with disabilities. While benefits and eligibility may vary from state to state, states must provide certain mandatory benefits, such as inpatient and outpatient hospital services, physicians' services, laboratory and X-ray services, and home health services, among others. Other services are optional, such as prescription medications, case management, physical therapy and occupational therapy. More information can be found on the official Medicaid website (https://www.medicaid.gov/medicaid/benefits/index.html).

Patient-centred medical home (PCMH): Also referred to as medical homes in some countries, PCMHs are a team-based health care delivery model led by a health care provider who coordinates a patient's care on their behalf. The provider takes on responsibility for the patient's care and coordinates and arranges the necessary services with other providers and across settings. Originating in the United States, the core elements of the model include primary care access, continuity, comprehensiveness and coordination of care, as well as office practice innovations and reimbursement reform *(4)*.

More specifically, the key principles of the delivery model tend to include:

- patients developing an ongoing relationship with a physician;
- a physician-led team taking responsibility for the patient's care;
- a whole-person orientation;
- coordinated care or integrated care, or both;
- continual improvements to quality and safety;
- enhanced access to care for patients when needed offered through a variety of measures, such as the use of health information technology and by offering after-hours access;
- payment recognizing the added value of the care beyond the FFS.

However, there is substantial variation in terms of characteristics and service delivery among PCMHs, with practices varying in size, composition, the type of patients they serve (e.g. general or certain vulnerable groups), among others.

In the USA, eligibility criteria for being recognized as a PCMH were developed by the National Committee for Quality Assurance in 2008 and subsequently revised in 2011 *(4)*. Other accreditations also exist across the country. In practice, there are many versions of the PCMH, with many arrangements incorporating elements of the PCMH without any official accreditation.

Purchasers of PCMH services can be either public or private. Most PCMHs voluntarily choose to become certified or recognized through an official national accreditation body, health plan or state agency and, therefore, initiate the process. In a more limited number of cases, a purchaser may require providers to become an accredited PCMH practice, such as major payers (e.g. Medicaid, private insurance). For example, the US Centers for Medicare and Medicaid Innovation has implemented alternative payment models aimed at supporting PCMH implementation.

The PCMH model is widespread across the USA with more than 10 000 practices and 50 000 physicians recognized by the most common accreditation programme, the National Committee for

Quality Assurance's PCMH Recognition programme. While the prominence of PCMHs varies across states, New York has a high number of PCMHs, with previous figures finding one quarter of all PCPs recognized as PCMHs *(5)*.

Shared savings and shared risk: Shared-savings models tend to comprise a group of providers that collectively agree to be responsible for the care provided to a defined population of patients. In terms of reimbursement, a target is set for the budget, and if the costs of care are less than the budget, the providers are eligible to share the savings with the purchaser. This rate of award based on the savings may depend on certain quality or performance measures being achieved.

Similarly, a shared-risk model requires providers to be accountable for any overspending in the budget. If providers spend in excess of the budget, they may either have to be responsible for the excess costs entirely or face a penalty and share the additional costs with the payer. The penalty associated with overspending may also be based on measures of quality. These models can vary in terms of the types of services and providers included within the budget.

References

1. Moy HP, Giardino AP, Varacallo M. Accountable care organization [website]. Treasure Island (FL): StatPearls; 2023 (https://www.statpearls.com/point-of-care/34821, accessed 31 July 2023).

2. Haggerty JL, Beaulieu MD, Pineault R, Burge F, Lévesque JF, Santor DA, et al. Comprehensiveness of care from the patient perspective: comparison of primary healthcare evaluation instruments. Healthc Policy. 2011;7:154-66. PMID:23205042.

3. Giardino AP, De Jesus O. Managed care [website]. Treasure Island (FL): StatPearls; 2021 (https://www.statpearls.com/point-of-care/106947, accessed 31 July 2023).

4. Arend J, Tsang-Quinn J, Levine C, Thomas D. The patient-centered medical home: history, components, and review of the evidence. Mt Sinai J Med. 2012;79(4):433–50. doi:10.1002/msj.21326.

5. National Committee for Quality Assurance. The future of patient-centered medical homes: foundation for a better health care system. New York: John A Hartford Foundation; 2014 (https://www.johnahartford.org/images/uploads/reports/The_Future_of_PCMH-NCQA_White_Paper.pdf, accessed 31 July 2023).

Executive summary

This scoping review aimed to identify alternative purchasing arrangements (apart from fee for service [FFS] and pay for performance [P4P]) and their associated service delivery models that countries have used to improve the quality of chronic care, as well as to assess how these purchasing arrangements have impacted the quality of care.

A scoping review of academic and grey literature published between 2013 and 2022 was carried out on alternative purchasing arrangements and their impact on the quality of chronic care using causal inference methods. The analytical approach consisted of clustering articles around similar purchasing arrangements or programmes, summarizing the evidence across the clusters, and highlighting the service delivery aspects and institutional design features that potentially impacted the quality of care.

The scoping review included 51 reports and revealed 5 distinct types of purchasing arrangements: capitation and global budgets (*n* = 22), bundled payments (*n* = 2), shared savings or shared risk (*n* = 8), pay-for-coordination (*n* = 12) and blended capitation (*n* = 4). Three articles compared different types of purchasing arrangements. Some articles focused on particular programmes, including in the global budget cluster, in the United States, the Massachusetts-based Alternative Quality Contract (*n* = 6) and the Maryland All-Payer Model (*n* = 4); and in the blended capitation cluster, in Canada, the Ontario, Family Health Organization (*n* = 4). Most articles focused on purchasing arrangements in the United States (*n* = 36), followed by those in China (*n* = 5), Canada (*n* = 4), Germany (*n* = 3) and Netherlands (Kingdom of the) (*n* = 2). Articles varied in terms of the type of purchaser involved, with most focusing on a single public purchaser (*n* = 18), followed by multiple payers (*n* = 17) and single private purchasers (*n* = 13), while the purchaser was unclear in three articles. Nearly 70% of purchasing arrangements stipulated the dual aims of improving the quality of care and reducing expenditures. Quality indicators frequently measured service utilization, chronic disease management and prevention, and health outcomes. In all clusters a primarily positive impact was found on the quality of care; however, the certainty of evidence was generally low due to selection bias among providers and patients' participation.

The highest certainty of evidence of a positive impact was found for capitation and global budget arrangements, shared savings and shared risk, and pay-for-coordination models, which were often implemented as part of accountable care organizations, patient-centred medical homes and integrated care models, as well as comprehensive care provided by general practitioners. Where

Scoping review on alternative purchasing arrangements: impact of purchasing, service delivery and institutional design on quality of chronic care

impact was limited across purchasing arrangements, the literature suggested this related to incentives that were incompatible across settings and levels, limited risks to providers at the level of decision-making, unequal requirements for public versus private purchasers, and insufficient value of rewards to incentivize providers. Population-based purchasing arrangements, even in the absence of explicit performance incentives, may incentivize providers to think through their decisions more carefully, to collaborate across settings, to take on a gatekeeping role and to improve the continuity of care to reduce future costs.

This research provided information to the WHO and OECD joint publication *Purchasing for quality chronic care: summary report*.

1

Background

1

Scoping review on alternative purchasing arrangements: impact of purchasing, service delivery and institutional design on quality of chronic care

1.1 Increasing prevalence of noncommunicable diseases, and purchasing as a potential mechanism for improving quality of care

Trends in healthy life expectancy indicate that people are living longer lives but in poorer health, with noncommunicable diseases (NCDs) increasingly contributing a larger share of the burden of years lived with disability *(1)*. Today, chronic diseases remain the number one cause of death and disability in the world *(2)*. As mortality due to communicable diseases has declined over the years, deaths attributed to NCDs have comprised an increasing share of all deaths globally, as high as 74% annually *(2)*. Of deaths caused by NCDs, about 80% are caused by four disease clusters, which include cardiovascular diseases, cancers, chronic respiratory diseases (e.g. chronic obstructive pulmonary disease) and diabetes *(3)*. Chronic diseases particularly pose a challenge for low- and middle-income countries, with more than three quarters of deaths attributed to NCDs occurring in these countries *(2)*. Although the prevalence has remained stable during the past decade, multimorbidity has also become prevalent over the past two decades *(4)*, and will increasingly pose problems for delivering quality care in fragmented health systems that are geared towards treating single diseases.

Not only can NCDs be financially burdensome to households having to pay out of pocket for care, particular in low- and middle-income countries *(5)* but also they are costly in terms of public health care expenditure, with some major clusters of NCDs (such as cardiovascular diseases) comprising as much as 16.5% of national health care budgets *(6)*. Chronic diseases can also severely impact one's quality of life and contribute to the loss of opportunities for individuals in terms of human capital. Given the increasing relevance of chronic diseases globally and their associated societal, economic and health expenditure costs, countries are posed the challenge of managing public spending on health while ensuring the delivery of high-quality care. With this challenge, alongside population ageing and escalating demands for care, countries have faced increasing pressure to reform the delivery of health care and the financing of health care systems to become more efficient, sustainable and value-driven, all while improving or maintaining the quality of care. Countries must, therefore, reorganize health care delivery systems in a financially sustainable way that better meets the needs of patients with chronic disease.

Organizational structures and financial mechanisms that incentivize stakeholders to take positive action have been recognized as key levers for enabling and improving the quality of care *(7, 8)*. Thus,

new payment methods in high-income countries have been implemented to improve the coordination of services across providers, reduce fragmentation and improve the value of care *(9)* by rewarding providers based on the quality and value of care and the integration of services, with the aim of improving the quality and efficiency of care. The success of these arrangements has relied strongly on creating appropriate incentives (financial and nonfinancial), organizational structures and measures that encourage stakeholders to work together to provide quality care for chronic diseases *(10)*. Several financial and nonfinancial mechanisms have been implemented in different countries to deliver of health care with the aim of improving quality. Redesigning how funds are allocated from purchasers to providers has been one major way through which purchasers have tried to promote quality in care. Another way is through the incorporation of contract terms that include certain quality standards and requirements or the use of quality specifications for the performance or the delivery of care that must be achieved to receive payment *(11)*. Other mechanisms include nonfinancial incentives, such as through quality monitoring efforts, performance profiling, public reporting and the use of other contractual obligations *(10, 12)*.

1.2 Aim of this scoping review

Despite the vast amount of literature about alternative purchasing arrangements and their impacts on care quality more generally, evidence is lacking about the mechanisms through which these instruments can improve the quality of chronic care specifically and, therefore, on their applicability to other contexts. The overarching aim of this study is to provide an overview of the evidence about purchasing arrangements and the accompanying service delivery models that have been used to improve the quality of care for patients with chronic diseases.

We define purchasing arrangements as any institutional arrangement designed to allocate pooled funds to health care providers for the services they offer *(13)*. We define chronic diseases as NCDs according to the World Health Organization's (WHO's) definition, which refers to diseases of long duration and generally slow progression that cannot be passed from person to person and are the result of genetic, physiological, environmental and behavioural factors *(2)*. The four main groups of NCDs are cardiovascular diseases (e.g. heart attack, stroke), cancers, respiratory diseases (e.g. chronic obstructive pulmonary disease, asthma) and diabetes. Our definition also extends to communicable diseases, such as HIV and AIDS, for which survival rates have improved substantially during the past decades, as well as mental

disorders and disabilities *(14)*. We also recognize that many people live with more than one chronic disease.

We define quality, in line with WHO's definition, as health services that are effective, safe and people-centred, as well as timely, equitable, integrated and efficient *(15)*. In practice, quality can be measured in a variety of ways, most commonly attributed to Donabedian's three pillars *(16)* of structure, process and outcome measures. We recognize that while improved health and clinical outcomes are an indication of improved care quality, not all quality improvements lead to improved outcomes and are validated in having a defined pathway towards improving health outcomes *(17)*. For example, an increase in preventive services and screenings has the potential to lead to improved health outcomes, but these are not guaranteed. In light of this, we take a broad view of quality and do not impose restrictions on the type of quality indicators studied. Instead, we report on the quality of health services for chronic conditions as operationalized and measured by the papers included in the review. As a result, in this study quality is captured through a combination of indicators based on utilization, process (e.g. chronic disease management, preventive services) and patient-perceived measures. This broad view of quality also means that purchasing arrangements aiming to incentivize the coordination and integration of services, or strengthen primary care, are also relevant, even in the absence of explicit quality or performance-based financial incentives, as they may lead to improved care quality.

In order to achieve the aim of this study, we address the questions highlighted in Box 1.

Box. 1. Research questions to be answered through this scoping review

1. Apart from fee-for-service and pay-for-performance schemes, what alternative purchasing arrangements have countries used to improve the quality of chronic care?
2. How have these purchasing arrangements influenced the quality of chronic care?
3. What service delivery, institutional design features and other supportive elements have been implemented alongside the purchasing arrangements?

1

Background

Despite the vast amount of literature about purchasing arrangements, authors have seldom focused on service delivery in conjunction with payment reforms. This gap in the literature ignores the fact that purchasing reforms are often implemented in conjunction with service delivery reforms or are implemented with the intention of influencing service delivery in a particular way. A purchasing arrangement cannot be considered as an isolated instrument but rather as a key component within a multipronged approach to affecting care delivery *(12)*. Therefore, one key objective of this review, which sets it apart from past literature reviews, is its focus on extracting details about the service delivery model in order to highlight the elements potentially impacting the quality of chronic care. The focus of the scoping review was also to examine more integrated approaches to purchasing, given that much of the literature about purchasing arrangements and their impact on the quality of chronic care has been related to pay-for-performance (P4P) approaches and much less on other alternative financing mechanisms *(12)*.

2

Methods

2

Scoping review on alternative purchasing arrangements: impact of purchasing, service delivery and institutional design on quality of chronic care

2.1 Approach

The results presented in this report are based on a scoping review of the academic and grey literature about purchasing arrangements and an examination of their relationship with service delivery models and the quality of chronic care delivered. The scoping review aims to provide a summary of the evidence about how countries have used alternative purchasing arrangements, beyond fee-for-service (FFS) and individualized P4P measures to strengthen service models to deliver care for people with NCDs. Table 1 highlights the scope of this review, following the Patient, Intervention, Comparison and Outcome (PICO) approach to systematic reviews.

Table 1. Patient, Intervention, Comparison and Outcome (PICO) framework used for this scoping review

Category	Description
Patient, population or problem	People living with or receiving treatment and care for NCDs or chronic illness
Intervention or exposure	Types of purchasing arrangements for NCDs, design of purchasing arrangement, rewards (e.g. financial and nonfinancial, governance structures), service delivery, institutional design features, supportive elements
Comparison	Fee-for-service model, financing model prior to implementation of purchasing arrangement, similar purchasing arrangements in other locations
Outcome	Quality of chronic care received by patient

NCDs: noncommunicable diseases.

The protocol for this scoping review was developed based on guidance from the Joanna Briggs Institute (JBI) to ensure adherence to quality standards, and follows the AMSTAR (A MeaSurement Tool to Assess systematic Reviews) checklist, as far as criteria for scoping and systematic reviews align *(18, 19)*. The final research protocol was published and is accessible via the Open Science Framework Registries *(20)*. A more detailed overview of the methodology is available in Annex 1.

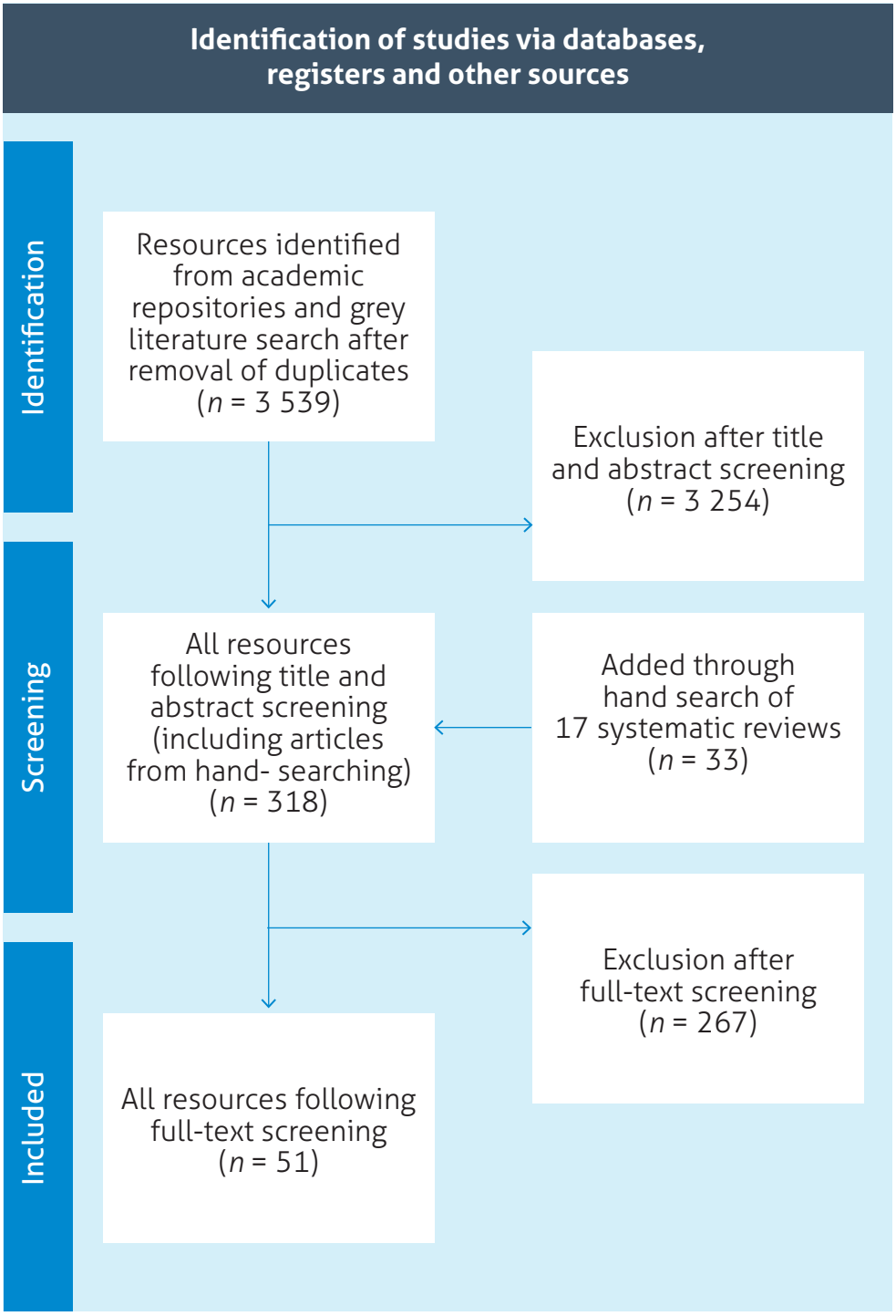
2.2 Search strategy

We adopted a three-stage strategy to identify relevant literature: (i) a systematic search of five academic repositories – PubMed, Web of Science, Cochrane Database of Systematic Reviews, Scopus and Google Scholar; (ii) a targeted search of the grey literature using Google incognito mode to reduce potential biases associated with the researchers' geographical location and their previous searches, and searches of relevant organizations' repositories (i.e. Organisation for Economic Co-operation and Development, the World Bank, WHO); and (iii) a hand-search of the reference lists of all relevant systematic reviews identified through our searches for relevant single-study papers. All searches were conducted in November 2022.

2.3 Screening process

After the removal of duplicate records, the searches of the academic and grey literature resulted in 3539 resources (Fig. 1). Before starting screening of the full title and abstract of each record, two researchers screened 20 titles using the criteria described in Table 2 finding that they agreed on more than 75% of the resources trialled *(18)* and then further clarified the inclusion and exclusion criteria. All titles and abstracts were individually screened by two researchers for alignment with the inclusion and exclusion criteria. Disagreements regarding inclusion or exclusion were resolved by consensus. The title and abstract screening resulted in the exclusion of 3254 resources.

Fig. 1. Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA) 2020 flow diagram for the scoping review, which includes searches of databases and other sources (21)



The full text of the 318 reports meeting the inclusion criteria were obtained; 6 of the reports eligible for inclusion could not be obtained. A second test was undertaken for the full-text screening (*n* = 5). Two researchers independently conducted the full-text screening and documented the decision for each resource in an Excel spreadsheet. After comparison, disagreements over the inclusion or exclusion of a report were resolved by consensus. The full-text screening resulted in the exclusion of 267 reports and inclusion of 51 in the analysis.

2

Methods

We included literature that examined the impact of purchasing arrangements on the quality of care and provided information about the design of the purchasing arrangement. Thus, we focused on studies that presented alternative financing arrangements beyond individual FFS arrangements and P4P measures. To be included, articles must have analysed a purchasing arrangement that moved away from per-service financing. In practice, this would mostly mean they were population-based or episode-based models, or other models requiring some form of collaboration or integration of financing or delivery, or both, across providers. However, while P4P was not the main focus of our review, we also included literature on any such alternative purchasing arrangement in which P4P mechanisms were one component of the purchasing arrangement in conjunction with other mechanisms (e.g. capitation with performance-related incentives). We excluded literature that did not provide specific details about the purchasing arrangement. Articles were also excluded if they detailed a purchasing arrangement that had the primary purpose of containing health expenditures. Therefore, the only purchasing arrangements included were those that aimed to improve the quality of care or that included explicit incentives for maintaining or improving quality. Further details about decisions to include or exclude studies are highlighted in Annex 1.

To ensure the robustness of the evidence, we limited the studies to those using causal inference methods (e.g. randomized controlled trials [RCTs], regression-based analysis, control groups), thus excluding studies that were observational or descriptive in nature (e.g. case studies). Studies included those that compared the purchasing arrangement with the prior arrangement and also those that made comparisons among other relevant groups, for example, in cases in which alternative purchasing arrangements were defined by a specific region or group of patients. Table 2 provides additional details about the inclusion and exclusion criteria.

Table 2. Overview of inclusion and exclusion criteria for the scoping review

Category	Inclusion criteria	Exclusion criteria
Themes	– Studies examining alternative purchasing arrangements that aimed at improving care quality beyond pay-for-performance for noncommunicable diseases and chronic diseases – Studies examining the impact on quality outcomes	– Studies not providing sufficient detail about the alternative purchasing arrangement – Studies focusing exclusively on individualized pay-for-performance mechanisms – Studies looking at quality improvements for one specific disease or condition
Time frame	– Published from 2013 until November 2022	– Published before 2013
Types of resources	– Peer-reviewed scientific publications – Grey literature (e.g. government publications, working papers, research reports) – Systematic reviews used for hand-searching	– Bachelor's and Master's theses
Types of studies	– Studies using causal inference methods (e.g. regression-based analysis, controlled studies) – Studies comparing a new purchasing arrangement with a prior purchasing arrangement; purchasing arrangements compared across several locations or with other types of purchasing arrangements	– Case studies – Descriptive or observational studies
Language	– English	– Other than English
Geographical areas	– Global	

2.4 Data extraction and quality appraisal

Two researchers entered data for the included records onto a prepared extraction sheet developed in line with the Cochrane Effective Practice and Organisation of Care (known as EPOC) resources group *(22)* and containing a priori defined categories deemed relevant for the analysis of the data and the research questions posed for this review. Data extraction was trialled with two researchers entering information for two selected research papers to ensure consistency.

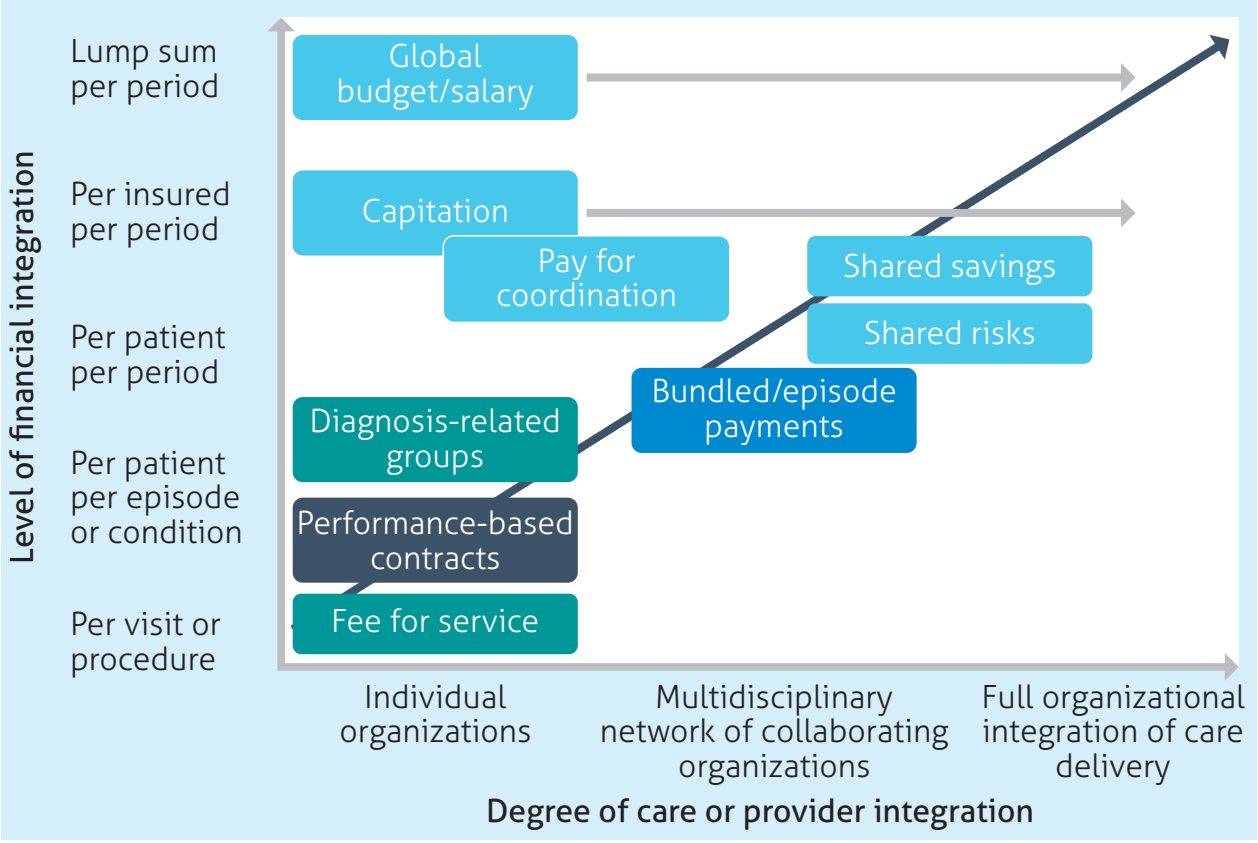
As there are no quality appraisal techniques recommended specifically for scoping reviews, we accounted for the quality of studies in two ways: (i) we applied the critical appraisal checklists developed by the JBI for different types of research to assess the risk of bias for each article (https://jbi.global/critical-appraisal-tools); (ii) we amended and applied the Grading of Recommendations Assessment, Development and Evaluation (GRADE) approach, taking into consideration the wide range of interventions captured in the scoping review. On this latter point, we assessed the certainty of the evidence of the outcomes reported in each article according to the type of study (e.g. RCT versus controlled study), the risk of bias (based on the JBI critical appraisal checklists) and the effect size. More specifically, we used the following criteria to assign a level of certainty to the evidence for the quantitative outcomes reported:

- high certainty –
  - an RCT with no bias or other issues AND a large effect;
- moderate certainty –
  - an RCT with some bias or issues OR a small effect
  - a controlled study with no bias or issues AND a large effect;
- low certainty –
  - an RCT with some bias or issues AND a small effect
  - a controlled study with some bias or issues OR a small effect;
- very low certainty –
  - a controlled study with some bias or issues AND a small effect.

2.5 Data analysis

Our analysis used a three-step approach. To address research question 1 (Box 1), we first summarized the articles included in the scoping review according to different relevant features, such as the type of purchasing arrangement, country, methodology used, types of outcomes measured and impact on the quality of care. To address questions 2 and 3, we clustered articles together based on similar purchasing arrangements in terms of the underlying base payment to providers. This classification is informed by Fig. 2, based on a figure by Tsiachristas *(23)*, which classifies purchasing arrangements based on their level of financial integration and the level of care integration across providers and settings. This classification is useful because it recognizes that some types of purchasing arrangements vary in terms of the extent of provider integration, while others are more succinctly defined. While we take this figure as a starting point for clustering the purchasing arrangements, we also note that not all purchasing arrangements fit perfectly within one category and thus may overlap several categories along the continuum of financial and care integration.

Fig. 2. Clustering of purchasing arrangements according to level of financial and care integration



Source: Adapted from a figure by Tsiachristas *(23)*. Additional categories have been added to the figure while others have been amended and condensed, in line with the focus of this study.

Within clusters, in cases in which four or more articles focused on the same purchasing arrangement programme, we created subclusters of articles, with the view that they were the most comparable and would allow us to draw concrete conclusions about that particular arrangement's impact on the quality of care and the service delivery model implemented. Altogether these clusters amounted to: (i) capitation and global budget arrangements, with the US Maryland All-Payer Model of a global budget and the Massachusetts Alternative Quality Contract (AQC) as relevant subclusters); (ii) bundled payments; (iii) shared-savings and shared-risk models; (iv) pay-for-coordination; (v) blended capitation; and (vi) a cluster comparing different types of purchasing arrangements across the various clusters (see the Glossary for additional explanation).

For each cluster of purchasing arrangements, we summarize the results across articles in terms of the arrangement's impact on the quality of care. We discuss the service delivery models implemented in conjunction with the purchasing arrangements that positively affected quality, as well as highlight other relevant accompanying institutional design and supportive elements, if prevalent within the cluster, relating to governance, information systems, quality initiatives and regulations. Throughout this summary, we use the certainty of evidence assigned during the data extraction phase as a guiding element.

3

Results

3.1 Summary of articles

In the following sections, we summarize the descriptive details of the literature included in the scoping review, as well as the types of purchasing arrangements studied, the aims of the arrangements, methods used, measures of quality used and the certainty of evidence.

3.1.1 Context

In terms of country representation, most articles were based on purchasing arrangements in the United States (*n* = 36), followed by those in China (*n* = 5), Canada (*n* = 4), Germany (*n* = 3) and Netherlands (Kingdom of the) (*n* = 2). One additional article focused on a comparison of purchasing arrangements across Europe and therefore covered multiple countries. Of the 51 included articles on purchasing arrangements, 17 analysed arrangements that involved multiple payers (see Annex 2 for further details about these), 18 looked at those covered by a single public purchaser, 13 by a single private purchaser (i.e. a commercial plan) and, finally, 3 provided no clear information.

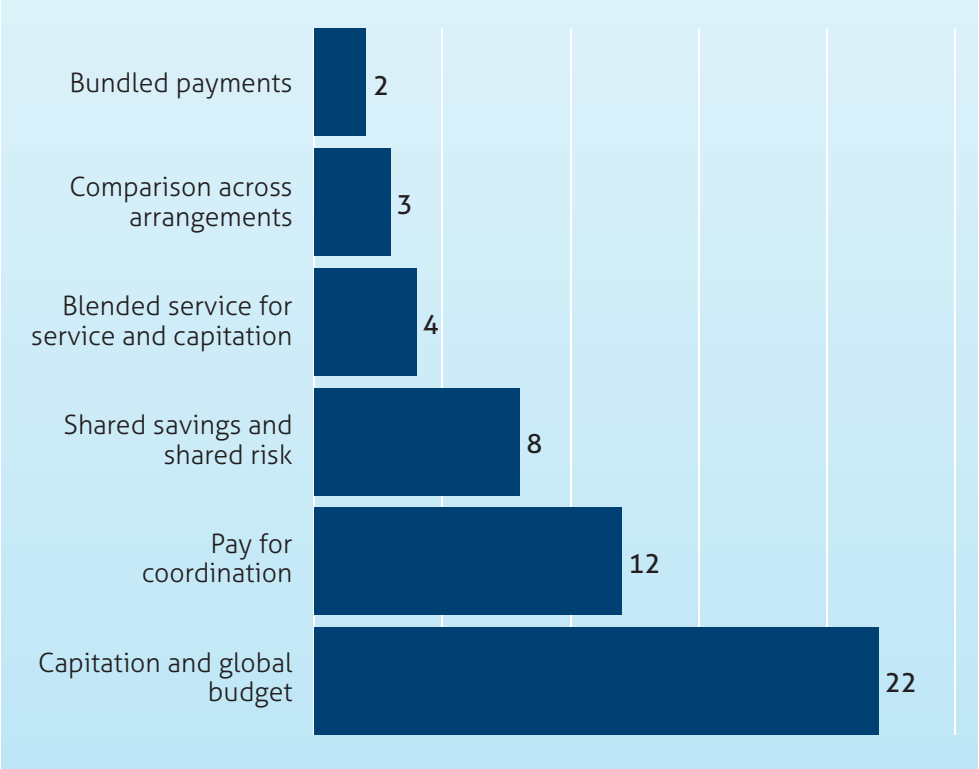
As the aim was to look at arrangements that targeted chronic diseases and broader population health rather than specific types of diseases, this is reflected in the target population of the purchasing arrangements reviewed. Most purchasing arrangements targeted the general population (*n* = 24), followed by those that targeted Medicaid and Medicare patients in the United States (*n* = 9) and patients enrolled in commercial health plans (*n* = 13). Five articles indicated the target population was people with chronic diseases.

3.1.2 Types of purchasing arrangements

The largest share of articles (*n* = 22) discussed capitation or population-based purchasing arrangements in which providers received a per-member fee for those covered by their services (*n* = 6) or a global budget to cover all included patients, as determined by certain criteria (*n* = 16) (Fig. 3). Next, some articles (*n* = 12) discussed pay-for-coordination arrangements in which providers received a per-member fee per enrolled patient (or a care management fee) to cover the coordination and integration of their services with those of other providers or services. A smaller number of articles discussed shared-savings programmes (*n* = 5) or shared savings combined with shared risk (*n* = 3). Only a couple of articles evaluated bundled payments or similar arrangements (*n* = 2) in which a bundled amount covered all services provided within an episode of care for certain chronic conditions or diseases. A small number of articles looked at blended capitation models (*n* = 4),

while three articles compared different types of purchasing arrangements.

Fig. 3. Types of purchasing arrangements in the 51 articles evaluated



A handful of articles focused on well-established purchasing arrangements (i.e. the Massachusetts AQC programme, the Maryland All-Payer Model and the Ontario Family Health Organizations and Family Health Groups), allowing us to compare a wide range of quality measures for particular programmes.

3.1.3 Aims of purchasing arrangements

Nearly all the purchasing arrangements stipulated the dual aims of improving the quality of care or strengthening service delivery alongside cost containment (*n* = 35), although the specification or operationalization of quality and service delivery varied. In terms of quality and strengthening service delivery, most purchasing arrangements explicitly stated the aim of improving the quality of care more broadly (*n* = 20), while others stated more specifically the aim was to improve care outcomes or population health (*n* = 10), strengthen primary care (*n* = 10) or promote coordinated care (*n* = 9), or a combination of these. Two articles specified that within the realm of improving quality, the aim was to improve patient-centred care; two specified that it was to improve access to services; and two others specified it was to improve the integration of care or reduce the fragmentation of services provided across settings.

Individuals with chronic diseases were occasionally referenced as one target of purchasing reforms (*n* = 5), whether in terms of improving chronic disease management or providing better access or better preventive care. In several cases within the quality-improvement realm, the aim was to correct some deficiency or disincentives in the previous system, such as shifting services across settings (*n* = 2), reducing overtreatment associated with FFS systems (*n* = 7) or overcoming inequality in access to care between populations (*n* = 2). The aim of several purchasing arrangements across articles centred around improving the value or efficiency of care (*n* = 12), implying that the quality and cost of care needed to be weighed together. Only one article detailed a purchasing arrangement that had the goal of supporting older people ageing in place for as long as possible.

3.1.4 Types of methods used

A majority of the articles (*n* = 30) used a difference-in-differences retrospective cohort approach. Nine articles used linear regressions, with time and fixed effects, while three used interrupted time series (i.e. regression discontinuity). Three high-quality articles were RCTs. The remainder of the articles used a difference-in-differences approach via a natural experiment (*n* = 3), triple difference-in-differences approach (*n* = 1) or linear or logistic regression without time effects (*n* = 2).

3.1.5 Quality indicators measured

Two thirds of the articles investigated service utilization as one impact of purchasing arrangements on the quality of care, such as the use of primary care, hospital-based care (e.g. the emergency department [ED], inpatient or outpatient care) and specialist care (Fig. 4). Around half of the included articles used measures of chronic disease management and prevention to measure the quality of care (e.g. cancer screening, diabetes control measures, immunization measures). Although cost was not a focus of this scoping review and the results are not shown here, it is important to note that many articles inevitably looked at cost as an outcome, given that most purchasing arrangements had the dual aim of managing costs while maintaining or improving care quality.

Fig. 4. Quality measures used in the 51 articles reviewed

Health outcomes were also frequently used to discern the impact on quality, (e.g. mortality, self-rated health, ED visits, hospital readmissions). Less frequently used measures included continuity of care (*n* = 5) (e.g. follow up after hospital discharge, receipt of postacute care services), patients' perceptions of the quality of their care (*n* = 5) and access to care (*n* = 4) (e.g. patients' perceived access to all necessary services, including to specialists). A small number of articles also used composite measures of quality with a larger number of different quality measures (*n* = 3) (e.g. scores based on chronic disease measures), prescription-related indicators (*n* = 3), coordination of care measures (*n* = 2) (e.g. patients' perceptions, number of uncoordinated contacts between services) and indicators of the quality of communication (*n* = 2). Further detail on the precise indicators measured can be found in the results for each cluster. In many cases, the measures of quality used within the articles were not aligned with the quality measures used to award incentives, but rather they were used as proxies to measure the impacts on quality.

3.1.6 Overall impact on quality and certainty of the evidence

The evidence on the impact of the purchasing arrangements studied on the quality of care is mixed, and there are varying degrees of certainty (Table 3), with some clusters consisting of more articles indicating a positive impact with higher certainty than others. These clusters are discussed in detail in the following sections.

Table 3. Impact of purchasing arrangements on care quality and certainty of the evidence[a]

Purchasing arrangement	Impact on quality	Certainty of evidence (reference)
Capitation or global budget	Positive	Moderate *(24)* Moderate/very low *(25)* Low *(26-28, 31)* Very low *(32-35)*
	None	Moderate *(29)* No certainty attributed[b] *(30)*
Massachusetts Alternative Quality Contract model	Positive	Low *(38-40)* Very low *(36-37)*
	Unclear[c]	Very low *(41)*
Maryland All-Payer Model	Mixed	Low/very low *(45)*
	Limited/ unclear	Moderate *(42)* Very low *(43-44)*
Bundled payments	Positive	Low *(46-47)*
Shared savings and shared risk	Positive	Moderate/low *(48)* Low *(50-53)* Low/very low *(49)*
	Mixed	Moderate/low/very low *(55)* Low *(54)*
Pay-for-coordination[d]	Positive	High *(64)* Low (*56, 59, 61, 66)* Very low *(57, 58, 62, 63)*
	Mixed	Moderate/low *(65)*
	None/ negative	No certainty attributed[2] *(60, 67)*
Blended capitation[e]	Positive	Low *(69)* Very low *(68, 70)*
	Unclear	Low *(71)*

[a] The table contains only articles looking at clusters with individual types of purchasing arrangements; it does not include the cluster of articles comparing different purchasing arrangements.

[b] No statistical significance was found by the authors, so the certainty of evidence could not be attributed.

[c] This article measures spillover effects and therefore not the impact of the Alternative Quality Contract on the target population.

[d] All articles in this cluster are based on arrangements in the United States.

[e] All articles in this cluster are based on arrangements in Ontario, Canada.

3.2 Impact of purchasing and service delivery aspects on quality of care

3.2.1 Capitation and global budgets

Within this cluster, the purchasing arrangements represent population-based payment models in which payments are allocated to providers for a specified population of patients, based on either a per-patient amount or on a lump sum based on previous years' budgets, the expected level of spending or diseases typically treated (Table 4). Variations in these models exist in terms of the financial responsibility and risk taken on by providers, as well as in terms of the services covered. For example, in full-risk capitation models, providers not only take on the full financial risk and accountability for their patients but also they are accountable for the performance and services offered by other providers in their network, which they cover through their budget.

Table 4. Summary of the articles evaluating capitation and global budgets[a]

Capitation and global budgets	Description
Summary *(24–35)*	– Providers receive a per-member per-month fee for patients or a lump sum based on different criteria – Most capitation models also incorporate some additional financing incentive, related either to savings or quality or performance indicators
Countries	China *(24-26, 29, 31)*, Germany *(33, 44)*, USA (*27, 28, 31, 32, 35*)
Setting	Primary care *(26, 30, 32-34)*, hospitals *(24, 25)*, multiple services or settings *(27-29, 31, 35)*
Impact on quality	**Overall impact** Positive (*n* = 10) with certainty rated as moderate *(24)*, moderate or very low *(25)*, low *(26-28, 31)* and very low *(32-35)* No impact (*n* = 2) with moderate certainty *(29)* or no certainty attributed *(30)* **Positive impact** – Improved continuity and coordination of care: discharge bridged with postacute care (low certainty) *(31)*, increased follow-up care for vulnerable groups (low certainty) *(26)*, improved coordination of services (low certainty for 1 study, very low for 1) *(26, 34)* – Improved process outcomes: increased care services or utilization for general population and people with chronic diseases (low certainty for 1 study, very low for 1) *(28, 35)*, improved process outcomes for disease-specific management or control indicators (moderate certainty) *(24)* – Improved health outcomes: lower risk of hospitalization or fewer visits for general and vulnerable groups (low certainty for 2 studies, very low for 2) *(27, 28, 32, 33)*, increased survival rates (low certainty for 1 study, very low for 1) *(28, 34)* – Improved efficiency by shifting services to other settings (moderate certainty for 1 study) *(25)* **No impact** – Impact on oral medication use and intravenous infusions deemed to result from insurance package redesign rather than purchasing reform (moderate certainty) *(29)* – No impact on rational prescribing of medications (no certainty attributed) *(30)*
Service delivery and supportive elements	**Positive impact** Integrated care model *(35)*, managed care organization *(31)*, Medicare Advantage organization *(28)*, patient-centred medical home *(32)*, accountable care organization *(27)*, comprehensive care in primary care *(33, 34)*, primary care setting *(26)*, hospital-based care *(24, 25)* **No impact** Part of insurance reforms *(29)*, care provided across township health clinics and village clinics *(30)*

[a] See the Glossary for details about the service delivery models mentioned in this table.

Some of these arrangements incorporate an additional financial incentive based on quality or performance measures to encourage providers to improve the quality of care. The arrangements without additional financial incentives assume that capitation or a global budget will incentivize better coordination of care services, thus leading to better quality outcomes. Of the articles in this cluster, 10 found a positive impact of the purchasing arrangements on the quality of care, albeit with a mixed certainty of evidence (n = 2 with moderate certainty, n = 4 with low certainty, n = 4 with very low certainty), while 2 articles found no impact on quality (n = 1 with moderate certainty, n = 1 with no certainty attributed).

Of the 10 articles that found a positive impact on the quality of care, 8 arrangements contained additional financial incentives or stipulations in addition to the capitation or global budget (e.g. per capita award, performance- or quality-based incentive, shared savings). One, which incentivized quality through a performance assessment based on international treatment guidelines and medical outcomes, found an improvement in treatment adhering to guidelines for chronic conditions (i.e. pneumonia and acute myocardial infarction; moderate certainty) *(24)*. Another evaluation found an improvement in the allocation and technical efficiency of services by shifting hospital admissions from out of county to within county, with no reductions in quality, as the result of higher tariffs for averted out-of-county admissions (moderate certainty) *(25)*. In another case, a capitation arrangement with a per capita award per enrolled patient led to improvements in patients' perceptions of outcomes related to the coordination and continuity of care (low certainty) *(26)*. Another global budget with a P4P component based on 17 performance measures (e.g. primary care, chronic disease management, prevention, ED utilization) showed a reduction in hospitalizations among women aged 15–44 (low certainty) *(27)*. A full-risk capitation model with shared savings led to increased preventive care utilization and survival rates alongside reduced use of ED services (low certainty) *(28)*. Of the two articles that found no effect, Powell-Jackson et al. *(29)* suggested that the payment reform was a slow transition and may not have taken full effect at the time of their study, while Sun et al. *(30)* suggested that the limited impact resulted from penalties for poor performance not being strong enough to incentivize a change in providers' behaviour.

Of the five articles that evaluated financial incentives relating to performance or managing chronic care, only three explicitly looked at outcomes that evaluators or regulators would use to measure performance as part of the financial incentive *(26, 27, 30)*. The remaining two arrangements did not contain additional financial incentives, albeit a positive impact on the quality of care was still found in terms of follow-up care in a Medicaid managed care contracting situation (low certainty) *(31)* and in reduced ED visits for

patients with chronic conditions in a single commercial payer context (very low certainty) *(32)*.

In terms of service delivery and institutional design features, capitation and global budgets that resulted in positive impacts on quality were implemented across a wide range of settings and in conjunction with various service delivery models. Continuity of care as an element incentivized through a purchasing arrangement or as a characteristic of the service delivery model was highlighted across numerous articles that found a positive impact on care quality. This was either through the forging of longer-term provider–patient relationships through patient enrolment *(26)*, scheduling patients with chronic conditions at more regular intervals after hospitalization *(28)* or mandatory enrolment of Medicaid beneficiaries within a geographically defined provider network *(27)*.

Other institutional elements that had a positive impact and were highlighted in service delivery models alongside purchasing arrangements included requiring comprehensive care to be provided by general practitioners (GPs) in primary care *(26, 32)*, GPs being encouraged to participate in structured disease management programmes *(33, 34)* and GPs being required to act as gatekeepers to care by specialists and other providers *(26, 33, 34)*. Health information technology and electronic health records were highlighted as supportive elements in several articles in which a positive impact on quality was found *(26, 32-34)*. Finally, in several cases, the articles highlighted key values or expected impacts of the service delivery model implemented, although they did not provide details about how this was operationalized in practice *(27, 32, 35)*. Conversely, in other cases authors specifically tied the purchasing arrangement to the impacted behaviour of providers. Munnich and Richards *(31)* speculate that managed care organizations may have led providers either to prioritize managing chronically ill patients given their high costs or to improve transitions from acute care to postacute care to reduce future costs (low certainty) and that additional competition for participating insurers may have maintained providers' behaviour. Similarly, Mandal et al. *(28)* stated that changing payments to full-risk capitation with shared savings incentivized risk stratification of prescribed services, which shifted service delivery and improved clinical outcomes. Other studies highlight the diverse actions taken by different providers and networks that cannot be accounted for that may have impacted the results on quality *(25, 28)*.

3.2.2 Massachusetts Alternative Quality Contract

Table 5. Summary of literature about the Blue Cross Blue Shield of Massachusetts Alternative Quality Contract

Alternative Quality Contract	Description
Summary *(36-41)*	– Risk-adjusted global budget, shared savings, shared risk, quality bonuses
Country	USA
Setting	Primary care, specialist care, hospitals
Impact on quality	**Overall impact** Positive (*n* =3) with low or very low certainty *(36, 38, 39)* Mixed impact *(n = 3)* with low or very low certainty *(37, 40, 41)* **Positive impact** – Improvements in care and treatment: adult preventive care (low and very low certainty) *(36, 37, 40)*, paediatric care (low and very low certainty) *(36-38, 40)*, reductions in the use of low-value health care services (very low certainty) *(38)*, treatment of upper respiratory infection (low certainty) *(39)*, reduction in emergency room visits (very low certainty) *(38)*, mental health care (low certainty) *(40)* – Improvements in screening and testing: chlamydia screening (low certainty) *(38)*; pharyngitis testing (low certainty) *(39)*; reductions in prescriptions, tests and imaging (very low certainty) *(38)*; colorectal cancer screening (very low certainty) *(38)* **Mixed results (partly positive effects, partly no effect)** – General chronic disease management (very low certainty) *(36, 37, 40)*, cardiovascular care (very low certainty) *(37, 39, 41)*, diabetes care (low and very low certainty) *(39, 41)* – Benefits for groups with lower socioeconomic status (low certainty) *(40)*, spillover effects for Medicare beneficiaries (very low certainty) *(42)* – Breast cancer screening (very low certainty) *(37, 41)* **No effect** – Hospital readmissions *(41)*, asthma care *(39)*, care for people with attention deficit–hyperactivity disorder *(39)* **Negative effect** – Increase in inpatient admissions (very low certainty) *(37)*, reduction in office visits (very low certainty) *(38)*
Service delivery and supportive elements	Organizations were responsible for managing a population budget and providing all services Large variation among participating providers in terms of service delivery Efforts to improve service delivery included changing how patients were referred, engagement in case management for high-risk patients, redesign of care processes to eliminate waste

Six articles analysed how the Blue Cross Blue Shield of Massachusetts AQC impacted the quality of care (Table 5). The AQC is a single-payer purchasing arrangement that was implemented by Blue Cross Blue Shield of Massachusetts, USA, in 2009 against the backdrop of a multipayer system in the state. While participation was voluntary for providers, by 2012 about 85% of physicians working in the Blue Cross Blue Shield network in the state had joined, including primary care and specialist care providers, as well as affiliated hospitals. In terms of purchasing, the AQC combines a risk-adjusted global budget with a two-sided shared-savings and shared-risk contract and quality bonuses. The contract stipulated that Blue Cross Blue Shield would reward physicians for costs that were below the risk-adjusted budget, but it also required them to share in deficits with insurer for spending that exceeded the budget. Furthermore, shared savings and shared risk were also tied to performance quality, with higher quality ensuring that a larger share of the savings was conveyed and there were lower deficits. Quality bonuses and the distribution of savings were based on 64 performance measures (including process measures for different services, testing, prevention, disease-specific outcomes, patients' experience and hospital measures) and defined on a per-member per-month basis.

Regarding the impact of the AQC on the quality of care, all articles reported some positive effects; the certainty of evidence, however, was low or very low across the articles. Assessing the impact on the entire population, Song et al. *(36)* reported very low certainty that the AQC had a positive effect on chronic disease management, adult preventive care and paediatric care. These results are confirmed by another study by Song et al. *(37)* that additionally found, with very low certainty, improvements in cancer screening, as well as reductions in the use of some health care services, but also, for example, an increase in inpatient admissions. The remaining articles assessed the impact of the AQC on specific patient groups: Chien et al. *(38)* showed with low certainty that the AQC had small positive impacts on the quality of paediatric care, particularly for children with special health care needs. However, only measures that were tied to performance incentives were affected. The AQC was also associated with significant but small improvements in care quality and utilization for patients with chronic diseases and co-occurring mental health issues (low certainty evidence) *(39)*. In comparing the AQC's impact on people in different socioeconomic groups, Song et al. *(40)* found that people living in areas of lower socioeconomic status experienced greater improvements in process quality, preventive care and paediatric care (all low certainty evidence) and, therefore, the arrangement may have contributed to reducing disparities. Furthermore, McWilliams et al. *(41)* analysed whether the implementation of the AQC had any spillover effects for

Medicare beneficiaries, and although some indicators improved, the authors concluded that overall the AQC did not lead to such effects (very low certainty).

In terms of service delivery, studies reported large variation among participating providers, including, for example, to what degree different levels and sites of care provision (i.e. primary, specialist and hospital) were integrated. Furthermore, physician groups differed in terms of whether a centralized administrative structure to support them in improving quality was available. Strategies employed to improve care coordination included engaging in high-risk case management across multiple payers in the system (i.e. with other payers apart from Blue Cross Blue Shield), changing referral patterns and redesigning care processes to eliminate waste *(41)*. Additionally, participating providers received periodic reports from the insurer about their performance in terms of cost and quality to identify areas for improvement. Furthermore, providers exchanged experiences with one another and provided peer support. Also, both AQC and non-AQC small practices received quality improvement support from the physician organizations to which they belonged.

3.2.3 Maryland All-Payer Model

Table 6. Summary of literature about the Maryland All-Payer Model

All-Payer Model	Description
Summary *(42-45)*	– Global budgets for hospitals (including inpatient department, emergency department and hospital outpatient department services) in combination with pay-for-performance and shared savings and shared risk
Country	USA
Setting	Hospitals
Impact on quality	**Overall impact** Limited-to-no impact with moderate, low and very low certainty *(42-45)* **Positive impact** – Greater reduction in hospital services for vulnerable groups of patients (low or very low certainty)*(45)*, reduced utilization of outpatient services in rural hospitals (moderate certainty) *(42)* **Mixed effects**[a] – Various measures relating to ischaemic stroke (very low certainty) *(44)*, acute myocardial infarction (very low certainty) *(44)* and congestive heart failure (very low certainty) *(44)* – Hospital admissions (moderate certainty) *(43)* – Outpatient visits (moderate certainty) *(43)* **No effect** – No reductions in use of hospital services or no increase in use of primary care for fee-for-service Medicare beneficiaries (very low certainty) *(43)* – No impact on the quality of inpatient care for patients with cardiovascular conditions (very low certainty) *(44)* – No effect on visits to emergency departments (moderate and very low certainty) *(42, 43)*, inpatient days (moderate certainty) *(42)*, follow up in primary care after hospital stays (very low certainty) *(43)* – Hospital readmissions (moderate certainty) *(42)*
Service delivery and supportive elements	– Implementation of care management strategies to reduce use of emergency departments, length of stay and unplanned readmissions – Implementation of care coordination with community providers – Implementation of more efficient referral pathways and follow-up visits after hospital discharge

[a] Note: Mixed results here refer to cases where some indicators are found to be positive, while other related measures find either a negative impact, or no impact.

Four articles from the United States analysed the impact of the Maryland All-Payer global budget model on the quality of health care provision (Table 6). This hospital-based model was implemented against the backdrop of a multipayer arrangement that included Medicaid, Medicare and commercial insurers, whereby all public and private insurers were required to pay the same rates for services (e.g. for inpatient care, ED visits and outpatient department visits). Thus the new model replaced a payment system that combined elements of FFS and case-based reimbursement; under the new model, hospitals were required to remain within the global budget assigned to them and, therefore, unit prices for services were adjusted depending on the aggregate volumes of patients served. The goal of the model was to reduce hospital expenditures, while improving care and health outcomes. The arrangement also included a shared-savings component if providers' costs were under budget, and the shared-risk component was applied as a penalty of up to 50% of any overspending. To protect against the underprovision of services and to ensure care quality, the purchasing arrangement also contained a P4P component in that the budget was adjusted based on a quality-based reimbursement programme (e.g. related to patients' satisfaction, rates of preventable admissions and readmissions, hospital risk-adjusted mortality and hospital-acquired conditions).

The articles suggested that this programme had a limited impact on the quality of care, with a mix of certainty of evidence (n = 1 with moderate certainty, n = 1 with low and very low certainty, and n = 1 with very low certainty). Done et al. *(42)* found that the programme positively affected some outpatient measures in rural hospitals, as it was associated with a decrease in admissions not originating from the ED (moderate certainty) and a decrease in all non-ED visits (moderate certainty). These positive impacts, however, were accompanied by increases in nondeferrable admissions (moderate certainty) and admissions from the ED (moderate certainty). In terms of Medicare beneficiaries, Roberts et al. *(43)* did not find any evidence that the model was associated with reductions in hospital stays or increases in primary care visits for this group (very low certainty). Concerning patients with one of three cardiovascular conditions, the reform did not have a positive impact on inpatient outcomes and quality measures (very low certainty) *(44)*. Instead, the length of stay for congestive heart failure increased (very low certainty), and hospitalizations for acute myocardial infarction and ischaemic stroke increased (very low certainty), as did procedural volumes for coronary artery bypass grafting (very low certainty). Masters et al. *(45)* found that the model led to nonuniform impacts on Medicare patients. The programme saw greater reductions in the utilization of hospital services for vulnerable groups (e.g. patients eligible for both Medicare and Medicaid, patients with disabilities

and patients with multiple chronic conditions) than for other patient groups, which the authors ascribe to the prioritization of high-cost, high-need patients as providers transformed their care delivery practices.

In terms of service delivery, hospitals were expected to respond to global budget incentives by implementing care management strategies to avoid admitting patients through the ED, to reduce patients' length of stay and to shift away from unplanned readmissions. Further expectations were that hospitals would improve care coordination with community providers, implement more efficient referral pathways and increase the number of follow-up visits after hospital discharge. The financial incentives aimed to encompass these expectations by incorporating an assessment of preventable conditions, patient safety outcomes and readmission programmes. Hospitals' budgets were monitored monthly via data transmitted to Maryland's Health Services Cost Review Commission.

Roberts et al. *(43)* and Done et al. *(42)* suggested that the limited impact of the model may have been the result of inconsistent incentives offered to hospitals and physicians and the limited influence and risk placed on physicians. Given that the global budget was at the hospital level and, therefore, hospitals bore the risks, the structure of the programme did not incentivize physicians in hospitals to reduce volumes. Done et al. *(42)* also posited that the financial incentives may have had more of an impact on outpatient services, given the flexibility of decisions that could be made in these services. Masters et al. *(45)* also suggested that there may have been limited coordination between hospitals and community providers, given that the latter were not included in the global budget.

3.2.4 Bundled payments

Table 7. Summary of articles about bundled payments

Bundled payments	Description
Summary *(46, 47)*	– Reimbursement of providers for a defined bundle of services for a defined episode of care based on expect costs, covering a range of services – Target populations: individuals with certain chronic conditions admitted to hospitals, elderly individuals living in the community with complex needs
Countries	Netherlands (Kingdom of the) *(47)*, USA *(46)*
Setting	Primary care and home-based care *(47)*, hospital-based care *(46)*
Impact on quality	**Overall impact** Mostly positive, with low certainty *(46, 47)* **Positive impact** – Improved patient-centredness of care (low certainty), improved patients' experiences at most time points (low certainty) *(47)* – Increased healthy days at home (low certainty), increased share of those eligible for both Medicare and Medicaid comprising treated patients (low certainty), reduction in frail patients (low certainty) *(46)* **Mixed impact** – Worse ratings for autonomy and burden of medication (low certainty) *(47)*
Service delivery and supportive elements	**Positive impact for integrated and chronic care programme *(47)*** – Programme to integrate primary care with community-based care for older people, specifically providing care for frail elderly people – Person-centred care, comprehensive assessments, care coordination, individualized care planning, adaptations based on patients' wishes and needs, case management and multidisciplinary team meetings – Information and communications technology platform to share information across professionals **Positive impact for hospital-based model (inpatient, outpatient, emergency, postacute) *(46)*** – Implied coordination across services but not explicitly discussed in article

Bundled payments are reimbursement for services based on the expected costs of a clinically defined episode of care for a certain illness or for a certain package of services (Table 7). Two papers examined the impact of bundled payment arrangements on the quality of care *(46, 47)*. The first examined the Bundled Payments for Care Improvement – Advanced (BPCI-A) model in the United States (see the Glossary for further details) and found there was no impact on the adverse selection of patients and outcomes were not worse among patients with vulnerabilities treated in BPCI-A hospitals versus those treated in non-BPCI-A hospitals (low certainty of evidence), and there was an increased number of healthy days at home for all vulnerable groups (low certainty) *(46)*.

The second study examined the Care Chain Frail Elderly (CCFE) programme in Netherlands (Kingdom of the), which targeted the 1% most frail elderly people living at home who had complex care needs *(47)*. This article found that the programme did not improve the health outcomes of the participants, but it led to a sustained improvement in patient perceived person-centredness (low certainty of evidence). At the same time, the programme scored worse on autonomy and burden of medication, according to patients, and care costs were higher than those for control patients, reflecting the greater resources required to support frail individuals with complex needs (low certainty).

Both bundled payment arrangements were accompanied by different models of service delivery. The BPCI-A took place in hospitals, skilled-nursing facilities and home health service settings, and included EDs, outpatient and inpatient services, and postacute care *(46)*. Care episodes were triggered when patients were admitted to inpatient or outpatient services. While care coordination and integration between providers in different settings may be assumed in bundled payment arrangements, the authors did not describe any further aspects of service delivery, nor did they provide an explanation of the potential mechanism leading to the increased number of healthy days at home (i.e. the total time the patient spends alive and at home after a hospitalization or clinical intervention) for vulnerable groups.

Conversely, Hoedemakers et al. *(47)* highlighted that the bundled payment from CCFE was part of a programme to integrate primary care and community care, using multidisciplinary teams of health care professionals. As the aim was to allow older patients with complex needs to live at home for as long as possible, these patient-centred approaches included comprehensive assessments of an individual's needs and drafting individualized care plans based on the individual's needs and personal goals. Multidisciplinary meetings took place between all professionals and included informal caregivers to develop and discuss the care plan. A case

manager was the main contact point and monitored the implementation of the care plan, adapting it as it progressed, according to the patient's needs. A secured information and communications technology platform was also used to share information across providers.

3.2.5 Shared savings and shared risk

In a shared-savings model, accountable providers are eligible to receive a portion of the cost savings if their total expenditure, based on FFS payments, is less than a previously agreed benchmark, given the population of patients covered. Another variation of this model also includes shared risk in which providers are held accountable for overspending if their costs exceed the benchmark (see Glossary for more information). The aim behind these types of arrangements is to encourage providers to be efficient in their delivery of services while improving the quality of care. These types of arrangements differ from bundled payments and payments based on episodes of care in that they tend to include a broader set of services delivered outside of hospital settings and across a network of providers for outpatient services (Table 8).

Table 8. Summary of articles about shared savings and shared risk

Shared-savings and shared-risk programmes	Description
Summary *(48-55)*	– Shared-savings programmes *(48-51, 52)*: providers receive a portion of the cost savings if total spending is less than a benchmark – Shared-savings and shared-risk programmes *(52, 53, 55)*: similar to shared-savings programmes except providers are also responsible for covering part of any overspending if they exceed a benchmark – The extent of savings depends on performance and quality targets or quality scores *(48, 50-54)*
Countries	Germany *(49)*, Netherlands (Kingdom of the) *(54)*, USA *(48, 50-53, 55)*,
Setting	Across primary and hospital-based care *(48, 53, 55)*; across primary, hospital-based care, social care and pharmacies *(49)*; primary care setting *(50, 51, 52)*; full continuum of care across all settings *(54)*

Shared-savings and shared-risk programmes	Description
Impact on quality	**Overall impact** Overall, the impact was positive (*n* = 6), with moderate or low certainty in 1 study *(49)*, low certainty in 4 studies *(50-53)*, and low or very low certainty in 1 study *(49)*. There was a mixed impact (*n* = 2) with low certainty in 1 study *(54)* and moderate, low and very low certainty in 1 study *(55)*. **Positive impact** – Improved process outcomes for disease-specific management or control indicators (e.g. diabetes), cancer screening, preventive care, improved care quality for chronic diseases (moderate or low certainty for 1 study) *(48)*; low certainty for 3 studies *(50-52)* – Improved health outcomes reflected in reduced visits to emergency department (low certainty for 3 studies) *(48, 50, 53)*, reduced hospitalizations (low certainty for 1 study) *(50)*, reduced acute inpatient days (low certainty for 1 study) *(53)*, increased survival (very low certainty for 1 study) or age at death (low certainty for one study) *(49)* – Improved continuity of care reflected in increased follow up after hospital discharge (low certainty for 1 study) *(53)* **Mixed or unclear impact** – Improvement in some disease-specific quality measures with a deterioration in others (moderate, low and very low certainty) *(55)* – Reduction in the number of people enrolled in chronic care programme for diabetes (low certainty) *(54)* – Decreased visits to primary care provider (low certainty for 2 studies) *(52, 55)*, increased visits to emergency department (low certainty) *(55)*
Service delivery and supportive elements	**Positive impact** – For an accountable care organization *(48, 49, 53)*: improved care coordination and management, cross-sectoral cooperation, preventive and health promotion programmes, data-driven approaches, monitoring and evaluation, patient registries and electronic medical records – For patient-centred medical homes *(50, 51)*: better care management through patient outreach, care coordination, improved care transitions, learning collaboratives, use of electronic medical records, preventive services – Population-based primary care model *(52)*: general practitioners responsible for continuum of care, shift to remote care, tracking of quality and performance **Mixed impact** – For an accountable care organization as part of a health maintenance organization *(55)* – Primary care centres *(54)*

Of the eight articles in this cluster, the authors of six studies were primarily positive in their evaluation of the impact of the shared-savings and shared-risk arrangements, with mixed certainty (moderate and low certainty in one study, low certainty in four, and low and very low certainty in one). Of the six studies that found a positive impact, four looked only at shared-savings arrangements and two looked at shared savings and shared risk. In all of the studies indicating a positive impact, shared savings were conditional on performance or quality measures, ranging from measures related to chronic disease management to preventive care to patient-centred medical home (PCMH) accreditation (see Glossary), patient satisfaction and more. Several articles did not specify the measures through which shared savings were measured *(48, 49)*.

The arrangements had positive impacts on various process measures, including improvements in measures of chronic disease care (e.g. for diabetes and increased cancer screening) (moderate or low certainty of evidence) *(48, 50-52)*. Positive impacts on health outcomes were also seen, as evidenced by shorter hospitalizations and fewer visits to the ED (low certainty) *(48, 50)*, higher age at death (low certainty) and longer survival time (very low certainty) *(49)*. Purchasing arrangements consisting of shared savings and shared risk were also mainly positive, with a reduction in visits to EDs, increased home health visits and increased follow-up visits after hospital discharge (low certainty) *(53)*. Navathe et al. *(52)* also found that implementation of the Population-based Payments for Primary Care programme (i.e. a shared-savings and shared-risk programme in Hawaii) led to an overall increase in a composite score indicating that quality goals had been achieved; there was also an improvement in process measures assessing advanced care planning and body mass index assessments (low certainty). A few authors also identified reductions in certain quality measures *(51, 52)*. These included a reduction in services offered by physicians (e.g. imaging, tests.) (low certainty) *(53)* and a reduction in eye examinations for people with diabetes (low certainty) and colorectal cancer screening (low certainty) *(51)*. These reductions in services may be either an indication of skimping as providers deliver fewer services or of providers making more conscious decisions about which services are necessary.

Of the six articles looking at shared savings that were conditional on performance or quality measures, three explicitly assessed the impact on measures that evaluators or regulators would use to determine the rate of shared savings. These included preventive measures (e.g. diabetes control, cancer screening, cholesterol screening), hospitalization, adherence to policies for prescription medicines and patient satisfaction *(50, 51, 54)*. The remainder provided insufficient detail to determine whether the outcomes measured were aligned with the design of the payments.

Two articles found that shared savings and shared risk, both with single commercial payers, had either no or conflicting impacts on quality, with low or very low certainty of evidence *(54, 55)*. These two studies, based on arrangements in Netherlands (Kingdom of the) *(54)* and California, USA, *(55)* reported mixed findings. Shared savings in Netherlands (Kingdom of the) that depended on quality metrics combined with bundled payments for chronic care delivery led to an increase in diabetes-related testing and control, but also to a reduction in quality measures for chronic obstructive pulmonary disease, as well as a decline in the number of individuals enrolled in a chronic care programme for people with diabetes *(54)*. This latter finding was the result of physicians not wanting to enrol patients with high prior no-show rates into the chronic care programme, as this would affect their score on the process indicators on which their shared savings rate was based. As pointed out by Kicinger et al. *(51)* in another article, these mixed findings may be the result of choice overload for providers in the indicators included in the evaluations of quality and performance used to determine their portion of the shared savings. Too many indicators may result in providers choosing particular indicators to invest in improving. Similarly, in a single-payer commercial shared-savings and shared-risk arrangement without quality incentives (the only one in this cluster), Zhang et al. *(55)* found mostly positive impacts in that some preventive processes improved (e.g. breast and cervical cancer screening, immunizations, diabetes control) (moderate or very low certainty), and visits to specialists increased (moderate certainty). At the same time, scores on other measures decreased, such as for visits to PCPs (low certainty) and colorectal cancer screening (moderate certainty), while ED visits increased (low certainty). In this latter case and in the absence of explicit quality incentives built into the arrangement, the authors suggest that the ACO (see Glossary for details) shared-risk model realigned financial incentives to encourage PCPs to refer patients with chronic conditions to lower-priced specialists within the network rather than to more costly out-of-network specialists to help manage chronic conditions more efficiently.

A majority of the shared-savings and shared-risk purchasing arrangements were implemented in conjunction with a service delivery model that included collaboration and coordination across providers and settings. For example, articles found a positive impact when shared savings and shared risk were implemented within ACO arrangements. ACOs are provider-led networks of health care professionals who work across different settings (e.g. physicians, or in hospitals and other settings) that agree to be accountable for the costs and quality of care for their assigned population of patients; therefore, they directly take on the financial risk and reward of caring for these patients. Care improvement activities included, for example,

the implementation of a disease-specific registry that identified patients with certain conditions who needed follow up to ensure continued and preventive care *(48)* and the use of preventive and health promotion programmes for specific conditions *(49)*. The use of digital health interventions – such as electronic medical records, automated alert systems and data-driven approaches to decision-making based on internal monitoring and external evaluations – were also incorporated into these arrangements *(48, 49)*.

In two cases, shared savings were implemented as part of a PCMH model and this had a positive impact on the quality of care *(50, 51)*. Also referred to as medical homes in some countries, PCMHs are models of team-based health care delivery led by a health care provider who coordinates the patient's care (see Glossary for more information). The health care provider takes on responsibility for the patient's care and coordinates and arranges the necessary services with other providers and across settings. The core attributes of the model include comprehensive and patient-centred care, and care management. In the CareFirst Blue Cross Blue Shield model in Maryland, USA, some of the quality components related to service delivery were financially incentivized through measures that assessed PCPs' engagement with the PCMH programme and care plans, certain information technology capacities, whether there were office hours on weekends and medical home certification *(51)*. Similarly, the northeastern Pennsylvania Chronic Care Initiative in the United States, implemented by two commercial providers, developed capabilities in care management, patient outreach, care coordination, referrals to community services and follow-up care. Both the model in Maryland and the model in Pennsylvania incorporated digital health tools, but with different aims: some were based on cost containment (e.g. electronic prescribing, online appointment scheduling, electronic test ordering), while others were based on quality improvement initiatives, such as in Pennsylvania's Chronic Care Initiative, which implemented a web-based disease registry to generate quality reports. Both models were required to obtain medical home certification, which meant they had to achieve certain requirements relating to care management.

3.2.6 Pay-for-coordination of care

Table 9. Summary of the literature about pay-for-coordination models

Pay-for-coordination	Description
Summary *(56-67)*	– Coordination fees in addition to fee-for-service arrangements *(56-60)* – Coordination fees and performance incentives in addition to fee-for-service arrangements *(61-67)*
Country	USA
Setting	Primary care
Impact on quality	**Coordination fees** **Overall impact** Positive or mixed impact with low or very low certainty *(56, 57, 58)*, mixed impact with low certainty *(59)*, and no impact (no certainty of evidence able to be attributed) *(60)* **Positive impact** – Improvements in care and treatment: reduction in use of some low-value services (low certainty) *(56)*, higher likelihood of patients with multiple chronic conditions receiving recommended services (very low certainty) *(58)*, hypertension care (very low certainty) *(58)* – Improvements in utilization: reduction in use of some low-value services (low certainty) *(56)*, reduced use of secondary and tertiary care (very low certainty) *(57)*, decrease in hospitalization for patients with both hypertension and cancer together (low certainty) *(59)*, reduction in inpatient and outpatient utilization (low certainty) *(56)* – Hospital readmission (very low certainty) *(57)* **Mixed impact** – Care and treatment: diabetes care (low and very low certainty) *(56-58, 60)*, care for mental health issues (very low certainty) *(58)*, asthma care (very low certainty) *(57, 58)*, paediatric care (low certainty) *(56)* – Utilization: number of primary care visits (very low certainty) *(57, 60)*, hospital admissions (low and very low certainty) *(57, 59, 60)* – Emergency department visits (very low certainty) *(56, 57, 59, 60)*, specialist visits (very low certainty) *(57, 60)* – Screening: breast cancer (low certainty) *(56, 57)*, cervical cancer (low certainty) *(56, 60)* **No impact** – Medication adherence and emergency department visits for patients with chronic conditions *(59)* – Colorectal cancer screening *(57, 60)*

Pay-for-coordination	Description
Impact on quality *(continued)*	**Coordination fees and performance incentives** **Overall impact** Positive impact with moderate, low and very low certainty *(61, 63, 65)*; mixed impact with high, low and very low certainty *(62, 64, 66)*; and no impact (no certainty of evidence could be attributed) *(67)* **Positive impact** – Utilization: increase in outpatient care (low certainty) *(65)*, reduction in preventable emergency department visits (low certainty) *(61)*, decrease in emergency department visits sensitive to ambulatory care (low certainty) *(66)* **Mixed impact** – Care and treatment: cardiovascular care (high certainty) *(63, 65)*, diabetes care (high, low and very low certainty) *(62, 64-67)* – Utilization: office-based visits (low certainty) *(65, 66)*, emergency department visits (high, moderate, low and very low certainty) *(61-66)*, inpatient admissions (moderate certainty) *(62, 65, 66)*, preventable inpatient admissions (low certainty) *(61, 62)* – Screening: breast cancer (high and very low certainty) *(62, 64, 66, 67)* **No impact** – Screening: cervical and colorectal cancer *(62, 66, 67)*, chlamydia *(62, 64)* – Hospital readmission *(61, 62)*
Service delivery and supportive elements	– Primary care settings: pay-for-coordination as part of implementation of patient-centred medical home or similar model – Patient-centred medical home elements: shift from care provided mainly by physicians to care provided by a team of health care professionals, including nonmedical staff, with more emphasis on prevention and chronic disease management instead of on acute care; combination of an individual patient and population perspective; increased emphasis on care coordination across providers and settings; implementation of health information technologies; more emphasis on evidence-based programmes and guidelines; specific programmes for defined patient groups (e.g. patients with diabetes, patients with chronic pain); participation of providers in peer-learning and quality-improvement activities

The fourth cluster included 12 articles that analysed pay-for-coordination arrangements, which, in most cases, paid providers in addition to baseline FFS arrangements, although not all articles provided information about the baseline payment (Table 9). Within this cluster, we differentiate between (i) arrangements that combine FFS payments with additional coordination fees and (ii) arrangements that also included financial performance incentives (i.e. P4P elements). In most of the articles analysed, coordination fees – with or without additional performance incentives – were introduced as part of the implementation of PCMHs or similar

models. Therefore, the findings presented in these articles and the authors' conclusions refer to the impact of PCMHs on quality and not only on the underlying purchasing arrangement.

Regarding the five studies that assessed the impact on the quality of care of models that included pay-for-coordination fees but no additional performance incentives, four reported positive effects, although the impact was modest in most cases, and the certainty of the evidence was generally low or very low. Jones et al. *(56)* report a positive impact of Vermont's medical home model on the quality of care for the whole population, including a reduction in the use of some low-value services, higher screening rates for cancer and better diabetes care (low and very low certainty). Kern et al. *(57)* found that a PCMH model in the Hudson Valley region of New York led to modest improvements in diabetes care, higher rates of primary care visits and lower rates of secondary and tertiary care utilization (very low certainty). In the Community Care of North Carolina model, a PCMH for Medicaid beneficiaries, enrolment for patients with multiple chronic conditions was associated with an increased likelihood of them receiving certain recommended services (very low certainty) *(58)*. For patients with hypertension and an additional cancer diagnosis, enrolment decreased hospitalizations (low certainty) *(59)*. The one article reporting no effects on the quality of care and the utilization of services assessed a PCMH pilot in New Hampshire, USA *(60)*.

Seven articles analysed models that combined pay-for-coordination elements with performance incentives, and all but one reported at least some positive effects on the quality of care, with the certainty of evidence ranging from high to very low. Although little detail was provided about how performance was measured, performance incentives were usually tied to meeting certain quality and utilization thresholds. In some cases, providers received additional per-member-months performance-based payments *(61, 62)*; in other cases, additional payments tied to performance were paid annually *(63)*; and in some cases, the amount of per-member per-month care management fee was bound to performance *(64, 65)*. For a PCMH in New York, USA, Fifield et al. *(64)* reported a reduction in ED visits, as well as an increase in blood pressure control among patients with hypertension and an increase in breast cancer screening. While the certainty of evidence was high, these were just a few of the indicators assessed, and most of the others did not improve. In an assessment of the Quality Blue Primary Care programme implemented by Blue Cross Blue Shield in Louisiana, USA, Shi et al. *(65)* found that enrolment was associated with an increase in outpatient care (low certainty), better diabetes care (low certainty) and a decrease in inpatient admissions (moderate certainty), but also with an increase in ED visits (moderate certainty). This last finding contrasts with most other articles in this cluster, which

reported decreases in ED visits in connection with the implementation of PCMH models *(61, 63)* (low and very low certainty of evidence). Modest improvements were also reported for a PCMH model in Cincinnati, Ohio, USA, which included a reduction in ambulatory care–sensitive ED visits and improvements in diabetes care *(66)* (low certainty); a PCMH model in New Jersey, USA, increased rates of mammography and nephropathy screening *(62)*. While most articles in this subgroup found some positive effects of coordination fees and performance incentives on the quality of care, for a Michigan-based PCMH model, no positive effects could be detected *(67)*.

The pay-for-coordination arrangements with and without additional performance incentives analysed in the literature discussed in this cluster have mainly been implemented as part of PCMH models and in connection with changes to care delivery processes, with the intention of fostering and sustaining these processes. Even though there were model- and site-specific differences, typically these changes included an extension of care teams (to include different health care professions and nonmedical staff, such as social workers, health educators and care coordinators), which were then better suited to provide comprehensive care and coordinate care across health care settings. Furthermore, many of the models focused on chronic disease management instead of on acute care, combined individual-patient and population-health perspectives, implemented evidence-based programmes and guidelines, and implemented specific programmes for defined patient groups (e.g. patients with diabetes, patients with chronic pain). Furthermore, the combined individual-patient and population-health perspectives included assessments of barriers to treatment adherence, such as health literacy, psychosocial issues, and living and working environments *(65)*, or the coordination of care for high-utilizing and high-risk patients *(62)*. In many cases, this combined perspective was supported by the use of individual- and population-level health data and health information technologies. Some of the PCMHs also received additional initial funding or assistance in the form of technical support, practice redesign and organizational support. Further elements that were introduced by transitioning to a PCMH model were participation in quality improvement activities, such as learning collaboratives, and continuing medical education programmes.

Because transformation to a PCMH model often comprise a combination of changes to infrastructure and service delivery, the literature suggests that it is difficult to attribute quality improvements to individual components. Additionally, PCMHs often differ in their design, even if they are part of the same initiative *(56)*. Furthermore, the literature suggests that the time between implementation and evaluation plays an important role in explaining

differences in quality outcomes and that the effects of PCMHs take years to unfold *(57)*. In addition, differences in outcomes may also be related to PCMHs targeting different populations *(61, 66)*. However, it has been suggested that particularly relevant to improving the quality outcomes of PCMHs and similar models are the incorporation of care managers as part of the health care team (*64)*, the additional support offered during transformation *(56)* and higher coordination fees than are currently paid out *(62)*. However, it is worth mentioning that the articles analysed in this cluster suggested that pay-for-coordination fees in PCMH models had a positive impact on quality, regardless of whether there was also an additional performance incentive.

The two articles that found no effects suggested that this may be due to the short time since the PCMH was implemented, the low value of the financial incentives, the relatively healthy population targeted and the lack of support for PCMH transformation *(60, 67)*. Werner et al. *(62)* suggested that coordination fees of 2% of the average underlying payment were insufficient to incentivize change. Markovitz et al. *(67)* also suggested that limited impact was the result of unequal requirements across public and private payers in that the latter were not required to adopt the incentive structure (in this case, the combination of pay-for-coordination and P4P elements) that was created for Medicare beneficiaries.

3.2.7 Blended capitation models

In this cluster, we examine a blended capitation model compared with a blended FFS model in primary care, as seen in Ontario, Canada. Further details about the two models are shown in Table 10 (and in the Glossary). In the blended capitation model, family practitioners in Family Health Organizations received a majority of their earnings from capitation based on a predefined bundle of services provided to enrolled patients, with a minor part based on FFS and P4P. The blended FFS model known as Family Health Groups is based mostly on FFS with only minor capitation and P4P components. Both models consist of targeted quality-related financial incentives based on measures of illness prevention (e.g. influenza vaccination, colorectal screening) and chronic disease management (e.g. diabetes management, smoking cessation counselling, heart failure management). In addition, financial incentives are based on whether after-hours care is provided, patients with severe mental illness are enrolled and on aftercare provided following hospital discharge. Additionally, the blended capitation model provides bonuses to physicians if they provide services that are part of the bundle to enrolled patients, thereby improving access for patients; the bonuses are reduced if enrolled patients receive these services from other physicians *(68)*. The impact of the blended capitation model relative to blended FFS is

primarily positive (low certainty in one study, very low certainty in two studies), while one article had no conclusive conclusion (low certainty).

Table 10. Summary of articles about blended capitation models

Blended capitation	Description
Summary *(68-71)*	– Family Health Groups: blended FFS (i.e. more FFS than capitation) plus financial incentives for preventive care and disease management – Family Health Organizations: blended capitation (mostly capitation with some FFS) plus financial incentives for preventive care and disease management
Country	Canada
Setting	Primary care
Impact on quality	**Overall impact** Overall, the impact was positive with low certainty in 1 study *(69)* and with very low certainty in 2 studies *(68, 70);* there was an unclear impact in 1 study with low certainty *(71)* **Positive impact** – The blended capitation model improved team-based care and process measures related to diabetes care and cancer screening (low certainty) *(69)* – Blended capitation increased after-hours care related to mental health, reduced visits to emergency departments and improved care for mental health during regular hours (very low certainty) *(68)* – Blended capitation reduced psychiatric hospitalizations relative to blended FFS (very low certainty) *(69)* **Mixed or unclear impact** – Switching to blended capitation led to decreased provision of capitated services and services to non-enrolled patients, but it also increased the provision of services outside of the capitated basket; unclear impact on quality (low certainty) *(71)*
Service delivery and supportive elements	– Patient-centred medical home–like model: physician-led governance, formal patient enrolment, disease management, after hours care – Blended capitation model also provided opportunity for additional funding to develop interdisciplinary team-based practice across different providers – Capitation models with team-based practice had the best outcomes relative to blended FFS and non-team-based blended capitation

FFS: fee for service.

All four articles investigated the impact of switching from blended FFS to blended capitation, but they focused on different aspects. The articles found that the blended capitation model (the Family Health Organizations) positively impacted certain quality outcomes relative to the blended FFS model (the Family Health Groups) (low certainty for one study, very low for two studies), but otherwise had a mixed impact, given that the design of the model incentivized physicians to provide services not included in the capitated bundle (low certainty). Kiran et al. *(69)* found that shifting to blended capitation improved some process measures related to diabetes care (low certainty) and for cervical cancer screening (low certainty). Two articles focused on the impact on mental health services. One found that switching to blended capitation had a positive impact as it led to a decrease in ED visits for mental health reasons and a decrease in the number of mental health services provided during regular hours, which was, however, compensated for by an increase in after-hours mental health services (very low certainty). The authors reported no change in referrals to psychiatrists *(68)*, which they attributed potentially to a reduction in service overprovision. Similarly, Vu et al. *(70)* found that switching to blended capitation reduced psychiatric hospitalizations (very low certainty), but had no effect on follow-up visits within 14 days of a psychiatric hospitalization. Physicians' characteristics, such as age and gender, moderated the impact of switching between models on the utilization of mental health services *(68)*.

However, one article found a seemingly negative impact in that switching led physicians to increase the services provided outside of the bundle of services, while reducing capitated services *(71)*. Switching led to a reduction in total services, capitated services and services provided to non-enrolled patients, but at the same time led to increased access to services through an increase in after-hours services and nonincentivized services *(71)*. Switching also resulted in a reduction in the total number of patients per physician, with enrolled patients decreasing and non-enrolled patients increasing. The impact of these changes on the quality of care is unclear in the absence of additional information, but it suggests that the relative change in the price of a service as determined by its inclusion in a capitated bundle can influence physicians' behaviour. In the pay-for-coordination cluster, Markovitz et al. *(67)* suggested that this mechanism could be used beneficially to incentivize the provision of important services, such as for chronic care, by leaving these services outside of the capitated bundle.

Both the blended capitation and blended FFS models are models of primary care service delivery that require physicians to provide comprehensive primary care. Both the blended FFS and blended capitation models require formal patient enrolment and a commitment to after-hours care to reduce the use of other clinics

outside of regular hours and to promote the continuity of care. Additionally, both models allow for bonuses to be paid that are related to chronic disease management and prevention services, and both are physician-led in terms of their governance. However, the Family Health Organization model (blended capitation) contains additional aspects that encourage collaboration and the coordination of care across providers, such as additional funding available to PCPs to form an interdisciplinary team-based practice with other professionals (e.g. nurses, dietitians, social workers, pharmacists, and specialists in mental health and addiction care). As evidenced in Kiran et al. *(69)*, having this additional team-based care component alongside capitation improved outcomes related to diabetes care more than that seen in a non-team-based blended capitation model or blended FFS. This suggests that team-based care implemented alongside capitation can help facilitate chronic disease management.

3.2.8 Comparisons across arrangements

Table 11. Summary of literature comparing different types of arrangements

Comparisons of different arrangements	Description
Summary *(72-74)*	– CCOs: global budget with bonus payments and full financial risk for providers; RCCOs: FFS combined with per-member per-month and bonus payments *(72)* – PSNs and POPs: shared savings and monthly case-management fees in addition to FFS; HMOs: capitated payments made by the state, full financial risk, with providers reimbursed in different ways (e.g. FFS or capitation) *(73)* – Three types of integrated payment: P4P, pay-for-coordination with coordination fees, and all-inclusive payments *(74)*
Countries	20 European countries *(74)*, USA *(72, 73)*
Setting	Primary care

Comparisons of different arrangements	Description
Impact on quality	**Overall impact** Positive or mixed impact, with moderate, low and very low certainty *(72, 73)*; no impact, with very low certainty*(74)* **Positive impact** – CCOs: reduction in avoidable emergency department visits and preventable hospitalizations, and increase in preventive care as compared with RCCOs (low certainty) *(72)* – PSNs and POPs: better ratings of doctors and plans as compared with HMOs (very low certainty) *(73)* – POPs: better rating for speciality care, making specialist appointments and access to care and tests as compared with HMOs (very low certainty) *(73)* – PSNs: better ratings for finding a personal doctor, receiving help and receiving prescriptions as compared with HMOs (very low certainty) *(73)* **Mixed effects** – Paediatric care and reduction in use of low-value services in CCOs (moderate and low certainty) *(72)* – Access to primary care in P4P types and to health care generally in all-inclusive types (very low certainty) *(74)* **No effect** – In CCOs and RCCOs: preventable inpatient days *(72)*, emergency department visits *(73)*, asthma care *(73)*, overall rating of health care in PSNs and POPs *(74)* – Self-reported health (in all types) *(74)* **Negative impact** – CCOs and RCCOs: decrease in primary care visits significantly better in CCOs (low certainty) *(72)* – Limitations in activities of daily living in P4P types (very low certainty) *(74)* – Quality of life in P4P types (very low certainty) *(74)*
Service delivery and supportive elements	**Positive impact** – PSN and POPs: managed care arrangements, provision of all services, care coordination, utilization management, health improvement activities; POPs focus on paediatric care *(73)* **Mixed impact** – CCOs and RCCOs: focus on primary care, emphasis on coordination of care, programmes for high-utilizers, reduction in emergency department visits and more support for social services; CCOs also include oral and mental health, hospital-to-home transition programme, as well as additional funding for implementation, administrative staff, data infrastructure and training *(72)*

CCO: coordinated care organization; FFS: fee for service; HMO: health maintenance organization; P4P: pay for performance; POP: paediatric-only plan; PSN: provider service network; RCCO: regional collaborative care organization.

Three articles compared different alternative purchasing arrangements (Table 11). Because they analysed different arrangements, overall conclusions cannot be drawn, so we provide a short summary of each of the articles.

McConnell et al. *(72)* compared two different Medicaid ACO models, implemented in Oregon and Colorado, United States. The Oregon model was a coordinated care organization (CCO) that combined a global budget with bonus payments and full financial risk for providers, whereas the Colorado model was a regional collaborative care organization (RCCO) based on a FFS model combined with per-member per-month and bonus payments in which providers did not face any risk. In terms of quality, the authors showed that the CCO model compared with the RCCO model, which was more limited in scope, reduced avoidable visits to the ED and preventable hospitalizations and increased access to some health care services, such as adolescent well-care visits, the only measure for which the certainty of evidence was moderate. However, while both models focused on primary care, primary care visits decreased in both models and were significantly lower in the CCO model compared with the RCCO model. In terms of service delivery, both models focused on primary care settings, aimed to improve the coordination of care and implemented programmes for high utilizers as well as to reduce visits to EDs; they also provided support for social services. However, the CCO model was more comprehensive and in terms of service delivery also included oral and mental health. Furthermore, the CCO model included hospital-to-home transition programmes, and providers received funding for implementation, administrative staff, data infrastructure and training. Also, governing boards were installed that included representatives from the health care delivery system and patients. In the CCO model, centralized data repositories were implemented to track and report clinical performance.

Hall et al. *(73)* analysed how parents rated their children's care in two types of Medicaid MCOs in Florida, USA, and compared these with HMOs. Provider service networks (PSNs) provide all services for a defined population and are also charged with care coordination, utilization management, and implementation of health improvement activities. In terms of purchasing, they combine shared savings with monthly case-management fees on top of FFS. Paediatric-only plans (POPs) are similar to PSNs but focus on providing coordinated care to children. For POPs, it was also specified that they employed care coordinators to work with physicians, families and social workers to schedule and follow up on health care services. HMOs, in contrast, were funded by the state through capitated payments. They assumed full financial risk and reimbursed providers in different ways (e.g. capitation or FFS payments). In terms of quality, parents of children enrolled in PSNs and POPs rated their doctors and their plans better than parents whose children were enrolled in HMOs.

PSNs and POPs also scored better on a few more specific indicators (all very low certainty). The authors concluded that there is some evidence that PSNs and POPs might result in better experiences and greater satisfaction with care.

Hayes et al. *(74)* compared different purchasing arrangements and their impact on health and health care use in persons with multimorbidity across 20 different European countries. In particular, they analysed purchasing arrangements that were not integrated and compared these with three types of integrated payment: P4P, pay-for-coordination and all-inclusive payments (i.e. bundled payments and risk-adjusted global payments). Despite the theoretical predictions and the rhetoric in many policy documents, the authors found little effect of the different payment methods on key outcomes for multimorbid individuals (low certainty). The particular service delivery models and supportive elements differed widely across countries, and the authors concluded that the success of any payment method likely relies on the specific design of incentives and their implementation. These include the responsiveness of providers to the scheme, the size of the incentives, the different patient groups targeted, and how the aspects of care that are incentivized interact with those that are disincentivized.

3.3 Summary of findings

This scoping review identified five distinct clusters of purchasing arrangements, with most clusters of articles providing evidence of a positive impact on the quality of care for people with chronic illness. Table 12 summarizes the findings and highlights the design and service delivery aspects in each cluster that authors have attributed to an arrangement's success or lack of success, as well as the common factors found across articles in which the impact on care quality was positive.

Table 12. Summary of findings and lessons learned across the different types of purchasing arrangements

Type of purchasing arrangement	Findings	Description
Capitation and global budget plus quality or performance incentives	Positive	– **Purchasing design**: Baseline model of capitation and global budget may incentivize improvements in some dimensions of care quality, such as continuity of care. The addition of quality or performance incentives can facilitate improvements in certain aspects of care quality. – **Service delivery**: This arrangement required improving the continuity of care and incorporating the use of health information technology. It also included strengthening the roles of general practitioners (e.g. in terms of gatekeeping, or providing comprehensive care and disease-management programmes).
	Potential barriers and lessons learned	– **Purchasing design**: Penalties based on performance must be designed with enough incentive to potentially cause a negative impact on providers if targets or goals are not achieved. – Payment reforms take time to have an effect.
	Example: Maryland All-Payer model (USA)	– **Lack of impact**: A lack of impact was attributed to the use of inconsistent incentives between hospitals and physicians. The programme's structure, with hospitals assuming all of the risk, did not incentivize physicians in hospitals to reduce volumes. – Financial incentives may have had more of an impact on outpatient services, given the flexibility of decisions that could be made in these services. – There was limited coordination between hospitals and community providers, since the latter were not included in the global budget.
	Example: Massachusetts Alternative Quality Contract (USA)	– **Service delivery**: There was substantial variation across providers, with varying degrees of integration between different levels and sites of care provision (i.e. primary, specialist, hospital care) and whether a centralized administrative structure to support quality improvement was available. – **Strategies**: These included improving care coordination by engaging in case management across multiple payers for patients with high-risk conditions, changing referral patterns and redesigning care processes to eliminate waste. – **Common elements**: These arrangements used periodic reports about quality performance to identify areas for improvement; providers received support to improve quality and engaged in peer-learning activities.

Type of purchasing arrangement	Findings	Description
Bundled payments	Positive	– **Purchasing design**: Bundled payments can improve some measures of patients' experience for certain target groups when part of an integrated, person-centred delivery model. – **Supportive elements**: The elements found to support implementation included ensuring there was comprehensive assessments of an individual's needs, developing individualized care plans based on the individual's needs, ensuring there was a case manager role, holding multidisciplinary meetings that included health care professionals and any informal carers.
Pay-for-coordination	Positive	– **Purchasing design**: Fees for service with additional fees for coordination can have some positive impact on the quality of care and on service utilization. Higher coordination fees can incentivize providers. – **Service delivery**: Pay-for-coordination implemented with a PCMH model can have a positive impact on the quality of care. The presence of care managers and additional support during the transformation to a PCMH model may improve the likelihood of positive impact.
	Potential barriers and lessons learned	– **Purchasing design**: If there are different requirements for public and private purchasers to adopt an incentive structure, then these may disincentivize private purchasers from participating; coordination fees that are about 2% of the average underlying payment may be too little to incentivize changes in providers' behaviour.
Shared savings and shared risk	Positive	– **Purchasing design**: Shared-savings and shared-risk arrangements with rates that are conditional on meeting performance and quality measures have the potential to positively impact the quality of care, particularly in terms of the management and control of chronic diseases and preventive care. In ACO arrangements with shared savings and shared risk, primary care physicians may be incentivized to refer patients with chronic conditions to lower-priced specialists or services within their network as a way to avoid high costs for out-of-network providers and to better manage chronic conditions. – **Service delivery**: ACO and PCMH service delivery models include activities related to care coordination, care management and patient-centred care. – **Supportive elements**: The elements found to support implementation include establishing a disease-specific registry, offering preventive care and health promotion programmes, and requiring certification as a medical home.

Type of purchasing arrangement	Findings	Description
Shared savings and shared risk	Potential barriers and lessons learned	– **Purchasing design**: Choice overload may lead providers to prioritize some measures of quality or performance over others.
Blended capitation	Positive	– **Service delivery**: Team-based care with blended capitation can help facilitate the management of patients with chronic disease relative to blended capitation or blended fee-for-service models without team-based care.
	Potential barriers and lessons learned	– **Purchasing design**: Providers may be incentivized to overprovide services not included within capitated bundles if the fees are set high enough.

ACO: accountable care organization; PCMH: patient-centred medical home.

Within the capitation and global budget cluster, the articles suggest that these arrangements may incentivize improvements in some dimensions of the quality of care, such as improving the continuity and coordination of care, process outcomes (e.g. ensuring increased care services for people with chronic diseases and improving indicators of disease management or control for chronic diseases), health outcomes and the efficiency of services. These impacts were seen regardless of whether there were additional quality or performance incentives, although when additional incentives were used the research showed that they can facilitate an improvement in certain process measures related to the control or management of chronic diseases. In terms of capitation and global budgets encompassing a defined set of services, some evidence from this scoping review suggests that the particular design of the bundle of services can impact providers' behaviour. Excluding services from a capitated bundle can incentivize the provision of services outside of the bundle, given their relatively higher price.

Capitation and global budgets implemented under a wide range of service delivery models were successful in positively impacting the quality of care. Among these delivery models were integrated care models, PCMHs, ACOs, MCOs, comprehensive care in primary care, and broader models of primary care and hospital-based care. Common elements across these delivery models included improving the continuity of care through patient enrolment and by scheduling patients at regular intervals after hospitalization, the use of health information technology for quality improvement and decision-making, and strengthening GPs' roles (e.g. for gatekeeping and comprehensive care, and in disease-management programmes). In cases in which the impact on quality was mixed or limited, authors concluded that penalties based on performance must be designed with enough incentive to potentially cause a negative impact on providers who do not achieve the targets and that reforms take time to become established.

Numerous articles focused on the AQC in Massachusetts and the Maryland All-Payer Model, both of which are global budget programmes implemented in the USA. The former was an ACO-like model, encompassing primary and specialty care that was reimbursed by a global budget with shared savings and shared risk based on selected measures of care quality. The articles looking at the AQC indicated there was primarily a positive impact on quality in terms of improving chronic disease management and increasing preventive services, although there was only a low or very low certainty of evidence. While strategies varied substantially across providers, key service delivery elements included making changes to referral patterns, encouraging peer support, engaging in case management for high-risk patients, redesigning processes to eliminate waste and improve quality, and ensuring contracting

support from physicians' organizations. The articles about the Maryland global budget All-Payer Model, which covered hospital services, were more mixed in their conclusions, which had low or very low certainty. The model had a positive impact on some outpatient measures, but effects were positive on only a few inpatient measures and outcomes, particularly for three major cardiovascular conditions. The programme had a nonuniform impact on vulnerable groups and also had a limited impact on reducing hospital use or primary care among Medicare beneficiaries.

The cluster that evaluated pay-for-coordination arrangements found a primarily positive impact on quality, with varying certainty of evidence. These arrangements positively impacted quality in terms of the management of chronic diseases (i.e. diabetes, hypertension, asthma), screening for cancer, increased primary care visits and reduced ED visits. These positive impacts were seen regardless of whether the arrangement included performance and quality incentives in addition to the coordination fees. Only a few articles found no impact on quality, which the authors attributed to an insufficient value of the financial incentives, a lack of consistency in that many private payers did not also adopt the incentive structure, and a lack of support for the transformation to the new payment model. Nearly all of the pay-for-coordination models studied were implemented as part of a PCMH model, which deployed team-based care that included care managers, focused on chronic disease management and implemented evidence-based programmes, among other changes. Where the impact was positive, providers received sufficient support in terms of financing and consulting during the transformation of the service delivery model. Quality improvement activities, such as implementation of learning collaboratives, continuing medical education programmes and the use of health information technology, were also commonly part of arrangements that positively impacted the quality of care.

The cluster examining shared savings and shared risk also provided evidence of a positive impact on quality in terms of improved process outcomes for indicators of disease-specific management, cancer screening and preventive care, as well as for improved health outcomes and continuity of care. In cases in which rates of savings were conditional on performance and quality measures, these tended to have a positive impact on the quality of care, particularly in terms of the management and control of chronic diseases and for preventive care. However, a few articles found an unclear impact in which some chronic disease management measures improved in some areas while they deteriorated in others. In cases in which the impact was positive, shared savings and shared risk were implemented as part of PCMHs, ACOs and population-based primary care models that promoted activities related to care coordination, care management and patient-centred care. Disease-specific

registries, preventive and health promotion programmes, and requirements for medical home certification were some common elements seen in programmes that had a positive impact.

The bundled payment and blended capitation clusters also offered evidence of a positive impact on the quality of care, although the evidence was of low and very low certainty, and there were fewer articles from which to draw conclusions. Bundled payments, as implemented in Netherlands (Kingdom of the) as an integrated model and in the USA as a hospital-based model, showed some positive impacts on person-centredness and healthy days at home, but there were also some negative impacts, such as reduced patient autonomy. Evaluation of the blended capitation model in Ontario, Canada, also found positive impacts related to chronic care management, some screening measures and health outcomes, but there was also a shift in services (e.g. from occurring during regular hours to after-hours) and a change in the composition of services, based on whether they were included in the capitated bundle. This latter finding indicates that the relative price of services can impact providers' behaviour. A key conclusion from these two arrangements is that purchasing positively impacted the quality of care when these arrangements were implemented in conjunction with a service delivery model that incorporated multidisciplinary team-based care delivered to patients. The CCFE programme for frail adults in Netherlands (Kingdom of the) particularly exemplified this because integrated care was provided based on individualized care plans for patients with complex needs. Similarly, blended capitation combined with the team-based delivery of services had the most impact on improving the quality of care in the Family Health Organization model in Ontario, Canada.

4

Discussion

This scoping review provides evidence, mostly from the United States but also from China and Europe, that purchasing arrangements and their accompanying service delivery features have positively impacted the quality of chronic care. The clusters examining capitation and global budgets, shared savings and shared risk, and pay-for-coordination found a mostly positive impact of these arrangements, some of which was of moderate or high certainty. Only a small number of articles within these clusters indicated that the purchasing arrangements had a mixed impact (i.e. both positive and negative), a limited impact according to the authors or no impact at all, with the evidence being of moderate certainty. Consolidating this evidence, the literature suggests that while these three types of purchasing arrangements can have a positive impact on quality, this is not always the case and rather the particular design and structure of the purchasing arrangement and the associated service delivery model are key to having an impact on quality.

The articles on the Massachusetts AQC indicated that this arrangement had primarily a positive impact on quality, suggesting that purchasing arrangements implemented by private commercial insurers can also improve the quality of care when shared savings and shared risk are tied to quality performance. The articles about the Maryland All-Payer Model had more mixed conclusions, which the authors attributed to different incentives offered to hospitals and physicians in which all of the risk was placed on hospitals and there were no incentives for physicians to change their behaviour. There was also limited coordination with community providers given that they were not included in the global budget.

The literature on blended capitation and bundled payments also indicated they can improve the quality of care, although the evidence is more limited. Despite expectations within the bundled payments cluster that providers would underprovide services or adversely select patients, this was not the case. In the case of the BPCI-A programme in the USA, the use of quality incentives may have reduced the potential for the adverse selection of patients. A key conclusion from these two clusters is that purchasing arrangements positively impacted the quality of care when they were implemented in conjunction with a service delivery model that incorporated multidisciplinary team-based care delivered to patients, as opposed to implementing the purchasing on its own. Furthermore, literature from the blended capitation cluster based on a PCMH model in Ontario, Canada, highlighted the key finding that the design of the capitated bundle of services can impact providers' behaviour. Specifically, providers may be incentivized to overprovide services that are not included within the capitated bundle if the fee for out-of-bundle services is set high enough. Similarly, providers may be incentivized to underprovide services

within the bundle. In such cases, ensuring minimum standards by using quality incentives may be beneficial to prevent the over- and underprovision. Excluding services from a capitated bundle can incentivize the provision of services outside of the bundle, given their relatively higher price. This finding can be used to incentivize the provision of important services, such as those related to chronic disease management and prevention.

Findings from the shared-savings and shared-risk cluster also suggested that both of these types of arrangements can improve the quality of care. While literature has previously highlighted the potential for two-sided models (i.e. shared savings and shared risk used together) to have more impact on providers' behaviour than one-sided models (i.e. shared savings only), the literature included in this review suggest that both arrangements can provide sufficient incentives for providers to improve the quality of care.

This scoping review provides evidence that across the different types of purchasing arrangements, the use of quality- or performance-based measures as a condition for awarding or calculating the rate of a financial incentive can incentivize improvements in quality measures. This was seen in the capitation and global budget clusters, for single providers and for all care across providers, and also in the shared-savings and shared-risk, pay-for-coordination and blended capitation clusters. In particular, some moderate certainty evidence shows that tying financial incentives to the management of chronic diseases or preventive measures (e.g. cancer screening, immunizations) can improve how often these programmes are provided. At the same time, a few articles highlighted the potential issue of choice overload – that is, including too many measures of quality or performance may result in providers focusing more on a small subset of measures and less on other measures. This finding makes the case for focusing on a more concise bundle of measures based on prioritized areas.

However, in the absence of additional incentives requiring providers to meet or surpass a threshold of quality or performance, within these same clusters the quality of care still improved, suggesting that elements of the underlying baseline payment can also contribute to improving the quality of care, such as through improving care continuity. This was the case in the capitation, global budget and shared-savings and shared-risk clusters, despite expectations that providers might underprovide services as a way to stay within their budget and maximize their revenue, therefore compromising the quality of care. Population-based purchasing arrangements that integrate payments across providers based on the care to be provided to a particular population may still incentivize providers to think through their decisions more carefully and to take on a gatekeeping role as a way to manage costs.

Population-based models without additional quality incentives may also incentivize providers to improve their efforts in areas in which costs are likely to be higher, namely for patients requiring chronic care, by ensuring follow up of care, among others. However, omitting certain providers from the budget may not incentivize coordination across settings (e.g. in community and postacute services).

This was also the case in the pay-for-coordination cluster in which providers received per-patient fees to promote the coordination of care, which can also improve the quality of care even in the absence of additional quality incentives. This result is likely linked to the PCMH service delivery model in that this purchasing arrangement was usually implemented, which requires the provision of team-based, patient-centred care. However, in cases in which there was no positive impact within pay-for-coordination arrangements, the authors suggested that coordination fees may sometimes be too little to incentivize a change in providers' behaviour and to compensate for the increased workload of coordinating care. In one particular case, coordination fees that were 2% of the average underlying payment were assumed to be too little by the authors *(62)*. Similarly, other authors suggested that unequal requirements for public and private purchasers can disincentivize private purchasers from participating in these arrangements *(67)*.

In situations in which purchasing arrangements were found to have a positive impact on quality, service delivery models such as ACOs, PCMHs, integrated care models and comprehensive care by GPs were commonly seen, suggesting that combining these purchasing arrangements with collaborative, team-based and patient-centred care may contribute to improving the quality of care. More generally, a number of service delivery aspects and supportive elements may be used as instruments to positively affect the quality of care. Table 12 summarizes these findings across the clusters. The use of health information technology to inform prescribing, to share information across providers, to create disease registries and to schedule visits online was seen across purchasing arrangements that had positive impacts. A number of purchasing arrangements with a positive impact had aimed to strengthen the role of GPs through requiring them to act as gatekeepers controlling the access of patients to other services, providing comprehensive care and taking a lead in coordinating the care of patients. Several arrangements with positive impacts on quality implemented disease-management programmes, whether through a structured programme with GPs opting to take part or through disease registries to ensure patients with chronic diseases are regularly seen. Ensuring there was sufficient financial and organizational support to implement new service delivery models was also seen to correspond with positive impacts on the quality of care.

5

Limitations and gaps

In line with the finding that the success of purchasing arrangements often depends on the particular details of the arrangement, one clear gap found in the literature is the limited attention paid to the particular organizational details of purchasing arrangements. For example, very few articles contextualized the amount of the financial incentives relative to providers' or organizations' earnings to better understand the strength of the incentive on providers' behaviour. This is relevant information because the size of the incentive likely contributes to the probability of positively impacting service provision and behaviour. Similarly, a majority of articles investigated the impact of reforms to purchasing on the quality of care without attempting to disentangle how exactly the financing mechanisms may incentivize (or disincentivize) changes to service provision and providers' behaviour. Another gap is the lack of information across articles about how funds are distributed when services are contracted as a group, such as in the case of global budgets and capitation covering all care. However, a few exceptions to this exist, which we have highlighted throughout the results.

In addition, while this review aimed to highlight service delivery models implemented in conjunction with or as a result of the purchasing arrangements, this information was often vaguely described in the articles. For example, while many articles mentioned coordination across providers, further details describing how coordination was pursued were often not presented. Similarly, generic terms such as patient-centred care or team-based care were often used but not clarified in terms of how they were operationalized in practice. Furthermore, it was often unclear which services the purchasing and service delivery models included. Related to this, none of the articles attempted to separate the impact of service delivery from the impact of purchasing arrangements on the quality of care. Very few articles attempted to quantitatively capture and incorporate any service delivery aspects into their regression modelling.

While the literature included in this scoping review provided a wide range of evidence about the impact of purchasing arrangements and their service delivery models on the quality of care, the certainty of evidence was generally quite low. Among some of the most prevalent issues concerning the certainty of evidence is the selection bias often present in studies in which participation in the arrangement is voluntary. Providers who have the capacity to implement the purchasing arrangement and will likely succeed in fulfilling the requirements and maximizing their potential revenue are more likely to voluntarily participate in a reform, thus there is likely upward biasing of the positive impact of purchasing arrangements. Similarly, patients with worse health may be biased to participate in pilot interventions. Another issue is the optimistic conclusions derived by authors. For example, often improvement in

one or two measures was sufficient to warrant a positive evaluation. In other cases, the purchasing arrangement had both positive and negative impacts, yet the authors concluded that the impact was positive.

Given that most of the articles included in this review used a difference-in-differences approach, another issue is the failure to verify that there were parallel trends or no differences between the control and intervention groups prior to the intervention, thus making it uncertain whether differences between groups were related to the intervention and not a time or group effect. On this latter point, authors tried to address this by using propensity score matching, although unobserved differences may still not be controlled for between groups (e.g. preferences for types health care). Another uncertainty is how to isolate spillover effects from the implementation of other programmes or incentives introduced at around the same time. One article *(42)* exemplified this by presenting the impact of Blue Cross Blue Shield of Massachusetts' AQC programme on Medicare beneficiaries not covered by the programme.

In addition, in many cases, the measures of quality used in articles were not aligned with the quality measures used to award incentives, but rather were used as proxies to measure the impact on care quality. As a result, the conclusions of articles about the impact of the purchasing arrangement on the quality of care are inherently shaped by whether and what kind of data were available to researchers. This is likely the result of data limitations. In some cases, the precise measures of quality used to evaluate performance and award incentives were not outlined, making this issue difficult to judge. Where possible and where the data permit, in order to develop this area of literature about how purchasing arrangements can incentivize providers' behaviour, researchers should strive to use measures that are as close as possible to those used to evaluate providers within the arrangement, and to clarify their choice of indicators.

It is also important to note that most of the studies included in this review come from the United States. The complex and highly decentralized nature and fragmentation of the regulation, financing and provision of health care, as well as of health insurance coverage, in the United States may limit some of the findings' applicability to other countries (see Annex 2). Thus, arrangements in the United States with private purchasers may not always be applicable to settings where purchasing is carried out by a public health authority.

6

Conclusions

6

Scoping review on alternative purchasing arrangements: impact of purchasing, service delivery and institutional design on quality of chronic care

This scoping review aimed to provide a map of robust evidence about purchasing arrangements and their service delivery models that have been used to improve the quality for care for people with chronic diseases. To do this, we evaluated a wide range of literature that analysed the impacts of various purchasing arrangements, their service delivery models and supportive elements on the quality of care. The results presented highlight several findings to be considered when designing purchasing arrangements and their service delivery models, particularly in regards to how these arrangements can impact providers' behaviour. The consolidated literature provides evidence from a wide range of arrangements that can be used by policy-makers in countries wishing to improve the provision of care for people with chronic diseases.

References

1. GBD 2019 Diseases and Injuries Collaborators. Global burden of 369 diseases and injuries in 204 countries and territories, 1990–2019: a systematic analysis for the Global Burden of Disease Study 2019. Lancet. 2020;396:1204–22. doi:10.1016/S0140-6736(20)30925-9.

2. Noncommunicable diseases [website]. Geneva: World Health Organization; 2023 (https://www.who.int/news-room/fact-sheets/detail/noncommunicable-diseases#:~:text=Noncommunicable%20diseases%20(NCDs)%2C%20also,physiological%2C%20environmental%20and%20behavioural%20factors, accessed 4 April 2023).

3. Global status report on noncommunicable diseases 2014. Geneva: World Health Organization; 2014 (https://www.who.int/publications/i/item/9789241564854, accessed 4 April 2023).

4. Chowdhury SR, Chandra Das D, Sunna TC, Beyene J, Hossain A. Global and regional prevalence of multimorbidity in the adult population in community settings: a systematic review and meta-analysis. Clin Med. 2023;57:101860. doi:10.1016/j.eclinm.2023.101860.

5. Kazibwe J, Tran PB, Annerstedt KS. The household financial burden of non-communicable diseases in low- and middle-income countries: a systematic review. Health Res Policy Syst. 2021;19(1):96. doi:10.1186/s12961-021-00732-y.

6. Muka T, Imo D, Jaspers L, Colpani V, Chaker L, van der Lee SJ, et al. The global impact of non-communicable diseases on healthcare spending and national income: a systematic review. Eur J Epidemiol. 2015;30(4):251–77. doi:10.1007/s10654-014-9984-2.

7. Nolte E, Knai C, Saltman RB. Assessing chronic disease management in European health systems: concepts and approaches. Copenhagen: WHO Regional Office for Europe, European Observatory on Health Systems and Policies; 2014 (https://iris.who.int/handle/10665/326351, accessed 18 October 2023).

8. Delivering quality health services: a global imperative for universal health coverage. Geneva: World Health Organization, Organisation for Economic Co-operation and Development, International Bank for Reconstruction and Development; 2018 (https://iris.who.int/handle/10665/272465, accessed 18 October 2023).

9. Chee TT, Ryan AM, Wasfy JH, Borden WB. Current state of value-based purchasing programs. Circulation. 2016;133(22):2197–205. doi:10.1161/CIRCULATIONAHA.115.010268.

10. Custers T, Klazinga NS, Brown AD. Increasing performance of health care services within economic constraints: working towards improved incentive structures. Z Arztl Fortbild Qualitatssich. 2007;101(6):381–8. doi:10.1016/j.zgesun.2007.05.004.

11. Figueras J, Robinson R, Jakubowski E. Purchasing to improve health systems performance. Buckingham (UK): Open University Press; 2005.

References

12. Jakab M, Evetovits T, McDaid D. Health financing strategies to support scale-up of core noncommunicable disease interventions and services. In: Jakab M, Farrington J, Borgermans L, Mantingh F, editors. Health systems respond to noncommunicable diseases: time for ambition. Copenhagen: WHO Regional Office for Europe; 2018:200–23 (https://iris.who.int/handle/10665/342223, accessed, 13 November 2023).

13. Mathauer I, Dale E, Jowett M, Kutzin J. Purchasing of health services for universal health coverage: how to make it more strategic? Geneva: World Health Organization; 2019 (https://www.who.int/publications/i/item/WHO-UCH-HGF-PolicyBrief-19.6, accessed 4 April 2023).

14. Busse R, Blümel M, Scheller-Kreinsen D, Zentner A. Tackling chronic disease in Europe: strategies, interventions and challenges. Copenhagen: WHO Regional Office for Europe; 2010 (https://iris.who.int/handle/10665/326484, accessed 18 October 2023).

15. Quality of care [website]. Geneva: World Health Organization; 2023 (https://www.who.int/health-topics/quality-of-care, accessed 22 May 2023).

16. Donabedian A. Evaluating the quality of medical care. Milbank Q. 2005;83(4):691-729. doi:10.1111/j.1468-0009.2005.00397.x.

17. Hammermeister KE, Shroyer AL, Sethi GK, Grover FL. Why it is important to demonstrate linkages between outcomes of care and processes and structures of care. Med Care. 1995;33(10 Suppl.):OS5-16. doi:10.1097/00005650-199510001-00002.

18. Peters M, Godfrey C, McInerney P, Munn Z, Trico A, Khalil H. Scoping reviews. In: Aromatatis E, Munn Z, editors. JBI manual for evidence synthesis. Melbourne: JBI; 2020 (https://synthesismanual.jbi.global, accessed 4 April 2023).

19. Shea BJ, Reeves BC, Wells G, Thuku M, Hamel C, Moran J, et al. AMSTAR 2: a critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both. BMJ. 2017;358:j4008. doi:10.1136/bmj.j4008.

20. Leichsenring K, Lorenz-Dant K, Pot M, Rosenberg M, Simmons C. Innovative purchasing arrangements beyond pay-for-performance measures to strengthen service delivery models to deliver chronic care: protocol for a scoping review [website]. Charlottesville (VA): Center for Open Science; 2023. doi:10.17605/OSF.IO/G4DH9.

21. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. Int J Surg. 2021;88:105906. doi:10.1016/j.ijsu.2021.105906.

22. Cochrane Effective Practice and Organisation of Care. EPOC resources for review authors [website]. London: Cochrane.org; 2021 (https://epoc.cochrane.org/resources/epoc-resources-review-authors, accessed 22 October 2022).

23. Tsiachristas A. Financial incentives to stimulate integration of care. Int J Integr Care. 2016;16(4):8. doi:10.5334/ijic.2532.

24. Zhou W, Jian W, Wang Z, Pan J, Hu M, Yip W. Impact of global budget combined with pay-for-performance on the quality of care in county hospitals: a difference-in-differences study design with a propensity-score-matched control group using data from Guizhou province, China. BMC Health Serv Res. 2021;21(1):1296. doi:10.1186/s12913-021-07338-8.

25. Zhang H, Zhang L, Xu R, Pan J, Hu M, Jian W, et al. Can a global budget improve health care efficiency? Experimental evidence from China. Health Econ. 2022;31(8):1676–94. doi:10.1002/hec.4531.

26. Yin J, Wei X, Li H, Jiang Y, Mao C. Assessing the impact of general practitioner team service on perceived quality of care among patients with non-communicable diseases in China: a natural experimental study. Int J Qual Health Care. 2016;28(5):554–60. doi:10.1093/intqhc/mzw075.

27. Yoon J, Oakley LP, Luck J, Harvey SM. Can accountable care divert the sources of hospitalization? Am J Manag Care. 2019;25(10):e296–303. PMID:31622069.

28. Mandal AK, Tagomori GK, Felix RV, Howell SC. Value-based contracting innovated Medicare advantage healthcare delivery and improved survival. Am J Manag Care. 2017;23(2):e41–9. PMID:28245661.

29. Powell-Jackson T, Yip WC-M, Han W. Realigning demand and supply side incentives to improve primary health care seeking in rural China. Health Econ. 2015;24(6):755–72. doi:10.1002/hec.3060.

30. Sun X, Liu X, Sun Q, Yip W, Wagstaff A, Meng Q. The impact of a pay-for-performance scheme on prescription quality in rural China. Health Econ. 2016;25(6):706–22. doi:10.1002/hec.3330.

31. Munnich EL, Richards MR. Treatment flows after outsourcing public insurance provision: evidence from Florida Medicaid. Health Econ. 2020;29(11):1343–63. doi:10.1002/hec.4135.

32. Salzberg AC, Bitton A, Lipsitz SR, Franz C, Shaykevich S, Newmark LP, et al. The impact of alternative payment in chronically ill and older patients in the patient-centered medical home. Med Care. 2017;55(5):483–92. doi:10.1097/mlr.0000000000000694.

33. Sawicki OA, Mueller A, Klaaßen-Mielke R, Glushan A, Gerlach FM, Beyer M, et al. Strong and sustainable primary healthcare is associated with a lower risk of hospitalization in high risk patients. Sci Rep. 2021;11(1):4349. doi:10.1038/s41598-021-83962-y.

34. Wensing M, Szecsenyi J, Kaufmann-Kolle P, Laux G. Strong primary care and patients' survival. Sci Rep. 2019;9(1):10859. doi:10.1038/s41598-019-47344-9.

35. Wright BJ, Royal N, Broffman L, Li H-F, Dulacki K. Oregon's coordinated care organization experiment: are members' experiences of care actually changing? J Healthc Qual. 2019;41(4):e38–46. doi:10.1097/JHQ.0000000000000178.

36. Song Z, Rose S, Safran DG, Landon BE, Day MP, Chernew ME. Changes in health care spending and quality 4 years into global payment. N Engl J Med. 2014;371(18):1704–14. doi:10.1056/NEJMsa1404026.

37. Song Z, Ji Y, Safran DG, Chernew ME. Health care spending, utilization, and quality 8 years into global payment. N Engl J Med. 2019;381(3):252–63. doi:10.1056/NEJMsa1813621.

38. Chien AT, Song Z, Chernew ME, Landon BE, McNeil BJ, Safran DG, et al. Two-year impact of the alternative quality contract on pediatric health care quality and spending. Pediatrics. 2014;133(1):96–104. doi:10.1542/peds.2012-3440.

39. Barry CL, Stuart EA, Donohue JM, Greenfield SF, Kouri E, Duckworth K, et al. The early impact of the 'Alternative Quality Contract' on mental health service use and spending in Massachusetts. Health Aff. 2015;34(12):2077–85. doi:10.1377/hlthaff.2015.0685.

40. Song Z, Rose S, Chernew ME, Safran DG. Lower- versus higher-income populations in the alternative quality contract: improved quality and similar spending. Health Aff. 2017;36(1):74–82. doi:10.1377/hlthaff.2016.0682.

41. McWilliams JM, Landon BE, Chernew ME. Changes in health care spending and quality for Medicare beneficiaries associated with a commercial ACO contract. JAMA. 2013;310(8):829–36. doi:10.1001/jama.2013.276302.

42. Done N, Herring B, Xu T. The effects of global budget payments on hospital utilization in rural Maryland. Health Serv Res. 2019;54(3):526–36. doi:10.1111/1475-6773.13162.

43. Roberts ET, McWilliams JM, Hatfield LA, Gerovich S, Chernew ME, Gilstrap LG, et al. Changes in health care use associated with the introduction of hospital global budgets in Maryland. JAMA Intern Med. 2018;178(2):260. doi:10.1001/jamainternmed.2017.7455.

44. Viganego F, Um EK, Ruffin J, Fradley MG, Prida X, Friebel R. Impact of global budget payments on cardiovascular care in Maryland: an interrupted time series analysis. Circ Cardiovasc Qual Outcomes. 2021;14(3):e007110. doi:10.1161/CIRCOUTCOMES.120.007110.

45. Masters SH, Rutledge RI, Morrison M, Beil HA, Haber SG. Effects of global budget payments on vulnerable Medicare subpopulations in Maryland. Med Care Res Rev. 2022;79(4):535–48. doi:10.1177/10775587211052748.

46. Joynt Maddox KE, Orav EJ, Zheng J, Epstein AM. Medicare's Bundled Payments for Care Improvement Advanced Model: impact on high-risk beneficiaries. Health Aff. 2022;41(11):1661–9. doi:10.1377/hlthaff.2022.00138.

47. Hoedemakers M, Karimi M, Leijten F, Goossens L, Islam K, Tsiachristas A, et al. Value-based person-centred integrated care for frail elderly living at home: a quasi-experimental evaluation using multicriteria decision analysis. BMJ Open. 2022;12(4):e054672. doi:10.1136/bmjopen-2021-054672.

48. Pope G, Kautter J, Leung M, Trisolini M, Adamache W, Smith K. Financial and quality impacts of the Medicare physician group practice demonstration. Medicare Medicaid Res Rev. 2014;4(3):mmrr2014-004-03-a01. doi:10.5600/mmrr.004.03.a01.

49. Pimperl A, Schulte T, Mühlbacher A, Rosenmöller M, Busse R, Groene O, et al. Evaluating the impact of an accountable care organization on population health: the quasi-experimental design of the German Gesundes Kinzigtal. Popul Health Manag. 2017;20(3):239–48. doi:10.1089/pop.2016.0036.

50. Friedberg MW, Rosenthal MB, Werner RM, Volpp KG, Schneider EC. Effects of a medical home and shared savings intervention on quality and utilization of care. JAMA Intern Med. 2015;175(8):1362–8. doi:10.1001/jamainternmed.2015.2047.

51. Kicinger IM, Cuellar A, Helmchen LA, Gimm G, Want J, Kells BJ, et al. Quality of care and preventive screening use in the CareFirst patient-centered medical home program. J Healthc Qual. 2019;41(6):339–49. doi:10.1097/JHQ.0000000000000169.

52. Navathe AS, Emanuel EJ, Bond A, Linn K, Caldarella K, Troxel A, et al. Association between the implementation of a population-based primary care payment system and achievement on quality measures in Hawaii. JAMA. 2019;322(1):57–68. doi:10.1001/jama.2019.8113.

53. Nyweide DJ, Lee W, Cuerdon TT, Pham HH, Cox M, Rajkumar R, et al. Association of pioneer accountable care organizations vs traditional Medicare fee for service with spending, utilization, and patient experience. JAMA. 2015;313(21):2152. doi:10.1001/jama.2015.4930.

54. Hayen A, van den Berg MJ, Struijs JN, Westert Gert GP. Dutch shared savings program targeted at primary care: reduced expenditures in its first year. Health Policy. 2021;125(4):489–94. doi:10.1016/j.healthpol.2021.01.013.

55. Zhang H, Cowling DW, Graham JM, Taylor E. Five-year impact of a commercial accountable care organization on health care spending, utilization, and quality of care. Med Care. 2019;57(11):845–54. doi:10.1097/MLR.0000000000001179.

56. Jones C, Finison K, McGraves-Lloyd K, Tremblay T, Mohlman MK, Tanzman B, et al. Vermont's community-oriented all-payer medical home model reduces expenditures and utilization while delivering high-quality care. Popul Health Manag. 2016;19(3):196–205. doi:10.1089/pop.2015.0055.

57. Kern LM, Edwards A, Kaushal R. The patient-centered medical home and associations with health care quality and utilization: a 5-year cohort study. Ann Intern Med. 2016;164(6):395–405. doi:10.7326/M14-2633.

58. Swietek KE, Domino ME, Beadles C, Ellis AR, Farley JF, Grove LR, et al. Do medical homes improve quality of care for persons with multiple chronic conditions? Health Serv Res. 2018;53(6):4667–81. doi:10.1111/1475-6773.13024.

59. Spees LP, Wheeler SB, Zhou X, Amin KB, Baggett CD, Lund JL, et al. Changes in chronic medication adherence, costs, and health care use after a cancer diagnosis among low-income patients and the role of patient-centered medical homes. Cancer. 2020;126(21):4770–9. doi:10.1002/cncr.33147.

60. Flieger SP. Impact of a patient-centered medical home pilot on utilization, quality, and costs and variation in medical homeness. J Ambul Care Manage. 2017;40(3):228–37. doi:10.1097/JAC.0000000000000162.

61. Cole MB, Galárraga O, Wilson IB. The impact of Rhode Island's multipayer patient-centered medical home program on utilization and cost of care. Med Care. 2019;57(10):801–8. doi:10.1097/mlr.0000000000001194.

62. Werner RM, Duggan M, Duey K, Zhu J, Stuart EA. The patient-centered medical home: an evaluation of a single private payer demonstration in New Jersey. Med Care. 2013;51(6):487–93. doi:10.1097/MLR.0b013e31828d4d29.

63. Shane DM, Nguyen-Hoang P, Bentler SE, Damiano PC, Momany ET. Medicaid Health Home reducing costs and reliance on emergency department: evidence from Iowa. Med Care. 2016;54(8):752–7. doi:10.1097/MLR.0000000000000555.

64. Fifield J, Forrest DD, Burleson JA, Martin-Peele M, Gillespie W. Quality and efficiency in small practices transitioning to patient centered medical homes: a randomized trial. J Gen Intern Med. 2013;28(6):778–86. doi:10.1007/s11606-013-2386-4.

65. Shi Q, Yan TJ, Lee P, Murphree P, Yuan X, Shao H, et al. Evaluation of the Quality Blue Primary Care program on health outcomes. Am J Manag Care. 2017;23(12):e402–8. PMID:29261245.

66. Rosenthal MB, Alidina S, Friedberg MW, Singer SJ, Eastman D, Li Z, et al. Impact of the Cincinnati Aligning Forces for Quality Multi-Payer Patient Centered Medical Home pilot on health care quality, utilization, and costs. Med Care Res Rev. 2016;73(5):532–45. doi:10.1177/1077558715618566.

67. Markovitz AA, Murray RC, Ryan AM. Comprehensive Primary Care Plus did not improve quality or lower spending for the privately insured. Health Aff. 2022;41(9):1255–62. doi:10.1377/hlthaff.2021.01982.

68. Vu T, Anderson KK, Devlin RA, Somé NH, Sarma S. Physician remuneration schemes, psychiatric hospitalizations and follow-up care: evidence from blended fee-for-service and capitation models. Soc Sci Med. 2021;268:113465. doi:10.1016/j.socscimed.2020.113465.

69. Kiran T, Kopp A, Moineddin R, Glazier RH. Longitudinal evaluation of physician payment reform and team-based care for chronic disease management and prevention. CMAJ. 2015;187(17):E494–502. doi:10.1503/cmaj.150579.

70. Vu T, Anderson KK, Somé NH, Thind A, Sarma S. Mental health services provision in primary care and emergency department settings: analysis of blended fee-for-service and blended capitation models in Ontario, Canada. Adm Policy Ment Health. 2021;48(4):654–67. doi:10.1007/s10488-020-01099-y.

71. Somé NH, Devlin RA, Mehta N, Zaric GS, Sarma S. Stirring the pot: switching from blended fee-for-service to blended capitation models of physician remuneration. Health Econ. 2020;29(11):1435–55. doi:10.1002/hec.4145.

72. McConnell KJ, Renfro S, Chan BKS, Meath THA, Mendelson A, Cohen D, et al. Early performance in Medicaid accountable care organizations: a comparison of Oregon and Colorado. JAMA Intern Med. 2017;177(4):538–45. doi:10.1001/jamainternmed.2016.9098.

73. Hall AG, Landry AY, Lemak CH, Boyle EL, Duncan RP. Reported experiences with Medicaid managed care models among parents of children. Matern Child Health J. 2014;18(3):544–53. doi:10.1007/s10995-013-1270-5.

74. Hayes H, Stokes J, Kristensen SR, Sutton M. The effect of payment method and multimorbidity on health and healthcare utilisation. J Health Organ Manag. 2021;35(3):382–405. doi:10.1108/jhom-05-2020-0208.

Annexes

Annex 1. Overview of methodology

This section outlines in greater detail the methodology underpinning the scoping review on which this report is based.

A1.1 Research protocol

We started by developing a research protocol, setting out in detail the methods to be used in the scoping review, including the research questions, search and screening strategies and the analytical framework. The protocol for this scoping review has been developed based on guidance from the Joanna Briggs Institute (JBI) to ensure adherence to quality standards, and it follows the AMSTAR (A MeaSurement Tool to Assess systematic Reviews) checklist, as far as criteria for scoping and systematic reviews align. Draft versions of the research protocol were shared with the World Health Organization (WHO) Centre for Health Development, Kobe, Japan, and amended according to the feedback received. The final research protocol was published and is accessible via the Open Science Foundation Registries *(1)*.

A1.2 Search strategy

We adopted a three-stage strategy to identify relevant literature. This included, first, a systematic search of the following five academic repositories: PubMed, Web of Science, Cochrane Database of Systematic Reviews, Scopus and Google Scholar. Second, we conducted a targeted search of the grey literature using Google incognito searches as well as searches of dedicated repositories of relevant organizations (i.e. the Organisation for Economic Co-operation and Development, the World Bank, WHO). Google incognito mode was used to reduce the risk that internet searches would be influenced by the researchers' previous searches or by their geographical location. A combination of these repositories has also been used in other relevant reviews. Third, we hand-searched the reference lists of all relevant systematic reviews that were identified through our searches for relevant single-study papers.

We identified key search terms for purchasing arrangements, noncommunicable diseases and outcomes from an initial search and from published reviews on the topic; we then refined these through trial searches. The final search terms were (in both British and American spelling):

- for noncommunicable diseases – noncommunicable, NCD, chronic care, chronic disease, multimorbidity.
- for purchasing arrangements – accountable care organization, alternative payment model, bundled payment, capitation,

comprehensive primary care, coordinated care model, disease-based model, global budget, global payment, episode-based payment, health care financing, health maintenance organization, integrated delivery system, integrated financing, managed care organization, network-level payment, outcome-based payment, patient-aligned care team, patient centred medical home, pay-for-coordination, per member per month, population-based payment, purchasing arrangement, quality-based purchasing, risk-based payment, shared savings, shared gain, value-based payment, value-based purchasing;

All searches were conducted in November 2022.

A1.3 Screening process

After the removal of duplicate records, the searches of the academic and grey literature resulted in 3 539 resources. We uploaded all resources into Mendeley reference management software (Elsevier, New York, USA). To screen the identified reports, we followed a two-step approach, consisting of screening the title and abstract, followed by subsequent full-text screening. Before starting the title and abstract screening, as a trial, two researchers screened 20 titles using the criteria described in Section A1.4) finding that they agreed on more than 75% of the resources trialled. After comparing and discussing their decisions, they further clarified and specified the inclusion and exclusion criteria. Following the trial, all titles and abstracts were individually screened by two researchers for alignment with the inclusion and exclusion criteria. Decisions about inclusion and exclusion (including reasons for exclusion) were recorded in an Excel spreadsheet and subsequently compared. Disagreements regarding inclusion and exclusion were resolved by consensus. The title and abstract screening resulted in the exclusion of 3 254 resources.

A1.4 Inclusion and exclusion criteria

We included literature that examines purchasing arrangements and their impact on the quality of care for people with chronic diseases, service delivery and other structural mechanisms. We focused on studies that presented alternative financing arrangements beyond individual fee-for-service (FFS) arrangements and exclusive pay-for-performance (P4P) measures. However, in recognition that performance-based financial incentives are often implemented as one component of a broader structure, we also included literature on any such financing arrangements in which P4P mechanisms were part of the wider underlying programme. In other words, we focused more on alternative financing arrangements that consist of integrated or collaborative financing and delivery approaches across providers or sectors rather than on incentive schemes targeted

towards individual providers. We excluded literature that addressed alternative purchasing arrangements but did not provide any specific details.

We limited the time frame to literature published from 2013 to November 2022, in line with the launch of WHO's Global Action Plan for the Prevention and Control of NCDs 2013–2020 *(2)*. We included peer-reviewed scientific publications, grey literature (e.g. government publications, working papers, research reports) and systematic reviews for hand-searching. We included studies that used causal inference methods in their analysis to reach their conclusions (e.g. randomized controlled trials [RCTs], regression-based analysis, control groups), thus excluding studies that were observational or descriptive in nature (e.g. case studies). We limited studies to those that used causal inference methods to be able to attribute the effects of the purchasing arrangements on strengthening of the service delivery models and the delivery of quality services for chronic care. This scope includes comparisons of new purchasing arrangements with those in place prior to the introduction of the innovation, and also includes comparisons with relevant other groups, for example, in cases in which alternative purchasing arrangements were confined to specific regions or patient groups. Taking a global perspective to ensure the inclusion of resources from high-, middle- and low-income countries, we did not limit inclusion to literature from specific geographical locations or particular settings (e.g. hospital care, care in the community, residential care, primary care, specialist care). However, we included only literature published in English.

At later stages in the screening process, other relevant reasons for exclusion included insufficient information about the purchasing arrangement, a specific focus on only one particular disease (e.g. diabetes) or intervention (e.g. breast cancer screening, joint replacement), a focus on only one specific measure of quality, insufficient information about the methods used, insufficient consideration of the outcome's impact on the quality of care and limited focus on chronic conditions.

A1.5 Data extraction and quality appraisal

Two researchers entered data for the included records onto a prepared extraction sheet developed in line with the Cochrane Effective Practice and Organisation of Care (known as EPOC) resources group that contained a priori defined categories deemed relevant for the analysis of the data and the research questions posed for this review. This included information about the type of purchasing arrangement (i.e. the design, provider participation, reward structures, payment structures), service delivery models, impact on outcomes (e.g. process outcomes, health outcomes, and

measures of continuity of care, integration of services and coordination) and other institutional design and supportive elements (e.g. governance, information systems, quality initiatives and regulations). As with the screening process, data extraction was trialled with two researchers entering information for two selected research papers to ensure consistency. After clarifying definitions and slightly adapting the extraction sheet, we divided the remaining studies included in this scoping review and independently entered the relevant data.

As there are no quality appraisal techniques recommended specifically for scoping reviews, quality appraisal tools developed for systematic reviews were largely unsuitable. This is also where we deviated from the AMSTAR checklist, as the studies considered for this scoping review required a broader quality appraisal. Therefore, we applied the critical appraisal checklists developed by the JBI for different types of research to assess the risk of bias in articles (https://jbi.global/critical-appraisal-tools); we mainly used the checklist for quasi-experimental studies, but also the checklist for randomized controlled trials in a few instances. As scoping reviews do not usually include an evaluation of effect or impact, there are – to the best of our knowledge – no tools that specifically facilitate an assessment of the quality and certainty of evidence for this type of review.

Given that we were carrying out a scoping review of many different types of interventions (i.e. purchasing arrangements) and their impact on several different outcomes across different target populations, we amended the Grading of Recommendations Assessment, Development and Evaluation (GRADE) approach for our purposes to allow us to maintain transparency in synthesizing the evidence. The GRADE framework, developed by Guyatt and colleagues *(3)*, is a systematic approach to assessing the certainty of evidence when carrying out a systematic review and other evidence syntheses, and it helps to guide the drawing of conclusions in a transparent way. The approach rates the certainty of evidence according to four levels based on a combination of factors, such as the type of study, inconsistency, indirectness, imprecision and publication bias.

Building on the GRADE approach, we assessed the certainty of evidence for the outcomes reported in each article according to the type of study (e.g. RCT versus controlled study), the risk of bias (based on the JBI critical appraisal checklists) and the effect size. More specifically, we used the following criteria to assign a level of certainty to the evidence for the quantitative outcomes reported:

- high certainty –
  - an RCT with no bias or other issues AND a large effect;
- moderate certainty –
  - an RCT with some bias or issues OR a small effect
  - a controlled study with no bias or issues AND a large effect;
- low certainty –
  - an RCT with some bias or issues AND a small effect
  - a controlled study with some bias or issues OR a small effect;
- very low certainty –
  - a controlled study with some bias or issues AND a small effect.

This level of certainty of evidence assigned across outcomes was later a guiding element in drawing conclusions across articles about the impact of purchasing arrangements on the quality of care.

References

1. Leichsenring K, Lorenz-Dant K, Pot M, Rosenberg M, Simmons C. Innovative purchasing arrangements beyond pay-for-performance measures to strengthen service delivery models to deliver chronic care: protocol for a scoping review [website]. Charlottesville (VA): Center for Open Science; 2023. doi:10.17605/OSF.IO/G4DH9.
2. Global action plan for the prevention and control of noncommunicable diseases 2013–2020. Geneva: World Health Organization; 2013 (https://www.who.int/publications/i/item/9789241506236, accessed 4 April 2023).
3. Guyatt GH, Oxman AD, Vist GE, Kunz R, Falck-Ytter Y, Alonso-Coello P, et al. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. BMJ. 2008;336:924-6. doi:10.1136/bmj.39489.470347.AD.

Annex 2. Overview of multipayer structures in health care

A2.1 Germany

In Germany, there is nearly universal population coverage of health care, with more than 99% of individuals enrolled in a health insurance plan; public statutory health insurance consists of 105 sickness funds and covers 87% of the population, and private health insurance consists of 42 private health insurance companies and covers 11% of the population *(1)*. The latter is available to those with income above a certain threshold who would like to opt out of the statutory plans, as well to those belonging to certain professional groups, such as self-employed people and civil servants. The remaining 2% are covered under special programmes (e.g. soldiers). Statutory health insurance covers a wide range of services and benefits and is managed by a central reallocation pool, which allocates pooled funds to the different statutory health insurance, or sickness, funds. Sickness funds then assign a budget to providers for services for people who are covered by the statutory health insurance in that region. Patients with private health insurance typically have to pay for services directly and are later reimbursed by their insurance provider.

A2.2 Netherlands (Kingdom of the)

In Netherlands (Kingdom of the), the health care system is primarily financed by a combination of compulsory contributions collected by the government and premiums paid to health insurance plans *(2)*. While all individuals are covered for a basic package of benefits under the Health Insurance Act (known as Zvw), individuals aged 18 and older are required to purchase a health insurance plan from an insurer to whom they pay a premium. Together, these contributions cover a basic package of benefits, namely for curative services (e.g. care by a general practitioner, maternity care, inpatient care). Health insurers contract with health care providers for services included in their plan, which are covered only after individuals have reached their deductible.

A2.3 United States

In the United States, a majority of health care is provided through voluntary health insurance managed by a private insurer, primarily through employer-sponsored plans. The next largest share of health care is publicly provided, through either Medicare or Medicaid *(3)*. A sizeable portion of the population, estimated at around 10%, is uninsured. The types of services covered by health plans vary depending on the insurance. In practice, health care facilities (e.g.

hospitals, general practitioners' offices, clinics) often provide services to patients with different types of insurance coverage; however, providers can decide which types of insurance to accept. Some types, such as Medicare and Medicaid, have a network of providers within which patients are covered, while care provided outside of the network is not covered. The multipayer structure of the American health care system is important to consider, given that a body of literature has found that Medicaid beneficiaries face barriers related to accessing services compared with privately insured individuals *(4)*.

References

1. Blümel M, Spranger A, Achstetter K, Maresso A, Busse R. Germany: health system review. Copenhagen: WHO Regional Office for Europe, European Observatory on Health Systems and Policies; 2020 (https://iris.who.int/handle/10665/341674, accessed 18 October 2023).

2. Kroneman M, Boerma W, van den Berg M, Groenewegen P, de Jong J, van Ginneken E. Netherlands: health system review. Copenhagen: WHO Regional Office for Europe, European Observatory on Health Systems and Policies; 2016 (https://iris.who.int/handle/10665/330244, accessed 18 October 2023).

3. Rice T, Rosenau P, Unruh LY, Barnes AJ. United States of America: health system review. Copenhagen: WHO Regional Office for Europe, European Observatory on Health Systems and Policies; 2020 (https://iris.who.int/bitstream/handle/10665/338880/HiT-22-4-2020-eng.pdf?sequence=1, accessed 31 July 2023).

4. Hsiang WR, Lukasiewicz A, Gentry M, Kim C-Y, Leslie MP, Pelker R, et al. Medicaid patients have greater difficulty scheduling health care appointments compared with private insurance patients: a meta-analysis. Inquiry. 2019;56:46958019838118. doi:10.1177/0046958019838118.

World Health Organization
Centre for Health Development
(WHO Kobe Centre – WKC)
Kobe, Japan

www.who.int/kobe_centre/en/