CURRENT OPINION



Is Economic Evaluation and Care Commissioning Focused on Achieving the Same Outcomes? Resource-Allocation Considerations and Challenges Using England as a Case Study

Matthew Franklin¹ • Sebastian Hinde² • Rachael Maree Hunter³ • Gerry Richardson² • William Whittaker⁴

Accepted: 8 February 2024 © The Author(s) 2024

Abstract

Commissioning describes the process of contracting appropriate care services to address pre-identified needs through preagreed payment structures. Outcomes-based commissioning (i.e., paying services for pre-agreed outcomes) shares a common goal with economic evaluation: achieving value for money for relevant outcomes (e.g., health) achieved from a finite budget. We describe considerations and challenges as to the practical role of relevant outcomes for evaluation and commissioning, seeking to bridge a gap between economic evaluation evidence and care commissioning. We describe conceptual (e.g., what are 'relevant' outcomes) alongside practical considerations (e.g., quantifying and using relevant endpoint or surrogate outcomes) and pertinent issues when linking outcomes to commissioning-based payment mechanisms, using England as a case study. Economic evaluation often focuses on a single endpoint health-focused maximand, e.g., quality-adjusted lifeyears (QALYs), whereas commissioning often focuses on activity-based surrogate outcomes (e.g., health monitoring), as easier-to-measure key performance indicators that are more acceptable (e.g., by clinicians) and amenable to being linked with payment structures. However, payments linked to endpoint and/or surrogate outcomes can lead to market inefficiencies; for example, when surrogates do not have the intended causal effect on endpoint outcomes or when service activity focuses on only people who can achieve prespecified payment-linked outcomes. Accounting for and explaining direct links from commissioners' payment structures to surrogate and then endpoint economic outcomes is a vital step to bridging a gap between economic evaluation approaches and commissioning. Decision-analytic models could aid this but they must be designed to account for relevant surrogate and endpoint outcomes, the payments assigned to such outcomes, and their interaction with the system commissioners purport to influence.

- Matthew Franklin matt.franklin@sheffield.ac.uk
- Sheffield Centre for Health and Related Research (SCHARR), Division of Population Health, School of Medicine and Population Health, The University of Sheffield, Regent Court, 30 Regent Street, Sheffield S1 4DA, UK
- ² Centre for Health Economics (CHE), University of York, Heslington, York YO10 5DD, UK
- Research Department of Primary Care and Population Health, Royal Free Medical School, University College London, Royal Free Campus, Rowland Hill Street, London NW3 2PF, UK
- Division of Population Health, Health Services Research & Primary Care, Alliance Manchester Business School, Institute for Health Policy and Organisation, Oxford Road, Manchester M13 9PL, UK

1 Introduction

Commissioning describes the process of assessing the needs of people in an area then contracting appropriate care services to address these needs through pre-agreed payment structures. Commissioning of care services is commonly achieved at local/regional levels rather than national levels. In England, the Health and Care Act 2022 embedded joint working between the health and social care systems, with the centrepiece being integrated care systems (ICSs), i.e., geographical area-based agencies responsible for planning local services to improve health and reduce inequalities [4, 5]. Although this is one example of evolving government infrastructure, the role of local government and associated commissioning practices to address achieving relevant

Key Points for Decision Makers

Much has been written regarding how outcomes-based commissioning fits with a value-based healthcare framework, however little relates to economic evaluation, despite it being a key method for judging the value for money of care interventions.

The commissioning landscape involves outcome frameworks for monitoring and evaluating services but these are often not the basis by which services are prospectively commissioned. Instead, activity tends to be the main commissioning focus to which service payments are linked. This is seemingly at odds with how patientcentred value is quantified within economic evaluation that focuses on a health maximand typically quantified using quality-adjusted life-years.

Care commissioners tend to link payment structures to surrogate outcomes (often based on healthcare activity) as an expenditure-based policy instrument to influence the system within their jurisdiction, and then rely on causal mechanisms to achieve the relevant (health) outcome that economic evaluations can quantify and report on explicitly.

Those conducting economic evaluation need to develop better ways to account for and communicate links from commissioning payment structures to surrogates and endpoint outcomes in their analyses (e.g., via decisionanalytic models).

outcomes (e.g., health) and reducing associated inequalities has long been recognised as important internationally [6–8].

Outcomes-based commissioning is a set of arrangements whereby a service is defined and remunerated based on preagreed outcomes, associated with similar concepts such as outcomes-based 'reimbursement' or 'contracts' proposed for medicine implementation [9–13]. Outcomes-based commissioning is akin to 'payment by results' (PbR), which gained traction in National Health Service (NHS) England following the 2011 'Open Public Services: white paper' [14]. PbR is a system for paying NHS healthcare providers a standard national price/tariff for each patient seen or treated; however, this scheme pays for activity undertaken rather than outcomes achieved [14–17]. Thus, outcomes-based commissioning requires shifting from a framework by which services are purchased and resources allocated for units of activity (e.g., hours/days/weeks of service provision) for

predefined needs to what is needed to ensure service users' predefined outcomes are achieved [10–13].

Economic evaluation frameworks quantify the difference in both costs and outcomes between two or more alternative courses of action. All forms of economic evaluation are related to value for money, with costs representing an integral aspect of the evaluation process owing to the resultant opportunity costs from resources not being available for other purposes [18]. Cost-effectiveness analysis (CEA) as a form of economic evaluation is operationalised with the normative stance that a relevant outcome from healthcare should be health, often quantified using quality-adjusted life-years (QALYs) [Appendix S1 describes other frameworks]. Incremental cost-per-QALY comparisons is the basis by which health technology assessment (HTA) agencies internationally, such as the National Institute for Health and Care Excellence (NICE) for England and Wales, suggest whether a new health technology is cost effective relative to any alternative(s) [19, 20]. As such, outcomes-based commissioning shares a common goal with economic evaluation and HTA processes, i.e., achieving value for money for relevant outcomes (e.g., health) achieved from a finite budget. However, even though HTA guidance for conducting economic evaluation seeks to be transparent and accessible, focus on HTA guidance for conducting economic evaluation may not be appropriate in all circumstances [7, 8, 19, 21]. For example, as commissioning is often based on activity undertaken not outcomes achieved, a disconnect potentially emerges between the evidence produced to inform HTA processes and that to inform commissioning, leading to potentially inefficient use of finite finances when trying to achieve differential outcomes [7, 8, 22, 23].

The current article is part of a trilogy of articles published in Applied Health Economics and Health Policy. Each article has explored the potential disconnect between economic evaluation processes and evidence, which have often been driven by guidance developed for HTA processes at a national level, compared with the needs and objectives of local and national government agents [7, 8, 19]. Our previous article by Howdon et al. [8] focused particularly on costs, whereas the article by Hinde et al. [7] focused on (health) inequality. The current article focuses specifically on outcomes, as although much has been written in regard to how outcomes-based commissioning fits with a value-based healthcare framework, less has been written on how it relates to economic evaluation [24]. As such, our aim is to describe considerations and challenges as to the practical role of relevant outcomes for economic evaluation and commissioning, to bridge a gap between the evidence generated and required in the two settings, using England as a case study.

2 Outcomes and Activity-Based Commissioning: An Overview Using England as a Case Study

In the UK under the New Labour government from 1997 to 2010, local commissioners were tasked with implementing priorities set by central government and complying with national standards, enforced through the setting of targets and use of performance management frameworks [25]. Subsequently, key performance management systems and associated standards became increasingly focused on outcomes and outcomes-based commissioning [12, 25]. Different test beds of outcomes-based commissioning have occurred across public services; for example, programmes to support troubled families and to help people who are long-term unemployed back into work [12, 26, 27]. These concepts fed into health and social care, with some local areas introducing outcomes-based contracts across physical and mental health services, and adult social care [13, 16]; however, often such contracts are linked with activity (e.g., PbR) rather than outcomes. For example, the NHS's Quality Outcomes Framework (QOF; see Box 1) refers to 'outcomes', but its dominant focus is activity associated with good quality care which can also be considered on the causal pathway to better outcomes; that is, healthcare activity as a surrogate for achieving health outcomes (Sect. 4 describes the nature and use of surrogates) [1, 3].

Outcome frameworks have been developed for services that could form the basis of outcomes-based commissioning, but have not been operationalised as such; for example, the Adult Social Care Outcomes Framework (ASCOF) in England (see Box 2). In spite of the ASCOF's use both locally and nationally to set priorities for care and support, alongside measuring progress and strengthening transparency and accountability, it has not gained traction for outcomes-based commissioning [2]. Instead, what is generally observed in the commissioning landscape is an interest in outcomes and outcome frameworks for monitoring and evaluating services, but such outcomes are not the basis by which services are prospectively commissioned. There are however good reasons why commissioning remains focused on activity rather than directly based on endpoint outcomes, which includes defining and quantifying relevant outcomes (Sect. 3); use of activity-based surrogate outcomes compared with endpoint outcomes (Sect. 4); and issues when linking outcomes to commissioning payment structures (Sect. 5).

Box 1 National Health Service's Quality Outcomes Framework for general medical services: an overview

The NHS's QOF, introduced as part of the GMS contract in 2004, is a pay-for-performance scheme designed to remunerate GP practices for providing good quality patient care and to fund further work to improve the quality of care delivered. Although the QOF was part of a revised contract for GPs, participation in the QOF is voluntary and NHS Digital states that QOF "... is not about performance management but resourcing and rewarding good practice" [1]. The QOF is based on the use of indicators across five domains: Clinical, Public Health. Public Health—Additional Services, Public Health— Vaccination and Immunisation, and Quality Improvement. Indicators are agreed as part of yearly GP contract negotiations, with indicators having associated points that can be rewarded to GP practices based on the activityrelated indicators, for which there is published guidance, e.g., for 2021/2022 [3]. Furthermore, as the QOF is a voluntary 'reward and incentive' scheme, it is not the basis by which the service is commissioned compared with a voluntary extra; however, under the GPs' contract, the QOF is a key area where GPs can make a difference to their income.

NHS National Health Service, QOF Quality Outcomes Framework, GMS general medical services, GP General Practitioner

Box 2 The Adult Social Care Outcomes Framework in England: four key outcomes

The ASCOF in England, published in 2011, specified four key outcomes [2]:

- 1. enhancing the quality of life for people with care and support needs;
- 2. delaying and reducing the need for care and support;
- 3. ensuring that people have a positive experience of care and support;
- 4. safeguarding adults whose circumstances make them vulnerable and protecting them from avoidable harm.

In spite of the ASCOF's use both locally and nationally to set priorities for care and support, alongside measuring progress and strengthening transparency and accountability, the ASCOF has not gained traction for outcomes-based commissioning [2].

ASCOF Adult Social Care Outcomes Framework

3 Defining and Quantifying Relevant Outcomes from Healthcare

While reflecting on health expenditure and equity, Culyer [28] suggests that a key outcome of healthcare is health, thus health should be a relevant outcome of interest for policymakers. Undoubtedly there are other complementary perspectives to this normative stance, e.g., the capability approach [29], and health and social care systems also try to achieve other complementary objectives, such as improving clinical care, service management, patient focus, and external focus [30]. However, an issue is how to quantify 'health' and use that quantification to inform decision making [31].

Health, among other constructs, can be defined and quantified in different ways, which has in part been done using patient-reported outcome measures (PROMs) [32]. Health-focused PROMs can have a generic (e.g., EQ-5D) or condition-specific (e.g., Patient Health Questionnaire-9 for depression [PHQ-9]) health focus [31, 33-35]. PROMs have also been developed to capture other relevant outcomes and for specific uses within or outside of healthcare, such as the capability-based ICECAP or Adult Social Care Outcomes Toolkit (ASCOT) measures [36–39]. However, PROMs will always be a necessarily restricted perspective of the construct it purports to represent, e.g., health or other relevant constructs. Despite limitations associated with the use of PROMs, there has been an evolution in their development and use; for example, to capture key performance metrics for health services such as the NHS's national PROMs initiative (which used the EQ-5D) and mental health services such as NHS England's Talking Therapies for anxiety and depression services (which uses the PHQ-9, among other PROMs) [40–42]. As such, perhaps routinely collected PROMs could form the basis of outcomes frameworks to support outcomes-based commissioning. For example, Porter's value-based healthcare arguments stress a focus on an increase in patient-centred outcomes, with value defined as the health outcomes achieved for each dollar spent [43]. Value-based healthcare incentives have become an area where PROMs have been suggested to have a role, although not all value-based incentives focus on patient-reported outcomes or, when they do, capture outcomes that are important to all patients [44].

However, despite the growing and sustained use of PROMs to capture relevant endpoint outcomes, there is sparce evidence of outcome measures being directly used to inform localised commissioning decisions. A key restricting consideration around outcomes-based commissioning is if the approach will alleviate pressures on the finite budgets of localised commissioners or increase costs due to the additional resources required to quantify and monitor the relevant predefined outcomes [16]. As such, data collection

itself presents an opportunity cost and poor data quality could have negative impacts/consequences for outcomes-based commissioning. Therefore, due to the resources required to constantly and appropriately capture such endpoints, this has in part lead to the requirement of using surrogate outcomes [45].

4 Surrogate and Endpoint Outcomes: Conceptual and Practical Considerations

Surrogate outcomes are those that may correlate, predict, or causally impact the endpoint outcome, typically occur earlier than the endpoint outcome of interest (e.g., at intermediate points), and tend to be easier to observe/quantify than the endpoint outcome of interest [45–47]. An issue with surrogate outcomes is that they do not have a guaranteed relationship with the endpoint outcome and therefore there are risks associated with their use [48, 49]. However, when an appropriate relationship is evidence-based, using surrogate outcomes have many benefits, including for commissioning and economic evaluation, e.g., see Box 3 [50–52].

Adopting easier-to-capture surrogate measures from health and care administrative data (e.g., as key performance indicators) has practical benefits and may be better understood or accepted by the clinical teams who collect this information as part of 'service as usual', e.g., measures of service uptake, activity, and/or disease incidence/ prevalence. For example, capturing prescription activity for statins is relatively easy (Box 3) in comparison with capturing cholesterol levels (which requires a blood test) and change to cardiovascular events (which has challenges in terms of defining relevant events and potential longevity of the impact).

The use of surrogate outcomes can also form the basis of decision-analytic modelling approaches, which have been used to inform decision-making processes [53, 54]. Decision-analytic modelling involves mathematical analysis to define the potential consequences of a set of alternative options (e.g., treatments or policies), drawing on estimates of transition probabilities among other parameters (e.g., costs or utilities) from a range of potential sources [55]. Compared with statistical analysis of direct endpoints (e.g., PROMs), decision-analytic modelling building on surrogates to estimate endpoint outcomes can be populated based on existing evidence, as well as allowing an extrapolation to the longer-term, e.g., what prescribing statins now could mean to reduced cardiovascular events and death rates in the future [55].

Importantly, surrogates do not necessarily have intrinsic value, but rather their importance is due to their associated or causal impact on patient-centred outcomes and longer-term endpoints. As such, their use relies on commissioners

understanding and having suitable evidence as to the anticipated effect of their chosen surrogate on the actual outcomes of interest, and also retaining an overall understanding that such surrogates are distinct from the endpoint of interest.

Box 3 Example use of surrogate outcomes: prescribing statins and body mass index monitoring

Prescribing statins: A sufficient evidence-base suggests that the activity of prescribing statins reduces blood cholesterol, which subsequently reduces the risk of associated cardiovascular events in those with cardiovascular risk factors [2]. This does not mean that every individual prescribed a statin will avoid a cardiovascular event, as this depends on a range of behavioural (e.g., adhering to taking the prescribed statins) and biological (i.e., statins and their impact on low-density lipoprotein cholesterol in the blood) factors; however, on average, prescribing statins reduces the risk of cardiovascular events because people (often) adhere to their use, and biologically they work well [2].

BMI monitoring: The NHS England's Health Checks for all 40- to 74-year-olds included BMI monitoring as a means of reducing health risks, not because the health service is interested in a person's body fat *per se* but rather because of the extensive body of empirical literature that suggests BMI is used as a surrogate of body fat-related disease risk (e.g., heart disease and diabetes) such that monitoring BMI is a logical and evidence-based activity [5, 6].

BMI body mass index, NHS National Health Service

5 Linking Outcomes to Payment Structures: A Major Challenge

A key issue for commissioning is at what point in the care pathway should payments be linked, e.g., should payments be linked to endpoint outcomes achieved, as may be preferred for outcomes-based commissioning, or at an intermediate point in the pathway, such as healthcare activity (i.e., as a surrogate). Linking payments to non-activity-based outcomes, such as a reduction in body mass index (BMI; rather than just the activity of monitoring BMI) for example (see Box 3), can lead to 'gaming the system', where clinicians may be more likely to take on patients they consider more likely to achieve this outcome. There is evidence that this has occurred in other parts of the system, such as when PbR was introduced into drug treatment services in pilot sites in England [56]. Linking payments to health outcomes may also perpetuate inequalities in health outcomes, as poorer outcomes are often linked with patients from lower socioeconomic areas, and hence services with a higher concentration of these patients may struggle to achieve the same outcomes. However, there is the potential for risk-adjusted payment and performance schemes, e.g., risk-adjusted PbR [57, 58]. For example, using bundled base payments that allow for efficient resource allocation as an objective of traditional PbR schemes, but alongside bonus payments that can directly discourage low-value services and encourage activities that promote clinical quality, patient well-being, and satisfaction [57, 58]. Such bonus 'risk adjustment' payments can avoid potential negative consequences associated with traditional PbR schemes, but with additional data collection and monitoring costs [57, 58].

There are situations that even when evidence-based treatment is provided, the patient's final endpoint outcome is outside the control of the health system; thus, focusing on surrogate outcomes can be preferred to focusing on endpoint outcomes. For example, a clinician presented with two identical stroke patients to whom they provide the same evidence-based treatment may end up with a disparate set of outcomes. Although results can and will be averaged out over patients, clinicians may feel unfairly penalised for things that are beyond their control when undertaking evidence-based activities. There is evidence that clinicians are resistant to payment systems linked to endpoint (health) outcomes compared with payments for undertaking the recommended, evidence-based activity [24].

There are also clinical areas where patients are unlikely to achieve a change in a measurable outcome or one that can be linked to payment structures. Mental health services are one area where efforts have been made to link payments with outcomes, with most countries focusing on using the Health of the Nation Outcome Scales (HoNOS) to cluster patients and hence link patient complexity with costs. However, studies have found that there is limited evidence for HoNOS scores predicting changes in activity and associated costs within secondary care services [59]. Mental health represents a tricky patient group, as many patients may never see an improvement in outcomes, or current outcomes are not sensitive to the 'needs' of the patients, thus linking payments to such outcomes represents a complex consideration. Linking payments to relevant outcomes that are unlikely to be wholly quantifiable is a complex area (with mental health being one example), alongside working out a reasonable price/payment for that outcome; however, there is ongoing work to better facilitate dynamic pricing of pharmaceutical innovations when we cannot observe the outcome(s) in time that perhaps could inform dynamic payment models for commissioned services in the future [60].

Although payment systems do exist and are being used, there is mixed evidence about the effectiveness of such pay-for-performance schemes; e.g., long-running schemes such as the QOF have limited cost-effectiveness evidence supporting their use [61]. There are also lessons that can and need to be learnt before introducing payment structures linked to surrogate or endpoint outcomes, including careful consideration as to what happens if such payment schemes were ever removed [62]. Overall, if payments are to be linked to surrogate or endpoint outcomes, care needs to be taken when choosing which outcomes to measure and subsequently link to payments, as there are potentially unintended consequences. For example, commissioning based on surrogate or endpoint outcomes may unintentionally exacerbate inequalities [63]. Given these complexities, both commissioners and economic evaluation must attempt to account for the influence of payments on both surrogate and endpoint outcomes to enable an efficient (and potentially equitable) use of finite resources.

6 Better Alignment of Economic Evaluation and Care Commissioning: Accounting for Outcome Payments

As commissioners deal with complex systems that often necessitate short-termism, a focus on payments at intermediate points along the system as surrogates to influence the endpoint outcome is perhaps inevitable [8]. As such, if economic evaluation is to be useful for commissioners, it must account for surrogate and endpoint outcomes, alongside any associated resource use and costs, including payments for achieving such outcomes; decision-analytic models can aid with this aspect. Decision-analytic models can directly represent the surrogates (e.g., activity) commissioners are focused on and how this relates to the economic evaluation, and also allows commissioners to see a quantification of how such surrogates are leading to potentially (in)efficient and/or (in)equitable changes in endpoint outcomes. By extension, this also facilitates an assessment of the potential suitability of that surrogate for achieving endpoint outcomes of interest.

However, decision-analytic tools for HTA often focus on cohort-based disease and/or care pathway models (e.g., whole-disease models) that are not necessarily sufficient to represent the complex systems within which commissioning decisions occur [64, 65]. In comparison, more complex models such as discrete event simulations (DES) and agent-based models can better represent individuals and the system within which they reside, but require additional knowledge and data to develop [66, 67]. Despite their complexity, DES and agent-based models are growing in use [66–69].

Additionally, decision-analytic models and economic evaluations in general need to better account for commissioning payment structures. Such payment structures are expenditure-based policy instruments that are within commissioners' control to influence the system within their jurisdiction [70]. As such, not accounting for the cost and

nature of such payments is in essence missing a key aspect of keen interest to such local decision makers [8]. As such, we propose four key considerations for accounting for these payment structures within an economic evaluation:

- 1. the monetary value of the total payment (e.g., the additional activity-based payment, not just the cost of the activity);
- the influence that payment has on the surrogate and/ or endpoint outcome and if the outcome is done as 'needed' rather than just to gain the payment, e.g., by gaming the system;
- 3. the elasticity of the payment's influence on the outcome (e.g., how an increase/decrease in payment has an influence on the activity);
- 4. the subsequent influence on the transition probabilities of a patient moving onto the next step (e.g., next surrogate outcome) up to the endpoint outcome.

Accounting for these aspects will not only better quantify how commissioning payment structures are influential in the economic evaluation but also to what extent payments linked to such surrogates are appropriate when also accounting for endpoint outcomes (e.g., population health). Although such considerations are accounted for in some public health policy models, such as those focused on the minimum unit price for alcohol (albeit focusing on the publics' actions to such cost changes, not commissioning payment structures), the relevance of such aspects have not been considered as relevant for all commissioner-focused modelling activities [71, 72]. For example, limited work has been done in calculating willingness-to-accept when estimating a suitable monetary value for the potential reimbursement payment, with one German study finding that a tenfold increase in payment almost doubled potential participation in a pay-forperformance scheme (from 28 to 50%); whether this additional payment is an efficient use of finite resources though would be a necessary next-step before implementing such a higher payment [73].

Even if an appropriate model is designed, and the appropriate payments and their influence accounted for, the adoption of economic evaluation into local decision making still requires accessibility and acceptability of what is produced [74]. There is always a trade-off between simplicity/understandability and complexity/accuracy, with the middle ground perhaps being sufficiency: what is sufficient to guide decision making? Arguably, a simpler model that can be used by commissioners is better than a more complex model that commissioners cannot or do not use; however, if decision-analytic models become common to inform commissioners, this will aid with broader understanding of the use of models and facilitate the use of more complex models over time. Although decision-analytic models are often used

to guide local and national decision makers, the extent to which they are designed appropriately and understood with certainty is still questionable. The negative consequence though is that if local decisions are focused on surrogates with little consideration of their influence on the endpoint outcome, such decisions are potentially leading to inefficiencies and inequalities. Even a simple model can aid commissioners understand links from surrogates/activity to endpoint outcomes, but striving for more complexity can potentially improve precision and accuracy with public benefits.

7 Discussion

Within this article, we have described the idea and basis of what are or could be considered relevant endpoint or surrogate outcomes for economic evaluation that are consistent with care commissioning. Although commissioners and those who conduct economic evaluations would agree that a relevant outcome from the healthcare system is an endpoint outcome associated with health, which is often the key headline result presented as part of the HTA-focused economic evaluation evidence-base (e.g., the cost per QALY was £ \times); a key difference is that whereas economic evaluations explicitly quantify the relevant outcome of health, often in the form of QALYs, healthcare commissioners are faced with a complicated task, which often means that activity-based surrogates are the actual key focus for commissioners compared with the endpoint outcome of health. It would be unreasonable to suggest that it is possible or even beneficial to fully align economic evaluation with commissioning processes, such as by having both processes focus on a single QALY outcome as a means of outcomes-based commissioning, for example. However, if the two do not align then choices made concerning what interventions/services should be funded based on HTA-focused economic evaluation evidence may conflict with what commissioners may choose based on their objectives, which may not necessarily align; thus, this has the potential to result in inefficient commissioning of services relative to the outcomes achieved. Therefore, we need to consider more carefully what needs to change to better align care commissioning with economic evaluation, and if there are methods/approaches that already exist that can bridge this gap. Debatably, such methods do exist (e.g., appropriate decision-analytic modelling techniques) and economic evaluation methods are particularly apt at linking costs (e.g., commissioning payment structures) with outcomes (e.g., surrogates up to endpoint outcomes); therefore, there is the potential to bridge this gap between current economic evaluation evidence that has often focused on HTA requirements, and the needs of commissioners whose focus on surrogate outcomes need to be better accounted for to understand if an efficient and equitable endpoint outcome (e.g., population health) can be better achieved in a way that is value for money.

7.1 International and Generalisable Considerations Beyond our English Case Studies

The notion of 'local-level economic evaluation' is a consideration internationally, such as suggested within a recent publication that sought to describe the nature, value, and sustainability of local-level economic evaluation in the context of Australia [75]. Internationally, local governing agents (e.g., commissioners) have their own names, roles, structures, and nuances between countries and even within countries. The institutional context of healthcare at the local level is more diverse and complex than the concept of the national healthcare state [70]. However, all health systems need to identify ways that payments into the healthcare system can lead to efficient (and, where desirable, equitable) outcomes by aligning such payments with endpoint outcomes directly or through the use of surrogates.

Without fully accounting for healthcare system and payment structure nuances internationally, a key consideration for those conducting economic evaluations to inform resource allocation within those systems is that payments into the care system should be seen as an expenditure-based policy instrument that should be accounted for in the economic evaluation [76]. If an economic evaluation is not accounting for such payments and their influence, then the economic evaluation is potentially missing a key part of the system it purports to represent. These payments are key aspects care commissioners are particularly interested in understanding, given they directly relate to their budgetary spending and agent-based objectives such as achieving better population health (i.e., an endpoint outcome) through influencing activity within the system (i.e., as a surrogate outcome) [7, 8].

7.2 Other Considerations

We have purposely chosen to ignore other relevant considerations within this article, such as the broader role of costs and (health) inequality, as we have addressed these relevant considerations elsewhere [7, 8], although it is worth (re)stating that the economic evaluation frameworks described in this article and Appendix S1 focus on efficiency rather than equality or equity considerations. While framework extensions are possible to incorporate equity considerations, such as distributional CEA (DCEA), the relevance of these approaches to the commissioning landscape have been discussed elsewhere [7, 77, 78]. In essence, maximising health outcomes comes with a trade-off with balancing equality or improving

equity [79, 80]. With health inequalities a recurring theme as a barometer for a well-functioning healthcare system, and equitable access a dominant aim of many health settings, there are pressures for commissioning to address such imbalances [81–84]. Thus, a future disconnect to bridge may be that between DCEA (and its associated evidence-base) and commissioning.

An additional disconnect between commissioning and HTA-focused economic evaluation is that of the relevant opportunity costs, i.e., what is being displaced for something else. For HTA, the opportunity cost is in relation to a new health technology compared with current/standard care, with the QALY-based cost-effectiveness threshold and associated monetary value of health also being a related (and debated) consideration [85–88]. The opportunity cost when commissioning services is different, as commissioners are trying to allocate their budgets across currently existing services, resulting in investment/disinvestment decisions potentially across sectors (e.g., for ICSs across health and social care) when something new is to be introduced, thus the relevant comparison might be quite different. In the case of disinvestment decisions for example, the decision might not just be an enhanced mental health service compared with current mental health service only, whereby we are displacing the current and potentially less efficient service for the newer more efficient but more expensive service; rather, due to finite and fixed budgets, the broader decision is if we spend more money on a new, more expensive, albeit more efficient, enhanced mental health service at all, we have to disinvest in an existing service beyond that mental health service to balance the budget, e.g., partial or complete disinvestment in a drugs rehabilitation service, in order to balance the decision-makers' budget across their jurisdiction. This disinvestment decision represents a broader consideration of disinvestment than current HTA processes directly account for within an economic evaluation, although the issue of disinvestment compared with investment has been discussed and debated also in relation to HTA [89–91].

8 Conclusion

It is our perspective that economic evaluations and care commissioners are not necessarily focused on achieving different outcomes, e.g., from the healthcare system perspective, the maximisation of health. However, care commissioners must necessarily use surrogate outcomes (often based on activity) and rely on causal mechanisms to achieve the relevant outcome that economic evaluations can quantify and report on explicitly. One way to bridge the gap could be to instate outcomes-based

commissioning based on the QALY: QALYs as a metric of health-related quality and length of life is almost certainly an outcome of interest to the healthcare system and therefore could be the basis of commissioning services if the system was amenable to its use. However, the commissioning landscape seems to have issues/concerns with outcomes-based commissioning (not just if there was an attempt to link payments to QALYs gained), such that linking payments to endpoint outcomes compared with activity-based surrogates can lead to inefficiencies and inequalities rather than enabling commissioning of cost-effective care.

Overall, it seems necessary for economic evaluations and commissioning to be focused on different primary outcomes (i.e., health compared with healthcare activity), but commissioners and researchers need to have a shared understanding that this is for practical reasons. Fundamentally, those conducting economic evaluations, as well as care commissioners, are all interested in trying to improve health outcomes from a constrained budget. However, we need to develop better ways to account for and communicate the link from commissioning payment structures to surrogate and endpoint outcomes to achieve a shared understanding of how economic evaluation evidence is relevant to care commissioning, when properly designed and conducted, alongside the influence that focusing on surrogates has on endpoint outcomes, such as when the overall relevant endpoint outcome is health.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s40258-024-00875-3.

Acknowledgement The idea for this article stems from discussions within the Study Steering Committee (SSC) meetings of the NIHR Public Health Research programme-funded (NIHR award identifier: 133634) 'Unlocking data to inform public health policy and practice' project: MF and SH led aspects of the project, whereas WW and GR were part of the SSC. As such, the authors would like to thank all the co-applicants and SSC members of the 'Unlocking data to inform public health policy and practice' project, as without that project and associated discussions, we may not have conceived and dedicated our time to writing this article.

Author Contributions Concept and design: MF, SH, WW, GR, RMH. Drafting of the manuscript: MF, SH, RMH, WW, GR. Critical revision of the paper for important intellectual content: MF, SH, RMH, WW, GR. Obtaining funding: MF, SH, WW, GR.

Declarations

Conflict of Interest Matthew Franklin, Sebastian Hinde, William Whittaker, Gerry Richardson, and Rachael Maree Hunter declare no conflicts of interest. No other disclosures were reported.

Funding This project was funded by the National Institute for Health and Care Research (NIHR) Applied Research Collaboration Yorkshire

and Humber (ARC-YH; NIHR award identifier: 200166) and Applied Research Collaboration Greater Manchester (ARC-GM; NIHR award identifier: 200174). It is also a spin out from the NIHR Public Health Research programme-funding (NIHR award identifier: 133634) 'Unlocking data to inform public health policy and practice' project.

Role of the Funder/Sponsor The NIHR had no role in the design and conduct of the study; collection, management, analysis and interpretation of the data; preparation, review or approval of the manuscript; and decision to submit the manuscript for publication. The views expressed are those of the authors and not necessarily those of the NIHR or the Department of Health and Social Care. The funding agreement ensured the authors' independence in developing the purview of the manuscript, and writing and publishing the manuscript.

Ethical Approval Not applicable as the manuscript is an opinion piece.

Consent to Participate Not applicable as the manuscript is an opinion piece.

Consent for Publication (from Patients/Participants) Not applicable as the manuscript is an opinion piece.

Availability of Data and Material Not applicable as the manuscript is an opinion piece.

Code Availability Not applicable as the manuscript is an opinion piece.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License, which permits any non-commercial use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by-nc/4.0/.

References

- NHS Digital. QOF 2021-22 results. 2023. https://qof.digital.nhs. uk/. Cited 23 Feb 2023.
- NHS Digital. Measures from the Adult Social Care Outcomes Framework. 2022 [cited 20 Feb 2023]. Available at: https://digital. nhs.uk/data-and-information/publications/statistical/adult-social-care-outcomes-framework-ascof. Cited 20 Feb 2023.
- NHS England. Quality and Outcomes Framework 2021/22; NHS England; 2021.
- The Health Foundation. Integrated care systems: what do they look like? 2022: https://www.health.org.uk/publications/longreads/integrated-care-systems-what-do-they-look-like?gclid= CjwKCAiA0cyfBhBREiwAAtStHLdMthxHD8XKHzcth7Zje UXW7hF_ZCm5R1hdR5vSaVVYuidXaudqCBoCBs4QAvD_ BwE. Cited 20 Feb 2022.
- Legislation.gov.uk. Health and Care Act 2022. 2022. https://www.legislation.gov.uk/ukpga/2022/31/contents/enacted. Cited 20 Feb 2023.

- World Health Organization (WHO). Addressing the social determinants of health: the urban dimension and the role of local government: World Health Organization, Regional Office for Europe; 2012.
- Hinde S, Howdon D, Lomas J, Franklin M. Health inequalities: to what extent are decision-makers and economic evaluations on the same page? An english case study. Appl Health Econ Health Policy. 2022;20(6):793–802.
- 8. Howdon D, Hinde S, Lomas J, Franklin M. Economic evaluation evidence for resource-allocation decision making: bridging the gap for local decision makers using English case studies. Appl Health Econ Health Policy. 2022;20(6):783–92.
- Bohm N, Bermingham S, Grimsey Jones F, Gonçalves-Bradley DC, Diamantopoulos A, Burton JR, et al. The challenges of outcomes-based contract implementation for medicines in Europe. Pharmacoeconomics. 2022;40(1):13–29.
- Connect to Commissioning Support. Outcome-based commissioning. 2023 https://commissioning.connecttosupport.org/s4s/Where ILive/Council?pageId=1943. Cited 20 Feb 2023.
- Moss I. The state of commissioning: preparing Whitehall for outcomes-based commissioning: Institute for Government; 2010.
- Robertson R, Ewbank L. Thinking differently about commissioning. London: King's Fund; 2020.
- 13. Taunt R, Allcock C, Lockwood A. Need to nurture: outcomesbased commissioning in the NHS. Health Foundation; 2015.
- 14. GOV.UK. Open public services: white paper; 2011.
- 15. NHS Digital. Payment by Results. 2019: https://digital.nhs.uk/data-and-information/data-tools-and-services/data-services/hospital-episode-statistics/payment-by-results#:~:text=Payment%20by%20Results%20(PbR)%20is,each%20patient%20seen%20or%20treated. Cited 20 Feb 2023.
- Bolton J. Emerging practice in outcome-based commissioning for social care. Oxford: Institute of Public Care, Oxford Brookes University; 23 April 2015.
- Feng Y, Kristensen SR, Lorgelly P, Meacock R, Sanchez MR, Siciliani L, et al. Pay for performance for specialised care in England: strengths and weaknesses. Health Policy. 2019;123(11):1036–41.
- Franklin M, Lomas J, Walker S, Young T. An educational review about using cost data for the purpose of cost-effectiveness analysis. Pharmacoeconomics. 2019;37(5):631–43.
- 19. NICE. NICE health technology evaluations: the manual 2022.
- Rowen D, Azzabi Zouraq I, Chevrou-Severac H, van Hout B. International regulations and recommendations for utility data for health technology assessment. Pharmacoeconomics. 2017;35(Suppl 1):11–9.
- Franklin M, Lomas J, Richardson G. Conducting value for money analyses for non-randomised interventional studies including service evaluations: an educational review with recommendations. Pharmacoeconomics. 2020;38(7):665–81.
- 22. Frew E, Breheny K. Health economics methods for public health resource allocation: a qualitative interview study of decision makers from an English local authority. Health Econ Policy Law. 2020;15(1):128–40.
- Frew E, Breheny K. Methods for public health economic evaluation: a Delphi survey of decision makers in English and Welsh local government. Health Econ. 2019;28(8):1052–63.
- Leao DL, Cremers H-P, van Veghel D, Pavlova M, Groot W. The impact of value-based payment models for networks of care and transmural care: a systematic literature review. Appl Health Econ Health Policy. 2023;21(3):441–66.
- Bovaird T, Davies R. Outcome-based service commissioning and delivery: does it make a difference? New steering concepts in public management. Emerald Group Publishing Limited; 2011. p. 93–114.

- National Audit Office. The Troubled Families Programme: Update. National Audit Office; 2016.
- National Audit Office. The work programme. National Audit Office: 2014.
- Culyer AJ. Health, health expenditures and equity. University of York Centre for Health Economics; 1991.
- Coast J, Smith RD, Lorgelly P. Welfarism, extra-welfarism and capability: the spread of ideas in health economics. Soc Sci Med. 2008;67(7):1190–8.
- Dawson J, Rigby-Brown A, Adams L, Baker R, Fernando J, Forrest A, et al. Developing and evaluating a tool to measure general practice productivity: a multimethod study. Health Serv Deliv Res. 2019;7(13):1–184.
- 31. Brazier J, Ratcliffe J, Saloman J, Tsuchiya A. Measuring and valuing health benefits for economic evaluation. Oxford: Oxford University Press; 2017.
- 32. Karimi M, Brazier J. Health, health-related quality of life, and quality of life: what is the difference? Pharmacoeconomics. 2016;34:645–9.
- $33. \ \ Bowling \ A. \ Measuring \ health. \ McGraw-Hill \ Education; \ 2004.$
- Bowling A. Measuring disease. Buckingham: Open University Press; 2001.
- Longworth L, Yang Y, Young T, Mulhern B, Hernández Alava M, Mukuria C, et al. Use of generic and condition-specific measures of health-related quality of life in NICE decision-making: a systematic review, statistical modelling and survey. Health Technol Assess. 2014. https://doi.org/10.3310/hta18090.
- Al-Janabi H, Flynn T, Coast J. Development of a self-report measure of capability wellbeing for adults: the ICECAP-A. Qual Life Res. 2012;21:167–76.
- Coast J, Flynn TN, Natarajan L, Sproston K, Lewis J, Louviere JJ, et al. Valuing the ICECAP capability index for older people. Soc Sci Med. 2008;67(5):874–82.
- Flynn TN, Chan P, Coast J, Peters TJ. Assessing quality of life among British older people using the ICEPOP CAPability (ICECAP-O) measure. Appl Health Econ Health Policy. 2011;9:317–29.
- 39. Netten A, Burge P, Malley J, Potoglou D, Towers A-M, Brazier J, et al. Outcomes of social care for adults: developing a preference-weighted measure. Health Technol Assess. 2012;16(16):1–166.
- Gyani A, Shafran R, Layard R, Clark DM. Enhancing recovery rates: lessons from year one of IAPT. Behav Res Ther. 2013;51(9):597–606.
- NHS Digital. Improving Access to Psychological Therapies (IAPT) data set reports. 2022: https://digital.nhs.uk/data-and-information/data-collections-and-data-sets/data-sets/improving-access-to-psychological-therapies-data-set/improving-access-to-psychological-therapies-data-set-reports. Cited 22 Feb 2023.
- NHS Digital. Patient Reported Outcome Measures (PROMs). 2022 https://digital.nhs.uk/data-and-information/data-tools-and-services/data-services/patient-reported-outcome-measures-proms. Cited 23 Feb 2023.
- 43. Porter ME. What is value in health care. N Engl J Med. 2010;363(26):2477-81.
- Zanotto BS, da Silva Etges APB, Marcolino MAZ, Polanczyk CA.
 Value-based healthcare initiatives in practice: a systematic review.
 J Healthc Manag. 2021;66(5):340.
- 45. DeMets DL, Psaty BM, Fleming TR. When can intermediate outcomes be used as surrogate outcomes? JAMA. 2020;323(12):1184–5.
- Hawkins N, Richardson G, Sutton AJ, Cooper NJ, Griffiths C, Rogers A, et al. Surrogates, meta-analysis and cost-effectiveness modelling: a combined analytic approach. Health Econ. 2012;21(6):742–56.

- 47. Ciani O, Taylor R. Surrogate, friend or foe? The need for case studies of the use of surrogate outcomes in cost-effectiveness analyses. Health Econ. 2013;22(2):251–2.
- Ciani O, Buyse M, Drummond M, Rasi G, Saad ED, Taylor RS. Time to review the role of surrogate end points in health policy: state of the art and the way forward. Value in Health. 2017;20(3):487–95.
- 49. Tassinari D. Surrogate end points of quality of life assessment: have we really found what we are looking for? Health Qual Life Outcomes. 2003;1(1):1–5.
- World Health Organization. The world health report 2002: reducing risks, promoting healthy life. World Health Organization; 2002.
- 51. Poirier P, Giles TD, Bray GA, Hong Y, Stern JS, Pi-Sunyer FX, et al. Obesity and cardiovascular disease: pathophysiology, evaluation, and effect of weight loss: an update of the 1997 American Heart Association Scientific Statement on Obesity and Heart Disease from the Obesity Committee of the Council on Nutrition, Physical Activity, and Metabolism. Circulation. 2006;113(6):898–918.
- 52. Brugts J, Yetgin T, Hoeks S, Gotto A, Shepherd J, Westendorp R, et al. The benefits of statins in people without established cardiovascular disease but with cardiovascular risk factors: meta-analysis of randomised controlled trials. BMJ. 2009;338: b2376.
- O'Donnell E, Atkinson J-A, Freebairn L, Rychetnik L. Participatory simulation modelling to inform public health policy and practice: rethinking the evidence hierarchies. J Public Health Policy. 2017;38:203–15.
- Hinde S, Bojke L, Richardson G, Retat L, Webber L. The costeffectiveness of population health checks: have the NHS Health Checks been unfairly maligned? J Public Health. 2017;25:425–31.
- Briggs A, Sculpher M, Claxton K. Decision modelling for health economic evaluation. Oxford: Oxford University Press; 2006.
- 56. Mason T, Whittaker W, Jones A, Sutton M. Did paying drugs misuse treatment providers for outcomes lead to unintended consequences for hospital admissions? Difference-in-differences analysis of a pay-for-performance scheme in England. Addiction (Abingdon, England). 2021;116(11):3082–93.
- Ash AS, Ellis RP. Risk-adjusted payment and performance assessment for primary care. Med Care. 2012;50(8):643.
- Cattel D, Eijkenaar F. Value-based provider payment initiatives combining global payments with explicit quality incentives: a systematic review. Med Care Res Rev. 2020;77(6):511–37.
- Twomey C, Prina AM, Baldwin DS, Das-Munshi J, Kingdon D, Koeser L, et al. Utility of the Health of the Nation Outcome Scales (HoNOS) in predicting mental health service costs for patients with common mental health problems: historical cohort study. PLoS ONE. 2016;11(11): e0167103.
- Woods B, Lomas J, Sculpher M, Weatherly H, Claxton K. Achieving dynamic efficiency in pharmaceutical innovation: identifying the optimal share of value, the payments required and evaluating pricing policies. Economic Methods of Evaluation in Health and Care Interventions (EEPRU); 2022.
- 61. Pandya A, Doran T, Zhu J, Walker S, Arntson E, Ryan AM. Modelling the cost-effectiveness of pay-for-performance in primary care in the UK. BMC Med. 2018;16:1–13.
- Minchin M, Roland M, Richardson J, Rowark S, Guthrie B. Quality of care in the United Kingdom after removal of financial incentives. N Engl J Med. 2018;379(10):948–57.
- Parkinson B, McManus E, Sutton M, Meacock R. Does recruiting patients to diabetes prevention programmes via primary care reinforce existing inequalities in care provision between general practices? A retrospective observational study. BMJ Qual Saf. 2023;32(5):274–85.
- Brennan A, Chick SE, Davies R. A taxonomy of model structures for economic evaluation of health technologies. Health Econ. 2006;15(12):1295–310.

- Jin H, Tappenden P, Ling X, Robinson S, Byford S. A systematic review of whole disease models for informing healthcare resource allocation decisions. PLoS ONE. 2023;18(9): e0291366.
- Günal MM, Pidd M. Discrete event simulation for performance modelling in health care: a review of the literature. J Simul. 2010;4:42–51.
- Cassidy R, Singh NS, Schiratti P-R, Semwanga A, Binyaruka P, Sachingongu N, et al. Mathematical modelling for health systems research: a systematic review of system dynamics and agent-based models. BMC Health Serv Res. 2019:19:1–24.
- Liu S, Xue H, Li Y, Xu J, Wang Y. Investigating the diffusion of agent-based modelling and system dynamics modelling in population health and healthcare research. Syst Res Behav Sci. 2018;35(2):203–15.
- 69. Zhang X. Application of discrete event simulation in health care: a systematic review. BMC Health Serv Res. 2018;18(1):1–11.
- Burau V, Blank RH. Comparing health policy: an assessment of typologies of health systems. In: Policy sectors in comparative policy analysis studies. London: Routledge; 2020. p. 46–59.
- Pryce R, Hollingsworth B, Walker I. Alcohol quantity and quality price elasticities: quantile regression estimates. Eur J Health Econ. 2019;20:439–54.
- 72. Holmes J, Meng Y, Meier PS, Brennan A, Angus C, Campbell-Burton A, et al. Effects of minimum unit pricing for alcohol on different income and socioeconomic groups: a modelling study. Lancet. 2014;383(9929):1655–64.
- Krauth C, Liersch S, Jensen S, Amelung VE. Would German physicians opt for pay-for-performance programs? A willingness-to-accept experiment in a large general practitioners' sample. Health Policy. 2016;120(2):148–58.
- Merlo G, Page K, Ratcliffe J, Halton K, Graves N. Bridging the gap: exploring the barriers to using economic evidence in healthcare decision making and strategies for improving uptake. Appl Health Econ Health Policy. 2015;13:303–9.
- Karnon J, Partington A, Gray J, Pincombe A, Schultz T. Local level economic evaluation: what is it? What is its value? Is it sustainable? Appl Health Econ Health Policy. 2023. https://doi. org/10.1007/s40258-023-00847-z. (Epub 18 Nov 2023).
- Capano G, Toth F. Health policy under the microscope: a micro policy design perspective. Front Public Health. 2023;11:1180836.
- 77. Asaria M, Griffin S, Cookson R. Distributional cost-effectiveness analysis: a tutorial. Med Decis Mak. 2016;36(1):8–19.

- Cookson R, Griffin S, Norheim OF, Culyer AJ, Chalkidou K. Distributional cost-effectiveness analysis comes of age. Value in Health. 2021;24(1):118–20.
- 79. Culyer AJ, Wagstaff A. Equity and equality in health and health care. J Health Econ. 1993;12(4):431–57.
- Wagstaff A. QALYs and the equity-efficiency trade-off. J Health Econ. 1991;10(1):21–41.
- Sir AD. Report of the independent inquiry into inequalities in health. London: Stationery office: 1998.
- Gray AM. Inequalities in health. The Black Report: a summary and comment. Int J Health Serv. 1982;12(3):349–80.
- 83. Marmot M, Allen J, Goldblatt P, Boyce T, McNeish D, Grady M, et al. The Marmot review: fair society, healthy lives. London: University College London; 2010.
- 84. Crombie IK, Irvine L, Elliott L, Wallace H. Closing the health and equalities gap: an international perspective. 2005.
- Towse A. Should NICE's threshold range for cost per QALY be raised? Yes. BMJ. 2009;338:b181.
- 86. Raftery J. Should NICE's threshold range for cost per QALY be raised? No. BMJ. 2009;338:b185.
- 87. Claxton K, Martin S, Soares M, Rice N, Spackman E, Hinde S, et al. Methods for the estimation of the National Institute for Health and Care Excellence cost-effectiveness threshold. Health Technol Assess (Winchester, England). 2015;19(14):1.
- Claxton K, Sculpher M, Palmer S, Culyer AJ. Causes for concern: is NICE failing to uphold its responsibilities to all NHS patients? Wiley Online Library; 2015. p. 1–7.
- Karnon J, Carlton J, Czoski-Murray C, Smith K. Informing disinvestment through cost-effectiveness modelling: is lack of data a surmountable barrier? Appl Health Econ Health Policy. 2009;7:1–9.
- Chambers JD, Salem MN, D'Cruz BN, Subedi P, Kamal-Bahl SJ, Neumann PJ. A review of empirical analyses of disinvestment initiatives. Value in Health. 2017;20(7):909–18.
- Polisena J, Clifford T, Elshaug AG, Mitton C, Russell E, Skidmore B. Case studies that illustrate disinvestment and resource allocation decision-making processes in health care: a systematic review. Int J Technol Assess Health Care. 2013;29(2):174–84.