HESG Summer 2019

Abstract Booklet

3rd-5th July 2019

Hosted by the University of East Anglia

University of East Anglia

HEALTH ECONOMICS GROUP

ID 665: Huajie Jin

Title: Using whole disease modelling to evaluate all key interventions for schizophrenia

Background:

As one of the most expensive mental health disorders, schizophrenia is estimated to cost society £11.8 billion a year in England. Economic models can be used to inform healthcare resource allocation decisions by comparing the cost-effectiveness of alternative interventions. However, a systematic review of published modelling studies for schizophrenia identified several problems of existing models: (1) lack of modelling evidence for non-drug interventions; (2) inconsistent conclusions reported by published studies (e.g. some papers reported that antipsychotic A is more cost-effective than antipsychotic B, while other papers reported that antipsychotic B is more cost-effective than antipsychotic A); and (3) most published models failed to capture the system interdependencies between different interventions (e.g. the cost-effectiveness of preventing schizophrenia depends on the cost-effectiveness of treating schizophrenia).

Aims of the Study:

The aim of this project is to build a Whole Disease Model (WDM), which can be used to inform resource allocation decisions for the prevention, diagnosis and treatment of schizophrenia.

Methods

A patient-level, discrete-event simulation model was developed using Simul8 software. This model simulates the disease and treatment pathways for people at high risk of psychosis, or people with a diagnosis of non-specific psychosis or schizophrenia. The model structure was informed by published literature and expert opinions. Two scenarios were modeled: a real-world scenario which simulates the current clinical pathway, and an ideal scenario which simulates the ideal care pathway recommended by the NICE schizophrenia guidelines. A life-time horizon was adopted. Most model parameters were obtained from meta-analysis, randomised controlled trial, utility and costing studies. Extensive verification and validation activities have been undertaken.

Results

The WDM was capable of evaluating the cost and effectiveness outcomes for 19 interventions for schizophrenia, and allows investment and disinvestment decisions to be made at the same time. Analysis of interactions between different interventions shows that, the change of existing system configurations altered the conclusion of the most cost-effective interventions for three of the five decision topics assessed.

ID 667: Sheheryar Banuri

Title: Encouraging Service Delivery to the Poor: Does Money Talk When Health Workers Are Pro-Poor?

Do service providers respond to pecuniary incentives to serve the poor? Service delivery to the poor is complicated by the extra effort required to deliver services to them and the intrinsic incentives of service providers to exert this effort. Incentive schemes typically fail to account for these complications. A lab-in-the-field experiment with nearly 400 health workers in rural Burkina Faso provides strong evidence that the interaction of effort costs, ability, and intrinsic and extrinsic incentives significantly influences service delivery to the poor. Health workers reviewed video vignettes of medical cases involving poor and nonpoor patients under a variety of bonus schemes. Bonuses to serve the poor have less impact on effort than bonuses to serve the nonpoor; health workers who receive equal bonuses to serve poor and nonpoor patients see fewer poor patients than workers who receive only a flat salary; and bonuses operate largely through their influence on the behavior of pro-poor workers. The paper also presents novel evidence on the selection effects of contract type: pro-poor workers prefer the flat salary contract to the variable salary contract.

ID 690: Javiera Cartagena Farias
Title: Understanding social care financial flows: A methodology for developing a social care price-based adjustment factor

The study of the underlying trends in social care expenditure in England is challenging. For instance, figures reflecting social care spending trends over time do not appropriately take into account the variability in the price of services across areas, differences in level of needs of their local population and quality of the services provided. Hence, it is crucial that robust approaches are developed for comparing local and national social care expenditure through time. This study explores the use of unit costs information provided by Local Authorities and publicly-available data from 2005 to 2017 to provide a more detailed picture on the pressures that the publicly funded social care system in England faces. We developed an approach for specifically discounting changes in social care unit costs (rather than general prices), constructed a social care price index and applied it to recent social care spending figures. This to enable a better comparison and interpretation of local and aggregate trends. Ordinary Least Squares (OLS) and Generalised Estimation Equation models (GEE) have been performed to reflect the local variability on the supply factors affecting social care services prices over time, together with controlling for area-level changes on population density, socio-demographic characteristics and level of needs. As expected, we found that when adjusting by social care specific prices, social care spending has not only decreased over time, but it has been much lower than when no adjustment is considered. The increase in social care prices over time was found to be mostly associated to changes in wages and house prices (around fifty percent of the increase in unit cost was given by an increase in these two factors). On the other hand, even when there was a reduction on unit cost variability across local authorities after adjusting by supply and demand factors, there were still important disparities across local authorities. We found some evidence that local authorities that may be under strong budget constrains have had higher incentives to better negotiate prices among their services suppliers compared to local authorities with healthier budgets.

Keywords: Social care spending, social care unit costs, social care expenditure, price index.

ID 691: Michael Shields
Title: Heterogeneity in the Impact of National and Local Economic Conditions on Morbidity in the UK

A substantive literature asks whether population health improves or worsens with changes in macroeconomic conditions. The workhorse model for many studies, including the seminal paper in economics by Ruhm (2000), is a linear regression of health on (un)employment with local area-fixed effects, and national and local area time trends. While this model controls for fixed and time-varying confounders, it is uninformative about (1) short versus long-run dynamics, (2) the extent of area-level heterogeneity in the relationship between the macroeconomy and health, (3) the optimal level of spatial disaggregation, and (4) does not provide estimates of the importance of national versus local area (un)employment on health. In this paper we develop a dynamic panel data model that informs on each of these issues and we apply the model to UK data on chronic health conditions over the period 2001-16.

ID 694: Peter Moffatt
Title: IMPACT AND DETECTION OF STRAIGHTLINING RESPONSE IN HEALTH SURVEYS

In the context of a stated-binary-choice experiment, or Discrete Choice Experiment (DCE), straightliners are defined as survey participants whose responses are not related to the attributes of the two alternatives, and can therefore be perceived as choosing at random in every choice task that they face. Straightlining may be seen as an extreme form of attribute non-attendance which has been studied extensively in the choice modelling literature. We apply recently developed econometric methodology to the MAUCa Australian EORTC QLU-C10D DCE data in order: (1) to investigate the extent of straightlining response in this data source; (2) to estimate preference parameters in ways that adjust for the presence of straightliners; (3) to investigate ways in which straightlining depends on respondent characteristics. We find a prevalence of straightliners in the survey, and we allow for their presence in the estimation of the preference parameters that are of central interest. We verify that estimates are seriously biased if their presence is neglected. For

each respondent in the survey, we obtain a posterior straightlining probability, which provides a simple means of identifying respondents who are most likely to be straightliners. We emphasise the usefulness of this as a justification for excluding identified straightliners from future surveys, leading to cost savings. Finally, we point out that the cross-country comparison of straightlining prevalence which is made possible by the MAUCa Consortium's program of research, is likely to be of considerable interest to the choice modelling community in general. Following analysis of the Australian QLU-C10D DCE data, results will be shared with other MAUCa members, and the interest of the custodians of QLU-C10D DCE data from other countries will be sought, with the aim of conducting cross-country comparisons of the prevalence of straightliners.

ID 695: Simon McNamara
Title: WHICH PREFERENCES? CHOICE ARCHITECTURE AND NORMATIVE HEALTH ECONOMICS

Every choice has an architecture: a design which shapes how we respond to that choice. Health economics is not immune to this phenomenon. Every preference elicitation exercise we conduct has an architecture, and that architecture inevitably influences how people respond to it. This raises a critical question for those who argue we should apply preferences in health-economic evaluation – if different, potentially legitimate, choice-architectures lead to different choices being made, and subsequently the inference of different preferences, which preferences should we apply? In this paper, we contextualise this question in the elicitation of the public's aversion to health inequalities. We first describe multiple, different, potentially legitimate choice-architectures that could reasonably be used to infer inequality aversion. We then make the case that these are likely to lead to different apparent preferences being elicited, as each is likely to be subject to differing behavioural effects, including compassion-fade, pseudoinefficacy, pseudoefficacy, choice-overload, and the focusing illusion. We then turn to the normative, and ask two questions: First, why are we interested in preferences? Second, do preference-centric normative positions provide guidance as to "which preferences" should be applied? We argue: (1) That traditional welfarism is not behaviorally sensitive, and does do not provide an answer to this second question; (2) That different types of behaviourally-sensitive welfarism could be called upon to justify use of a range of different preferences; (3) That Culyer's decision-maker centric "extra-welfarism" is sufficiently flexible to support any position advocated by decision-makers; (4) That different democracy-centric normative positions could be used to support application of a range of different preferences. Finally, we call for health economists who elicit preferences to be clear on why they are eliciting preferences, to reflect on which type of preferences they are aiming to elicit, and to reflect on how differing choice-architectures may, or may not, lead to those preferences being elicited.

ID 696: Sabina Sanghera
Title: How do patients with fluctuating health states complete EQ-5D, SF-12 and EORTC-QLQ-C30? A think-aloud study

Background:
Recurrent fluctuations in health states occur due to long-term conditions with episodic symptoms (e.g. multiple sclerosis) or side effects of cycles of treatment (e.g. chemotherapy for cancer). These fluctuations can have an important impact on a patient's quality of life. For example, chemotherapy causes cyclical side effects as each cycle has a treatment and rest period. It is not clear how patients with fluctuating health complete questionnaires. Quality of life may be over/underestimated due both to measure recall periods ('health today', 'past four weeks') and the timing of data collection. Therefore, data used in economic evaluations to inform decision-making could be inappropriate.

Aim:
Using chemotherapy treatment as a case study, we explore whether patients adhere to recall periods, construct an average, or recall the worst point of the chemotherapy cycle when completing EQ-5D-5L, SF-12 and EORTC-QLQ-C30.

Methods:
Adult patients receiving chemotherapy for urological, gynaecological or bowel cancers each took part in a think-aloud interview. Each patient: verbalised their thoughts while completing EQ-5D, SF-12 and EORTC-QLQ-C30; completed a pictorial task illustrating how quality of life changed over one chemotherapy cycle; and took part in a semi-structured interview. Transcripts were analysed using thematic analysis.

Results:
24 patients were interviewed between July and September 2018. Preliminary results suggest that patients were more likely to adhere to the recall period of EQ-5D ('health today'), but emphasised the recall period did not reflect their quality of life during chemotherapy. With respect to recall, patients seemed to have most difficulty completing SF-12 ('past four weeks') and could more easily complete EORTC ('past week'). However, across all questionnaires, patients attempted to provide averages, focused on the worst or best part of their cycle, or their most recent experience. Patients commonly used more than one of these approaches when completing even one questionnaire.

Discussion:
The results and implications for using current approaches to calculate quality-adjusted life-years when health fluctuates will be discussed. Since patients appear to provide inconsistent responses to questionnaires when health fluctuates, alternative ways of asking these questions should be explored to ensure meaningful evidence on quality of life is used to inform cost-effectiveness decisions.

ID 701: Mikyung Kelly Seo
Title: Economic evaluations of cancer biomarkers for targeted therapies: practices, challenges and policy implications

It is often argued that biomarkers for targeted therapies will eventually improve patient outcomes and help to achieve the resource allocation efficiency in healthcare. However, there is a widespread scepticism about the research and development (R&D) of biomarkers because the number of biomarkers successfully entering into clinical routine is very low compared to the number of biomarkers published.

It implies a significant time-lag between the pace of rapidly evolving medical technologies and the actual use in clinical practice. Such lagged integration may potentially delay the improvement of patient outcomes or may harm patients unresponsive to the corresponding therapies. Therefore, it is of public interest to ensure the timely integration of new technologies into clinical use through adequate levels of reimbursement and coverage.

However, it requires that test developers demonstrate robust evidence of health economic impact of biomarkers for targeted therapies. Biomarker characteristics captured in economic evaluations are often limited to the cost or the accuracy of tests. Often, only the costs of biomarker testing are modelled in current practice of assessing the cost-effectiveness of cancer biomarkers for targeted therapies. Clinical outcomes or clinical utilities are often difficult to include due to limited data generated by clinical trials.

This study reviews current methodological approaches and challenges in conducting economic evaluations of cancer biomarkers and the complexity of evidence generation faced by test developers without clear guidance on evidentiary standards and data requirements. It then discusses some implications for regulatory and reimbursement bodies. Although some of the issues could be better addressed by other fields (for example, clinical utility of biomarker testing), a consensus on methodological approaches and data requirements for economic evaluations of cancer biomarkers needs to be reached in the field of health economics. Currently, no agreement exists whether existing methods are sufficient to evaluate the health economic impact of biomarkers, or whether different methodological approaches might produce conflicting results regarding the cost-effectiveness of biomarkers.

ID 704: Haseeb Ahmed
Title: MOBILE MONEY AND HEALTHCARE USE: EVIDENCE FROM EAST AFRICA

Poor households rely on informal risk-sharing, which may be incomplete due to information asymmetries and transaction costs. Mobile money transfer technology (MMT) can increase risk-sharing by reducing the transaction costs associated with borrowing and lending. This paper uses a difference-in-difference framework to estimate the effects MMT on healthcare use in the face of negative health shocks. We use survey data from 2013-16 with quarterly observations on about 1,600 households of 10 villages in the Kisumu region of Western Kenya. The validity of our identification strategy rests on the assumption that health shocks are exogenous. We find that health shocks are uncorrelated with observables, suggesting that these shocks are exogenous and equally affect users and non-users of MMT. If this is the case, we will obtain unbiased estimates of the effect of MMT on consumption smoothing as long as unobserved differences between MMT user and non-users do not vary by health shock status.

We find that during times of illness, users of MMT spend 63% more per capita on health, are 50% more likely to buy medication, are almost two times more likely to pay consultation fees, and utilize more formal healthcare facilities than non-users of the technology. These results suggest that if households are liquidity constrained in the short-run when acute health shocks arise, MMT could help them overcome these constraints by availing their social networks for loans or gifts to pay for needed healthcare. Self-insurance against health shocks due to MMT may also have important longer-run consequences on household welfare if utilization of healthcare services decrease mortality and long-term illness.

We also test the mechanisms through which MMT can increase risk-sharing. We find that MMT users are 5 percentage points more likely to acquire loans than non-users during a negative health shock, indicating that borrowing may be the main mechanism. Lastly, we note that the quantity of maize received as a gift increases for MMT non-users but does not increase for MMT users under a health shock. These results are consistent with the notion that cash transfers via MMT substitute for in-kind transfers during periods of (health) distress.

ID 705: Olivier Wouters
Title: What is the price impact of tendering for medicines by therapeutic class? A difference-in-differences estimation

Context: Many new medicines offer little or no additional therapeutic benefit over existing therapies. One option to reduce spending on so-called "me-too" drugs, and to indirectly incentivise drug companies to develop truly innovative medicines, is for payers to tender for medicines by therapeutic class. Under a therapeutic tendering policy, the payer only buys the least expensive products in defined therapeutic classes. The aim of this study was to measure the association between implementation of therapeutic tendering and medicine prices in South Africa.

Methods: I carried out a retrospective observational study of changes in the pack prices of solid-dose medicines (n=138) over an 11-year period (2003-2014) before and after the introduction of a national therapeutic tendering system in South Africa and compared the findings with changes over the same period for solid-dose medicines which were not subjected to the policy (n=672). I ran various difference-in-differences models, with and without adjustments for product characteristics, to estimate the impact of the new policy on medicine prices and spending.

Findings: Compared to trends in the prices of solid-dose medicines not subjected to therapeutic tendering, implementation of therapeutic tendering was associated with a 33.0% (95% confidence interval [CI], 4.0%-53.3%) to 43.7% (95% CI, 20.7%-60.0%) reduction in the average price paid per medicine, depending on the statistical model used. Spending by the national government declined by an estimated 1.13 billion (95% CI, ZAR 138.01 million - ZAR 1.83 billion) to 1.50 billion rand (95% CI, ZAR 709.65 million - ZAR 2.06 billion) as a

result of therapeutic tendering. In robustness checks, the results remained significant and comparable in size to the estimates from the main models.

Conclusions: Therapeutic tendering was associated with large drops in the prices of solid-dose medicines, which likely saved the South African national government over one billion rand in 2014. Given that solid-dose medicines only accounted for around 16% of total drug spending in 2014, there is potential for even greater savings in other medicine categories. Future studies should re-examine therapeutic tendering systems with more data from post-intervention periods.

ID 709: Hannah Louise Bromley
Title: Valuing health states associated with breast cancer: a primary study exploring methods, challenges and the issue of overdiagnosis

Background: Utilities informing breast screening and treatment policy are limited in their ability to adequately capture the benefits and risks. Evaluations of breast cancer do not include the disutility of overdiagnosis in economic outcomes and may potentially bias the decision on screening. However, not all methods for deriving utility may account for uncertainty in disease progression and risk of overdiagnosis during the valuation process.

Objectives: The objective of this paper was to explore methods to derive the utilities associated with breast cancer, including the potential overdiagnosis and unnecessary treatment of low-risk disease. In doing so, the methodological challenges in health state valuation are explored, in particular, the uncertainties associated with valuing population screening outcomes, overdiagnosis and the impact of risk on utility.

Methods: We conducted a primary study interviewing 172 women with and without a history of breast cancer in Melbourne, Australia between April and September 2018. Participants were recruited from the Breast Cancer Network Australia and Lifepool registries. Women were presented with seven vignettes describing breast cancer treatments, which explicitly included the risk of overdiagnosis and unnecessary treatment in the description. Utilities were collected using the visual analogue scale, standard gamble and EQ-5D-5L questionnaire to explore the validity, reliability and feasibility of direct and indirect approaches in capturing the utility associated with the overdiagnosis and risk of potential unnecessary treatment.

Preliminary discussion: The results of the empirical work are currently being analysed. The findings will likely discuss the methodological challenges of valuing overdiagnosis using conventional valuation approaches and highlight further research necessary to address the issues identified. Whilst interpretation of the results is ongoing, we believe the values have the potential to advise policy decisions by explicitly capturing disutility in the evidence informing breast cancer screening programmes.

ID 710: Claryn Kung
Title: How well can childhood circumstances explain later life loneliness? Life-course evidence from two UK longitudinal surveys

A growing literature has documented a strong link between loneliness and health and mortality. Loneliness is particularly prevalent in old age, but recent studies have documented similar issues in adolescence and early adulthood. However, robust evidence on drivers over the life course is limited. The genetic heritability of loneliness has been found to be very low (<5%), but few studies have quantified the role that childhood conditions and circumstances might play in putting individual on a trajectory of future (chronic) loneliness. This study aims to fill this gap, using data from the English Longitudinal Study of Ageing (ELSA), and the Avon Longitudinal Study of Parents and Children (ALSPAC). Together with its eight biennial measurements of loneliness, the ELSA (aged 50+) administers a rich set of retrospective life history questions such as childhood socioeconomic status, health, household composition including parents' marital status, adverse events, and parenting style. We supplement our findings with the ALSPAC, a birth cohort study providing detailed prospective data on these childhood conditions, as well as loneliness captured from ages 10 to 23.

These two longitudinal surveys therefore span adolescence and early adulthood as well as mid-to-late-adulthood, enabling us to provide a unique life-course perspective and a comprehensive assessment of the potential importance of childhood in explaining later life loneliness.

To estimate the long-term relationship between childhood circumstances and loneliness, we use a variety of econometric techniques, including a dynamic panel finite fixture model that enables us to shed new light on the heterogeneity in response to childhood. Descriptive findings across these data sets reveal childhood mental health, maternal relationship, financial hardship, and abuse strongly predict future loneliness for older adults, with some differences by gender and age at which loneliness is measured. For instance, paternal relationship matters more for males, and the loneliness impacts of sexual abuse only emerges in adulthood. We further aim to examine the proportion of variation in adulthood loneliness explicable from fixed childhood variables, and how their impact evolves over time. We also aim to examine pathways via which these childhood variables impact later life loneliness, such as marital status, fertility and employment.

ID 714: Chidubem Okeke Ogwulu
Title: The valuation of outcomes for the temporary and chronic health states associated with Chlamydia trachomatis infection

Objective: Eliciting health-state utility values (HSUVs) for some diseases is complicated by the mix of associated temporary (THSs) and chronic health-states (CHSs). This study uses one such disease, chlamydia infection, to explore the challenges. The objectives were to:

1) Define a set of health-state descriptions related to chlamydia and
2) Derive HSUVs for the THSs and CHSs associated with chlamydia.

Methods: HSUVs were elicited for seven health-states (five THSs and two CHSs) depicting the symptoms of chlamydia, developed using evidence from the literature and clinical experts. Chained time trade-off (TTO) and visual analogue scale (VAS) were applied to THSs while conventional TTO was applied to CHSs. Ectopic pregnancy was used as a lower anchor for chained TTO. The study sampled from three different population groups and the survey was administered face-to-face.

Results: One hundred participants (37 students, 37 STI clinic attendees, and 26 university staff) were assessed with an interview completion rate of 100%. Mean TTO utilities were consistently higher than VAS scores. The aggregated mean chained TTO results for the THS ranged from 0.46 (SD, 0.24) for ectopic pregnancy, to 0.77 (SD 0.21) for cervicitis. However, some participants preferred other THSs to the anchor health-state.

Conclusions: Methodological challenges included the development of health-state descriptions, the selection, and duration of appropriate anchor state. Chained TTO was shown to be feasible in this population and the resulting HSUVs could have implications for economic evaluations for chlamydia prevention and control.

ID 716: Edward Burn
Title: Guiding innovation by estimating threshold prices: a case study of robot-assisted knee and hip replacement

BACKGROUND
While knee and hip replacements are both highly cost-effective interventions, there remains room for improvement in patient outcomes. Robot-assisted systems have been developed to improve component positioning, but their long-term effectiveness has not been established and the price which would be paid for such improvements is not known.

OBJECTIVE

To estimate threshold prices for potential improvements in knee replacement and hip replacement.

METHODS
A lifetime cohort Markov model was used as the framework for analysis. Model inputs were based on routinely collected data from NHS England hospital records covering 1997 to 2014. The model was run for the average patient profile and then re-run for three base case scenarios, representing best-case scenarios for robot-assisted systems: (a) health utility estimate increased by 5% for those unrevised following surgery, and (b) risk of revision reduced by 50%, (c) health utility estimate increased by 5% for those unrevised following surgery and risk of revision reduced by 50%. Threshold prices, a price at which the health benefits gained will be offset by equivalent health displaced due to opportunity cost (i.e. where the net health benefit from funding the intervention would be zero), were estimated for a cost-effectiveness threshold of £20,000 per additional QALY gained.

RESULTS
Threshold prices for a 5% improvement in post-operative health utility were £9853 (£9500 to £10194) for knee replacement and £10787 (£10388 to £11085) for hip replacement. Threshold prices for a 50% reduction in the risk of revision £1217 (£1099 to £1512) for knee replacement and £1392 (£1047 to £1816) for hip replacement. A combined 5% improvement in post-operative health utility and a 50% reduction in the risk of revision £11258 (£11046 to £11845) for knee replacement and £12383 (£12064 to £12695) for hip replacement.

CONCLUSIONS
This study has estimated threshold prices for best-case scenarios for robot-assisted knee and hip replacement. To be considered cost-effective at these prices, robotic systems will need to be costed at or under such prices and demonstrate their effectiveness. If the true effectiveness of robot-assisted surgery is less, threshold prices will also be lower.

ID 717: Francesco Manca
Title: Should we still be screening for HCV in Scotland and the UK?

Background: In the UK about 215,000 people are chronically infected with HCV, yet this is the 'tip of the iceberg' as the majority of people infected with HCV remain undiagnosed. Now that new curative antiviral agents for HCV are available, it is crucial to identify early stage and asymptomatic undiagnosed infections which if left untreated can cause serious and potentially life-threatening damage over many years, with escalating healthcare costs. In areas where HCV prevalence is not endemic, the main concern is to understand how to prioritise screening strategies and target underdiagnosed population groups effectively. This study aims to identify and assess the most cost-effective strategies for diagnosing HCV infection in 'under diagnosed populations' using Scotland as case study.

Methods: Two key populations were identified: intravenous drug users and targeted high risk patients among general population. A cost-effectiveness analysis was undertaken for each population comparing six and five alternative screening strategies, respectively. Each strategy, differing for point of care and targeted subpopulation, was compared against the standard care diagnostic pathways from the Scottish NHS perspective. Data on each strategy was obtained from observational sources and recently published pilot studies, then synthesised in a decision analytic model. A decision tree was developed to explore the incremental cost per additional positive patient detected, over a 1 year time horizon, and a previously published Markov model was adapted and employed to present lifetime outcomes in terms of incremental cost per QALYs gained. Scenario analysis was undertaken to explore impacts of reducing re-infection rate in a static framework, since evidence suggests that drug users have a higher risk of re-infection.

Results & Discussion: Preliminary results show that the most cost-effective strategy for drug users is providing tests at Harm Reduction Services, and that current practice of self-referral to GP practice for testing is the least cost-effective strategy. Access to testing is a significant obstacle for early diagnosis. The most cost-effective strategies are those targeting the highest risk subpopulations at early age of disease, yet there may be obstacles to implementation in practice. This work can help informing the Scottish Government's plans to achieve a hepatitis free Scotland.

ID 718: Rita Santos
Title: The impact of mergers on general practices quality

There has been a trend increase in the size of general practice and policy is directed towards encouraging larger practices and collaborations. There is limited and mixed evidence of the impact on practices outcomes. We examine whether mergers affect practice quality.

We find practices that merge and construct pre-merger characteristics for them as the size-weighted means of the merging practices. We matched merged practices with stable practices that did not merge using coarsened exact matching (CEM) on measures of performance and other characteristics in the period before the merged practice was created. We estimate the effect of mergers on both parametrically (using a difference in differences fixed effect linear panel models) and non-parametrically (comparing merged practice performance with the mean performance of their matched stable controls) performance.

We use the NHS England General Practitioner's membership of GP Practices data and the location of GP Practices, collected from NHS Choices and Connecting for Health (archive and current data files) to find mergers. Quality indicators include the proportions of patients that are very satisfied with overall care, with opening hours and who would recommend the practice as reported in GP Patient Survey (GPPS), overall clinical quality measures calculated from the Quality Outcome Framework (QOF), and rates of emergency admissions for Ambulatory Care Sensitive Conditions (ACSCs), prescription measures from NHS Digital and avoidable attendances to Accident and Emergency department from Hospital Episode Statistics (HES).

GP practice workforce and practice list data were collected from NHS Digital General and Personal Medical Services.
Small area socioeconomic characteristics were attributed to practices using the proportions of their patients in local small areas,

Merged practices have lower quality post-merger compared with their matched stable control practices. For example, the percentages of patients who were satisfied with opening hours, satisfied with their overall experience, and who would recommend their practice were 2%, 1.5%, and 5% smaller in merged practices post-merger.

ID 720: Dwayne Boyers
Title: External validity of Discrete Choice Experiments in dental care - the impact of ex-ante hypothetical bias correction methods

Background:
Stated preference methods such as discrete choice experiments (DCEs) rely on individual's responses to hypothetical questions. This raises the question if respondents' stated choices accurately reflect the choices they would make in reality. Research suggests that stated preference studies overestimate service uptake and willingness to pay (WTP).

Researchers have proposed three novel ex-ante methods to mitigate hypothetical bias by adding additional information or questions to the DCE survey: 1. Cheap-talk scripts, which explain the hypothetical bias problem to respondents; 2. Consequentiality scripts, which inform respondents about how the study results can affect policy; and 3. Honesty oaths, which ask respondents to commit to providing truthful answers.

Whilst they have shown promise within the environmental literature, only cheap talk has been tested in a health care DCE. No previous studies have directly compared the performance of the three methods. Furthermore there are concerns that the methods may 'over-correct' creating additional biases.

Aims and objectives:
To test the performance of ex-ante hypothetical bias mitigation methods in a healthcare DCE.

Methods:
We use a DCE eliciting general population preferences for dental care. The sample is randomised to: 1. cheap-talk script; 2. consequentiality script; 3. oath or 4. no mitigation. Data quality (consistency tests, dominance tests, opt-out decision, completion times, dropout rates, and successful completion of an in-built survey attention question), elicited WTP, and the external validity of service uptake predictions are compared across groups. External validity is tested by comparing the DCE predictions with service uptake estimated from revealed preference data.

Results:
We find no cross group differences in data quality or marginal WTP. We find hypothetical bias in terms of service uptake, but the magnitude of the bias is small. Nevertheless, we find that the consequentiality script and oath generate lower levels of hypothetical bias. However, we find that respondents are more likely to drop out of the survey when presented with the oath, which raises concerns about sample representativeness.

Conclusion:
We find that none of the methods eradicate hypothetical bias. We also find that ex-ante mitigation does not introduce additional biases.

ID 723: John O'Dwyer
Title: A Systematic Review of Self-reported Quality of Life Measures for Adults with Mild to Moderate Learning Disabilities

Around 1.2 million people in the UK have a mild or moderate learning disability, living on average 16 years less than the general population. Whilst there are clear health inequalities, research with this population is constrained. Evidence suggests that people with learning disabilities often have difficulty completing research materials. This systematic review brings together evidence on self-reported quality of life (QoL) measures that have been used in research involving adults with a mild or moderate learning disability and that have assessed their validity and reliability, to identify potential adaptations that might be made to the NICE recommended HRQoL measure, the EQ-5D, in order to reduce completion difficulty.

We included studies, in clinical and non-clinical settings, which contain self-reported QoL data of adults (>16 years) who have a mild or moderate learning disability. Searches were completed on Medline, EMBASE, Cochrane Library and other databases for English-language articles, reporting data from English-language QoL measures, published between 1990 and 2018. Searches included published and unpublished trials. Screening of titles and abstracts, full-text assessments and data extraction was completed by one reviewer and a random sample (10%) reviewed and validated by a second reviewer. Data extraction included QoL measure(s); methods and adaptations for data collection; validity; and reliability. A narrative synthesis was completed.

From 22,291 unique citations identified, 218 studies formed the final dataset which included 47 self-reported QoL measures; 18 have been validated for use by adults with mild to moderate learning disabilities. These generic measures include various domains of QoL. Adaptations such as pictograms, contextualisation of language and longer completion times are reported. There is no preference-based QoL measure validated for this population.

The range of generic self-reported QoL measures used with this population is wide; comparison of measures is difficult as definitions of QoL, and hence domains examined, differ. The NICE reference case recommends using preference-based measures of HRQoL, with EQ-5D being their preferred measure in adults. Neither the EQ-5D nor any other preference-based measure has been validated for this population; Potential adaptations of EQ-5D to remedy this issue are discussed.

Prospero Registration: CRD42018092423
Funding: NIHR DRF 2017-10-159

ID 724: Alessio Gaggero
Title: Does Lifestyle Matter for Healthy Ageing?

In this paper, we use novel approach based on biomarkers data, collected through blood samples, to examine the effect of a lifestyle intervention on healthy ageing. Exploiting the fact that medical guidelines cause a discontinuity in the probability individuals are assigned to an intensive lifestyle intervention, as a function of exact blood test cut-off, we use a regression discontinuity design (RDD) to estimate the impacts of a change in lifestyle on both physical and mental health, for a sample of individuals aged 50 and above. Using data from the English Longitudinal Study of Ageing (ELSA), our findings provide compelling evidence that lifestyle does matter for healthy ageing. Our results suggest that individuals assigned to a lifestyle intervention report significantly better physical and mental health scores than their counterpart. We provide a wide variety of evidence on the validity of our results.

ID 727: Jaesh Naik
Title: Comparing results from a pathway of idelalisib and venetoclax modelled using consistent survival assumptions to results produced from the two interventions modelled independently.

Introduction: The National Institute for Health and Care Excellence (NICE) provides recommendations on the use of pharmaceuticals within the National Health Service through its technology appraisal process. There is no certain method for ensuring manufacturer submissions for drugs across different treatment lines are consistent in terms of their approach. In the case study of TA359 (idelalisib) and TA487 (venetoclax) for chronic lymphocytic leukaemia, both companies made markedly different assumptions regarding idelalisib post-progression survival (PPS).

Aims: To compare cost-effectiveness results when consistent survival assumptions are used within the idelalisib-venetoclax pathway to results for the interventions modelled independently. To subsequently determine whether one or more of the NICE recommendations was incorrect. To highlight the importance to decision making of a consistent modelling approach across intrinsically linked pathways.

Methods: The idelalisib and venetoclax cost-effectiveness models were reconstructed, using publicly available NICE documentation and a supporting systematic review of the literature, to ascertain the incremental cost-effectiveness ratios (ICERs) for which each drug received its recommendation. As venetoclax is recommended for patients who have failed idelalisib, overall survival in the comparator arm of the venetoclax model was amended to equal idelalisib PPS and cost-effectiveness results estimated.

Results: The study indicates that the ICER for idelalisib falls between £28,560 (when the venetoclax company estimate of idelalisib PPS is used) and £37,798 (when the NICE estimate of idelalisib PPS is used). The venetoclax ICER ranges from £52,017 (when the NICE estimate of idelalisib PPS is used) to £77,287 (when the idelalisib company estimate of PPS is used).

Conclusions: The study shows that, whilst venetoclax should have be appraised at the higher end-of-life willingness-to-pay threshold of £50,000, it was not a cost-effective treatment option for routine use. Idelalisib was cost-effective at the end-of-life threshold. Ensuring that idealsisib PPS was consistent across models had a significant impact on results. In practice, more accurately modelling treatment pathways by ensuring consistency across lines may influence the price discount required to achieve cost-effectiveness and overall reimbursement decisions. The results of the study should be approached with caution due to the uncertainty surrounding accurate model replication.

ID 728: Mandana Zanganeh
Title: An assessment of the construct validity of the CHU-9D in school-aged children: Evidence from a Chinese trial

Background: The CHU-9D is a newly developed paediatric utility measure for application in the economic evaluation of prevention and treatment interventions. Although there is emerging evidence regarding the psychometric properties of the CHU-9D instrument, more evidence is required with respect to its validity for use in different age groups and country settings considering different tariffs. The aim of this study was to examine the construct validity of the CHU-9D instrument in 6-7 year old Chinese children. To facilitate this assessment, the CHU-9D was directly compared to the PedsQL instrument: a validated generic health-related quality of life (HRQoL) measure in children.

Methods: HRQoL and demographic data were examined using data collected from 1539 children recruited as part of the Chinese randomised controlled trial (CHIRPY DRAGON). Utility-based HRQoL was measured using the Chinese-translated version of the CHU-9D instrument. General HRQoL was measured using the validated Chinese version of PedsQL instrument. We calculated utility scores using both UK and Chinese tariffs. The relationship between sociodemographic variables (gender, mother/father education level, weight status) and HRQoL, using either instrument, was examined through a series of descriptive and multivariate analysis. We examined the construct validity of the CHU-9D instrument by reporting specifically on the discriminant and convergent validity which was directly compared to the PedsQL instrument.

Preliminary results: The direction of the relationships between each instrument and sociodemographic variables was very similar. However, only CHU-9D was able to discriminate between children by gender (Kruskal–Wallis p =0.003 and p =0.004 using UK and Chinese tariffs respectively). The mean PedsQL total scores decreased mostly linearly with increasing levels of severity on each dimension of the CHU-9D (p = <0.001). The correlation between the CHU-9D utility values (using both UK and Chinese tariffs) and PedsQL total scores showed a statistically significant moderate positive correlation (Rs = 0.5221, p = <0.001) and (Rs = 0.5316, p = <0.001) respectively.

Conclusions: The findings from this study support the construct validity of the CHU-9D within Chinese children aged 6-7 years.
Trial registration number: ISRCTN11867516

ID 731: Laia Maynou
Title: The creep of the robots

Medical technology is one of the leading factors contributing to health care expenditure growth. We consider the sequential uptake and diffusion patterns of three interrelated surgical technologies within the English NHS: robotic, laparoscopic and open prostatectomy. This is an interesting empirical question that allows us to quantify the elasticity of substitution between these surgical technologies, with newer technologies bringing efficiency improvements and initial better patient outcomes. Our descriptive analysis shows that open prostatectomy is substituted by laparoscopic prostatectomy, but that robotic prostatectomy then rapidly substitutes for laparoscopic. In light of the lack of guidance from NICE on which

technology is to be used, it is important not only to quantify the rates of robotic diffusion but also any efficiency changes brought by the use of new technologies. Using data from Hospital Episodes Statistics (HES) for the period 2000-2014, we analyse 206 trusts, who perform prostatectomy, of which 41 undertake robotic surgery. Our empirical analysis first identifies substitution effects across technologies: firstly, laparoscopic procedures to replace open; secondly, robotic procedures to replace laparoscopic. We then quantify robotic diffusion through the potential creep of robotic surgery to other surgical specialties (i.e. rectum and colon, kidney, bladder and uterus procedures). Finally, we analyse the effect on LOS of the adoption of these new technologies. While diffusion and creep are analysed at the provider level, the effects on LOS are analysed at the patient level. Our diff-in-diff results suggest there is a one-day reduction in LOS when comparing robotic to open and laparoscopic prostatectomy, but that there is no significant effect when comparing robotic and laparoscopic alone. Through focusing on the degree of substitution across different forms of surgical input in treating similar cases, the adoption pathway of two new technologies is quantified with precision.

ID 732: Claire Simons
Title: Stratification of optimal treatment allocation by disease severity measures: an application to treatment for sleep apnoea.

Cost-effectiveness analyses show whether an intervention is cost-effective for the population 'on average'. However, cost-effectiveness may differ between population groups, leading to different optimal interventions for subgroups of the population. Exploration of cost-effectiveness within subgroups can lead to more efficient resource allocation. Despite obvious benefits of stratifying treatment, this is not regularly implemented.

This work explores issues in the potential stratification of treatment decisions by a continuous measure of disease severity, applied to a case study of treatment for patients with sleep apnoea. This includes the selection of appropriate data, modelling how treatment effects vary with severity, full quantification of uncertainty, and efficient calculation of the stratified decision and value of further research.

Literature on the use of Mandibular Advancement Devices (MADs) and Continuous Positive Airway Pressure (CPAP) as treatments for sleep apnoea were reviewed. Evidence was obtained on the Epworth Sleepiness Scale (ESS), a measure of disease severity, and the impact of treatment on this score. Both aggregate data (n=39) and accessible individual participant data (n=2) were extracted and combined using Bayesian model-based meta-regression.

The parameter estimates and posterior distributions from the meta-regressions were included in a cost-effectiveness analysis to quantify the benefit of stratifying the treatment decision. Using a regression approximation, the optimal treatment decision can be determined for any value of the stratifying variable without re-running the cost-effectiveness model. Value of Information ideas were used to quantify the economic benefits of stratification and potential research to improve stratified decision-making.

The meta-regressions found that more severe disease, indicated by higher baseline ESS, is related to a higher absolute treatment effect. The strength of this relationship differs between interventions. When incorporating the meta-regression results into the cost-effectiveness analysis, there is some evidence that MAD is the optimal treatment for those with lower ESS values and CPAP for those with higher ESS. In addition, collecting additional information on specific populations using particular study designs was estimated to be beneficial.

Despite limited data, groups of the population who would benefit from alternative interventions have been identified. However, Value of Information analysis found further data would be useful to guide stratification.

ID 733: Francesco Longo
Title: Does Perspective Matter in the Economic Evaluation of Social Care Services? An Application to Vision Rehabilitation Services in England

Objectives: The number of economic evaluations of social care (SC) interventions continues to expand. For interventions with a SC focus, NICE guidance in the UK offers multiple choices as to the cost and outcome perspective for consideration. This might result in methodological diversity in practice across analysts as to the perspective/s taken. This study investigated whether the perspective taken matters. We undertook a cost-effectiveness analysis of in-house vs contracted-out vision rehabilitation (VR) services under a SC and an integrated social and health care (S&HC) perspective. In England, VR services are provided by local authorities through two dominant models, i.e. in-house and contracted-out VR services featuring either no or some integration with the NHS. Regardless of how VR is delivered, it aims to promote users' health and wellbeing, and support users' independence in daily living at home and within the community.

Methods: We collected data on VR users' outcomes and costs, based on a prospective cohort study with a 6-month follow-up. Under the SC perspective, we analysed SC-QALYs obtained using ASCOT, and SC costs. Under the integrated S&HC perspective we undertook two evaluations, (i) using SC-QALYs and (ii) using HC-QALYs obtained using EQ-5D-5L. Both (i) and (ii) used S&HC costs. Incremental outcomes and costs were estimated using regression analysis, accounting for user and LA characteristics as potential sources of confounding. In addition, we used multiple imputation to deal with missing data.

Results: Our findings showed that, from a SC perspective, in-house VR services had a high probability (greater than 80% vs contracted-out VR services) of being cost-effective. Under an integrated S&HC perspective, however, in-house VR services were found to have a reduced probability (less than 45%) of being cost-effective, whether based on SC-QALYs or HC-QALYs. Irrespective of the outcome being considered, results were driven mainly by higher use of hospital services by in-house VR users compared to contracted-out VR users.

Conclusions: Our results have the potential to inform different decision makers, i.e. those in a non-integrated SC system, and those in an integrated S&HC system. Focusing on a narrow SC perspective may introduce inefficiency by ignoring broader impacts.

ID 739: Mona Aghdaee
Title: CART Analysis: A new approach to mapping patient reported outcome measures to multi-attribute utility instruments

Introduction: Patient Reported Outcome Measures (PROMs) are gaining attention as healthcare system funders increasingly seek value-based care. One instrument used to collect PROMs is the Patient-Reported Outcomes Measurement Information System (PROMIS) tool. While PROMIS is used in healthcare systems, its results cannot be used to estimate utilities, making it less relevant for economic evaluations. Mapping PROMIS to a multi-attribute utility instrument (MAUI) enables estimation of utilities. Previous studies have mapped PROMIS to EQ-5D-3L, but not to the new EQ-5D-5L. Furthermore, one approach that has not been explored in the mapping literature to date is classification and regression trees analysis (CART). CART is a non-parametric, machine learning method, based on splitting the data into increasingly homogenous categories.

Objective: To map the PROMIS Global 10 to EQ-5D-5L using CART analysis and traditional methods of mapping, in order to derive utilities from PROMIS, and to compare the accuracy of different methods.

Methods: An online survey was conducted to collect responses to PROMIS Global 10 and EQ 5D 5L from the Australian general population (N=2,032). Data from enrolees of a coordinated care program were combined with survey data to increase the sample size of patients with lower utilities (N=120).

Estimation approaches included: CART analysis, linear regression, Tobit, generalised linear model, quantile regression, censored least absolute deviations(CLAD) and Multinomial logistic regression. Predictive accuracies of models were compared using MAE and MSE based on averaged 10-fold cross-validation.

Results: All the approaches reported similar MAE and MSE. CART was superior in predicting full health, but inferior in predicting utilities<0. The direct approach using quantile regression was considered the best approach based on reporting lowest MAE and MSE in cross-validation (0.097, 0.028) and ability in predicting lower utilities.

Conclusion: This study sheds light on the strengths and weaknesses of applying CART analysis in mapping exercises. In particular, its major advantage is the flexibility in terms of model pre-specification and its intuitive interpretation.
The main challenges of the current study were that the sample was relatively small and may have over-represented the healthy population. Only 0.02% of the respondents in the full sample reported negative utilities.

ID 743: Daliya Kaskirbayeva
Title: The Impact of Kazakhstani Provider Payment Reform on Hospital Outcomes: An Interrupted Time Series Approach

Background: In 2012 Kazakhstan introduced a case-based payment system based on Diagnostic Related Groups (DRGs) to replace fee-for-service (FFS) remuneration for public hospitals. Prior to nationwide implementation, 20 hospitals were piloted for a period of 17 weeks starting in September 2011.

Objective: This paper examines whether the introduction of DRGs was associated with changes in inpatient, day cases, surgery cases for two age groups (adults and children), and average length of stay (ALOS) to capture hospital activity and hospital standardised mortality rate as a proxy for quality of care.

Methods: We used hospital episode records of 2011/2012 and 2012/2013. We used interrupted time series across three levels (regional, city and rural) with matched controls to allow for structural changes in outcomes due to payment reform.

Results: In regional hospitals, the surgery cases for children increased by 1.3 cases per week and hospital standardised mortality rate decreased by 0.06 points during the piloting stage. In city hospitals, average length of stay decreased by 0.12 days during the piloting stage. After the nationwide DRG introduction the inpatient cases increased by 1.3 cases per week, day cases decreased by 0.3 cases per week, while the fall in hospital standardised mortality rate was negligible (0.02 points). The results show no effect in rural hospitals during either period. The study findings suggest that despite the slight reduction in length of stay and increased hospital activity, the quality of care was not affected by the implementation of DRGs. In general, the piloted hospitals adapted during roll-out period faster compared to others.

Conclusion: This is the first study evaluating the shift from FFS to DRGs in the Post Soviet region. Despite some evidence of an association between DRGs and hospital activity and quality of care, the principal goal of reform – an increase in day cases and reduced ALOS – was not observed. Possible reasons for the moderated outcomes of the reform are lack of support for hospitals and general unpreparedness of the health care system for the ad hoc implementation of the new payment system.

ID 745: Philip Clarke
Title: When do we die in the long-run? The relative and absolute survival of politicians over the 20th Century.

As Thomas Piketty has demonstrated, compiling pragmatic measures of income distribution (e.g. top income approach) over long periods of time, and across many countries, provides valuable insights on the long-term evolution of income inequalities. To date there are no comparable statistics available for evaluating the evolution of income-related health inequalities.

The aim of our study is to develop a new approach to the measurement of long-run health inequalities, by comparing the relative and absolute survival rates of those on top incomes with the general population.

To operationalize, we focus on a large cohort of politicians, as they are generally high-income earners, and there is information in many countries that contain the information required to estimate their survival (i.e. date of birth/death and date they attained office). We will estimate relative and absolute survival differences by matching each politician by year at risk, age at risk and sex to the life table of the country the individual represented. We will examine cross country trends in health inequalities and compare them with trends in income inequalities.

Where data are available, we will explore health inequalities within politicians based on their educational attainment.

Our results will be drawn from a database we have compiled on 106,092 politicians from 12 developed countries (Australia, Austria, Canada, Netherlands, Finland, France, Germany, Italy, Ireland, New Zealand, Norway, Switzerland, United Kingdom, United States) that in some cases contains data dating back more than three centuries.

We will conclude with a discussion of limitations and propose an agenda for future research.

ID 746: Michelle Tew
Title: Quality of Life Trajectories in Total Knee Replacement Patients: What can they tell us?

There is growing evidence that patient-reported quality-of-life (QoL) is an important predictor of outcomes such as hospitalisation and mortality. Therefore, understanding patients' QoL trajectories can reveal important information on disease progression and outcomes. The aim of this study was to employ latent class growth analysis (LCGA) to identify specific QoL trajectories following total knee replacement (TKR) and examine patient characteristics of identified trajectories.

Data on all patients who had TKR between January 2006 and December 2011 were extracted from the St. Vincent's Melbourne Arthroplasty Outcomes Registry which captures patient demographics, co-morbidities and patient reported QoL (SF12). Utilities values at baseline (pre-surgery), year 1, 2, 3, 4 and 5 were derived from SF12 responses using published algorithm.

LCGA was conducted using utility scores to categorise patients based on their QoL trajectories over 5 years post-surgery. Assessment to identify the optimum number of classes was considered using goodness of fit measures and posterior probability diagnostics. Multinomial logistic regression was used to determine independent predictors of class membership to understand the patient characteristics that contribute to the heterogeneity in QoL trajectories.

1,553 patients contributed to the analysis. Preliminary analysis indicates the presence of 6 distinct QoL trajectories; no QoL improvement following surgery (18%), low baseline QoL with moderate QoL improvement (30%), slow gradual improvement (9%), significant improvement with sustained effect (19%), significant improvement without sustained effect (18%) and high baseline QoL with moderate improvement (7%). Age and presence of co-morbidities appear to be important predictors of declining trajectories.

Patient-reported outcomes can reveal important unobserved heterogeneity among TKR patients suggesting greater scope for such measures in clinical practice. QoL trajectories also reveal variable gains in QALYs across different trajectory groups. This heterogeneity matters when estimating to cost-effectiveness.

ID 747: Gowokani Chijere Chirwa
Title: The impact of Ghana's National Health Insurance on Psychological Distress

Background and Objective
Poor mental health (psychological distress) is amongst the growing non-communicable diseases in low-and middle-income (LMICs). Despite mental ill-health accounting for an already considerable, growing burden of disease in many LMICs, policy action to confront the challenge has been limited, both at the international and national level. Recently, several LMICs have embarked on the journey towards universal health coverage, by expanding their public health insurance provision, with the ultimate objective of improving population health, in addition to other health system objectives. While mental ill-health interventions typically may not have been specifically covered in the publicly funded benefit package, this raises the question as to whether, and if so, by how much the expansion of public health insurance may have contributed – directly or indirectly – to improved mental health. In this paper, we assess the mental health impact of Ghana's implementation of national health insurance (NHIS) in 2003.

Methods
The study uses the first wave of the 2009/2010 Ghana Social Economic Panel survey, covering 10,007 respondents. We employ an instrumental variable regression and propensity score matching methods, separately and jointly, to estimate the causal impact of health insurance on psychological distress, measured by the Kessler Psychological Distress Scale (K10).

Results
The median K10 score in Ghana is 16 ($p<0.001$), with a minimum of 10 ($p<0.001$) and a maximum of 45 ($p<0.001$). The results from the instrumental variable, without matching, indicate that the K10 score for the insured is 11.6% lower ($p<0.001$) than that of the insured. Likewise, after running an IV regression on the matched sample, we find that the insured have a K10 score which is 10.6% ($p<0.001$) lower than the uninsured. Similarly, the estimates based on the propensity score indicates that the insured have a lower K10 score (-0.023; $p<0.05$). The findings are robust to the various estimation methods.

Conclusion
This study suggests that having health insurance has reduced psychological distress and hence improved mental health, even though mental illness treatment or prevention were at best partially covered by the public health insurance in Ghana.

ID 750: Apostolos Tsiachristas
Title: A retrospective propensity score matched analysis using administrative data of hospital-at-home for older people in Scotland

Objectives: To compare the characteristics of populations admitted to hospital-at-home services with the population admitted to hospital and assess the association of these services with healthcare costs and mortality.

Design: In a retrospective observational cohort study of linked patient level data, we used propensity score matching in combination with regression analysis.

Participants: Patients aged 65 years and older admitted to hospital-at-home or hospital.

Interventions: Three geriatrician-led admission avoidance hospital-at-home services in Scotland.

Outcome measures: Healthcare costs and mortality.

Results: Patients in hospital-at-home were older and more socioeconomically disadvantaged, had higher rates of previous hospitalization, and there was a greater proportion of women and people with several chronic conditions compared with the population admitted to hospital. The cost of providing hospital-at-home varied between the three sites from £628 to £2928 per admission. Hospital-at-home was associated with 18% lower costs during the follow-up period in site one (ratio of means 0.82; 95%CI: 0.76-0.89). Limiting the analysis to costs during the 6 months following index discharge, patients in the hospital-at-home cohorts had 27% higher costs (ratio of means 1.27; 95%CI: 1.14-1.41) in site one, 9% (ratio of means 1.09; 95%CI: 0.95-1.24) in site two and 70% in site three (ratio of means 1.70; 95%CI: 1.40-2.07) compared with patients in the control cohorts. Admission to hospital-at-home was associated with an increased risk of death during the follow-up period in all three sites (1.09, 95%CI: 1.00-1.19 site one; 1.29, 95%CI: 1.15-1.44 site two; 1.27, 95%CI: 1.06-1.54 site three).

Conclusions: Our findings indicate that in these three cohorts, the populations admitted to hospital-at-home and hospital differ. We cannot rule out the risk of residual confounding, as our analysis relied on an administrative data set and we lacked data on disease severity and type of hospitalised care received in the control cohorts.

ID 754: Chris Bojke
Title: Plus ca Change: Why, when and how we should routinely accommodate patient heterogeneity in economic evaluation

In economic evaluation there is a common misconception that it is generally okay to ignore patient-level heterogeneity where data are generated by RCTs unless there is a specific aim of identifying subgroups or equity considerations. This work challenges that view. We draw on an existing statistical literature dating back 30+ years to show that the use of any non-linear models to extrapolate beyond RCT data generates biased estimates when heterogeneity is present but ignored. The problem may be particularly acute for survival analysis where two forms of omitted bias may be present, including negative dependence duration. We use simulation and real evaluation data to identify the extent of the problem and the scenarios in which bias is most likely to become an issue – predictably they are related to the extent of heterogeneity, the nature of the individual survival functions and the degree of censoring. We identify the common statistical solutions and examine the limitations which may prevent their application in economic evaluation. We finish this exploratory work by proposing a population-weighted cohort solution which avoids the need for individual patient-level simulation models.

We conclude that there is an increased role for exploring heterogeneity in economic evaluation even when the evaluation is based on RCT data. Our proposed solutions allow us to avoid any bias that occurs in using non-linear models, especially survival models, in extrapolating outcomes beyond the periods observed in the trial. The methods are also naturally amenable to addressing the common internal and external validity issue where the economic model may need to cover a wider population than that contained in the trial and/or explain variation between RCT and Real World evidence. As such we argue that there is a case to be made for including the exploration of the degree of heterogeneity as part of any HEAP where extrapolation is required even if only to demonstrate that the conditions for bias are not present.

ID 755: Robert Heggie
Title: Should health economists care about implementation?

Background: In assessing complex interventions, in both clinical and population health context, it is critical to account for implementation. Current Medical Research Council (MRC) guidance has highlighted four phases for the assessment of complex interventions in a “cyclical sequence”: development, feasibility/piloting, evaluation and implementation. We argue that implementation should not be

considered as a distinct phase, but as one of the interacting components within a single framework incorporating implementation and economic evaluation throughout of the assessment.

Methods: A systematic review of National Institute for Health Research (NIHR) monographs over the last five years, to examine how implementation has been assessed and evaluated within HTAs. We will describe and critique the approaches that have been used. In particular, we will focus on how implementation has been incorporated within the economic evaluation.

Results: Our preliminary results showed that there is little consistency in how implementation has been incorporated into economic evaluations. The issues of economic evaluation and implementation are typically considered in isolation – with implementation factors only considered after the economic evaluation has taken place. Several studies estimated the value of implementation, focusing on the trade-off between directing resources towards further research or implementation activities, but not on the value of considering implementation early within the economic evaluation. Other studies have presented qualitative data alongside quantitative results, without being explicitly integrated within the economic evaluation or without an interpretation of these mixed method results.

Conclusion/Discussion: The failure to consider implementation factors within the economic evaluation of complex interventions may lead to suboptimal allocation decisions. Ideally, economic evaluation and implementation should be considered throughout the assessment of complex interventions – i.e. not only to evaluate, but also to inform the development and feasibility of an intervention. Where full implementation of interventions has non-trivial budget implications, the inclusion of implementation factors within the economic evaluation can allow for the prioritisation of subgroups with the greatest potential to benefit from treatment. Further research is necessary to recommend which type of data will be required to inform economic evaluation and how this can best be captured.

ID 756: Andrea Gabrio
Title: Adjusting for partially observed baseline utilities and costs in trial-based cost-effectiveness analysis: a comparison of different methods and their performance

In trial-based cost-effectiveness analysis, individual utility and cost values at each time point in the study are typically derived using self-reported questionnaires (e.g. EQ-5D for the utilities) in combination with some external evidence (e.g. EQ-5D tariffs). These variables are then used to calculate health (e.g. QALYs) and total cost measures over the duration of the trial. The economic analysis is performed on these aggregated outcomes, often using linear regression methods to control for the potential imbalance in baseline utilities and costs between treatment groups, with estimates that are evaluated at the mean of the baseline variables.

When some individuals in the trial have their baseline utility and cost values missing, restricting the analysis only to the complete cases is inefficient and can result in biased estimates. In addition, mean baseline adjustment can be implemented by evaluating the mean of the baseline variables on either the complete or all observed utilities/costs, possibly leading to different estimates and cost-effectiveness conclusions. While different naïve methods exist in the literature to adjust for partially observed baseline variables in randomised trials, they usually focus on the case where missingness is confined to the covariates and are therefore not adequate in cost-effectiveness analysis, where individuals with missing baseline utility and cost values will also have their outcome values missing.

Using data from two randomised studies and a small simulation exercise, this paper compares the relative performance of alternative methods to deal with partially observed baseline utilities and costs in cost-effectiveness analysis. Attention is paid to the performance of the methods when the data are not missing at random. The results suggest that standard regression methods fitted to the complete cases are potentially dangerous and can mislead the decision-making process. Alternative approaches such as

multiple imputation or Bayesian methods, which impute the missing values and take into account the variability between imputations under a range of assumptions about missingness, should be preferred. Analysts should therefore consider methods that can explicitly incorporate missing data assumptions and assess the robustness of the results to a range of plausible alternatives.

ID 757: Markus Haacker
Title: Assessing Effectiveness and Cost-Effectiveness of HIV Prevention Interventions: Interactions with Assumptions on Treatment Coverage

In the analysis of effectiveness and cost-effectiveness of alternative HIV policies, and contributions of specific interventions, treatment coverage is frequently considered a policy variable. Treatment coverage, however, is endogenous, reflecting past and current HIV incidence and transitions to treatment. In the analysis of HIV prevention interventions, or of policies with a strong HIV prevention component, this raises the possibility of biased results on effectiveness and cost-effectiveness, if the induced decline in HIV incidence – which directly affects treatment coverage through the denominator – is complemented "under the hood" by an implausible reduction in transition to treatment to keep assumed treatment coverage unchanged. The paper assesses the relevance and magnitude of such bias.

The practice assessed in the paper includes the dominant model used in global HIV policy analysis (Spectrum/Goals), underlying the UNAIDS global policy targets and deployed widely in providing policy advice to countries on HIV investment and spending allocations.

We built a dynamic model of HIV disease progression, featuring an early and late stage of the disease, transition to treatment, and mortality depending on stage of disease and treatment status. HIV incidence depends on a force-of-infection parameter and on treatment coverage. The model was parameterized to match output across Spectrum files for 56 countries with HIV prevalence above 0.5 percent, supplied by UNAIDS, and assessed for different levels of treatment coverage.

We find substantial bias in results on effectiveness and cost-effectiveness of HIV prevention interventions. For a one-off HIV prevention intervention, the practice described underestimates the impact on HIV infections (between 3% for low and 9% for high treatment coverage), AIDS-related deaths (30%-55%), and overstates resulting savings in treatment costs (between 115% and 30%), over a time horizon of 15 years. For a lasting decline in HIV transmission, the impact is underestimated for HIV infections (5%-10%) and AIDS deaths (61%-80%), and overestimated for resulting savings in treatment costs (between 56% and 222%).

Results from this family of models, with regard to cost-effectiveness of alternative policies/interventions or projected resource needs, are thus unreliable and biased unless they explicitly account for the implications of reduced incidence for treatment coverage targets.

ID 759: Jon Sussex
Title: Measuring public preferences between health and social care funding options

Background and objectives:
Additional funding will be needed to meet the growing demand for health and social care in the UK. What is the most acceptable way to raise it? Options range from taxation to mandatory insurance, voluntary insurance and user charges. We sought to analyse the preferences of the UK general public.

Methods:
A quantitative survey embedded with a DCE was undertaken with a representative sample of 2,756 members of the public in England, Northern Ireland, Scotland and Wales, recruited from a survey panel. The DCE was designed on the basis of detailed background research including focus groups and cognitive interviews. The survey also collected information on respondents', age, health state, experience of health and social care, income, employment status and education; and tested respondents' knowledge and

awareness of NHS and social care funding levels and sources. From the DCE data we developed models to understand the influence that differences in attributes had on the propensity to choose funding mechanisms.

Results:
From the scaled nested logit model results, all sections of the public – across age groups, income groups, employment status, health status and countries of the UK – would like additional funding for social care to be raised in the same way as additional NHS funding. Specifically, the public prefer a collective rather than individualistic approach to raising additional funds; and preferably a progressive system. All age groups prefer that contributions should not differ by age per se. Raising additional funds should for preference be by a public, not a private, organisation. There is support for earmarking the funds raised to only be used for health care or social care. Preferences are very similar across the four UK countries, once age and socioeconomic characteristics are controlled for.

Conclusions:
This research provides novel evidence to help policy makers understand the relative public acceptability of different options for raising additional funds for health and social care in the UK. We find uniformity of preferences across the UK countries and across sub-groups of the population.

ID 760: Katie Spencer
Title: Outcome based pricing of stereotactic radiotherapy for bone metastases

Background
Palliative radiotherapy (cEBRT) is the standard of care for localised pain due to metastatic cancer. Treatment is not associated with improved survival and the median survival of patients treated in a routine setting approximately 6 months. High dose stereotactic radiotherapy (SABR) may deliver improvements in quality and durability of pain relief, although at significantly increased cost. Costs show a learning curve effect and whilst initial trial data may suggest cost-effectiveness this may not be sustained when treatment diffuses into a population with poorer prognosis. This study aims to assess whether SABR can be cost-effective in the NHS. Beyond this it estimates an outcome based price (OBP) to support the use of routinely available survival data in commissioning cost-effective care.

Methods
A Markov model is used to assess the cost-effectiveness of SABR. This model is extended to incorporate a threshold sensitivity analysis for survival probability beyond the expectation of the observed population. Using probabilistic sensitivity analyses for these varying cohorts an outcome based SABR price is determined at varying willingness to pay thresholds. The observed 30 day mortality in routine UK practice is determined using data from the National radiotherapy dataset.

Results
Base-case cost-effectiveness outcomes for SABR are presented in comparison to cEBRT. An OBP for SABR at varying levels of 30 day mortality is presented graphically with levels of uncertainty. The results of the OBP analysis are considered in the context of the variation in early mortality observed using the routine population level data.

Conclusions
We demonstrate a possible role for outcome based pricing in the implementation of a novel radiotherapy technique where variable treatment costs are observed over time and diffusion into practice may result in treatment no longer being cost-effective. We discuss the challenges of this approach resulting from uncertainty in cost-effectiveness and clinical outcomes, in addition to the inherent challenges of modelling cost-effectiveness in a population with limited survival. Using existing national data sources this method offers promise as a means to guide and monitor implementation of a novel technique, helping to ensure

patients have access to potentially beneficial treatment whilst ensuring the intervention remains cost-effective.

ID 762: Kathleen Boyd
Title: Health Economic Analysis & Modelling Plans: To HEAP or not to HEAP? Why should we bother?

The last decade has seen a rising number of published Health Economic Analysis and Modelling Plans (HEAPs), as distinct entities from overall study protocols. Recent research in this area has indicated an appetite for a guideline on developing HEAPs for economic evaluations of Randomised Controlled Trials (RCTs). While work is ongoing to find a clear consensus on who the audience of a HEAP is and what goes into one, there are numerous methodological issues and broader concerns relating to the development and design of economic analysis plans which to date have been overlooked. For instance, the current focus of a HEAPs value is that it supports the evaluation alongside an RCT. Given the increasing prominence of economic analyses using observational studies and natural experiments, we argue there is a much wider remit for HEAPs. We also advocate a place for modelling within a HEAP, in the form of decision analytic modelling as well as conceptual modelling and qualitative research to inform the design of and compliment the evaluation.

A systematic review of published health economics analysis and modelling plans from 2008-2019 was conducted in Medline, Embase, Web of Science and Cochrane library databases. A content analysis of the HEAPs was undertaken to examine the type of studies, type and form of economic evaluation, disease or health care area, and inclusion of content. The findings from this review are summarised and discussed narratively. This white paper has been developed to (i) summarise the current use and examine the content of HEAPs in the health care setting (ii) outline the multiple purposes of and value in developing such a document (iii) highlight and address outstanding methodological issues that have been overlooked, and (iv) suggest a means to deal with these issues using case study examples from research projects. We propose that these four aspects need to be set out and discussed prior to establishing a standardised guideline for developing a HEAP. This will provide clarity and advice for the health technology assessment community and our clinical and multidisciplinary collaborators on the use and appropriateness of HEAPs.

ID 765: Xuemin Zhu
Title: The stability of physicians' risk attitudes

Background
Risk attitudes are known to influence physician's decision-making including the use of laboratory services and admission of patients. Research on risk attitudes of physicians is therefore important, for example, to assist a better understand practice variations. However, little is known about the stability of risk attitude in physicians. Standard economics assumes that an individual's risk attitude is stable over time. This assumption has been challenged in recent years. The aim of this paper is to explore the stability of risk attitude in physicians over time and to examine the impact of negative personal life events.

Methods
This paper uses data from Medicine in Australia: Balancing Employment and Life (MABEL), a prospective panel study of Australian doctors. Waves 6 to 9 (2013-2016) included risk attitude questions in the financial, career and professional, and clinical domains. The survey also collects information on a range of personal life events such as serious personal injury or illness, death of family member and being named as defendant in a medical negligence claim. Data are available for over 5000 physicians. We examine the mean level changes in risk attitude between wave 6 and 9.

Results
Preliminary results show that the unconditional mean level changes in risk attitude are very small. The stability of risk attitude is similar across genders and type of physician but varies by age. Controlling for

systematic panel attrition, we find that risk attitude is not sensitive to negative personal life events. Full results will be available at the conference.

Conclusion
The results of this study show that risk attitudes are relatively stable among physicians. This may be due to the relatively short time frame examined (3 years). It suggests that risk attitudes can be assumed to be stable when examining the relationship between risk attitude and clinical decision making.

ID 767: Afschin Gandjour
Title: A novel standard-gamble approach for calculating QALYs

Background/aim: Quality-adjusted life years (QALYs) combine life years and strength of preference for different health states. Strength of preference is measured on a scale that is anchored by 0 representing death and 1.0 representing perfect health. Strength of preference can be measured directly using standard-gamble (SG), time-trade-off, or visual analogue scale methods. Yet, only the SG method is able to incorporate risk. Still, SG validity hinges upon the assumption of risk neutrality with respect to life years. The purpose of this study is to develop a SG approach that is able to circumvent this assumption.

Methods: A SG procedure is developed that obeys the continuity axiom of the von Neumann Morgenstern utility theory and thus establishes cardinal utility but without the need to impose the assumption of risk neutrality. After defining the upper and lower bound of the utility scale, a search of the medical literature is conducted to describe the bounds more precisely.

Results: The upper bound of the scale (1.0) represents absence of negative thoughts about current health. While a zero utility score is still assigned to dead individuals for the purpose of calculating QALYs, the dead state itself is not used for dividing health states into those better or worse than death. Instead, the new SG method exploits the fact that a zero utility score can also be assigned to people who are alive but have a preference for death. Specifically, the new SG method defines the zero point as a state of suicidal ideation (wish) due to the health state in question. The zero point is more concretely described in terms of duration, frequency, and strength of suicidal ideation. Given that payoffs of the gamble need to be defined from a short-term perspective due to the inherent acuteness of suicidal ideation, the assumption of risk neutrality with respect to life years is avoided. Disease states worse than death must not be formally measured but can be included in the zero point as the real bottom of the scale, thus providing another advantage over the conventional SG approach.

Conclusion: The SG method developed in this study avoids the assumption of risk neutrality with respect to life years and therefore allows for a more valid assessment of preferences for health states under risk. The exclusive focus of the method on (subjective) mental states is discussed.

ID 768: Iryna Schlackow
Title: Model of cardiovascular disease in people with reduced kidney function developed using routine healthcare data

Long-term disease models are used widely to evaluate the net benefits and cost-effectiveness of preventive interventions. Increasingly such models are required to present results relevant for distinct categories of patients in an effort to inform decisions for individual patients. At the same time, routine healthcare data is becoming more available and potentially an important source to inform long-term cost-effectiveness models. In this paper we illustrate the use of large UK routine primary healthcare data (Clinical Practice Research Datalink (CPRD)) linked with secondary healthcare (Hospital Episodes Statistics) and mortality registry data to derive a large longitudinal cohort of patients with reduced kidney function and use the data to develop and validate a long-term cardiovascular disease model in this population.

We developed a Markov state-transition model using information on patients' baseline characteristics, including sociodemographic and clinical factors, renal function and cardiovascular events over time to project risks of myocardial infarction, stroke, hospital admission for heart failure and vascular death. A series of gender-specific survival risk equations were fitted were estimated in the cohort; external data informed annual healthcare costs and quality of life related to disease states in the model.

From the 5.5 million people registered with primary care practices in CPRD in January 2005, 1.1million people met the reduced kidney function inclusion criteria and included in the study cohort (45% men, 19% with prior CVD at entry, 74% with only mildly decreased eGFR (60-90mL/min/1.73m2)). 50% of participants were observed in the cohort for at least 4.9years and 25% for at least 7.8years. Age, eGFR, and CVD events were key determinants of evolving cardiovascular disease in this population. The model-projected CVD risks corresponded well to those observed risks in categories of patients and over the follow up period.

In this paper we discuss the advantages and disadvantages of using routine healthcare data rather than prospective studies' data for the development of cost-effectiveness disease models including the selection of target patient population, selectivity of data with respect to disease conditions, testing and other diagnostic data. We demonstrate the use of the model to inform management strategies for this population in primary care.

This research was funded by the National Institute for Health Research (NIHR) Programme Grants for Applied Research programme (Reference: RP-PG-1210-12003) and supported by the NIHR Biomedical Research Centre, Oxford. The views expressed are those of the authors and not necessarily those of the NHS, the NIHR or the Department of Health and Social Care.

ID 769: Farasat Bokhari
Title: Promotions in Alcohol Sales: Lessons from the Scottish Experiment

Motivated by health concerns, Scotland imposed a ban on front of the aisle and multi-buy promotion of alcohol in retail stores in October 2011. Using take-home purchasers scan records of a representative sample of households living in Scotland, England and Wales collected before and after the ban, we evaluate the effect of the policy. We find that there are differential effects of the ban which vary by segment of alcohol as well as by type of consumers. First, retailers appear to have responded by replacing multi-buys with simple price reduction promotions which have lowered real average price per unit in some segments. Second, those classified as low or moderate purchasers either had no impact on purchase patterns or in some segments reduced consumption slightly (for instance in Spirits segment which did not have multi-buy promotions but were affected by front of aisle ban). On the other hand, those classified as heave purchasers and responsible for almost three quarter of purchases by volume among these three groups, had an overall increase in consumption of alcohol as well as in segments for beers and wines which were significantly affected by multi-buy bans. Third, the increase in consumption among heavy purchasers persists even after we account for any simultaneous price changes. Finally, we also find that number of shopping trips per week also increased among heave purchasers. We tentatively conclude that non-linear pricing was restraining heavy purchasers but once the ban came into effect, uniform pricing removed that constraint and both the frequency and amount of alcohol purchased increased.

ID 770: Andrew Sylvester
Title: Does providing Same Day Emergency Care reduce time spent in A&E?

Background
A&E operational performance is a key NHS priority, with waiting time standards enshrined in the NHS Constitution and used to manage trusts. Same Day Emergency Care (SDEC) is a transformational change in care delivery to manage more emergency patients same-day and avoid an overnight admission. There are clinical and operational benefits to reducing patients' time spent in hospital, with SDEC receiving policy backing in the NHS Long Term Plan. In addition to its impact on occupied beds, SDEC has the potential to

reduce time spent in A&E both for patients receiving SDEC and all other patients. Existing literature evaluating A&E performance is trust-level and does not estimate the impact of SDEC, which requires patient-level analysis.

Method
We linked the 16 million Type 1 A&E attendances in the Hospital Episode Statistics dataset for November 2017 to October 2018 with patients' hospital admission and outpatient records. Using a list of SDEC-amenable diagnoses, we identified patients in the linked dataset who received SDEC on their care pathway, and other amenable patients who did not receive it. The patient-level analysis also allows us to include new variables of interest, such as the length of the queue waiting in A&E at patient arrival. We also include established trust-level factors from the literature, such as bed occupancy. We then model the relationship between each factor and patient-level time in A&E.

Results
Patients who receive SDEC spend a shorter time in A&E, and a greater trust-level proportion of admissions treated in SDEC is associated with shorter times spent in A&E for all admitted patients. We also confirm the strong relationship between bed occupancy and A&E performance, finding a significant and accelerating effect on patient waiting time as occupancy increases. Finally, we find that the A&E queue has a significant effect on patients' time in A&E.

Conclusion
This is the first study to model A&E operational performance at the patient-level to estimate the impact of SDEC. The positive effects found will support ongoing national work on SDEC, and our findings on bed occupancy and A&E queues will inform broader policy debates about operational performance and resourcing.

ID 772: Jannis Stöckel
Title: The Dynamic Health Effects of Providing Informal Care in the UK

In the United Kingdom (UK) responsibility for non-health related long-term care is predominantly placed with the individual and their family, translating to a large demand for informal care. For caregivers, however, providing informal care can be challenging, potentially leading to mental and physical health problems. This study aims to generate insights into (i) the extent and persistence of this caregiving effect, and (ii) its relation to care-duration, by estimating the causal long-term and dynamic health effects of providing informal care in the UK.

Using a propensity score matching framework we address endogeneity concerns by modelling the caregiving decision based on an extensive set of variables covering the individual's potential care obligations, ability to care, and personality traits. The baseline analysis uses a static matching approach to estimate the caregiving effects for up to five years after the initial caretaking episode. The dynamic impact of caregiving, the marginal effects of providing additional years of care, is estimated by matching individuals on the defined variables at subsequent decision points.

We utilize all 8 waves (2009-2017) of the UK Household Longitudinal Study, a representative panel data survey covering the adult UK population. Among the 27,169 respondents in our sample, 1,905 initial and 3,146 continuing caregivers are identified, of which we can follow 3,351 for five years following care provision.

In the baseline analysis we find small negative mental health effects of caregiving, independent of care intensity, persisting for up to five years. A detailed subgroup analysis indicates a strong asymmetry of these average effects. Especially spousal caregivers experience significant and persistent losses in the mental, and

smaller, less persistent losses in the physical domain, seemingly driven by the high care-intensity among this subgroup. Further, the estimated dynamic effects suggest a small increase of these effects with the duration of care.

Our results provide insights for long-term care policymakers faced with the question what the associated costs are of meeting current and future care demand through informal care. The asymmetry of effects and their scarring nature for a subgroup of high-intensity caregivers suggests the need for targeted support to counteract health losses.

ID 773: Anne-Marie Konopka
Title: Geographical determinants of cervical cancer screening in France

Context: In France, cervical cancer is the 10th leading cause of death for women. There are approximately 3000 new cases and 1100 deaths each year. Cervical cancer screening is mainly based on individual screening performed by a health professional (gynaecologist, general practitioner or midwife) in a physician's office, health centre, family planning clinic or hospital. The French National Health Authority (Haute Autorité de Santé) recommends cervical cancer screening every three years (after two annual normal pap smears) for women between 25 to 65 years of age.

Objective: Since only 60% of French women aged 25 to 65 years old are screened for cervical cancer, the goal of this study is to analyse the drivers of geographical disparities in cervical cancer screening in France. Multilevel modelling will enable us to consider the potential correlation between women living in each department, and to examine the role of contextual characteristics on cervical cancer screening.

Methods: Based on administrative data from the MGEN (a French non-for-profit health insurance organisation providing both national health insurance and complementary health insurance coverage), a multilevel model with random effects was used to estimate the association between cervical cancer screening and individual and contextual variables. The studied sample includes women aged 25 to 65 years old, covered at least by NHI at MGEN, from 1st January 2012 to 31st December 2014.

Results: The multilevel model confirms geographical inequalities in cervical cancer screening. Moreover, it shows that the likelihood of being screened decreases with age, the price of the gynaecological visit and living in a deprived area, and increases with living in a couple, being covered by both NHI and CHI, being followed for contraception or pregnancy, being screened for breast cancer and the local density of health professionals.

Conclusion: Different policy levers will be discussed in relation to the major role played by health professionals in screening access, in regards to their unequal distribution and the existence of potential financial barriers. Indeed, most cervical screenings are performed by gynaecologists, who may be lacking in some municipalities and who can charge extra fees.

ID 775: Bernarda Zamora
Title: Reconciling ACEA and MCDA: Is There a Way Forward?

Background: Health plan coverage and reimbursement decisions for new medical technologies typically include efficiency (cost-effectiveness) criteria. Alongside cost-effectiveness, other criteria are sometimes considered, related to either clinical (e.g., burden of disease, benefit-risk impact) or broader societal (e.g. productivity, equity) considerations. However, methods to aggregate and measure the trade-offs between these different criteria are diverse and usually neither transparent nor replicable. The ISPOR Special Task Force on US value assessment frameworks recommended development of an augmented cost-effectiveness analysis (ACEA) to include additional elements of value subject to further research on their theoretical and

empirical validity. It recommended comparing the resource allocation implications of the use of a weighted cost-per-QALY efficiency criteria as compared with the use of Multi Criteria Decision Analysis (MCDA).

Methods: We explore the idea of what could be called a "QALY-anchored" MCDA and compared with a form of AECA which includes novel elements of value, such as value of knowing, insurance value, option value and value of hope. For economic evaluation, the two key MCDA attributes would be length of life and quality of life incorporated into the expected multi-attribute utility function. This same expected multi-attribute utility function is then used to incorporate new element of value or criteria both under ACEA and MCDA. On the one hand, ACEA includes new elements of value into the utility function, and the trade-offs among them depend on preferences over risk and the probability of disease. On the other hand, MCDA additively aggregates the elements of value or criteria with the trade-offs among them represented by weights which need to be pre-defined.

Results: The comparison of ACEA and MCDA under the expected multi-attribute utility framework allows testing the additivity assumption of different elements of value underlying the MCDA model. This is possible since ACEA includes different elements of value in a non-additively separable utility function (e.g., health gain and physical risk insurance value) from which the trade-offs can be implicitly estimated. Therefore, these estimated trade-offs can be tested as weights in an MCDA-oriented expected multi-attribute utility function to assess comparability and overlapping or double counting potentially caused from MCDA additivity.

ID 776: Koen Pouwels
Title: Estimating the causal effect of time-varying exposures on excess length of hospital stay: Can inverse probability-weighted survival curves improve our estimates?

Background: Length of hospital stay (LoS) is a key determinant of cost-effectiveness for many health care interventions, especially those aimed at reducing hospital-acquired infections (HAIs) and antimicrobial resistance. However, most studies that have estimated excess length of stay from HAIs have ignored their time-dependent nature. An increasingly popular approach to address this challenge is to use multi-state models. Unfortunately, studies using multi-state models often do not adjust for confounders measured at hospital admission, and time-dependent confounding is difficult to incorporate in these models. Here we present a methodology based on inverse probability-weighted survival curves to estimate the total excess LoS due to time-varying exposures such as HAI.

Methods: We used data from two general intensive care units (ICUs) at London teaching hospitals, with complete patient follow-up (3,159 admissions). A conventional Kaplan Meier curve was created on the sample including both patients that acquired bacteraemia and those remaining bacteraemia-free to obtain the daily probabilities of still being in the ICU, irrespective of patient status. Next, we modified the dataset by treating ICU-acquired bacteraemias as censoring events. The daily probability of being censored was calculated, taking into account baseline and time-dependent confounding using pooled logistic regression. Inverse probability-weighted Kaplan Meier curves estimated the daily probabilities of still being present in the ICU in the modified dataset without bacteraemias.

The difference in average LoS between both datasets was multiplied by the number of admitted patients to obtain the total excess LoS due to ICU-acquired bacteraemia. Results were compared with the total excess LoS estimated by combining cumulative incidence functions for ICU mortality and ICU discharge. These functions were obtained by combining cause-specific inverse probability-weighted cox-proportional hazards models.

Result and discussions: Preliminary results show that the method based on inverse-probability weighted Kaplan Meier curves is much quicker and is a less cumbersome approach to address time-varying

confounding. Moreover, confidence intervals can be obtained without relying on bootstrapping, and the approach does not rely on the proportional hazards assumption. Our discussion will describe the extent to which our approach is likely to be generalizable to estimating excess LoS in other health care.

ID 778: David G. Lugo-Palacios
Title: Did new care models strengthen links amongst NHS health care providers?

Background: In 2015, vanguards were established across England as part of the NHS new care models programme aiming to integrate health services to allow better decision making and more sustainable use of resources. By September 2016 there were a total of 50 vanguards classified in five types of models. The integrated primary and acute care systems (PACS) model aims to improve the health and wellbeing of its local population by strengthening the links amongst GPs, hospital, community and mental health services, but to date results are mixed.

Objective: To explore if the characteristics of the health care provider networks in three areas where PACS were implemented changed from before to after their implementation and to investigate the extent to which the variation in network metrics was associated with the rate of ambulatory care sensitive hospital admissions (ACSAs).

Data: English outpatient and inpatient Hospital Episode Statistics (HES) data for the period 2012/13-2017/18.

Methods: For each of the three vanguards studied, health care provider networks were constructed based on patient sharing 1) between GP practices and outpatient departments; 2) amongst outpatient departments within hospitals; 3) amongst outpatient departments between hospitals. Firstly, we compare node centrality, network density and network modularity before and after the vanguard implementation to see whether PACS did in fact strengthen provider interactions as intended. Secondly, we test whether network characteristics and changes in these characteristics are related to variation in ACSA performance.

Results: The analysis is currently in progress. Results will include descriptive network characteristics for participating and non-participating hospitals before and after Vanguards, and estimation results of how network characteristics are related to ACSA performance and how changes in network characteristics are related to changes in ACSA performance.

Conclusion: The new care models programme and specifically the PACS model rely on creating new links and strengthening current links amongst providers of prevention and integrated community care. If the policy was successful in this aim, we expect to see a change in the strength of ties between GP practices and outpatient specialists within and between hospitals, and that the stronger networks are associated with reductions in ACSAs.

ID 779: Vijay Gc
Title: How value of information methods are used in decision making to inform further research within an iterative process of analysis? A systematic review

Background
Decision making in healthcare is based on currently available information, and new information affecting the decision becomes available throughout the lifecycle of all technologies. An iterative framework to economic appraisal of health technologies has been suggested beginning with early indicative studies and progressing towards more rigorous assessment as data become available. The framework incorporates decision analytical modelling, probabilistic and value of information (VOI) analyses to inform the adoption and research priority setting decision on an iterative basis. Economic evaluations are increasingly using VOI methods to inform decision making. However, it is not clear how VOI methods within the iterative process have been applied in real life economic evaluation.

Objective
To identify literature on how VOI methods are used in decision making to inform further research within an iterative process of analysis.

Methods
A systematic literature review was conducted to identify the use of VOI methods in economic evaluation within the iterative framework of decision making. Identified studies were mapped within the iterative framework to identify where such studies fit. A cross reference search was conducted to locate any relevant publication that suggested further research following the VOI analysis.

Results
Various VOI methods have been applied to quantify decision uncertainty and to inform decision-making, research focus and/or trial design by determining optimal sample size, and setting future research priorities. When mapped within the iterative framework, only a few studies followed the steps of the iterative process sequentially.

Conclusion
Although the iterative framework supports a process of information gathering and reducing uncertainty in decision making, the applied examples indicate that the use of this approach in healthcare, in general, is still limited.

ID 780: Philip Kinghorn
Title: ROI, YSC, SCALY, Avoid? Overcoming the Challenges Associated with Conducting Economic Analysis in Social Care, and Identifying opportunities and priorities for future research

Background: In the UK there is both a managed market for social care and state provision of publicly funded social care. Local authorities are struggling to provide social care services across the life course, given an ageing population, increased complexity of need and significantly reduced budgets. Whilst economic evidence is needed, the fact that traditional methods for economic evaluation do not transfer seamlessly across from healthcare may explain the relative neglect of social care by economists.

Methods: Challenges to generating and using economic evidence in the context of social care are discussed; solutions are proposed (including lessons that can be learnt from the evaluation of public health interventions). A range of economic methods are currently available, and these are introduced and critiqued. Priorities are identified for future research.

Results: Challenges to conducting and using economic analysis broadly relate to: Economic methodology aligning to the objectives of social care; Research support and Infrastructure; Lack of data; General complexities associated with the funding and provision of social care; Dissemination of findings. Relevant economic methods are: Return on Investment; Generic and social care specific capability-based measures of well-being; MDCA; and Cost-consequence analysis. Promoting research infrastructure, collaboration and robust study design are priorities for the future (although there are also opportunities to exploit natural experiments).

Discussion: It is important to work with commissioners and service providers to understand their needs and the complexities of social care, whilst promoting rigorous methodologies and study designs that can generate meaningful and useful evidence. There is a potential trade-off between generating rapid evidence, working within the realms of what is feasible, and establishing a lasting evidence base that is generalizable. An increasing emphasis on social care by funding bodies such as NIHR presents opportunities and creates a need for capacity building in this complex and challenging area.

Conclusion: There is a history of research in social care, as well as opportunities to exploit learning from public health. Existing (non-QALY) methods are fit for purpose; the range of options is as much a strength as it is a weakness. Perhaps it is time for economists to embrace the 'different'?

ID 781: José Robles-Zurita
Title: Estimating the causal effect of treatment: using randomisation as an instrumental variable

In the context of randomised clinical trials (RCTs), the simple comparison of intervention vs. control average outcomes is called intention to treat (ITT) effect. This estimator accounts for differences between the planned and actual treatment received by patients. However, this approach does not necessarily estimates the real causal treatment effect (TE) because final allocation of patients to treatment is determined by non-compliance, in addition to randomisation. The causal TE is useful information for medical and patient decisions, and can help understand the mechanism of the effect of the planned intervention. The use of actual treatment (AT) received, as an explanatory variable, would lead to biased estimates of the TE due to plausible endogenous compliance; compliance could be related to patient's characteristics. Interestingly, RCTs give us an opportunity to explore the TE by using randomisation as an instrumental variable (IV) to predict actual treatment received by patients. In this paper, we exploit a data set (6065 patients) from the Short Course Oncology Therapy (SCOT) trial (Iveson, et al. 2018) where late stage colorectal cancer (CRC) patients were allocated to either 3 months or 6 months duration of adjuvant chemotherapy. Despite randomisation, the actual time on treatment (ToT) differed for each patient due to doctor-patient decisions as per standard practice in CRC patient monitoring. All patients were followed up for a minimum of 3 years, up to 8 years post randomisation. Resource use data and EQ5D answers (1832 patients) were collected. In this analysis, we estimate the TE by using three approaches: a) AT, using actual ToT as explanatory variable; b) ITT, planned ToT as explanatory variable; c) IV, instrumented ToT as the independent variable. In a first exploratory analysis, we found that the effect of one additional month of ToT on EQ5D utility (UK scores) is higher and more significant if estimated by IV (-0.008, p-value=0.012) than by IIT (-0.006, p-value=0.014) and by AT (-0.004, p-value=0.047). A similar analysis will study the estimated TE on costs under the three approaches. Interpretation of the three methods and their implications for cost-effectiveness and medical decisions are discussed.

ID 785: James Buchanan
Title: Using 'Big Data' in economic evaluations: Assessing the pitfalls and testing solutions in the 100,000 Genomes Project in England

Background
Routinely collected administrative datasets containing health data are increasingly being used to inform economic evaluations, and could improve research efficiency. However, these datasets pose several challenges for health economists, including concerns over data accuracy and time lags between events and data availability.

The use of big datasets to inform economic evaluations of genomic tests is of growing interest. In the 100,000 Genomes Project in England, whole genome sequencing (WGS) data has been linked to Hospital Episode Statistics data and could be used to evaluate the cost-effectiveness of using WGS for disease diagnosis.

There are, however, challenges with using big datasets in this way. A documented diagnosis may not accurately reflect a patient's disease, and an assigned diagnostic code may not precisely match a documented diagnosis. Analysts may also be unable to differentiate between resource use that would be routine even with an earlier diagnosis, and events associated with an unknown diagnosis. This could lead to inaccurate estimates of the costs of diagnostic journeys, impacting estimates of cost-effectiveness.

This paper describes these challenges, tests solutions using data for patients with rare diseases, and calculates mean secondary care costs per patient by disorder.
Methods
Our analyses use data on patients with rare diseases and their families from the 100,000 Genomes Project. We focus on a subgroup of 544 patients with renal disease, intellectual disability, immunodeficiency and epileptic encephalopathy. We will review the clinical pathways of 40 of these patients to develop a set of analytical rules to determine when a diagnosis has been made, and whether healthcare resource use before/after such a diagnosis is routine or attributable to late diagnosis. These rules will then be applied to a larger sample, and secondary care costs for these patients will be summarised.

Results and Discussion
Initial results indicate that WGS could be used to make a diagnosis 5-7 years earlier in patient pathways, reducing the cost of diagnosis. Our discussion will describe the consequences for healthcare resource allocation of resolving these big data challenges, and consider the generalisability of our analytical rules to other clinical areas and countries.

ID 788: Adam Martin
Title: Determinants of bicycle commuting and the effect of bicycle infrastructure investment in London: evidence from UK census microdata

Worldwide, concern about physical inactivity and excessive car dependence has encouraged ambitious targets and policies to promote cycling. But policy making is hindered by limited knowledge about why the prevalence of – and changes in – cycling varies between geographic areas (e.g. in London 15% in Hackney), and about the role of cycling infrastructure investments in explaining such variation. Individual-level data on 333,740 London commuters in the 2001 and 2011 UK Census were linked to geographic data, including over-time changes in cycling infrastructure investment at the Borough level (n=32).

First, regression models showed that older people, ethnic minorities, women and people of lower socioeconomic status were notably less likely to cycle and that, whilst cycle commuting increased in all groups over time, these inequalities in the likelihood of cycling also increased. Concentration curves and indices (CI) also showed that cycling became more pro-rich over time (CI rose from 0.051 to 0.172). The effect of these rising inequalities in cycling on known inequalities in health and accessibility in cities (i.e. being able to access desired destinations) warrants further study.

Second, geographic heterogeneity, in terms of hilliness, greenspace, footpaths and crime levels had a statistically significant impact on the likelihood of cycling in expected directions (after controlling for individual-level characteristics). However, the magnitude of the effects identified in this study were inconsistent with current Department for Transport models that place substantial weight on such geographic factors to predict the future potential for cycling in areas of England where cycling is uncommon.

Third, investment in cycling infrastructure was associated with a small statistically significant increase in commuter cycling, and reduced inequality in cycling. After controlling for over-time changes in population structure, a non-linear first-differences regression model identified a cost per additional commuter cyclist in London boroughs ranging from £2,000 to £20,000. Whereas other studies have examined the impact of small-scale cycle infrastructure policies, none has reported on the cost-effectiveness of area-level investment in cycle infrastructure. More research is necessary to determine whether such return on investment constitutes good value for money.

ID 792: Steven Paling

Title: The costs of faster cancer diagnosis: using patient-level information and costing system (PLICS) data to examine the drivers of cancer pathway costs in the NHS

Background

Increasing the speed of diagnosis is a major focus of the NHS Long Term Plan for improving cancer outcomes, including a new 28-day Faster Diagnosis standard to ensure most patients receive a definitive diagnosis (or ruling out) of cancer within 28 days of referral from a GP or from screening. The NHS is also expanding innovative diagnostic services such as Rapid Diagnostic Centres (RDCs), which aim to reduce diagnosis time through the use of "one stop clinics" where patients have multiple tests on the same day. While these may lead to improvements in outcomes, they may also be associated with increased costs for the NHS, at least in the short-run.

Method

We use 2017/18 PLICS data matched with Hospital Episode Statistics (HES) data to explore the key drivers of cancer pathway costs for patients who are urgently referred for suspected cancer by their GP. PLICS is a novel dataset which provides actual care costs data at the patient level. By combining PLICS and HES, we are able to construct estimates of total pathway costs as well as costs of first appointment, treatment and diagnosis. We then estimate econometric models examining how pathway characteristics, including "one stop clinics", affect pathway and diagnosis costs. We control for provider/individual level characteristics.

Results

We find that the presence of one-stop clinics increases total pathway costs, even after controlling for the number of appointments or the type of tests the patient had along their pathway. We also find that total pathway costs increase for patients who have to wait longer to have their first appointment after their urgent GP referral.

Conclusion

To the best of our knowledge, this is the first study to provide evidence on the short-term cost implications of shorter cancer pathways and one-stop clinics. Our study provides new evidence that can be used by policymakers and providers as they redesign services to achieve the goals of the NHS Long Term Plan. Our work also demonstrates the potential uses and benefits of the new PLICS dataset for wider health economics research.

ID 793: Hannah Forbes

Title: What is the effect of working at scale in general practice on quality of care and payments to general practice?

Introduction

English health policy encourages general practices to work in larger groups serving >30,000 patients. Some of the objectives of working at scale, as this is known, are to reduce costs and improve quality of care. However, in the UK and internationally, the evidence that any particular size or type of primary care organisation provides better care than another is limited. We have identified 210 practices, serving 2.7 million registered patients, that were, in late 2017, working at scale to populations of >30,000 patients for the purposes of delivering general medical services (rather than simply extended services over and above core general practice), sharing strategy and risk. We have also identified 3693 practices, serving 30 million patients, that were working at scale in looser collaborations such as federations. We aimed to examine the effects of working at scale on NHS payments to general practice and quality of care.

Data and Methods

We examined NHS payments to general practice using practice level data from 2017-18. We measured quality of care through number of ambulatory care sensitive condition admissions and accident and

emergency appointments using national Hospital Episodes Statistics (HES). Multivariate regression analysis was conducted comparing practices working at scale, those working in looser collaborations and those not working at scale.

Results
We found that practices that reported they were working at scale had significantly more ambulatory care sensitive admissions compared to those who were not working at scale. Practices who were working in looser collaborations had higher payments per patient compared to practices that were not working as part of a group. Overall practices working at scale did not have lower costs compared to practices who were not working at scale.

Implications
There is mixed evidence that working at scale offers higher quality care and reduces costs.

ID 794: Embarika Mostafa
Title: Monitoring Subnational Inequalities in Child Nutritional Health in Egypt

Over the last three decades, the child mortality rate decreased from 9% to 1.5% in Egypt. Therefore, the emphasis shifted from improving survival prospects to improving the child health. Importantly, the success of health programs should entail an equal access to healthcare services. The previous studies showed inequalities in child health in developing countries is remarkable with urban-rural gap. They argued those countries witness a rapid pace of urbanization accompanied with increasing rates in child malnutrition. Many studies conducted in Egypt claimed that households' heterogeneity, in terms of socioeconomic characteristics, could explain the differentiation in child health in Egypt. However, the basic question of among which region the child malnutrition is more prevalent in Egypt? Thus, this study is primarily directed to possibilities of enhancing child health across different regions in Egypt. The main objective is to explore the degree of inequalities at disaggregated geographical level. Hence, the analysis is conducted for four major administrative regions in Egypt and for 25 governorates. The study uses different methods, mainly Blinder-Oaxaca decomposition approach and Generalized Entropy class. Besides, the study quantifies the degree of socioeconomic inequalities in child health among different regions through concentration curve and concentration index. A comprehensive analysis of regional inequalities at more disaggregated level could provide a better understanding of the consequences of malnutrition has been evolving in Egypt. Hence, the study might provide evidence for policy development aiming to reduce inequalities in child health, improving nutritional health, and child well-being.

ID 795: Leonardo Koeser
Title: Handling measurement error in an economic evaluation of inpatient rehabilitation for people with psychosis based on data from electronic health records

Background: The use of routinely collected data is increasingly popular in health economic evaluations. However, such data often contains measurement errors (MEs), i.e. some of the measured values and their true quantity differ. MEs can potentially bias and/or increase the variance of statistical estimates. Yet, in health economic evaluations, MEs have received relatively little attention compared to other common statistical challenges such as missing data. Our motivating case study to examine this issue further is an analysis based on a database of electronic health records covering secondary psychiatric care in the south of London which allows access to patients' anonymised medical notes. This creates the possibility to reduce ME by reading these notes, but this is only feasible for a subset of patients and, usually, data cannot be derived from unstructured text in an automated fashion without some degree of ME.

Objective: To evaluate the cost-effectiveness of inpatient rehabilitation compared to treatment as usual for people with psychosis while making an informed choice between potential approaches to deal with ME

Methods: We identify two potential types of strategies to handle MEs: (1) Targeting observations for manual coding by using alternatives to simple random sampling; (2) modelling the ME mechanism. We discuss some of the advantages and disadvantages of these strategies and relevant identifiability assumptions. Based on this qualitative assessment, we choose a combination of matching, stratified sampling and multiple imputation approach to reduce the impact of MEs.

Results: The analysis is currently underway.

Conclusion: Missing and mismeasured data differ only by a matter of degree. Therefore, the handling of MEs in economic evaluations warrants attention. Some of the identified statistical strategies are well-known in economic evaluation but typically used for a different purpose, others appear underexplored.

ID 796: Baptiste Leurent
Title: A novel approach for addressing missing not at random data in health economic studies: the reference-based imputation.

Missing data is a recurrent problem in health economic studies, and approaches such as inverse-probability weighting and multiple imputation typically address missingness assuming "missing-at-random" (MAR). However, in these studies the missingness tends to be related to unobserved values, i.e. missing-not-at-random (MNAR); for example, the chances of completing a health survey are related to patient's (unobserved) health status. Sensitivity analysis approaches, such as pattern-mixture models, play a central role in assessing the impact of plausible MNAR mechanisms.

This paper presents a novel sensitivity analysis approach, the reference-based imputation (RBI). RBI's major strength is that it allows framing relevant, accessible missing data assumptions in a qualitative way through the use of reference groups. For example, a plausible MNAR mechanism in drug trials is to assume that participants in experimental arm who drop-out may stop taking the drug, and are expected to have similar outcomes to those in placebo arm (mechanism known as 'jump-to-reference'). We show how RBI can address key challenges faced by health economic studies, including non-monotone missingness and joint estimation of costs and outcomes.

We illustrate RBI in the CoBalT trial (20% drop-out) comparing cognitive behavioural therapy (CBT) with antidepressants for treating resistant depression. Assuming data were MAR, patients receiving CBT had higher mean QALY (0.087, 95% CI 0.033 to 0.141), higher mean cost (£1000; £816 to £1191), resulting in cost per QALY of £11,519 and suggesting CBT is cost-effective at typical NICE's thresholds (£20-£30,000/QALY). Informed by discussions with trial investigators, we considered a more realistic scenario where patients were assumed to maintain treatment benefits until drop-out and then follow the same trajectory of control arm (MNAR mechanism known as 'copy increments in reference'). More conservative MNAR mechanisms such as 'jump-to-reference' (described above), or 'last mean carried forward' (after drop-out, endpoints remain around mean value at last timepoint) were also considered. Overall, study's conclusion that CBT was effective and cost-effective was robust to alternative missing data assumptions, even when MNAR mechanisms were allowed to differ by endpoint or timepoint.

The RBI approach proposed here assumes data are normal, and it is being extended to handle non-normal cost and outcome data.

ID 797: Caitlin Daly
Title: A non-parametric approach for combining evidence on restricted mean progression free and overall survival time in the presence of non-proportional hazards

Background
Evaluating the cost effectiveness of cancer treatments may require synthesising evidence on time to event outcomes such as progression free survival (PFS) and overall survival (OS). The traditional approach of pooling hazard ratios from randomised controlled trials requires an often implausible assumption of proportional hazards. When this assumption is not reasonable, pooling the outcomes through a single parametric model may serve as an alternative approach; however, survival curves are unlikely to have the same shape across all trials. Our objective was to develop a novel non-parametric approach for synthesising evidence from Kaplan-Meier curves to provide inputs to economic models without requiring any parametric assumptions.

Methods
We developed a method to pool treatment effects as differences in restricted mean survival time (RMST), i.e., the mean survival time accrued from randomisation up to T years. RMST was estimated as the area under the survival curves (AUCs) for PFS and OS. The correlation between the AUCs of PFS and OS within trials was estimated using non-parametric bootstrap sampling techniques. To account for the relationship between the PFS and OS outcomes, we jointly modelled their corresponding AUCs, but the synthesis model was given to PFS and post-progression survival (PPS). This method is motivated by a network of trials evaluating three treatments for Stage IIIA-N2 Non-Small Cell Lung Cancer.

Discussion
This method allowed us to estimate treatment effects on PFS and PPS in a situation where the survival curve shapes were heterogeneous and without needing to assume proportional hazards. The model conformed to the natural constraint that OS is always greater than PFS, and was simple to implement.

Economic models require estimates of the total mean time spent progression free and post-progression. The pooled RMST provides the contribution to the total mean time up until time T. External evidence on survival beyond time T can be combined with the pooled RMST to obtain total mean PFS and PPS time on each treatment. We discuss how the NMA model may be further adapted in order to also incorporate discounting.

ID 799: Paul Peter Schneider
Title: SOCIAL TARIFFS, PREFERENCE HETEROGENEITY, AND COLLECTIVE CHOICE – IS THE EQ-5D TARIFF COMPATIBLE WITH LIBERAL DEMOCRACY?

In economic evaluations of health technologies, health outcomes are commonly measured in terms of QALYs, which are the product of time and health-related quality of life. Health-related quality of life, in turn, is informed by the public's preference over health states. In the UK, EQ-5D is the preferred instrument to collect health information, and the social tariff is the preferred method to translate them into QALYs.

The current social EQ-5D tariff is based on a generalised least squares model, which was fitted to the elicited health state preferences of a representative sample of the UK population: for each health states, it provides the predicted average utility value. In this study, we challenge the conceptual foundations of both, the method and the level of aggregation used in the social tariff. We demonstrate how the tariff ultimately translates individual preferences over health states into societal preferences over health technologies. Thereby, the tariff is – and should be understood as – an instrument for social choice.

The aggregation of individual preferences into a social welfare function is a fundamental problem in social choice and welfare economic theory. However, the available literature suggests that groups of individuals

have no preferences that are independent of the aggregation method: if the aggregation method is changed, the social preference changes too. Moreover, no aggregation method is unequivocally superior, and it thus remains a normative problem to select an adequate method.

In this paper, we discuss the normative implications of the use of the 'average' preference for societal decision making and contrast those with alternative methods, such as the 'median preference' and relative and weighted utilitarian welfare functions. Furthermore, we demonstrate how individual health state preferences could also be translated into individual preferences over health technologies, which could provide yet another method to derive a social choice.

Finally, we explore the theoretical basis for alternative levels of aggregation and how social tariffs on the level of smaller regions (e.g. Scotland), population subgroups (e.g. women), or even preference groups (e.g. groups with unusual health state preferences) could potentially support societal decision making.

ID 803: Charlotte Davies
Title: Competition and Innovation in the UK Hip Implant Market

We analyse competition, turbulence and innovation in the artificial hip implant market in England and Wales, using National Joint Registry data. In the first part, we find a highly concentrated, static market, with two manufacturers having a share of approximately 70% and very little movement of manufacturers or brands over the 12 year period. Viewed through the eyes of an anti-trust (competition) economist, such a structure might be indicative of weak competition. Therefore, in the second part we examine the market's record in innovation – a key indicator of competitive performance. We identify only limited evidence of the emergence of new brands of implants - the smaller manufacturers have had little success in introducing new brands and the main two manufacturers appear to be concentrating their efforts on updates of existing brands. We also find very little evidence that generic brands have been able to penetrate the market. Whilst these findings are not conclusive evidence of an anti-competitive market, they do suggest a rich agenda of competition issues which should be addressed in research.