organizational influences (e.g., size, profit status, location and population served, and most recently, organizational culture).

Methods We used a factorial experimental design in which the subjects, primary care doctors (n=192), viewed clinically authentic vignettes of “patients” presenting with identical signs and symptoms suggesting diabetes. They were stratified according to gender and level of experience. During an in-person interview, they were asked how they would diagnosis and manage the vignette ‘patient’.

Results After controlling for the first two levels of influence (patient and provider), each of which contributed 4.4 and 2% respectively, organizational culture significantly contributed to their behavior, accounting for 14.3% of the variation in clinical decision-making for diabetes. Considering nine different dimensions of practice culture, organizational trust and business emphasis contributed most to the variance in treatment for diabetic foot neuropathy.

Conclusion Attempts to reduce health care variations continues to focus on the levels of patient attributes and physician characteristics (e.g. improved educational efforts). Findings from this experiment suggest a need to further investigate the contribution of organizational factors (third generation) and suggest appropriate interventions at that level. As important however, is the consideration of a **fourth generation**: cognitive aspects of physician decisions. Qualitative methods (i.e. “think aloud”) and newly developing methods to objectively measure unconscious bias (e.g., the Implicit Association Test) can provide more robust assessment of what goes on “inside the doctors’ head,” which eventually produce healthcare disparities.

HSR: General

OP69 THE USE OF NHS ROUTINE DATA TO ANALYSE THE EFFECTS OF HEALTHCARE INTERVENTIONS

doi:10.1136/jech-2012-201753.069

¹E V Cecil, ¹M Soljak, ²D Osipovic, ³S Peckham, ¹A Majeed. ¹Primary Care and Public Health, Imperial College, London, UK; ²Health Services Research and Policy, LSHTM, London, UK

Background In 2007, the strategy ‘Healthcare for London’ planned the establishment of polyclinics in each PCT to help improve the primary care infrastructure. These polyclinics serve as a hub for a group of GP practices which combined form a polysystem. We aimed to investigate the impact of services provided in a polysystem on unplanned (emergency) admissions, using routine NHS activity data. In one polysystem case study, a pulmonary rehabilitation service for COPD was established. In another, a diabetes patient education programme was offered.

Methods We used routine Hospital Episodes Statistics inpatient data from 2004/5 to 2009/10 covering around 1500 London practices. Differences in adjusted trends in emergency admission rates were compared between the polysystem GP practices case studies and other London practices. Multilevel regression models assessed standardized rate ratios, controlling for year variation, GP practice and population factors. We incorporated interaction terms to allow the effects of the polysystem to vary independently by year.

Results Over the study period COPD emergency admission rates fell across London with a yearly admission rate ratio of (95% CI) 0.98 (0.97, 0.99, p<0.001). The rates in the study polysystem did not differ from the rest of London at baseline and there was not enough evidence to suggest that the introduction of the polysystem service in May 2007 had any effect on the admission rates. A total of 12% of the registered COPD population was seen as new contacts in the first year of the service.

Diabetes emergency admission rates have been falling across London over the study period with a yearly admission rate ratio (95% CI) of 0.98 (0.96, 0.99, p=0.001). In the first year of the diabetes

intervention, the rate of emergency admissions for diabetes fell by 80% in patients from the polysystem practices compared with London, with an interaction factor (95% CI) of 0.20 (0.13, 0.31) p<0.001 and this fall was maintained in the following year. Intervention in this London Polysystem covered 70% of the diabetic population. The power to detect an effect was increased due to the coverage.

Conclusion NHS routine data can be used to assess the impact of health service interventions that are aimed at reducing admissions. Commissioners must be aware that to assess the impact of interventions, the implementation needs to be on a large scale and that medium term follow up is required in order to study the trends.

OP70 PLACE OF CANCER DEATHS IN ENGLAND, 2001-2010: TIME TRENDS AND DETERMINANTS

doi:10.1136/jech-2012-201753.070

¹W Gao, ¹Y Ho, ²J Verne, ³M Glickman, ¹IJ Higginson. ¹Palliative Care, Policy and Rehabilitation, King's College London, London, UK; ²South West Public Health Observatory, Bristol, UK; ³Health and Life Events Division, Office for National Statistics, Newport, UK

Background Place of death is an important quality indicator for end of life care. Understanding where patients with cancer die and what factors are associated with place of death is important for improving end of life care and has significant health care cost implications. There is, however, limited up-to-date evidence to guide practice.

Methods This is a whole population-based observational study. All deaths with cancer as the underlying cause of death (ICD-10 codes: C00-C97), which occurred in the period 2001–2010, were extracted from death registrations in England. The time trends in place of death (coded as hospital, home, hospice, nursing home and elsewhere) were described and tested by using the Cochran-Armitage Trend Test. Factors associated with hospital death (age, gender, year of death, cancer site, index of multiple deprivation (IMD), region) were evaluated using the generalised linear model. The clustering effect within region was accounted for using an exchangeable working correlation structure.

Results A total of 1,270, 071 cancer deaths were recorded over the 10 years period, of which 47.4% occurred in hospitals, 24.1% in homes, 17.1% in hospices, 6.6% in nursing homes and 4.8% elsewhere. In the study period, deaths in hospital decreased from 49.7% to 42.2% (annual decrease 0.8%; $P_{\chi^2_{59,6}} < 0.001$), and home deaths increased from 22.3% to 27.6% (annual increase 0.4%; $P_{\chi^2_{48,7}} < 0.0001$). Independent risk factors for dying in hospitals were: being older (Adjusted OR_{75-84 vs 18-54} (AOR): 1.25; 95% confidence interval (95%CI): 1.24–1.26) and living in deprived areas (AORs: 1.06–1.31). Leukaemia (AOR_{ref = lung}: 3.66; 95%CI: 3.65–3.68), Non-Hodgkin's Lymphoma (AOR_{ref = lung}: 2.89; 95%CI: 2.88–2.90) and Bladder cancer (AOR_{ref = lung}: 1.46; 95%CI: 1.45–1.47) were the top three cancer deaths most likely to occur in hospitals. Men had a slightly higher chance than women of dying in hospitals (AOR: 1.13; 95%CI: 1.13–1.14). Cancer deaths in hospitals were more likely to occur in London than in South West (AORs range: 0.84 in South East Coast to 1.27 in London).

Conclusion We found an overall reducing trend in hospital deaths and an increasing trend in home deaths in the past ten years. However, significant inequality in place of death still exists. Future research needs to explore the underlying reasons.

OP71 DISCRIMINATING CLINICAL OUTCOME MODELS MAY DRIFT UNACCEPTABLY: EXAMPLE OF CARDIAC SURGERY MORTALITY

doi:10.1136/jech-2012-201753.071

¹K McAllister, ¹G Hickey, ^{1,2}S Grant, ^{1,2}B Bridgewater, ¹I Buchan. ¹Northwest Institute for Biohealth Informatics, University of Manchester, Manchester, UK; ²Department of Cardiac Surgery, University Hospital of South Manchester, Manchester, UK

Background Clinical prediction models are used for different purposes, but purpose-specific validation is not usually carried out. The ability of a model to discriminate between true positives and false positives has applications in clinical decision making, screening, and service evaluation. The calibration (goodness-of-fit) of a model is a key indicator of how well a model's predicted outcomes reflect those actually observed. Initial validation of models usually includes assessment of these features but re-evaluation over time might not be performed.

EuroSCORE is an adult cardiac surgery risk model which has been in use since 1998. It predicts in-hospital mortality and is used for clinical decision making and service evaluation. It is widely acknowledged to have demonstrated 'calibration drift', but this has not been formally evaluated in the UK population.

Methods We assessed the performance of EuroSCORE in the Central Cardiac Audit Database (CCAD), covering all NHS cardiac procedures in the UK. Discrimination was tested using the area under the Receiver Operator Characteristic (ROC) curve (AUC). Calibration was assessed with the Hosmer-Lemeshow goodness of fit test. In addition, we developed new models with longer-term outcomes using the data, and tested year-on-year model performance.

Results A total of 399,314 eligible procedures from 1st April 1998 to 31st March 2011 were included in the analysis. Assessing the discrimination of EuroSCORE by financial year showed consistency across the period (AUC values ranging from 0.788 to 0.818). Model calibration, however, drifted considerably with a cumulative mortality over-estimate of 10,801 deaths by the end of the period (increasing from 147 over-estimated deaths in 1998 to 1,500 in 2010). This represented a predicted overall mortality rate of 6.0% compared with the observed rate of 3.4%. We will also present findings relating to year-on-year performance of a panel of models tailored to longer-term outcomes in specific procedures.

Conclusion Models that retain accurate discrimination while undergoing calibration drift may be implemented in settings for longer than is appropriate. A model that maintains good discrimination may be useful in a subset of scenarios, but for most purposes good calibration is also crucial. For models developed for multiple applications, purpose-specific validation and recalibration should be considered. Model performance should be appraised in context and not by indicators in isolation.

OP72 PROSPECTIVE EXTERNAL VALIDATION OF RISK PREDICTION MODELS FOR ACUTE TRAUMATIC BRAIN INJURY IN UK CRITICAL CARE UNITS: THE RAIN STUDY

doi:10.1136/jech-2012-201753.072

¹DA Harrison, ¹KA Griggs, ²M Gomes, ³DK Menon, ¹KM Rowan. ¹Clinical Trials Unit, Intensive Care National Audit & Research Centre (ICNARC), London, UK; ²Department of Public Health & Policy, LSHTM, London, UK; ³School of Clinical Medicine, University of Cambridge, Cambridge, UK

Background Acute traumatic brain injury (TBI) is the leading cause of death and disability in adults aged under 40 years. Statistical models have been developed to predict the risk of mortality or unfavourable outcome (death or severe disability) at six months following acute TBI but to date these risk prediction models have only been validated using existing data sources. The Risk Adjustment In Neurocritical care (RAIN) Study aimed to validate these risk prediction models among adults with acute TBI admitted to UK critical care units.

Methods Ten risk prediction models were identified: four for mortality at six months (the Hukkelhoven model and IMPACT Core, Extended and Lab models); and six for unfavourable outcome at six months (as mortality plus CRASH Basic and CT models). Risk factor data were collected from 67 UK critical care units (including 90% of regional neuroscience centres) from August 2009 to March 2011. Patients were followed up to six months for mortality by

linkage with death registration and unfavourable outcome using the Glasgow Outcome Scale (Extended) administered by postal or telephone questionnaire.

The risk prediction models were validated for calibration (c index), discrimination (Hosmer-Lemeshow test and Cox calibration regression) and overall fit (Brier score). Missing data were handled with multiple imputation.

Results Data were collected for 2,975 eligible patients admitted to critical care following acute TBI. 97% of patients were followed-up for mortality and 81% for unfavourable outcome at six months. Following multiple imputation, mortality and unfavourable outcome at six months were 26% and 57%, respectively. Risk prediction models for mortality at six months had good discrimination (c index 0.75–0.78) and the Hukkelhoven and IMPACT Lab models were well calibrated, although the IMPACT Core and Extended models over-predicted mortality. The models for unfavourable outcome at six months had worse discrimination (c index 0.69–0.71) and all models substantially under-predicted risk of unfavourable outcome. The best performance overall was found for the IMPACT Lab model, which was the most complex model, incorporating laboratory measurements. Models of the next level of complexity (Hukkelhoven, CRASH CT, IMPACT Extended) all performed similarly.

Conclusion Risk prediction models for acute TBI had acceptable discrimination among a large, representative sample of patients admitted to UK critical care units. Calibration was good for mortality but poor for unfavourable outcome, and these models therefore require recalibration for use in this setting.

Friday 14 September 2012

Parallel Session D

Research Methods: Surveys and Use of Routine Data

OP73 EXPLORING IMPACTS OF SURVEY NON-RESPONSE USING RECORD-LINKAGE OF SCOTTISH HEALTH SURVEY DATA (2003 TO 2008)

doi:10.1136/jech-2012-201753.073

¹L Gray, ²G McCartney, ³IR White, ⁴L Given, ¹SV Katikireddi, ¹AH Leyland. ¹Social and Public Health Sciences Unit, MRC/CSO, Glasgow, UK; ²Health Scotland, NHS, Glasgow, UK; ³Biostatistics Unit, MRC, Cambridge, UK; ⁴Scottish Centre for Social Research, Nat-Cen, Edinburgh, UK

Background Inference on population health is commonly derived from health survey data, based on the assumption that they are representative of the target communities. Departure from representativeness may weaken external validity leading to biased estimates, with important implications for public health evidence, particularly in relation to prevalence and quantity estimates such as population alcohol consumption. A key aspect determining the extent to which surveys are representative is the level of participant response. Inverse probability weights based on a limited range of demographic variables are usually applied in an attempt to correct for non-response. We aimed to investigate whether weighted estimates of all-cause mortality and mortality from alcohol-related conditions derived from the 2003 Scottish Health Survey (SHeS) – with household response of 67% – reflect those in the population of Scotland.

Methods Baseline observations from SHeS were individually record-linked to mortality data for the 91% of respondents consenting to linkage, and directly age-standardised survey-weighted mortality rates were calculated for the 3117 men and 3980 women aged 20 years and older at interview. Equivalent mortality rates were calculated for the whole of Scotland in the same age group using