EMCA
Gestión de la Calidad Asistencial

Boletín de Artículos Científicos

Agosto 2014
Número 84

1. **ARTÍCULO Nº: 4287**

Wiener RS, Gould MK, Slatore CG, Fincke BG, Schwartz LM, Woloshin S. ***Resource use and guideline concordance in evaluation of pulmonary nodules for cancer: too much and too little care***. JAMA Intern.Med. 2014; 174(6): 871-880.

IMPORTANCE: Pulmonary nodules are common, and more will be found with implementation of lung cancer screening. How potentially malignant pulmonary nodules are evaluated may affect patient outcomes, health care costs, and effectiveness of lung cancer screening programs. Guidelines for evaluating pulmonary nodules for cancer exist, but little is known about how nodules are evaluated in the usual care setting. OBJECTIVE: To characterize nodule evaluation and concordance with guidelines. DESIGN, SETTING, AND PARTICIPANTS: A retrospective cohort study was conducted including detailed review of medical records from pulmonary nodule detection through evaluation completion, cancer diagnosis, or study end (December 31, 2012). The participants included 300 adults with pulmonary nodules from 15 Veterans Affairs hospitals. MAIN OUTCOMES AND MEASURES: Resources used for evaluation at any Veterans Affairs facility and guideline-concordant evaluation served as the main outcomes. RESULTS: Twenty-seven of 300 patients (9.0%) with pulmonary nodules ultimately received a diagnosis of lung cancer: 1 of 57 (1.8%) with a nodule of 4 mm or less, 4 of 134 (3.0%) with a nodule of 5 to 8 mm, and 22 of 109 (20.2%) with a nodule larger than 8 mm. Nodule evaluation entailed 1044 imaging studies, 147 consultations, 76 biopsies, 13 resections, and 21 hospitalizations. Radiographic surveillance (n = 277) lasted a median of 13 months but ranged from less than 0.5 months to 8.5 years. Forty-six patients underwent invasive procedures (range per patient, 1-4): 41.3% (19 patients) did not have cancer and 17.4% (8) experienced complications, including 1 death. Notably, 15 of the 300 (5.0%) received no purposeful evaluation and had no obvious reason for deferral, seemingly "falling through the cracks." Among 197 patients with a nodule detected after release of the Fleischner Society guidelines, 44.7% received care inconsistent with guidelines (17.8% overevaluation, 26.9% underevaluation). In multivariable analyses, the strongest predictor of guideline-inconsistent care was inappropriate radiologist recommendations (overevaluation relative risk, 4.6 [95% CI, 2.3-9.2]; underevaluation, 4.3 [2.7-6.8]). Other systems factors associated with underevaluation included receiving care at more than 1 facility (2.0 [1.5-2.7]) and nodule detection during an inpatient or preoperative visit (1.6 [1.1-2.5]). CONCLUSIONS AND RELEVANCE: Pulmonary nodule evaluation is often inconsistent with guidelines, including cases with no workup and others with prolonged surveillance or unneeded procedures that may cause harm. Systems to improve quality (eg, aligning radiologist recommendations with guidelines and facilitating communication across providers) are needed before lung cancer screening is widely implemented.

2. **ARTÍCULO Nº: 4288**

Tannenbaum C, Martin P, Tamblyn R, Benedetti A, Ahmed S. ***Reduction of inappropriate benzodiazepine prescriptions among older adults through direct patient education: the EMPOWER cluster randomized trial.*** JAMA Intern.Med. 2014; 174(6): 890-898.

IMPORTANCE: The American Board of Internal Medicine Foundation Choosing Wisely Campaign recommends against the use of benzodiazepine drugs for adults 65 years and older. The effect of direct patient education to catalyze collaborative care for reducing inappropriate prescriptions remains unknown. OBJECTIVE: To compare the effect of a direct-to-consumer educational intervention against usual care on benzodiazepine therapy discontinuation in community-dwelling older adults. DESIGN, SETTING, AND PARTICIPANTS: Cluster randomized trial (EMPOWER [Eliminating Medications Through Patient Ownership of End Results] study [2010-2012, 6-month follow-up]). Community pharmacies were randomly allocated to the intervention or control arm in nonstratified, blocked groups of 4. Participants (303 long-term users of benzodiazepine medication aged 65-95 years, recruited from 30 community pharmacies) were screened and enrolled prior to randomization: 15 pharmacies randomized to the educational intervention included 148 participants and 15 pharmacies randomized to the "wait list" control included 155 participants. Participants, physicians, pharmacists, and evaluators were blinded to outcome assessment. INTERVENTIONS: The active arm received a deprescribing patient empowerment intervention describing the risks of benzodiazepine use and a stepwise tapering protocol. The control arm received usual care. MAIN OUTCOMES AND MEASURES: Benzodiazepine therapy discontinuation at 6 months after randomization, ascertained by pharmacy medication renewal profiles. RESULTS: A total of 261 participants (86%) completed the 6-month follow-up. Of the recipients in the intervention group, 62% initiated conversation about benzodiazepine therapy cessation with a physician and/or pharmacist. At 6 months, 27% of the intervention group had discontinued benzodiazepine use compared with 5% of the control group (risk difference, 23% [95% CI, 14%-32%]; intracluster correlation, 0.008; number needed to treat, 4). Dose reduction occurred in an additional 11% (95% CI, 6%-16%). In multivariate subanalyses, age greater than 80 years, sex, duration of use, indication for use, dose, previous attempt to taper, and concomitant polypharmacy (10 drugs or more per day) did not have a significant interaction effect with benzodiazepine therapy discontinuation. CONCLUSIONS AND RELEVANCE: Direct-to-consumer education effectively elicits shared decision making around the overuse of medications that increase the risk of harm in older adults. TRIAL REGISTRATION: clinicaltrials.gov Identifier: NCT01148186.

3. **ARTÍCULO Nº: 4289**

Tosteson AN, Fryback DG, Hammond CS, Hanna LG, Grove MR, Brown M et al. ***Consequences of false-positive screening mammograms.*** JAMA Intern.Med. 2014; 174(6): 954-961.

IMPORTANCE: False-positive mammograms, a common occurrence in breast cancer screening programs, represent a potential screening harm that is currently being evaluated by the US Preventive Services Task Force. OBJECTIVE: To measure the effect of false-positive mammograms on quality of life by measuring personal anxiety, health utility, and attitudes toward future screening. DESIGN, SETTING, AND PARTICIPANTS: The Digital Mammographic Imaging Screening Trial (DMIST) quality-of-life substudy telephone survey was performed shortly after screening and 1 year later at 22 DMIST sites and included randomly selected DMIST participants with positive and negative mammograms. EXPOSURE: Mammogram requiring follow-up testing or referral without a cancer diagnosis. MAIN OUTCOMES AND MEASURES: The 6-question short form of the Spielberger State-Trait Anxiety Inventory state scale (STAI-6) and the EuroQol EQ-5D instrument with US scoring. Attitudes toward

future screening as measured by women's self-report of future intention to undergo mammographic screening and willingness to travel and stay overnight to undergo a hypothetical new type of mammography that would identify as many cancers with half the false-positive results. RESULTS: Among 1450 eligible women invited to participate, 1226 (84.6%) were enrolled, with follow-up interviews obtained in 1028 (83.8%). Anxiety was significantly higher for women with false-positive mammograms (STAI-6, 35.2 vs 32.7), but health utility scores did not differ and there were no significant differences between groups at 1 year. Future screening intentions differed by group (25.7% vs 14.2% more likely in false-positive vs negative groups); willingness to travel and stay overnight did not (9.9% vs 10.5% in false-positive vs negative groups). Future screening intention was significantly increased among women with false-positive mammograms (odds ratio, 2.12; 95% CI, 1.54-2.93), younger age (2.78; 1.5-5.0), and poorer health (1.63; 1.09-2.43). Women's anticipated high-level anxiety regarding future false-positive mammograms was associated with willingness to travel overnight (odds ratio, 1.94; 95% CI, 1.28-2.95). CONCLUSIONS AND RELEVANCE: False-positive mammograms were associated with increased short-term anxiety but not long-term anxiety, and there was no measurable health utility decrement. False-positive mammograms increased women's intention to undergo future breast cancer screening and did not increase their stated willingness to travel to avoid a false-positive result. Our finding of time-limited harm after false-positive screening mammograms is relevant for clinicians who counsel women on mammographic screening and for screening guideline development groups.

4. **ARTÍCULO Nº: 4290**

Chan JC, Sui Y, Oldenburg B, Zhang Y, Chung HH, Goggins W et al. ***Effects of telephone-based peer support in patients with type 2 diabetes mellitus receiving integrated care: a randomized clinical trial***. JAMA Intern.Med. 2014; 174(6): 972-981.

IMPORTANCE: In type 2 diabetes mellitus (T2DM), team management using protocols with regular feedback improves clinical outcomes, although suboptimal self-management and psychological distress remain significant challenges. OBJECTIVE: To investigate if frequent contacts through a telephone-based peer support program (Peer Support, Empowerment, and Remote Communication Linked by Information Technology [PEARL]) would improve cardiometabolic risk and health outcomes by enhancing psychological well-being and self-care in patients receiving integrated care implemented through a web-based multicomponent quality improvement program (JADE [Joint Asia Diabetes Evaluation]). DESIGN, SETTING, AND PARTICIPANTS: Between 2009 and 2010, 628 of 2766 Hong Kong Chinese patients with T2DM from 3 publicly funded hospital-based diabetes centers were randomized to the JADE + PEARL (n = 312) or JADE (n = 316) groups, with comprehensive assessment at 0 and 12 months. INTERVENTIONS: Thirty-three motivated patients with well-controlled T2DM received 32 hours of training (four 8-hour workshops) to become peer supporters, with 10 patients assigned to each. Peer supporters called their peers at least 12 times, guided by a checklist. MAIN OUTCOMES AND MEASURES: Changes in hemoglobin A(1c) (HbA(1c)) level (primary), proportions of patients with attained treatment targets (HbA(1c) <7%; blood pressure <130/80 mm Hg; low-density lipoprotein cholesterol <2.6 mmol/L [to convert to milligrams per deciliter, divide by 0.0256]) (secondary), and other health outcomes at month 12. RESULTS: Both groups had similar baseline characteristics (mean [SD] age, 54.7 [9.3] years; 57% men; disease duration, 9.4 [7.7] years; HbA(1c) level, 8.2% [1.6%]; systolic blood pressure, 136 [19] mm Hg; low-density lipoprotein cholesterol level, 2.89 [0.82] mmol/L; 17.4% cardiovascular-renal complications; and 34.9% insulin treated). After a mean (SD) follow-up period of 414 (55) days, 5 patients had died, 144 had at least 1 hospitalization, and 586 had repeated comprehensive assessments. On intention-to-treat analysis, both groups had similar reductions in HbA(1c) (JADE + PEARL, 0.30% [95% CI, 0.12%-0.47%], vs JADE, 0.29% [95% CI, 0.12%-0.47%] [P = .97])

and improvements in treatment targets and psychological-behavioral measures. In the JADE + PEARL group, 90% of patients maintained contacts with their peer supporters, with a median of 20 calls per patient. Most of the discussion items were related to self-management. CONCLUSIONS AND RELEVANCE: In patients with T2DM receiving integrated care, peer support did not improve cardiometabolic risks or psychological well-being. TRIAL REGISTRATION: clinicaltrials.gov Identifier: NCT00950716.

5. **ARTÍCULO Nº: 4291**

Vallejo-Gutierrez P, Baneres-Amella J, Sierra E, Casal J, Agra Y. ***Lessons learnt from the development of the Patient Safety Incidents Reporting an Learning System for the Spanish National Health System: SiNASP***. Rev.Calid.Asist. 2014; 29(2): 69-77.

OBJECTIVE: To describe the development process and characteristics of a patient safety incidents reporting system to be implemented in the Spanish National Health System, based on the context and the needs of the different stakeholders. DESIGN: Literature review and analysis of most relevant reporting systems, identification of more than 100 stakeholder's (patients, professionals, regional governments representatives) expectations and requirements, analysis of the legal context, consensus of taxonomy, development of the software and pilot test. RESULTS: Patient Safety Events Reporting and Learning system (Sistema de Notificacion y Aprendizajepara la Seguridad del Paciente, SiNASP) is a generic reporting system for all types of incidents related to patient safety, voluntary, confidential, non punitive, anonymous or nominative with anonimization, system oriented, with local analysis of cases and based on the WHO International Classification for Patient Safety. The electronic program has an on-line form for reporting, a software to manage the incidents and improvement plans, and a scoreboard with process indicators to monitor the system. CONCLUSIONS: The reporting system has been designed to respond to the needs and expectations identified by the stakeholders, taking into account the lessons learned from the previous notification systems, the characteristics of the National Health System and the existing legal context. The development process presented and the characteristics of the system provide a comprehensive framework that can be used for future deployments of similar patient safety systems.

6. **ARTÍCULO Nº: 4292**

Bennell KL, Egerton T, Martin J, Abbott JH, Metcalf B, McManus F et al. ***Effect of physical therapy on pain and function in patients with hip osteoarthritis: a randomized clinical trial***. JAMA. 2014; 311 (19): 1987-1997.

IMPORTANCE: There is limited evidence supporting use of physical therapy for hip osteoarthritis. OBJECTIVE: To determine efficacy of physical therapy on pain and physical function in patients with hip osteoarthritis. DESIGN, SETTING, AND PARTICIPANTS: Randomized, placebo-controlled, participant- and assessor-blinded trial involving 102 community volunteers with hip pain levels of 40 or higher on a visual analog scale of 100 mm (range, 0-100 mm; 100 indicates worst pain possible) and hip osteoarthritis confirmed by radiograph. Forty-nine patients in the active group and 53 in the sham group underwent 12 weeks of intervention and 24 weeks of follow-up (May 2010-February 2013) INTERVENTIONS: Participants attended 10 treatment sessions over 12 weeks. Active treatment included education and advice, manual therapy, home exercise, and gait aid if appropriate. Sham treatment included inactive ultrasound and inert gel. For 24 weeks after treatment, the active group continued unsupervised home exercise while the sham group self-applied gel 3 times weekly. MAIN OUTCOMES AND MEASURES: Primary outcomes were average pain (0 mm, no pain; 100 mm, worst pain possible) and physical function (Western Ontario and McMaster Universities Osteoarthritis Index,

0 no difficulty to 68 extreme difficulty) at week 13. Secondary outcomes were these measures at week 36 and impairments, physical performance, global change, psychological status, and quality of life at weeks 13 and 36. RESULTS: Ninety-six patients (94%) completed week 13 measurements and 83 (81%) completed week 36 measurements. The between-group differences for improvements in pain were not significant. For the active group, the baseline mean (SD) visual analog scale score was 58.8 mm (13.3) and the week-13 score was 40.1 mm (24.6); for the sham group, the baseline score was 58.0 mm (11.6) and the week-13 score was 35.2 mm (21.4). The mean difference was 6.9 mm favoring sham treatment (95% CI, -3.9 to 17.7). The function scores were not significantly different between groups. The baseline mean (SD) physical function score for the active group was 32.3 (9.2) and the week-13 score was 27.5 (12.9) units, whereas the baseline score for the sham treatment group was 32.4 (8.4) units and the week-13 score was 26.4 (11.3) units, for a mean difference of 1.4 units favoring sham (95% CI, -3.8 to 6.5) at week 13. There were no between-group differences in secondary outcomes (except greater week-13 improvement in the balance step test in the active group). Nineteen of 46 patients (41%) in the active group reported 26 mild adverse effects and 7 of 49 (14%) in the sham group reported 9 mild adverse events (P = .003). CONCLUSIONS AND RELEVANCE: Among adults with painful hip osteoarthritis, physical therapy did not result in greater improvement in pain or function compared with sham treatment, raising questions about its value for these patients. TRIAL REGISTRATION: anzctr.org.au Identifier: ACTRN12610000439044.

7. **ARTÍCULO Nº: 4293**

Shekelle PG. ***Updating practice guidelines***. JAMA. 2014; 311(20): 2072-2073.

8. **ARTÍCULO Nº: 4294**

Neuman MD, Goldstein JN, Cirullo MA, Schwartz JS. ***Durability of class I American College of Cardiology/American Heart Association clinical practice guideline recommendations***. JAMA. 2014; 311(20): 2092-2100.

IMPORTANCE: Little is known regarding the durability of clinical practice guideline recommendations over time. OBJECTIVE: To characterize variations in the durability of class I ("procedure/treatment should be performed/administered") American College of Cardiology/American Heart Association (ACC/AHA) guideline recommendations. DESIGN, SETTING, AND PARTICIPANTS: Textual analysis by 4 independent reviewers of 11 guidelines published between 1998 and 2007 and revised between 2006 and 2013. MAIN OUTCOMES AND MEASURES: We abstracted all class I recommendations from the first of the 2 most recent versions of each guideline and identified corresponding recommendations in the subsequent version. We classified recommendations replaced by less determinate or contrary recommendations as having been downgraded or reversed; we classified recommendations for which no corresponding item could be identified as having been omitted. We tested for differences in the durability of recommendations according to guideline topic and underlying level of evidence using bivariable hypothesis tests and conditional logistic regression. RESULTS: Of 619 index recommendations, 495 (80.0%; 95% CI, 76.6%-83.1%) were retained in the subsequent guideline version, 57 (9.2%; 95% CI, 7.0%-11.8%) were downgraded or reversed, and 67 (10.8%; 95% CI, 8.4%-13.3%) were omitted. The percentage of recommendations retained varied across guidelines from 15.4% (95% CI, 1.9%-45.4%) to 94.1% (95% CI, 80.3%-99.3%; P < .001). Among recommendations with available information on level of evidence, 90.5% (95% CI, 83.2%-95.3%) of recommendations supported by multiple randomized studies were retained, vs 81.0% (95% CI, 74.8%-86.3%) of recommendations supported by 1 randomized trial or observational data and 73.7% (95% CI, 65.8%-80.5%) of recommendations supported by opinion (P = .001). After accounting for guideline-level factors, the probability of being downgraded, reversed, or omitted was greater for

recommendations based on opinion (odds ratio, 3.14; 95% CI, 1.69-5.85; P < .001) or on 1 trial or observational data (odds ratio, 3.49; 95% CI, 1.45-8.41; P = .005) vs recommendations based on multiple trials. CONCLUSIONS AND RELEVANCE: The durability of class I cardiology guideline recommendations for procedures and treatments promulgated by the ACC/AHA varied across individual guidelines and levels of evidence. Downgrades, reversals, and omissions were most common among recommendations not supported by multiple randomized studies.

9. **ARTÍCULO Nº: 4295**

Oresanya LB, Lyons WL, Finlayson E. ***Preoperative assessment of the older patient: a narrative review.*** JAMA. 2014; 311(20): 2110-2120.

IMPORTANCE: Surgery in older patients often poses risks of death, complications, and functional decline. Prior to surgery, evaluations of health-related priorities, realistic assessments of surgical risks, and individualized optimization strategies are essential. OBJECTIVE: To review surgical decision making for older adult patients by 2 measures: defining treatment goals for elderly patients and reviewing the evidence relating risk factors to adverse outcomes. Assessment and optimization strategies for older surgical patients are proposed. EVIDENCE ACQUISITION: A review of studies relating geriatric conditions such as functional and cognitive impairment, malnutrition, facility residence, and frailty to postoperative mortality and complications (including delirium, discharge to an institution, and functional decline). Medline, EMBASE, and Web of Science databases were searched for articles published between January 1, 2000, and December 31, 2013, that included patients older than 60 years. RESULTS: This review identified 54 studies of older patients; 28 that examined preoperative clinical features associated with mortality (n = 1,422,433 patients) and 26 that examined factors associated with surgical complications (n = 136,083 patients). There was substantial heterogeneity in study methods, measures, and outcomes. The absolute risk and risk ratios relating preoperative clinical conditions to mortality varied widely: 10% to 40% for cognitive impairment (adjusted hazard ratio [HR], 1.26 [95% CI, 1.06-1.49] to 5.77 [95% CI, 1.55-21.55]), 10% to 17% for malnutrition (adjusted odds ratio [OR], 0.88 [95% CI, 0.78-1.01] to 59.2 [95% CI, 3.6-982.9]), and 11% to 41% for institutionalization (adjusted OR, 1.5 [95% CI, 1.02-2.21] to 3.27 [95% CI, 2.81-3.81]).) Risk ratios for functional dependence relating to mortality ranged from an adjusted HR of 1.02 (95% CI, 0.99-1.04) to an adjusted OR of 18.7 (95% CI, 1.6-215.3) and for frailty relating to mortality, ranged from an adjusted HR of 1.10 (95% CI, 1.04-1.16) to an adjusted OR of 11.7 (95% CI not reported) (P < .001). Preoperative cognitive impairment (adjusted OR, 2.2; 95% CI, 1.4-2.7) was associated with postoperative delirium (adjusted OR, 17.0; 95% CI, 1.2-239.8; P < .05). Frailty was associated with a 3- to 13-fold increased risk of discharge to a facility (adjusted OR, 3.16 [95% CI, 1.0-9.99] to 13.02 [95% CI, 5.14-32.98]). CONCLUSIONS AND RELEVANCE: Geriatric conditions may be associated with adverse surgical outcomes. A comprehensive evaluation of treatment goals and communication of realistic risk estimates are essential to guide individualized decision making.

10. **ARTÍCULO Nº: 4296**

Safford MM. ***Comparative effectiveness research and outcomes of diabetes treatment.*** JAMA. 2014; 311(22): 2275-2276.

11. **ARTÍCULO Nº: 4297**

Roumie CL, Greevy RA, Grijalva CG, Hung AM, Liu X, Murff HJ et al. ***Association between intensification of metformin treatment with insulin vs sulfonylureas and cardiovascular events and all-cause mortality among patients with diabetes.*** JAMA. 2014; 311(22): 2288-2296.

IMPORTANCE: Preferred second-line medication for diabetes treatment after metformin failure remains uncertain. OBJECTIVE: To compare time to acute myocardial infarction (AMI), stroke, or death in a cohort of metformin initiators who added insulin or a sulfonylurea. DESIGN, SETTING, AND PARTICIPANTS: Retrospective cohort constructed with national Veterans Health Administration, Medicare, and National Death Index databases. The study population comprised veterans initially treated with metformin from 2001 through 2008 who subsequently added either insulin or sulfonylurea. Propensity score matching on characteristics was performed, matching each participant who added insulin to 5 who added a sulfonylurea. Patients were followed through September 2011 for primary analyses or September 2009 for cause-of-death analyses. MAIN OUTCOMES AND MEASURES: Risk of a composite outcome of AMI, stroke hospitalization, or all-cause death was compared between therapies with marginal structural Cox proportional hazard models adjusting for baseline and time-varying demographics, medications, cholesterol level, hemoglobin A1c level, creatinine level, blood pressure, body mass index, and comorbidities. RESULTS: Among 178,341 metformin monotherapy patients, 2948 added insulin and 39,990 added a sulfonylurea. Propensity score matching yielded 2436 metformin + insulin and 12,180 metformin + sulfonylurea patients. At intensification, patients had received metformin for a median of 14 months (IQR, 5-30), and hemoglobin A1c level was 8.1% (IQR, 7.2%-9.9%). Median follow-up after intensification was 14 months (IQR, 6-29 months). There were 172 vs 634 events for the primary outcome among patients who added insulin vs sulfonylureas, respectively (42.7 vs 32.8 events per 1000 person-years; adjusted hazard ratio [aHR], 1.30; 95% CI, 1.07-1.58; P = .009). Acute myocardial infarction and stroke rates were statistically similar, 41 vs 229 events (10.2 and 11.9 events per 1000 person-years; aHR, 0.88; 95% CI, 0.59-1.30; P = .52), whereas all-cause death rates were 137 vs 444 events, respectively (33.7 and 22.7 events per 1000 person-years; aHR, 1.44; 95% CI, 1.15-1.79; P = .001). There were 54 vs 258 secondary outcomes: AMI, stroke hospitalizations, or cardiovascular deaths (22.8 vs 22.5 events per 1000 person-years; aHR, 0.98; 95% CI, 0.71-1.34; P = .87). CONCLUSIONS AND RELEVANCE: Among patients with diabetes who were receiving metformin, the addition of insulin vs a sulfonylurea was associated with an increased risk of a composite of nonfatal cardiovascular outcomes and all-cause mortality. These findings require further investigation to understand risks associated with insulin use in these patients.

12. **ARTÍCULO Nº: 4298**

Wallia A, Molitch ME. ***Insulin therapy for type 2 diabetes mellitus***. JAMA. 2014; 311(22): 2315-2325.

IMPORTANCE: The incidence and prevalence of type 2 diabetes mellitus are increasing. OBJECTIVE: To review currently available insulin therapy, as well as evidence on the use, application, initiation, and intensification of insulin in the outpatient setting. EVIDENCE REVIEW: Data sources included PubMed for trials and investigations in type 2 diabetes examining insulin use from January 1998 to April 2014. FINDINGS: The hemoglobin A1c target for most patients with type 2 diabetes is 7% but needs to be modified when there is increased risk of hypoglycemia, reduced life expectancy, extensive comorbidities, or reduced resources. Insulin therapy may be considered early or late in the disease course; adverse effects include weight gain and hypoglycemia. Basal insulin can be added to oral hypoglycemic agents (generally stopping sulfonylureas) initially, and later, prandial insulin can be added in a stepwise fashion. Insulin treatment must be individualized, and there are a number of challenges to insulin initiation and intensification. CONCLUSIONS AND RELEVANCE: Insulin can help achieve ideal hemoglobin A1c goals for patients with type 2 diabetes. Barriers such as adherence, patient preferences, clinician preferences, and resource allocation must be addressed.

13. **ARTÍCULO Nº: 4299**

Andrews MA, O'Malley PG. ***Diabetes overtreatment in elderly individuals: risky business in need of better management.*** JAMA. 2014; 311(22): 2326-2327.

14. **ARTÍCULO Nº: 4300**

Noest S, Ludt S, Klingenberg A, Glassen K, Heiss F, Ose D et al. ***Involving patients in detecting quality gaps in a fragmented healthcare system: development of a questionnaire for Patients' Experiences Across Health Care Sectors (PEACS).*** Int.J.Qual.Health Care. 2014; 26(3): 240-249.

OBJECTIVE: The purpose of this study was to develop and validate a generic questionnaire to evaluate experiences and reported outcomes in patients who receive treatment across a range of healthcare sectors. DESIGN: Mixed-methods design including focus groups, pretests and field test. SETTING: The patient questionnaire was developed in the context of a nationwide program in Germany aimed at quality improvements across the healthcare sectors. PARTICIPANTS: For the field test, 589 questionnaires were distributed to patients via 47 general practices. MAIN MEASUREMENTS: Descriptive item analyzes non-responder analysis and factor analysis (PCA). Retest coefficients (r) calculated by correlation of sum scores of PCA factors. Quality gaps were assessed by the proportion of responders choosing a response category defined as indicating shortcomings in quality of care. RESULTS: The conceptual phase showed good content validity. Four hundred and seventy-four patients who received a range of treatment across a range of sectors were included (response rate: 80.5%). Data analysis confirmed the construct, oriented to the patient care journey with a focus on transitions between healthcare sectors. Quality gaps were assessed for the topics 'Indication', including shared-decision-making (6 items, 24.5-62.9%) and 'Discharge and Transition' (10 items; 20.7-48.2%). Retest coefficients ranged from r = 0.671 until r = 0.855 and indicated good reliability. Low ratios of item-non-response (0.8-9.3%) confirmed a high acceptance by patients. CONCLUSIONS: The number of patients with complex healthcare needs is increasing. Initiatives to expand quality assurance across organizational borders and healthcare sectors are therefore urgently needed. A validated questionnaire (called PEACS 1.0) is available to measure patients' experiences across healthcare sectors with a focus on quality improvement.

15. **ARTÍCULO Nº: 4301**

Power M, Fogarty M, Madsen J, Fenton K, Stewart K, Brotherton A et al. ***Learning from the design and development of the NHS Safety Thermometer.*** Int.J.Qual.Health Care. 2014; 26(3): 287-297.

QUALITY ISSUE: Research indicates that 10% of patients are harmed by healthcare but data that can be used in real time to improve safety are not routinely available. INITIAL ASSESSMENT: We identified the need for a prospective safety measurement system that healthcare professionals can use to improve safety locally, regionally and nationally. CHOICE OF SOLUTION: We designed, developed and implemented a national tool, named the NHS Safety Thermometer (NHS ST) with the goal of measuring the prevalence of harm from pressure ulcers, falls, urinary tract infection in patients with catheters and venous thromboembolism on one day each month for all NHS patients. IMPLEMENTATION: The NHS ST survey instrument was developed in a learning collaborative involving 161 organizations (e.g. hospitals and other delivery organizations) using a Plan, Do, Study, Act method. EVALUATION: Testing of operational definitions, technical capability and use were conducted and feedback systems were established by site coordinators in each participating organization. During the 17-month pilot, site coordinators reported a total of 73,651 patient entries. LESSONS LEARNED: It is feasible to obtain national data through standardized reporting by site coordinators at the point of care. Some caution is required in interpreting data and work is required locally to ensure data

collection systems are robust and data collectors were trained. Sampling is an important strategy to optimize efficiency and reduce the burden of measurement.

16. **ARTÍCULO Nº: 4302**

Olomu AB, Stommel M, Holmes-Rovner MM, Prieto AR, Corser WD, Gourineni V et al. ***Is quality improvement sustainable? Findings of the American College of Cardiology's Guidelines Applied in Practice***. Int.J.Qual.Health Care. 2014; 26(3): 215-222.

OBJECTIVE: (i) To examine the sustainability of an in-hospital quality improvement (QI) intervention, the American College of Cardiology's Guideline Applied to Practice (GAP) in acute myocardial infarction (AMI). (ii) To determine the predictors of physician adherence to AMI guidelines-recommended medication prescribing. DESIGN: Prospective observational study. SETTING: Five mid-Michigan community hospitals. PARTICIPANTS: 516 AMI patients admitted consecutively 1 year after the GAP intervention. These patients were compared with 499 post-GAP patients. MAIN OUTCOME MEASURES: The main outcome was adherence to medication use guidelines. Predictors of medication use were determined using multivariable logistic regression analysis. RESULTS: 1 year after GAP implementation, adherence to most medications remained high. We found a significant increase in beta-blocker (BB) use in-hospital (87.9 vs. 72.1%, P < 0.001) whereas cholesterol assessment within 24 h (79.5 vs. 83.6%, P > 0.225) did not change significantly. However, discharge aspirin (83 vs. 90%, P < 0.018) and BB prescriptions (84 vs. 92%, P < 0.016) dropped to preintervention rates. Discharge angiotensin-converting enzyme inhibitor and treatment of patients with low-density lipoprotein of >/= 100 were unchanged. Predictors of receiving appropriate medications were male gender (for aspirin and BBs) and treatment with percutaneous coronary intervention compared with coronary artery bypass graft. Notably, prescription rates for discharge medications differed significantly by hospital. CONCLUSIONS: Early benefits of the Mid-Michigan GAP intervention on guideline use were only partially sustained at 1 year. Differences in guideline adherence by treatment modality and hospital demonstrate challenges for follow-up phases of GAP. Additional strategies to improve sustainability of QI efforts are urgently needed.

17. **ARTÍCULO Nº: 4303**

Renzi C, Asta F, Fusco D, Agabiti N, Davoli M, Perucci CA. ***Does public reporting improve the quality of hospital care for acute myocardial infarction? Results from a regional outcome evaluation program in Italy***. Int.J.Qual.Health Care. 2014; 26(3): 223-230.

OBJECTIVE: To evaluate whether public reporting of performance data was associated with a change over time in quality indicators for acute myocardial infarction (AMI) in Italian hospitals. DESIGN: Pre-post evaluation of AMI indicators in the Lazio region, before and after disclosure of the Regional Outcome Evaluation Program, and a comparative evaluation versus other Italian regions not participating in the program. SETTING/DATA SOURCES: Nationwide Hospital Information System and vital status records. PARTICIPANTS: 24 800 patients treated for AMI in Lazio and 39 350 in the other regions. INTERVENTION: Public reporting of the Regional Outcome Evaluation Program in the Lazio region. MAIN OUTCOME MEASURE: Risk-adjusted indicators for AMI. RESULTS: The proportion of ST-segment elevation myocardial infarction (STEMI) patients treated with percutaneous coronary interventions (PCI) within 48 h in Lazio changed from 31.3 to 48.7%, before and after public reporting, respectively (relative increase 56%; P < 0.001). In the other regions, the proportion increased from 51.5 to 58.4% (relative increase 13%; P < 0.001). Overall 30-day mortality and 30-day mortality for patients treated with PCI did not improve during the study period. The 30-day mortality for STEMI patients not treated with PCI in Lazio was significantly higher in 2009 (29.0%) versus 2006/07 (24.0%)

(P = .002). CONCLUSIONS: Public reporting may have contributed to increasing the proportion of STEMI patients treated with timely PCI. The mortality outcomes should be interpreted with caution. Changes in AMI diagnostic and coding systems should also be considered. Risk-adjusted quality indicators represent a fundamental instrument for monitoring and potentially enhancing quality of care.

18. **ARTÍCULO Nº: 4304**

Scholte M, Neeleman-van der Steen CW, Hendriks EJ, Nijhuis-van der Sanden MW, Braspenning J. ***Evaluating quality indicators for physical therapy in primary care.*** Int.J.Qual.Health Care. 2014; 26(3): 261-270.

OBJECTIVE: To evaluate measurement properties of a set of public quality indicators on physical therapy. DESIGN: An observational study with web-based collected survey data (2009 and 2010). SETTING: Dutch primary care physical therapy practices. PARTICIPANTS: In 3743 physical therapy practices, 11 274 physical therapists reporting on 30 patients each. MAIN OUTCOME MEASURES: Eight quality indicators were constructed: screening and diagnostics (n= 2), setting target aim and subsequent of intervention (n = 2), administrating results (n = 1), global outcome measures (n = 2) and patient's treatment agreement (n = 1). Measurement properties on content and construct validity, reproducibility, floor and ceiling effects and interpretability of the indicators were assessed using comparative statistics and multilevel modeling. RESULTS: Content validity was acceptable. Construct validity (using known group techniques) of two outcome indicators was acceptable; hypotheses on age, gender and chronic vs. acute care were confirmed. For the whole set of indicators reproducibility was approximated by correlation of 2009 and 2010 data and rated moderately positive (Spearman's rho between 0.3 and 0.42 at practice level) and interpretability as acceptable, as distinguishing between patient groups was possible. Ceiling effects were assessed negative as they were high to extremely high (30% for outcome indicator 6-95% for administrating results). CONCLUSION: Weaknesses in data collection should be dealt with to reduce bias and to reduce ceiling effects by randomly extracting data from electronic medical records. More specificity of the indicators seems to be needed, and can be reached by focusing on most prevalent conditions, thus increasing usability of the indicators to improve quality of care.

19. **ARTÍCULO Nº: 4305**

Taha M, Pal A, Mahnken JD, Rigler SK. ***Derivation and validation of a formula to estimate risk for 30-day readmission in medical patients.*** Int.J.Qual.Health Care. 2014; 26(3): 271-277.

OBJECTIVE: To create a simple readmission risk-prediction tool that can be generated easily at the bedside by physicians, nurses, care coordinators and discharge planners. DESIGN: Retrospective cohort study. SETTING: Tertiary academic medical center. PARTICIPANTS: Inpatients aged 18 and older on general internal medicine services. MEASURES: Predictor variables included age, prior hospitalization, high-risk diagnoses, high-risk medications, polypharmacy, depression, use of palliative care and a cumulative score summing these factors (readmission risk score-RRS). The main outcome measure was 30-day readmission. Predictive values were calculated. RESULTS: Readmission increased linearly from 4.9% of those whose RRS score was 0-37.5% of those with highest risk scores (P = 0.0002). We derived a simple formula for readmission risk as 8 and 4% more for each additional readmission risk factor. The positive predictive value for RRS >0 was low, while the negative predictive value for this cutoff was 95%. CONCLUSIONS: An easily calculated 7-point score can be used to estimate readmission risk. This tool may be particularly useful for identifying lower risk patients who may not require intensive intervention, thus aiding in appropriate targeting of resources.

20. **ARTÍCULO Nº: 4306**

Reznek MA, Barton BA. ***Improved incident reporting following the implementation of a standardized emergency department peer review process.*** Int.J.Qual.Health Care. 2014; 26(3): 278-286.

OBJECTIVE: Incident reporting is an important component of health care quality improvement. The objective of this investigation was to evaluate the effectiveness of an emergency department (ED) peer review process in promoting incident reporting. DESIGN: An observational, interrupted time-series analysis of health care provider (HCP) incident reporting to the ED during a 30-month study period prior to and following the peer review process implementation and a survey-based assessment of physician perceptions of the peer review process' educational value and its effectiveness in identifying errors. SETTING: Large, urban, academic ED. PARTICIPANTS AND INTERVENTIONS: HCPs were invited to participate in a standardized, non-punitive, non-anonymous peer review process that involved analysis and structured discussion of incident reports submitted to ED physician leadership. MAIN OUTCOME MEASURES: Monthly frequency of incident reporting by HCPs and physician perceptions of the peer review process. RESULTS: HCPs submitted 314 incident reports to the ED over the study period. Following the intervention, frequency of reporting by HCPs within the hospital increased over time. The frequencies of self-reporting, reporting by other ED practitioners and reporting by non-ED practitioners within the hospital increased compared with a control group of outside HCPs ($P = 0.0019$, $P = 0.0025$ and $P < 0.0001$). Physicians perceived the peer review process to be educational and highly effective in identifying errors. CONCLUSIONS: The implementation of a non-punitive peer review process that provides timely feedback and is perceived as being valuable for error identification and education can lead to increased incident reporting by HCPs.

21. **ARTÍCULO Nº: 4307**

Manias E, Williams A, Liew D, Rixon S, Braaf S, Finch S. ***Effects of patient-, environment- and medication-related factors on high-alert medication incidents.*** Int.J.Qual.Health Care. 2014; 26(3): 308-320.

OBJECTIVE: To measure the rate of medication incidents associated with the prescription and administration of high-alert medications and to identify patient-, environment- and medication-related factors associated with these incidents. DESIGN: A retrospective chart audit design was conducted of medical records for patient admissions from 1 January 2010 to 31 December 2010. SETTING: Five practice settings (cardiac care, emergency care, intensive care, oncology care and perioperative care) at a public teaching hospital in Melbourne, Australia. PARTICIPANTS: Patients were considered for inclusion if they were prescribed at least one high-alert medication and if they were admitted to one of five practice settings. MAIN OUTCOME MEASURES: High-alert prescribing and administering incidents were measured in each of the five practice settings. Generalized linear mixed modeling was used for data analysis. RESULTS: There were 6984 opportunities for high-alert medication incidents across the five clinical settings. The overall medication incident rate was 1934/6984 (27.69%). There were 1176 prescribing incidents (16.84%) and 758 administering incidents (10.85%). Statistical modeling showed that, in each of the five clinical settings, an increased number of ward transfers was associated with increased odds of prescribing incidents. In addition, statistical modeling demonstrated that an increased number of ward transfers was associated with increased odds of administering incidents in emergency care and perioperative care. CONCLUSIONS: Complex relationships were found in managing high-alert medications in specialty clinical settings. Employing measures to address patients' movements across ward settings can reduce high-alert medication incidents and improve quality of care.

22. **ARTÍCULO Nº: 4308**

Braithwaite J, Marks D, Taylor N. ***Harnessing implementation science to improve care quality and patient safety: a systematic review of targeted literature.*** Int.J.Qual.Health Care. 2014; 26(3): 321-329.

BACKGROUND: Getting greater levels of evidence into practice is a key problem for health systems, compounded by the volume of research produced. Implementation science aims to improve the adoption and spread of research evidence. A linked problem is how to enhance quality of care and patient safety based on evidence when care settings are complex adaptive systems. Our research question was: according to the implementation science literature, which common implementation factors are associated with improving the quality and safety of care for patients? METHODS: We conducted a targeted search of key journals to examine implementation science in the quality and safety domain applying PRISMA procedures. Fifty-seven out of 466 references retrieved were considered relevant following the application of exclusion criteria. Included articles were subjected to content analysis. Three reviewers extracted and documented key characteristics of the papers. Grounded theory was used to distil key features of the literature to derive emergent success factors. RESULTS: Eight success factors of implementation emerged: preparing for change, capacity for implementation-people, capacity for implementation-setting, types of implementation, resources, leverage, desirable implementation enabling features, and sustainability. Obstacles in implementation are the mirror image of these: for example, when people fail to prepare, have insufficient capacity for implementation or when the setting is resistant to change, then care quality is at risk, and patient safety can be compromised. CONCLUSIONS: This review of key studies in the quality and safety literature discusses the current state-of-play of implementation science applied to these domains.

23. **ARTÍCULO Nº: 4309**

Hogan H, Healey F, Neale G, Thomson R, Vincent C, Black N. ***Relationship between preventable hospital deaths and other measures of safety: an exploratory study.*** Int.J.Qual.Health Care. 2014; 26(3): 298-307.

OBJECTIVE: To explore associations between the proportion of hospital deaths that are preventable and other measures of safety. DESIGN: Retrospective case record review to provide estimates of preventable death proportions. Simple monotonic correlations using Spearman's rank correlation coefficient to establish the relationship with eight other measures of patient safety. SETTING: Ten English acute hospital trusts. PARTICIPANTS: One thousand patients who died during 2009. RESULTS: The proportion of preventable deaths varied between hospitals (3-8%) but was not statistically significant ($P = 0.94$). Only one of the eight measures of safety (Methicillin-resistant Staphylococcus aureus bacteraemia rate) was clinically and statistically significantly associated with preventable death proportion ($r = 0.73$; $P < 0.02$). There were no significant associations with the other measures including hospital standardized mortality ratios ($r = -0.01$). There was a suggestion that preventable deaths may be more strongly associated with some other measures of outcome than with process or with structure measures. CONCLUSIONS: The exploratory nature of this study inevitably limited its power to provide definitive results. The observed relationships between safety measures suggest that a larger more powerful study is needed to establish the inter-relationship of different measures of safety (structure, process and outcome), in particular the widely used standardized mortality ratios.

24. **ARTÍCULO Nº: 4310**

Bekelis K, Roberts DW, Zhou W, Skinner JS. ***Fragmentation of care and the use of head computed tomography in patients with ischemic stroke.*** Circ.Cardiovasc.Qual.Outcomes. 2014; 7(3): 430-436.

BACKGROUND: Computed tomographic (CT) scans are central diagnostic tests for ischemic stroke. Their inefficient use is a negative quality measure tracked by the Centers for Medicare and Medicaid Services. METHODS AND RESULTS: We performed a retrospective analysis of Medicare fee-for-service claims data for adults admitted for ischemic stroke from 2008 to 2009, with 1-year follow-up. The outcome measures were risk-adjusted rates of high-intensity CT use (>/=4 head CT scans) and risk- and price-adjusted Medicare expenditures in the year after admission. The average number of head CT scans in the year after admission, for the 327 521 study patients, was 1.94, whereas 11.9% had >/=4. Risk-adjusted rates of high-intensity CT use ranged from 4.6% (Napa, CA) to 20.0% (East Long Island, NY). These rates were 2.6% higher for blacks than for whites (95% confidence interval, 2.1%-3.1%), with considerable regional variation. Higher fragmentation of care (number of different doctors seen) was associated with high-intensity CT use. Patients living in the top quintile regions of fragmentation experienced a 5.9% higher rate of high-intensity CT use, with the lowest quintile as reference; the corresponding odds ratio was 1.77 (95% confidence interval, 1.71-1.83). Similarly, 1-year risk- and price-adjusted expenditures exhibited considerable regional variation, ranging from $31 175 (Salem, MA) to $61 895 (McAllen, TX). Regional rates of high-intensity CT scans were positively associated with 1-year expenditures (r=0.56; P<0.01). CONCLUSIONS: Rates of high-intensity CT use for patients with ischemic stroke reflect wide practice patterns across regions and races. Medicare expenditures parallel these disparities. Fragmentation of care is associated with high-intensity CT use.

25. **ARTÍCULO Nº: 4311**

Bufalino V, Bauman MA, Shubrook JH, Balch AJ, Boone C, Vennum K et al. ***Evolution of "the guideline advantage": lessons learned from the front lines of outpatient performance measurement***. Circ.Cardiovasc.Qual.Outcomes. 2014; 7(3): 493-498.

26. **ARTÍCULO Nº: 4312**

Steinberg BA, Kim S, Thomas L, Fonarow GC, Hylek E, Ansell J et al. ***Lack of concordance between empirical scores and physician assessments of stroke and bleeding risk in atrial fibrillation: results from the Outcomes Registry for Better Informed Treatment of Atrial Fibrillation (ORBIT-AF) registry***. Circulation. 2014; 129(20): 2005-2012.

BACKGROUND: Physicians treating patients with atrial fibrillation (AF) must weigh the benefits of anticoagulation in preventing stroke versus the risk of bleeding. Although empirical models have been developed to predict such risks, the degree to which these coincide with clinicians' estimates is unclear. METHODS AND RESULTS: We examined 10 094 AF patients enrolled in the Outcomes Registry for Better Informed Treatment of AF (ORBIT-AF) registry between June 2010 and August 2011. Empirical stroke and bleeding risks were assessed by using the congestive heart failure, hypertension, age >/=75 years, diabetes mellitus, and previous stroke or transient ischemic attack (CHADS2) and Anticoagulation and Risk Factors in Atrial Fibrillation (ATRIA) scores, respectively. Separately, physicians were asked to categorize their patients' stroke and bleeding risks: low risk (<3%); intermediate risk (3%-6%); and high risk (>6%). Overall, 72% (n=7251) in ORBIT-AF had high-risk CHADS2 scores (>/=2). However, only 16% were assessed as high stroke risk by physicians. Although 17% (n=1749) had high ATRIA bleeding risk (score >/=5), only 7% (n=719) were considered so by physicians. The associations between empirical and physician-estimated stroke and bleeding risks were low (weighted Kappa 0.1 and 0.11, respectively). Physicians weighed hypertension, heart failure, and diabetes mellitus less significantly than empirical models in estimating stroke risk; physicians weighted anemia and dialysis less significantly than empirical models when estimating bleeding risks. Anticoagulation use was highest among patients with high stroke risk, assessed by either empirical

model or physician estimates. In contrast, physician and empirical estimates of bleeding had limited impact on treatment choice. CONCLUSIONS: There is little agreement between provider-assessed risk and empirical scores in AF. These differences may explain, in part, the current divergence of anticoagulation treatment decisions from guideline recommendations. CLINICAL TRIAL REGISTRATION URL: http://www.clinicaltrials.gov. Unique identifier: NCT01165710.

27. **ARTÍCULO Nº: 4313**

Hess CN, Peterson ED, Neely ML, Dai D, Hillegass WB, Krucoff MW et al. ***The learning curve for transradial percutaneous coronary intervention among operators in the United States: a study from the National Cardiovascular Data Registry.*** Circulation. 2014; 129(22): 2277-2286.

BACKGROUND: Adoption of transradial percutaneous coronary intervention (TRI) in the United States is low and may be related to challenges learning the technique. We examined the relationships between operator TRI volume and procedural metrics and outcomes. METHODS AND RESULTS: We used CathPCI Registry data from July 2009 to December 2012 to identify new radial operators, defined by an exclusively femoral percutaneous coronary intervention approach for 6 months after their first percutaneous coronary intervention in the database and >/=15 total TRIs thereafter. Primary outcomes of fluoroscopy time, contrast volume, and procedure success were chosen as markers of technical proficiency. Secondary outcomes included in-hospital mortality, bleeding, and vascular complications. Adjusted outcomes were analyzed by using operator TRI experience as a continuous variable with generalized linear mixed models. Among 54 561 TRI procedures performed at 704 sites, 942 operators performed 1 to 10 procedures, 942 operators performed 11 to 50 procedures, 375 operators performed 51 to 100 procedures, and 148 operators performed 101 to 200 procedures. As radial caseload increased, more TRIs were performed in women, in patients presenting with ST-segment elevation myocardial infarction, and for emergency indications. Decreased fluoroscopy time and contrast use were nonlinearly associated with greater operator TRI experience, with faster reductions observed for newer (<30-50 cases) compared with more experienced (>30-50 cases) operators. Procedure success was high, whereas mortality, bleeding, and vascular complications remained low across TRI volumes. CONCLUSIONS: As operator TRI volume increases, higher-risk patients are chosen for TRI. Despite this, operator proficiency improves with greater TRI experience, and safety is maintained. The threshold to overcome the learning curve appears to be approximately 30 to 50 cases.

28. **ARTÍCULO Nº: 4314**

Anderson JL, Heidenreich PA, Barnett PG, Creager MA, Fonarow GC, Gibbons RJ et al. ***ACC/AHA statement on cost/value methodology in clinical practice guidelines and performance measures: a report of the American College of Cardiology/American Heart Association Task Force on Performance Measures and Task Force on Practice Guidelines.*** Circulation. 2014; 129(22): 2329-2345.

29. **ARTÍCULO Nº: 4315**

Hendel RC, Bozkurt B, Fonarow GC, Jacobs JP, Lichtman JH, Smith EE et al. ***ACC/AHA 2013 methodology for developing clinical data standards: a report of the American College of Cardiology/American Heart Association Task Force on Clinical Data Standards.*** Circulation. 2014; 129(22): 2346-2357.

30. **ARTÍCULO Nº: 4316**

Kaboli PJ, Mosher HJ. ***Using balanced metrics and mixed methods to better understand QI interventions.*** BMJ Qual.Saf. 2014; 23(6): 437-439.

31. **ARTÍCULO Nº: 4317**

Feltner C, Jones CD, Cene CW, Zheng ZJ, Sueta CA, Coker-Schwimmer EJ et al. ***Transitional care interventions to prevent readmissions for persons with heart failure: a systematic review and meta-analysis***. Ann.Intern.Med. 2014; 160(11): 774-784.

BACKGROUND: Nearly 25% of patients hospitalized with heart failure (HF) are readmitted within 30 days. PURPOSE: To assess the efficacy, comparative effectiveness, and harms of transitional care interventions to reduce readmission and mortality rates for adults hospitalized with HF. DATA SOURCES: MEDLINE, Cochrane Library, CINAHL, ClinicalTrials.gov, and World Health Organization International Clinical Trials Registry Platform (1 January 1990 to late October 2013). STUDY SELECTION: Two reviewers independently selected randomized, controlled trials published in English reporting a readmission or mortality rate within 6 months of an index hospitalization. DATA EXTRACTION: One reviewer extracted data, and another checked accuracy. Two reviewers assessed risk of bias and graded strength of evidence (SOE). DATA SYNTHESIS: Forty-seven trials were included. Most enrolled adults with moderate to severe HF and a mean age of 70 years. Few trials reported 30-day readmission rates. At 30 days, a high-intensity home-visiting program reduced all-cause readmission and the composite end point (all-cause readmission or death; low SOE). Over 3 to 6 months, home-visiting programs and multidisciplinary heart failure (MDS-HF) clinic interventions reduced all-cause readmission (high SOE). Home-visiting programs reduced HF-specific readmission and the composite end point (moderate SOE). Structured telephone support (STS) interventions reduced HF-specific readmission (high SOE) but not all-cause readmissions (moderate SOE). Home-visiting programs, MDS-HF clinics, and STS interventions produced a mortality benefit. Neither telemonitoring nor primarily educational interventions reduced readmission or mortality rates. LIMITATIONS: Few trials reported 30-day readmission rates. Usual care was heterogeneous and sometimes not adequately described. CONCLUSION: Home-visiting programs and MDS-HF clinics reduced all-cause readmission and mortality; STS reduced HF-specific readmission and mortality. These interventions should receive the greatest consideration by systems or providers seeking to implement transitional care interventions for persons with HF. PRIMARY FUNDING SOURCE: Agency for Healthcare Research and Quality.

32. **ARTÍCULO Nº: 4318**

La RP. ***Basic learning concepts in EBM: the bidimentional hierarchy of evidence***. Evid.Based.Med. 2014; 19(3): 83-84.

33. **ARTÍCULO Nº: 4319**

Simms RA, Yelland A, Ping H, Beringer AJ, Draycott TJ, Fox R. ***Using data and quality monitoring to enhance maternity outcomes: a qualitative study of risk managers' perspectives***. BMJ Qual.Saf. 2014; 23(6): 457-464.

INTRODUCTION: Risk management is a core part of healthcare practice, especially within maternity services, where litigation and societal costs are high. There has been little investigation into the experiences and opinions of those staff directly involved in risk management: lead obstetricians and specialist risk midwives, who are ideally placed to identify how current implementation of risk management strategies can be improved. METHODS: A qualitative study of consultant-led maternity units in an English region. Semistructured interviews were conducted with the obstetric and midwifery risk management leads for each unit. We explored their approach to risk management, particularly

their opinions regarding quality monitoring and related barriers/issues. Interviews were recorded, transcribed and thematically analysed. RESULTS: Twenty-seven staff from 12/15 maternity units participated. Key issues identified included: concern for the accuracy and validity of their local data, potential difficulties related to data collation, the negative impact of external interference by national regulatory bodies on local clinical priorities, the influence of the local culture of the maternity unit on levels of engagement in the risk management process, and scepticism about the value of benchmarking of maternity units without adjustment for population characteristics. CONCLUSIONS: Local maternity risk managers may provide valuable, clinically relevant insights into current issues in clinical data monitoring. Improvements should focus on the accuracy and ease of data collation with a need for an agreed maternity indicators set, populated from validated databases, and not reliant on data collection systems that distract clinicians from patient activity and quality improvement. It is clear that working relationships between risk managers, their own clinical teams and external national bodies require improvement and alignment. Further discussion regarding benchmarking between maternity units is required prior to implementation. These findings are likely to be relevant to other clinical specialties.

34. **ARTÍCULO Nº: 4320**

Boyce MB, Browne JP, Greenhalgh J. ***The experiences of professionals with using information from patient-reported outcome measures to improve the quality of healthcare: a systematic review of qualitative research.*** BMJ Qual.Saf. 2014; 23(6): 508-518.

OBJECTIVES: To synthesise qualitative studies that investigated the experiences of healthcare professionals with using information from patient-reported outcome measures (PROMs) to improve the quality of care. DESIGN: A qualitative systematic review was conducted by searching PubMed, PsycINFO and CINAHL with no time restrictions. Hand searching was also performed. Eligible studies were evaluated using the Critical Appraisal Skills Programme toolkit for qualitative studies. A thematic synthesis identified common themes across studies. Study characteristics were examined to explain differences in findings. SETTING: All healthcare settings. PARTICIPANTS: Healthcare professionals. OUTCOMES: Professionals' views of PROMs after receiving PROMs feedback about individual patients or groups of patients. RESULTS: Sixteen studies met the inclusion criteria. Barriers and facilitators to the use of PROMs emerged within four main themes: collecting and incorporating the data (practical), valuing the data (attitudinal), making sense of the data (methodological) and using the data to make changes to patient care (impact). CONCLUSIONS: Professionals value PROMs when they are useful for the clinical decision-making process. Practical barriers to the routine use of PROMs are prominent when the correct infrastructure is not in place before commencing data collection and when their use is disruptive to normal work routines. Technology can play a greater role in processing the information in the most efficient manner. Improvements to the interpretability of PROMs should increase their use. Attitudes to the use of PROMs may be improved by engaging professionals in the planning stage of the intervention and by ensuring a high level of transparency around the rationale for data collection.

35. **ARTÍCULO Nº: 4321**

Mueller JT, Wright AJ, Fedraw LA, Murad MH, Brown DR, Thompson KM et al. ***Standardizing central line safety: lessons learned for physician leaders.*** Am.J.Med Qual. 2014; 29(3): 191-199.

A comprehensive central venous catheter (CVC) safety program reduces mechanical and infectious complications and requires an integrated multidisciplinary effort. A multistate health care system implemented a discovery and diffusion project addressing CVC insertion, maintenance, and removal.

Process and outcome measures were collected before and after the intervention. The project was completed in 12 months. It was associated with statistically significant improvement in 6 process measures and reduction in the rate of ICU central line-associated bloodstream infection (from 1.16 to 0.80 infections/1000 catheter days; incidence rate ratio = 0.69; 95% confidence interval = 0.51, 0.93). A comprehensive CVC standardization project increased compliance with several established best practices, was associated with improved outcomes, produced a refined definition of discovery and diffusion project components, and identified several discrete leadership principles that can be applied to future clinical improvement initiatives.

36. **ARTÍCULO Nº: 4322**

Mull HJ, Borzecki AM, Chen Q, Shin MH, Rosen AK. ***Using AHRQ patient safety indicators to detect postdischarge adverse events in the Veterans Health Administration***. Am.J.Med Qual. 2014; 29(3): 213-219.

Patient safety indicators (PSIs) use inpatient administrative data to flag cases with potentially preventable adverse events (AEs) attributable to hospital care. This study explored how many AEs the PSIs identified in the 30 days post discharge. PSI software was run on Veterans Health Administration 2003-2007 administrative data for 10 recently validated PSIs. Among PSI-eligible index hospitalizations not flagged with an AE, this study evaluated how many AEs occurred within 1 to 14 and 15 to 30 days post discharge using inpatient and outpatient administrative data. Considering all PSI-eligible index hospitalizations, 11 141 postdischarge AEs were identified, compared with 40 578 inpatient-flagged AEs. More than 60% of postdischarge AEs were detected within 14 days of discharge. The majority of postdischarge AEs were decubitus ulcers and postoperative pulmonary embolisms or deep vein thromboses. Extending PSI algorithms to the postdischarge period may provide a more complete picture of hospital quality. Future work should use chart review to validate postdischarge PSI events.

37. **ARTÍCULO Nº: 4323**

Ryan J, Andrews R, Barry MB, Kang S, Iskandar A, Mehla P et al. ***Preventability of 30-day readmissions for heart failure patients before and after a quality improvement initiative***. Am.J.Med Qual. 2014; 29(3): 220-226.

The objective of this study was to estimate the frequency of heart failure (HF) readmissions that can be prevented through a quality improvement (QI) program. All HF patients at the University of Connecticut Health Center who had a readmission within 30 days of discharge in the year before (2008) and the year after (2011) a QI program were studied. Through chart review, the percentage of patients who had preventable readmissions in each year was estimated. Prior to the QI initiative, chart reviewers identified that 20% to 30% of readmissions were preventable. The decrease in readmissions after the QI program was similar at 28%. Fewer readmissions after the QI initiative were deemed preventable compared with before. In conclusion, this study found a percentage of preventable readmissions similar to the actual 28% reduction in readmissions after a QI program was launched. Preventable readmissions were less common after the QI program was in place.

38. **ARTÍCULO Nº: 4324**

Prentice JC, Davies ML, Pizer SD. ***Which outpatient wait-time measures are related to patient satisfaction?*** Am.J.Med Qual. 2014; 29(3): 227-235.

Long waits for appointments decrease patient satisfaction. Administrative wait-time measures are used by managers, but relationships between these measures and satisfaction have not been studied.

Data from the Veterans Health Administration are used to examine the relationship between wait times and satisfaction. Outcome measures include patient-reported satisfaction and timely appointment access. Capacity and retrospective and prospective time stamp measures are calculated separately for new and returning patients. The time stamp measures consist of the date when the appointment was created in the scheduling system (create date [CD]) or the date the appointment was desired as the start date for wait-time computation. Logistic regression models predict patient satisfaction using these measures. The new-patient capacity, new-patient time stamp measures using CD, and the returning-patient desired-date prospective measure were significantly associated with patient satisfaction. Standard practices can be improved by targeting wait-time measures to patient subpopulations.

39. **ARTÍCULO Nº: 4325**

Guyatt GH, Oxman AD, Schunemann HJ, Tugwell P, Knottnerus A. ***GRADE guidelines: a new series of articles in the Journal of Clinical Epidemiology***. J.Clin.Epidemiol. 2011; 64(4): 380-382.

The "Grades of Recommendation, Assessment, Development, and Evaluation" (GRADE) approach provides guidance for rating quality of evidence and grading strength of recommendations in health care. It has important implications for those summarizing evidence for systematic reviews, health technology assessment, and clinical practice guidelines. GRADE provides a systematic and transparent framework for clarifying questions, determining the outcomes of interest, summarizing the evidence that addresses a question, and moving from the evidence to a recommendation or decision. Wide dissemination and use of the GRADE approach, with endorsement from more than 50 organizations worldwide, many highly influential (http://www.gradeworkinggroup.org/), attests to the importance of this work. This article introduces a 20-part series providing guidance for the use of GRADE methodology that will appear in the Journal of Clinical Epidemiology.

40. **ARTÍCULO Nº: 4326**

Guyatt G, Oxman AD, Akl EA, Kunz R, Vist G, Brozek J et al. ***GRADE guidelines: 1. Introduction-GRADE evidence profiles and summary of findings tables***. J.Clin.Epidemiol. 2011; 64(4): 383-394.

This article is the first of a series providing guidance for use of the Grading of Recommendations Assessment, Development, and Evaluation (GRADE) system of rating quality of evidence and grading strength of recommendations in systematic reviews, health technology assessments (HTAs), and clinical practice guidelines addressing alternative management options. The GRADE process begins with asking an explicit question, including specification of all important outcomes. After the evidence is collected and summarized, GRADE provides explicit criteria for rating the quality of evidence that include study design, risk of bias, imprecision, inconsistency, indirectness, and magnitude of effect. Recommendations are characterized as strong or weak (alternative terms conditional or discretionary) according to the quality of the supporting evidence and the balance between desirable and undesirable consequences of the alternative management options. GRADE suggests summarizing evidence in succinct, transparent, and informative summary of findings tables that show the quality of evidence and the magnitude of relative and absolute effects for each important outcome and/or as evidence profiles that provide, in addition, detailed information about the reason for the quality of evidence rating. Subsequent articles in this series will address GRADE's approach to formulating questions, assessing quality of evidence, and developing recommendations.

41. **ARTÍCULO Nº: 4327**

Guyatt GH, Oxman AD, Kunz R, Atkins D, Brozek J, Vist G et al. ***GRADE guidelines: 2. Framing the question and deciding on important outcomes***. J.Clin.Epidemiol. 2011; 64(4): 395-400.

GRADE requires a clear specification of the relevant setting, population, intervention, and comparator. It also requires specification of all important outcomes--whether evidence from research studies is, or is not, available. For a particular management question, the population, intervention, and outcome should be sufficiently similar across studies that a similar magnitude of effect is plausible. Guideline developers should specify the relative importance of the outcomes before gathering the evidence and again when evidence summaries are complete. In considering the importance of a surrogate outcome, authors should rate the importance of the patient-important outcome for which the surrogate is a substitute and subsequently rate down the quality of evidence for indirectness of outcome.

42. **ARTÍCULO Nº: 4328**

Balshem H, Helfand M, Schunemann HJ, Oxman AD, Kunz R, Brozek J et al. ***GRADE guidelines: 3. Rating the quality of evidence***. J.Clin.Epidemiol. 2011; 64(4): 401-406.

This article introduces the approach of GRADE to rating quality of evidence. GRADE specifies four categories-high, moderate, low, and very low-that are applied to a body of evidence, not to individual studies. In the context of a systematic review, quality reflects our confidence that the estimates of the effect are correct. In the context of recommendations, quality reflects our confidence that the effect estimates are adequate to support a particular recommendation. Randomized trials begin as high-quality evidence, observational studies as low quality. "Quality" as used in GRADE means more than risk of bias and so may also be compromised by imprecision, inconsistency, indirectness of study results, and publication bias. In addition, several factors can increase our confidence in an estimate of effect. GRADE provides a systematic approach for considering and reporting each of these factors. GRADE separates the process of assessing quality of evidence from the process of making recommendations. Judgments about the strength of a recommendation depend on more than just the quality of evidence.

43. **ARTÍCULO Nº: 4329**

Guyatt GH, Oxman AD, Vist G, Kunz R, Brozek J, onso-Coello P et al. ***GRADE guidelines: 4. Rating the quality of evidence--study limitations (risk of bias)***. J.Clin.Epidemiol. 2011; 64(4): 407-415.

In the GRADE approach, randomized trials start as high-quality evidence and observational studies as low-quality evidence, but both can be rated down if most of the relevant evidence comes from studies that suffer from a high risk of bias. Well-established limitations of randomized trials include failure to conceal allocation, failure to blind, loss to follow-up, and failure to appropriately consider the intention-to-treat principle. More recently recognized limitations include stopping early for apparent benefit and selective reporting of outcomes according to the results. Key limitations of observational studies include use of inappropriate controls and failure to adequately adjust for prognostic imbalance. Risk of bias may vary across outcomes (e.g., loss to follow-up may be far less for all-cause mortality than for quality of life), a consideration that many systematic reviews ignore. In deciding whether to rate down for risk of bias--whether for randomized trials or observational studies--authors should not take an approach that averages across studies. Rather, for any individual outcome, when there are some studies with a high risk, and some with a low risk of bias, they should consider including only the studies with a lower risk of bias.

44. **ARTÍCULO Nº: 4330**

Guyatt GH, Oxman AD, Montori V, Vist G, Kunz R, Brozek J et al. ***GRADE guidelines: 5. Rating the quality of evidence--publication bias***. J.Clin.Epidemiol. 2011; 64(12): 1277-1282.

In the GRADE approach, randomized trials start as high-quality evidence and observational studies as low-quality evidence, but both can be rated down if a body of evidence is associated with a high risk of publication bias. Even when individual studies included in best-evidence summaries have a low risk of bias, publication bias can result in substantial overestimates of effect. Authors should suspect publication bias when available evidence comes from a number of small studies, most of which have been commercially funded. A number of approaches based on examination of the pattern of data are available to help assess publication bias. The most popular of these is the funnel plot; all, however, have substantial limitations. Publication bias is likely frequent, and caution in the face of early results, particularly with small sample size and number of events, is warranted.

45. **ARTÍCULO Nº: 4331**

Guyatt GH, Oxman AD, Kunz R, Brozek J, onso-Coello P, Rind D et al. ***GRADE guidelines 6. Rating the quality of evidence--imprecision***. J.Clin.Epidemiol. 2011; 64(12): 1283-1293.

GRADE suggests that examination of 95% confidence intervals (CIs) provides the optimal primary approach to decisions regarding imprecision. For practice guidelines, rating down the quality of evidence (i.e., confidence in estimates of effect) is required if clinical action would differ if the upper versus the lower boundary of the CI represented the truth. An exception to this rule occurs when an effect is large, and consideration of CIs alone suggests a robust effect, but the total sample size is not large and the number of events is small. Under these circumstances, one should consider rating down for imprecision. To inform this decision, one can calculate the number of patients required for an adequately powered individual trial (termed the "optimal information size" [OIS]). For continuous variables, we suggest a similar process, initially considering the upper and lower limits of the CI, and subsequently calculating an OIS. Systematic reviews require a somewhat different approach. If the 95% CI excludes a relative risk (RR) of 1.0, and the total number of events or patients exceeds the OIS criterion, precision is adequate. If the 95% CI includes appreciable benefit or harm (we suggest an RR of under 0.75 or over 1.25 as a rough guide) rating down for imprecision may be appropriate even if OIS criteria are met.

46. **ARTÍCULO Nº: 4332**

Guyatt GH, Oxman AD, Kunz R, Woodcock J, Brozek J, Helfand M et al. ***GRADE guidelines: 7. Rating the quality of evidence--inconsistency***. J.Clin.Epidemiol. 2011; 64(12): 1294-1302.

This article deals with inconsistency of relative (rather than absolute) treatment effects in binary/dichotomous outcomes. A body of evidence is not rated up in quality if studies yield consistent results, but may be rated down in quality if inconsistent. Criteria for evaluating consistency include similarity of point estimates, extent of overlap of confidence intervals, and statistical criteria including tests of heterogeneity and I(2). To explore heterogeneity, systematic review authors should generate and test a small number of a priori hypotheses related to patients, interventions, outcomes, and methodology. When inconsistency is large and unexplained, rating down quality for inconsistency is appropriate, particularly if some studies suggest substantial benefit, and others no effect or harm (rather than only large vs. small effects). Apparent subgroup effects may be spurious. Credibility is increased if subgroup effects are based on a small number of a priori hypotheses with a specified

direction; subgroup comparisons come from within rather than between studies; tests of interaction generate low P-values; and have a biological rationale.

47. **ARTÍCULO Nº: 4333**

Guyatt GH, Oxman AD, Kunz R, Woodcock J, Brozek J, Helfand M et al. ***GRADE guidelines: 8. Rating the quality of evidence--indirectness***. J.Clin.Epidemiol. 2011; 64(12): 1303-1310.

Direct evidence comes from research that directly compares the interventions in which we are interested when applied to the populations in which we are interested and measures outcomes important to patients. Evidence can be indirect in one of four ways. First, patients may differ from those of interest (the term applicability is often used for this form of indirectness). Secondly, the intervention tested may differ from the intervention of interest. Decisions regarding indirectness of patients and interventions depend on an understanding of whether biological or social factors are sufficiently different that one might expect substantial differences in the magnitude of effect. Thirdly, outcomes may differ from those of primary interest-for instance, surrogate outcomes that are not themselves important, but measured in the presumption that changes in the surrogate reflect changes in an outcome important to patients. A fourth type of indirectness, conceptually different from the first three, occurs when clinicians must choose between interventions that have not been tested in head-to-head comparisons. Making comparisons between treatments under these circumstances requires specific statistical methods and will be rated down in quality one or two levels depending on the extent of differences between the patient populations, co-interventions, measurements of the outcome, and the methods of the trials of the candidate interventions.

48. **ARTÍCULO Nº: 4334**

Guyatt GH, Oxman AD, Sultan S, Glasziou P, Akl EA, onso-Coello P et al. ***GRADE guidelines: 9. Rating up the quality of evidence***. J.Clin.Epidemiol. 2011; 64(12): 1311-1316.

The most common reason for rating up the quality of evidence is a large effect. GRADE suggests considering rating up quality of evidence one level when methodologically rigorous observational studies show at least a two-fold reduction or increase in risk, and rating up two levels for at least a five-fold reduction or increase in risk. Systematic review authors and guideline developers may also consider rating up quality of evidence when a dose-response gradient is present, and when all plausible confounders or biases would decrease an apparent treatment effect, or would create a spurious effect when results suggest no effect. Other considerations include the rapidity of the response, the underlying trajectory of the condition, and indirect evidence.

49. **ARTÍCULO Nº: 4335**

Shepperd S, Straus S. ***GRADE: complex decisions***. J.Clin.Epidemiol. 2011; 64(12): 1270-1271.

50. **ARTÍCULO Nº: 4336**

Guyatt GH, Oxman AD, Schunemann HJ. ***GRADE guidelines-an introduction to the 10th-13th articles in the series***. J.Clin.Epidemiol. 2013; 66(2): 121-123.

51. **ARTÍCULO Nº: 4337**

Brunetti M, Shemilt I, Pregno S, Vale L, Oxman AD, Lord J et al. ***GRADE guidelines: 10. Considering resource use and rating the quality of economic evidence***. J.Clin.Epidemiol. 2013; 66(2): 140-150.

OBJECTIVES: In this article, we describe how to include considerations about resource utilization when making recommendations according to the Grading of Recommendations Assessment, Development, and Evaluation (GRADE) approach. STUDY DESIGN AND SETTINGS: We focus on challenges with rating the confidence in effect estimates (quality of evidence) and incorporating resource use into evidence profiles and Summary of Findings (SoF) tables. RESULTS: GRADE recommends that important differences in resource use between alternative management strategies should be included along with other important outcomes in the evidence profile and SoF table. Key steps in considering resources in making recommendations with GRADE are the identification of items of resource use that may differ between alternative management strategies and that are potentially important to decision makers, finding evidence for the differences in resource use, making judgments regarding confidence in effect estimates using the same criteria used for health outcomes, and valuing the resource use in terms of costs for the specific setting for which recommendations are being made. CONCLUSIONS: With our framework, decision makers will have access to concise summaries of recommendations, including ratings of the quality of economic evidence, and better understand the implications for clinical decision making.

52. **ARTÍCULO Nº: 4338**

Guyatt G, Oxman AD, Sultan S, Brozek J, Glasziou P, onso-Coello P et al. ***GRADE guidelines: 11. Making an overall rating of confidence in effect estimates for a single outcome and for all outcomes***. J.Clin.Epidemiol. 2013; 66(2): 151-157.

GRADE requires guideline developers to make an overall rating of confidence in estimates of effect (quality of evidence-high, moderate, low, or very low) for each important or critical outcome. GRADE suggests, for each outcome, the initial separate consideration of five domains of reasons for rating down the confidence in effect estimates, thereby allowing systematic review authors and guideline developers to arrive at an outcome-specific rating of confidence. Although this rating system represents discrete steps on an ordinal scale, it is helpful to view confidence in estimates as a continuum, and the final rating of confidence may differ from that suggested by separate consideration of each domain. An overall rating of confidence in estimates of effect is only relevant in settings when recommendations are being made. In general, it is based on the critical outcome that provides the lowest confidence.

53. **ARTÍCULO Nº: 4339**

Guyatt GH, Oxman AD, Santesso N, Helfand M, Vist G, Kunz R et al. ***GRADE guidelines: 12. Preparing summary of findings tables-binary outcomes***. J.Clin.Epidemiol. 2013; 66(2): 158-172.

Summary of Findings (SoF) tables present, for each of the seven (or fewer) most important outcomes, the following: the number of studies and number of participants; the confidence in effect estimates (quality of evidence); and the best estimates of relative and absolute effects. Potentially challenging choices in preparing SoF table include using direct evidence (which may have very few events) or indirect evidence (from a surrogate) as the best evidence for a treatment effect. If a surrogate is chosen, it must be labeled as substituting for the corresponding patient-important outcome. Another such choice is presenting evidence from low-quality randomized trials or high-quality observational studies. When in doubt, a reasonable approach is to present both sets of evidence; if the two bodies of evidence have similar quality but discrepant results, one would rate down further for inconsistency. For binary outcomes, relative risks (RRs) are the preferred measure of relative effect and, in most instances, are applied to the baseline or control group risks to generate absolute risks. Ideally, the baseline risks come from observational studies including representative patients and identifying easily

measured prognostic factors that define groups at differing risk. In the absence of such studies, relevant randomized trials provide estimates of baseline risk. When confidence intervals (CIs) around the relative effect include no difference, one may simply state in the absolute risk column that results fail to show a difference, omit the point estimate and report only the CIs, or add a comment emphasizing the uncertainty associated with the point estimate.

54. **ARTÍCULO Nº: 4340**

Guyatt GH, Thorlund K, Oxman AD, Walter SD, Patrick D, Furukawa TA et al. ***GRADE guidelines: 13. Preparing summary of findings tables and evidence profiles-continuous outcomes.*** J.Clin.Epidemiol. 2013; 66(2): 173-183.

Presenting continuous outcomes in Summary of Findings tables presents particular challenges to interpretation. When each study uses the same outcome measure, and the units of that measure are intuitively interpretable (e.g., duration of hospitalization, duration of symptoms), presenting differences in means is usually desirable. When the natural units of the outcome measure are not easily interpretable, choosing a threshold to create a binary outcome and presenting relative and absolute effects become a more attractive alternative. When studies use different measures of the same construct, calculating summary measures requires converting to the same units of measurement for each study. The longest standing and most widely used approach is to divide the difference in means in each study by its standard deviation and present pooled results in standard deviation units (standardized mean difference). Disadvantages of this approach include vulnerability to varying degrees of heterogeneity in the underlying populations and difficulties in interpretation. Alternatives include presenting results in the units of the most popular or interpretable measure, converting to dichotomous measures and presenting relative and absolute effects, presenting the ratio of the means of intervention and control groups, and presenting the results in minimally important difference units. We outline the merits and limitations of each alternative and provide guidance for meta-analysts and guideline developers.

55. **ARTÍCULO Nº: 4341**

Thornton J, Alderson P, Tan T, Turner C, Latchem S, Shaw E et al. ***Introducing GRADE across the NICE clinical guideline program.*** J.Clin.Epidemiol. 2013; 66(2): 124-131.

OBJECTIVES: Grading of Recommendations Assessment, Development and Evaluation (GRADE) is a system for rating the confidence in estimates of effect and grading guideline recommendations. It promotes evaluation of the quality of the evidence for each outcome and an assessment of balance between desirable and undesirable outcomes leading to a judgment about the strength of the recommendation. In 2007, the National Institute for Health and Clinical Excellence began introducing GRADE across its clinical guideline program to enable separation of judgments about the evidence quality from judgments about the strength of the recommendation. STUDY DESIGN AND SETTING: We describe the process of implementing GRADE across guidelines. RESULTS: Use of GRADE has been positively received by both technical staff and guideline development group members. CONCLUSION: A shift in thinking about confidence in the evidence was required leading to a more structured and transparent approach to decision making. Practical problems were also encountered; these have largely been resolved, but some areas require further work, including the application of imprecision and presenting results from analyses considering more than two alternative interventions. The use of GRADE for nonrandomized and diagnostic accuracy studies needs to be refined.

56. **ARTÍCULO Nº: 4342**

Andrews J, Guyatt G, Oxman AD, Alderson P, Dahm P, Falck-Ytter Y et al. ***GRADE guidelines: 14. Going from evidence to recommendations: the significance and presentation of recommendations***. J.Clin.Epidemiol. 2013; 66(7): 719-725.

This article describes the Grading of Recommendations Assessment, Development, and Evaluation (GRADE) approach to classifying the direction and strength of recommendations. The strength of a recommendation, separated into strong and weak, is defined as the extent to which one can be confident that the desirable effects of an intervention outweigh its undesirable effects. Alternative terms for a weak recommendation include conditional, discretionary, or qualified. The strength of a recommendation has specific implications for patients, the public, clinicians, and policy makers. Occasionally, guideline developers may choose to make "only-in-research" recommendations. Although panels may choose not to make recommendations, this choice leaves those looking for answers from guidelines without the guidance they are seeking. GRADE therefore encourages panels to, wherever possible, offer recommendations.

57. **ARTÍCULO Nº: 4343**

Andrews JC, Schunemann HJ, Oxman AD, Pottie K, Meerpohl JJ, Coello PA et al. ***GRADE guidelines: 15. Going from evidence to recommendation-determinants of a recommendation's direction and strength***. J.Clin.Epidemiol. 2013; 66(7): 726-735.

In the GRADE approach, the strength of a recommendation reflects the extent to which we can be confident that the composite desirable effects of a management strategy outweigh the composite undesirable effects. This article addresses GRADE's approach to determining the direction and strength of a recommendation. The GRADE describes the balance of desirable and undesirable outcomes of interest among alternative management strategies depending on four domains, namely estimates of effect for desirable and undesirable outcomes of interest, confidence in the estimates of effect, estimates of values and preferences, and resource use. Ultimately, guideline panels must use judgment in integrating these factors to make a strong or weak recommendation for or against an intervention.

58. **ARTÍCULO Nº: 4344**

Mustafa RA, Santesso N, Brozek J, Akl EA, Walter SD, Norman G et al. ***The GRADE approach is reproducible in assessing the quality of evidence of quantitative evidence syntheses***. J.Clin.Epidemiol. 2013; 66(7): 736-742.

OBJECTIVE: We evaluated the inter-rater reliability (IRR) of assessing the quality of evidence (QoE) using the Grading of Recommendations, Assessment, Development, and Evaluation (GRADE) approach. STUDY DESIGN AND SETTING: On completing two training exercises, participants worked independently as individual raters to assess the QoE of 16 outcomes. After recording their initial impression using a global rating, raters graded the QoE following the GRADE approach. Subsequently, randomly paired raters submitted a consensus rating. RESULTS: The IRR without using the GRADE approach for two individual raters was 0.31 (95% confidence interval [95% CI] = 0.21-0.42) among Health Research Methodology students (n = 10) and 0.27 (95% CI = 0.19-0.37) among the GRADE working group members (n = 15). The corresponding IRR of the GRADE approach in assessing the QoE was significantly higher, that is, 0.66 (95% CI = 0.56-0.75) and 0.72 (95% CI = 0.61-0.79), respectively. The IRR further increased for three (0.80 [95% CI = 0.73-0.86] and 0.74 [95% CI = 0.65-0.81]) or four raters (0.84 [95% CI = 0.78-0.89] and 0.79 [95% CI = 0.71-0.85]). The IRR did not improve when QoE

was assessed through a consensus rating. CONCLUSION: Our findings suggest that trained individuals using the GRADE approach improves reliability in comparison to intuitive judgments about the QoE and that two individual raters can reliably assess the QoE using the GRADE system.

59. **ARTÍCULO Nº: 4345**

Guyatt G, Montori V, Schunemann H, onso-Coello P, Dahm P, Brito Campana JP et al. ***Letter reply to GRADE guidelines articles 14 and 15***. J.Clin.Epidemiol. 2014; 67(2): 240

60. **ARTÍCULO Nº: 4346**

McGregor M. ***The GRADE recommendations. weak recommendations are unhelpful to decision makers***. J.Clin.Epidemiol. 2014; 67(2): 239-240.

61. **ARTÍCULO Nº: 4347**

Navarro-Perez J, Peiro S, Brotons-Munto F, Lopez-Alcina E, Real-Romaguera A. ***[Quality of care indicators for benign prostatic hyperplasia. A qualitative study]***. Aten.Primaria. 2014; 46(5): 231-237.

OBJECTIVE: To assess quality of care indicators for benign prostatic hyperplasia (BPH), and to evaluate their strengths and weaknesses for incorporation into health information systems. DESIGN: Structured expert meeting, using procedures adapted from the nominal group techniques and the Rand consensus method. SETTING: Valencian School of Health Studies. PARTICIPANTS AND/OR CONTEXTS: Forty panellists (74% doctors, 70% from primary care settings) with experience in the management of BPH from 15 departments of the Valencia Health Agency. METHOD: Three workshops were held simultaneously (examination and diagnosis, drug therapy, and appropriateness and results), and the 15 quality indicators selected by the coordination group were assessed. RESULTS: Eleven of the 15 indicators scored in the range of high relevance. The 5 best rated were: the use of alpha-blockers + 5-alpha reductase inhibitor from certain severity level, digital rectal examination in the initial assessment, follow-up with the International Prostate Symptoms Score (IPSS), the rate of urgent catheterization in Hospital Accident & Emergency Units, initial assessment with the IPSS and the use of alpha-blockers prior to catheter removal for acute retention of urine. CONCLUSIONS: Some of the assessed indicators can be useful for incorporation into health information systems.

62. **ARTÍCULO Nº: 4348**

Nunez Montenegro AJ, Montiel LA, Martin AE, Torres VB, Lara MC, Gonzalez Correa JA. ***[Adherence to treatment, by active ingredient, in patients over 65 years on multiple medication]***. Aten.Primaria. 2014; 46(5): 238-245.

AIM: To assess the level of adherence, by active ingredient, to treatment and associated factors in polymedicated patients over 65 years-old. DESIGN: Observational, descriptive and cross-sectional study over polymedicated patients over 65 years of the Costa del Sol Health District and the North Malaga Health Area. The study was performed between January 2011 and September 2012 on 375 subjects obtained by simple random sampling from lists provided by each health centre. Data was collected by means of an interview with structured questions. Informed consent was given and signed by all patients before interview. STUDY VARIABLES: Main results variable adherence to treatment (Morisky-Green's test). PREDICTABLE VARIABLES: Prescription by active ingredient, socio-demographic variables, health care centre variables, and treatment associated variables. A descriptive analysis of variables was performed. Statistical inference was determined using univariate analysis (t test of Student or Mann-Whitney U, and Chi-squared), and controlling for confounding factors by multivariate analysis (linear and logistic regression). RESULTS: The result for therapeutic compliance was 51.7%. No

statistically significant differences were observed as regards sex and age. A relationship was found in those who resided in rural areas (P=.001), lived with family (P<.05), and were not at risk of suffering from anxiety (P=.046). CONCLUSIONS: We found similar patient adherence to treatment despite the prescribing generic drugs. Failure to therapeutic compliance was greater in those patients who lived by themselves, in a city close to the coast, or in those patients who were at risk of suffering from anxiety.

63. **ARTÍCULO Nº: 4349**

Kolozsvari LR, Orozco-Beltran D, Rurik I. ***Do family physicians need more payment for working better? Financial incentives in primary care.*** Aten.Primaria. 2014; 46(5): 261-266.

INTRODUCTION: Financial incentives are widely used in health services to improve the quality of care or to reach some specific targets. Pay for performance systems were also introduced in the primary health care systems of many European countries. OBJECTIVE: Our study aims to describe and compare recent existing primary care indicators and related financing in European countries. METHODS: Literature search was performed and questionnaires were sent to primary care experts of different countries within the European General Practice Research Network. RESULTS: Ten countries have published primary care quality indicators (QI) associated with financial incentives. The number of QI varies from 1 to 134 and can modify the finances of physicians with up to 25% of their total income. CONCLUSIONS: The implementations of these schemes should be critically evaluated with continuous monitoring at national or regional level; comparison is required between targets and their achievements, health gains and use of resources as well.

64. **ARTÍCULO Nº: 4350**

Martinez-Gonzalez NA, Djalali S, Tandjung R, Huber-Geismann F, Markun S, Wensing M et al. ***Substitution of physicians by nurses in primary care: a systematic review and meta-analysis.*** BMC.Health Serv.Res. 2014; 14: 214

BACKGROUND: In many countries, substitution of physicians by nurses has become common due to the shortage of physicians and the need for high-quality, affordable care, especially for chronic and multi-morbid patients. We examined the evidence on the clinical effectiveness and care costs of physician-nurse substitution in primary care. METHODS: We systematically searched OVID Medline and Embase, The Cochrane Library and CINAHL, up to August 2012; selected and critically appraised published randomised controlled trials (RCTs) that compared nurse-led care with care by primary care physicians on patient satisfaction, Quality of Life (QoL), hospital admission, mortality and costs of healthcare. We assessed the individual study risk of bias, calculated the study-specific and pooled relative risks (RR) or standardised mean differences (SMD); and performed fixed-effects meta-analyses. RESULTS: 24 RCTs (38,974 participants) and 2 economic studies met the inclusion criteria. Pooled analyses showed higher overall scores of patient satisfaction with nurse-led care (SMD 0.18, 95% CI 0.13 to 0.23), in RCTs of single contact or urgent care, short (less than 6 months) follow-up episodes and in small trials (N </= 200). Nurse-led care was effective at reducing the overall risk of hospital admission (RR 0.76, 95% CI 0.64 to 0.91), mortality (RR 0.89, 95% CI 0.84 to 0.96), in RCTs of on-going or non-urgent care, longer (at least 12 months) follow-up episodes and in larger (N > 200) RCTs. Higher quality RCTs (with better allocation concealment and less attrition) showed higher rates of hospital admissions and mortality with nurse-led care albeit less or not significant. The results seemed more consistent across nurse practitioners than with registered or licensed nurses. The effects of nurse-led care on QoL and costs were difficult to interpret due to heterogeneous outcome reporting, valuation of resources and the small number of studies. CONCLUSIONS: The available

evidence continues to be limited by the quality of the research considered. Nurse-led care seems to have a positive effect on patient satisfaction, hospital admission and mortality. This important finding should be confirmed and the determinants of this effect should be assessed in further, larger and more methodically rigorous research.

65. **ARTÍCULO Nº: 4351**

Forbes HJ, Bhaskaran K, Thomas SL, Smeeth L, Clayton T, Langan SM. ***Quantification of risk factors for herpes zoster: population based case-control study***. BMJ. 2014; 348: g2911

Objectives: To quantify the effects of possible risk factors for herpes zoster at different ages.
Design: Case-control study.
Setting: UK Clinical Practice Research Datalink primary care data.
Participants: 144 959 adults diagnosed with zoster between 2000 and 2011; 549 336 age, sex, and practice matched controls.
Main outcome measures: Conditional logistic regression was used to generate adjusted odds ratios to estimate the strength of association of each potential risk factor with zoster and assess effect modification by age.
Results: The median age of the cases and controls was 62 years. Factors associated with increased risk of zoster included rheumatoid arthritis (3111 (2.1%) *v* 8029 (1.5%); adjusted odds ratio 1.46, 99% confidence interval 1.38 to 1.55), inflammatory bowel disease (1851 (1.3%) *v* 5118 (0.9%); 1.36, 1.26 to 1.46), chronic obstructive pulmonary disease (6815 (4.7%) *v* 20 201 (3.7%); 1.32, 1.27 to 1.37), asthma (10 243 (7.1%) *v* 31 865 (5.8%); 1.21, 1.17 to 1.25), chronic kidney disease (8724 (6.0%) *v* 29 437 (5.4%); 1.14, 1.09 to 1.18), and depression (6830 (4.7%) *v* 22 052 (4.0%); 1.15, 1.10 to 1.20). Type 1, but not type 2, diabetes showed some association with zoster (adjusted odds ratio 1.27, 1.07 to 1.50). The relative effects of many assessed risk factors were larger in younger patients. Patients with severely immunosuppressive conditions were at greatest risk of zoster.for example, patients with lymphoma (adjusted odds ratio 3.90, 3.21 to 4.74) and myeloma (2.16, 1.84 to 2.53), who are not eligible for zoster vaccination.
Conclusions: A range of conditions were associated with increased risk of zoster. In general, the increased risk was proportionally greater in younger age groups. Current vaccines are contraindicated in people at the greatest risk of zoster, highlighting the need for alternative risk reduction strategies in these groups.

66. **ARTÍCULO Nº: 4352**

Dombrowski SU. ***Long term maintenance of weight loss with non-surgical interventions in obese adults: systematic review and meta-analyses of randomised controlled trials***. BMJ. 2014; 348: g2646

Objective: To systematically review and describe currently available approaches to supporting maintenance of weight loss in obese adults and to assess the evidence for the effectiveness of these interventions.
Design: Systematic review with meta-analysis.
Data sources: Medline, PsycINFO, Embase, and the Cochrane Central Register of Controlled Trials.
Study selection: Studies were identified through to January 2014. Randomised trials of interventions to maintain weight loss provided to initially obese adults (aged .18) after weight loss of .5% body weight with long term (.12 months) follow-up of weight change (main outcome) were included.
Study appraisal and synthesis: Potential studies were screened independently and in duplicate; study characteristics and outcomes were extracted. Meta-analyses were conducted to estimate the effects of interventions on weight loss maintenance with the inverse variance method and a random effects

model. Results are presented as mean differences in weight change, with 95% confidence intervals.
Results: 45 trials involving 7788 individuals were included. Behavioural interventions focusing on both food intake and physical activity resulted in an average difference of .1.56 kg (95% confidence interval .2.27 to .0.86 kg; 25 comparisons, 2949 participants) in weight regain compared with controls at 12 months. Orlistat combined with behavioural interventions resulted in a .1.80 kg (.2.54 to .1.06; eight comparisons, 1738 participants) difference compared with placebo at 12 months. All
orlistat studies reported higher frequencies of adverse gastrointestinal events in the experimental compared with placebo control groups. A dose-response relation for orlistat treatment was found, with 120 mg doses three times a day leading to greater weight loss maintenance (.2.34 kg, .3.03 to .1.65) compared with 60 mg and 30 mg three times a day (.0.70 kg, 95% confidence interval .1.92 to 0.52), P=0.02.
Conclusions: Behavioural interventions that deal with both diet and physical activity show small but significant benefits on weight loss maintenance.

67. **ARTÍCULO Nº: 4353**
Amir LH. ***Managing common breastfeeding problems in the community***. BMJ. 2014; 348: g2954

68. **ARTÍCULO Nº: 4354**
Reddy M, Gill SS. ***The effectiveness of pressure ulcer risk assessment instruments and associated intervention protocols remains uncertain***. Evid.Based.Med. 2014; 19(3): 93

69. **ARTÍCULO Nº: 4355**
Allinson JP, Donaldson GC. ***Long-term antibiotic therapy reduces exacerbation frequency in patients with COPD but it remains unclear which patients to target***. Evid.Based.Med. 2014; 19(3): 99

70. **ARTÍCULO Nº: 4356**
Shaikh N, Martin JM. ***Delayed prescription worsens reported symptoms and increases antibiotic use compared with clinical score with or without rapid antigen testing in patients with sore throat***. Evid.Based.Med. 2014; 19(3): 117

71. **ARTÍCULO Nº: 4357**
Schuetz P. ***Neither ibuprofen nor steam improves symptom control compared with paracetamol in patients with acute respiratory tract infections in primary care***. Evid.Based.Med. 2014; 19(3): 102

72. **ARTÍCULO Nº: 4358**
Brokel J. ***Evidence-based clinical decision support improves the appropriate use of antibiotics and rapid strep testing***. Evid.Based.Med. 2014; 19(3): 118

73. **ARTÍCULO Nº: 4359**
Pascarella L. ***Intermittent pneumatic compression is effective in reducing venous thromboembolism risk in hospitalised patients***. Evid.Based.Med. 2014; 19(3): 104

74. **ARTÍCULO Nº: 4360**
Wartolowska K, Judge A, Hopewell S, Collins GS, Dean BJ, Rombach I et al. ***Use of placebo controls in the evaluation of surgery: systematic review***. BMJ. 2014; 348: g3253

OBJECTIVE: To investigate whether placebo controls should be used in the evaluation of surgical interventions. DESIGN: Systematic review. DATA SOURCES: We searched Medline, Embase, and the Cochrane Controlled Trials Register from their inception to November 2013. STUDY SELECTION:

Randomised clinical trials comparing any surgical intervention with placebo. Surgery was defined as any procedure that both changes the anatomy and requires a skin incision or use of endoscopic techniques. DATA EXTRACTION: Three reviewers (KW, BJFD, IR) independently identified the relevant trials and extracted data on study details, outcomes, and harms from included studies. RESULTS: In 39 out of 53 (74%) trials there was improvement in the placebo arm and in 27 (51%) trials the effect of placebo did not differ from that of surgery. In 26 (49%) trials, surgery was superior to placebo but the magnitude of the effect of the surgical intervention over that of the placebo was generally small. Serious adverse events were reported in the placebo arm in 18 trials (34%) and in the surgical arm in 22 trials (41.5%); in four trials authors did not specify in which arm the events occurred. However, in many studies adverse events were unrelated to the intervention or associated with the severity of the condition. The existing placebo controlled trials investigated only less invasive procedures that did not involve laparotomy, thoracotomy, craniotomy, or extensive tissue dissection. CONCLUSIONS: Placebo controlled trial is a powerful, feasible way of showing the efficacy of surgical procedures. The risks of adverse effects associated with the placebo are small. In half of the studies, the results provide evidence against continued use of the investigated surgical procedures. Without well designed placebo controlled trials of surgery, ineffective treatment may continue unchallenged.

75. **ARTÍCULO Nº: 4361**

Kipping RR, Howe LD, Jago R, Campbell R, Wells S, Chittleborough CR et al. ***Effect of intervention aimed at increasing physical activity, reducing sedentary behaviour, and increasing fruit and vegetable consumption in children: active for Life Year 5 (AFLY5) school based cluster randomised controlled trial.*** BMJ. 2014; 348: g3256

OBJECTIVE: To investigate the effectiveness of a school based intervention to increase physical activity, reduce sedentary behaviour, and increase fruit and vegetable consumption in children. DESIGN: Cluster randomised controlled trial. SETTING: 60 primary schools in the south west of England. PARTICIPANTS: Primary school children who were in school year 4 (age 8-9 years) at recruitment and baseline assessment, in year 5 during the intervention, and at the end of year 5 (age 9-10) at follow-up assessment. INTERVENTION: The Active for Life Year 5 (AFLY5) intervention consisted of teacher training, provision of lesson and child-parent interactive homework plans, all materials required for lessons and homework, and written materials for school newsletters and parents. The intervention was delivered when children were in school year 5 (age 9-10 years). Schools allocated to control received standard teaching. MAIN OUTCOME MEASURES: The pre-specified primary outcomes were accelerometer assessed minutes of moderate to vigorous physical activity per day, accelerometer assessed minutes of sedentary behaviour per day, and reported daily consumption of servings of fruit and vegetables. RESULTS: 60 schools with more than 2221 children were recruited; valid data were available for fruit and vegetable consumption for 2121 children, for accelerometer assessed physical activity and sedentary behaviour for 1252 children, and for secondary outcomes for between 1825 and 2212 children for the main analyses. None of the three primary outcomes differed between children in schools allocated to the AFLY5 intervention and those allocated to the control group. The difference in means comparing the intervention group with the control group was -1.35 (95% confidence interval -5.29 to 2.59) minutes per day for moderate to vigorous physical activity, -0.11 (-9.71 to 9.49) minutes per day for sedentary behaviour, and 0.08 (-0.12 to 0.28) servings per day for fruit and vegetable consumption. The intervention was effective for three out of nine of the secondary outcomes after multiple testing was taken into account: self reported time spent in screen viewing at the weekend (-21 (-37 to -4) minutes per day), self reported servings of snacks per day (-0.22 (-0.38 to -0.05)), and servings of high energy drinks per day (-0.26 (-0.43 to -0.10)) were all reduced. Results from a series of sensitivity analyses testing different assumptions about missing data

and from per protocol analyses produced similar results. CONCLUSION: The findings suggest that the AFLY5 school based intervention is not effective at increasing levels of physical activity, decreasing sedentary behaviour, and increasing fruit and vegetable consumption in primary school children. Change in these activities may require more intensive behavioural interventions with children or upstream interventions at the family and societal level, as well as at the school environment level. These findings have relevance for researchers, policy makers, public health practitioners, and doctors who are involved in health promotion, policy making, and commissioning services. Trial registration Current Controlled Trials ISRCTN50133740.

76. **ARTÍCULO Nº: 4362**

Manzoli L, Flacco ME, D'Addario M, Capasso L, De VC, Marzuillo C et al. ***Non-publication and delayed publication of randomized trials on vaccines: survey***. BMJ. 2014; 348: g3058

OBJECTIVE: To evaluate the extent of non-publication or delayed publication of registered randomized trials on vaccines, and to investigate potential determinants of delay to publication. DESIGN: Survey. DATA SOURCES: Trials registry websites, Scopus, PubMed, Google. STUDY SELECTION: Randomized controlled trials evaluating the safety or the efficacy or immunogenicity of human papillomavirus (HPV), pandemic A/H1N1 2009 influenza, and meningococcal, pneumococcal, and rotavirus vaccines that were registered in ClinicalTrials.gov, Current Controlled Trials, WHO International Clinical Trials Registry Platform, Clinical Study Register, or Indian, Australian-New Zealand, and Chinese trial registries in 2006-12. Electronic databases were searched up to February 2014 to identify published manuscripts containing trial results. These were reviewed and classified as positive, mixed, or negative. We also reviewed the results available in ClinicalTrials.gov. MAIN OUTCOME MEASURES: Publication status of trial results and time from completion to publication in peer reviewed journals. DATA SYNTHESIS: Cox proportional hazards regression was used to evaluate potential predictors of publication delay. RESULTS: We analysed 384 trials (85% sponsored by industry). Of 355 trials (404,758 participants) that were completed, 176 (n=151,379) had been published in peer reviewed journals. Another 42 trials (total sample 62,765) remained unpublished but reported results in ClinicalTrials.gov. The proportion of trials published 12, 24, 36, and 48 months after completion was 12%, 29%, 53%, and 73%, respectively. Including results posted in ClinicalTrials.gov, 48 months after study completion results were available for 82% of the trials and 90% of the participants. Delay to publication between non-industry and industry sponsored trials did not differ, but non-industry sponsored trials were 4.42-fold (P=0.008) more likely to report negative or mixed findings. Negative results were reported by only 2% of the published trials. CONCLUSIONS: Most vaccine trials are published eventually or the results posted in ClinicalTrials.gov, but delays to publication of several years are common. Actions should focus on the timely dissemination of data from vaccine trials to the public.

77. **ARTÍCULO Nº: 4363**

Barr B, Bambra C, Whitehead M. ***The impact of NHS resource allocation policy on health inequalities in England 2001-11: longitudinal ecological study***. BMJ. 2014; 348: g3231

OBJECTIVE: To investigate whether the policy of increasing National Health Service funding to a greater extent in deprived areas in England compared with more affluent areas led to a reduction in geographical inequalities in mortality amenable to healthcare. DESIGN: Longitudinal ecological study. SETTING: 324 lower tier local authorities in England, classified by their baseline level of deprivation. INTERVENTION: Differential trends in NHS funds allocated to local areas resulting from the NHS resource allocation policy in England between 2001 and 2011. MAIN OUTCOME MEASURE: Trends in mortality from causes considered amenable to healthcare in local authority areas in England. Using

multivariate regression, we estimated the reduction in mortality that was associated with the allocation of additional NHS resources in these areas. RESULTS: Between 2001 and 2011 the increase in NHS resources to deprived areas accounted for a reduction in the gap between deprived and affluent areas in male mortality amenable to healthcare of 35 deaths per 100,000 population (95% confidence interval 27 to 42) and female mortality of 16 deaths per 100,000 (10 to 21). This explained 85% of the total reduction of absolute inequality in mortality amenable to healthcare during this time. Each additional pound10 m of resources allocated to deprived areas was associated with a reduction in 4 deaths in males per 100,000 (3.1 to 4.9) and 1.8 deaths in females per 100,000 (1.1 to 2.4). The association between absolute increases in NHS resources and improvements in mortality amenable to healthcare in more affluent areas was not significant. CONCLUSION: Between 2001 and 2011, the NHS health inequalities policy of increasing the proportion of resources allocated to deprived areas compared with more affluent areas was associated with a reduction in absolute health inequalities from causes amenable to healthcare. Dropping this policy may widen inequalities.

78. **ARTÍCULO Nº: 4364**

Michaelsson K. ***Surgeon volume and early complications after primary total hip arthroplasty***. BMJ. 2014; 348: g3433

79. **ARTÍCULO Nº: 4365**

Selak V, Elley CR, Bullen C, Crengle S, Wadham A, Rafter N et al. ***Effect of fixed dose combination treatment on adherence and risk factor control among patients at high risk of cardiovascular disease: randomised controlled trial in primary care***. BMJ. 2014; 348: g3318

OBJECTIVE: To evaluate whether provision of fixed dose combination treatment improves adherence and risk factor control compared with usual care of patients at high risk of cardiovascular disease in primary care. DESIGN: Open label randomised control trial: IMPACT (IMProving Adherence using Combination Therapy). SETTING: 54 general practices in the Auckland and Waikato regions of New Zealand, July 2010 to August 2013. PARTICIPANTS: 513 adults (including 257 indigenous Maori) at high risk of cardiovascular disease (established cardiovascular disease or five year risk >/= 15%) who were recommended for treatment with antiplatelet, statin, and two or more blood pressure lowering drugs. 497 (97%) completed 12 months' follow-up. INTERVENTIONS: Participants were randomised to continued usual care or to fixed dose combination treatment (with two versions available: aspirin 75 mg, simvastatin 40 mg, and lisinopril 10 mg with either atenolol 50 mg or hydrochlorothiazide 12.5 mg). All drugs in both treatment arms were prescribed by their usual general practitioners and dispensed by local community pharmacists. MAIN OUTCOME MEASURES: Primary outcomes were self reported adherence to recommended drugs (antiplatelet, statin, and two or more blood pressure lowering agents) and mean change in blood pressure and low density lipoprotein cholesterol at 12 months. RESULTS: Adherence to all four recommended drugs was greater among fixed dose combination than usual care participants at 12 months (81% v 46%; relative risk 1.75, 95% confidence interval 1.52 to 2.03, P<0.001; number needed to treat 2.9, 95% confidence interval 2.3 to 3.7). Adherence for each drug type at 12 months was high in both groups but especially in the fixed dose combination group: for antiplatelet treatment it was 93% fixed dose combination v 83% usual care (P<0.001), for statin 94% v 89% (P=0.06), for combination blood pressure lowering 89% v 59% (P<0.001), and for any blood pressure lowering 96% v 91% (P=0.02). Self reported adherence was highly concordant with dispensing data (dispensing of all four recommended drugs 79% fixed dose combination v 47% usual care, relative risk 1.67, 95% confidence interval 1.44 to 1.93, P<0.001). There was no statistically significant improvement in risk factor control between the fixed dose combination and usual care groups over 12 months: the difference in systolic blood pressure was -2.2 mm Hg (-4.5

v -2.3, 95% confidence interval -5.6 to 1.2, P=0.21), in diastolic blood pressure -1.2 mm Hg (-2.1 v -0.9, -3.2 to 0.8, P=0.22) and in low density lipoprotein cholesterol -0.05 mmol/L (-0.20 v -0.15, -0.17 to 0.08, P=0.46). The number of participants with cardiovascular events or serious adverse events was similar in both treatment groups (fixed dose combination 16 v usual care 18 (P=0.73), 99 v 93 (P=0.56), respectively). Fixed dose combination treatment was discontinued in 94 participants (37%). The most commonly reported reason for discontinuation was a side effect (54/75, 72%). Overall, 89% (227/256) of fixed dose combination participants' general practitioners completed a post-trial survey, and the fixed dose combination strategy was rated as satisfactory or very satisfactory for starting treatment (206/227, 91%), blood pressure control (180/220, 82%), cholesterol control (170/218, 78%), tolerability (181/223, 81%), and prescribing according to local guidelines (185/219, 84%). When participants were asked at 12 months how easy they found taking their prescribed drugs, most responded very easy or easy (224/246, 91% fixed dose combination v 212/246, 86% usual care, P=0.09). At 12 months the change in other lipid fractions, difference in EuroQol-5D, and difference in barriers to adherence did not differ significantly between the treatment groups. CONCLUSIONS: Among this well treated primary care population, fixed dose combination treatment improved adherence to the combination of all recommended drugs but improvements in clinical risk factors were small and did not reach statistical significance. Acceptability was high for both general practitioners and patients, although the discontinuation rate was high. TRIAL REGISTRATION: Australian New Zealand Clinical Trial Registry ACTRN12606000067572.

80. **ARTÍCULO Nº: 4366**

Gumbinger C, Reuter B, Stock C, Sauer T, Wietholter H, Bruder I et al. ***Time to treatment with recombinant tissue plasminogen activator and outcome of stroke in clinical practice: retrospective analysis of hospital quality assurance data with comparison with results from randomised clinical trials.*** BMJ. 2014; 348: g3429

OBJECTIVE: To study the time dependent effectiveness of thrombolytic therapy for acute ischaemic stroke in daily clinical practice. DESIGN: A retrospective cohort study using data from a large scale, comprehensive population based state-wide stroke registry in Germany. SETTING: All 148 hospitals involved in acute stroke care in a large state in southwest Germany with 10.4 million inhabitants. PARTICIPANTS: Data from 84,439 patients with acute ischaemic stroke were analysed, 10,263 (12%) were treated with thrombolytic therapy and 74,176 (88%) were not treated. MAIN OUTCOME MEASURES: Primary endpoint was the dichotomised score on a modified Rankin scale at discharge ("favourable outcome" score 0 or 1 or "unfavourable outcome" score 2-6) analysed by binary logistic regression. Patients treated with recombinant tissue plasminogen activator (rtPA) were categorised according to time from onset of stroke to treatment. Analogous analyses were conducted for the association between rtPA treatment of stroke and in-hospital mortality. As a co-primary endpoint the chance of a lower modified Rankin scale score at discharge was analysed by ordinal logistic regression analysis (shift analysis). RESULTS: After adjustment for characteristics of patients, hospitals, and treatment, rtPA was associated with better outcome in a time dependent pattern. The number needed to treat ranged from 4.5 (within first 1.5 hours after onset; odds ratio 2.49) to 18.0 (up to 4.5 hours; odds ratio 1.26), while mortality did not vary up to 4.5 hours. Patients treated with rtPA beyond 4.5 hours (including mismatch based approaches) showed a significantly better outcome only in dichotomised analysis (odds ratio 1.25, 95% confidence interval 1.01 to 1.55) but the mortality risk was higher (1.45, 1.08 to 1.92). CONCLUSION: The effectiveness of thrombolytic therapy in daily clinical practice might be comparable with the effectiveness shown in randomised clinical trials and pooled analysis. Early treatment was associated with favourable outcome in daily clinical practice, which underlines the importance of speeding up the process for thrombolytic therapy in hospital and

before admission to achieve shorter time from door to needle and from onset to treatment for thrombolytic therapy.

81. **ARTÍCULO Nº: 4367**

Ravi B, Jenkinson R, Austin PC, Croxford R, Wasserstein D, Escott B et al. ***Relation between surgeon volume and risk of complications after total hip arthroplasty: propensity score matched cohort study***. BMJ. 2014; 348: g3284

OBJECTIVES: To identify a cut point in annual surgeon volume associated with increased risk of complications after primary elective total hip arthroplasty and to quantify any risk identified. DESIGN: Propensity score matched cohort study. SETTING: Ontario, Canada. PARTICIPANTS: 37,881 people who received their first primary total hip arthroplasty during 2002-09 and were followed for at least two years after their surgery. MAIN OUTCOME MEASURE: The rates of various surgical complications within 90 days (venous thromboembolism, death) and within two years (infection, dislocation, periprosthetic fracture, revision) of surgery. RESULTS: Multivariate splines were developed to visualize the relation between surgeon volume and the risk for various complications. A threshold of 35 cases a year was identified, under which there was an increased risk of dislocation and revision. 6716 patients whose total hip arthroplasty was carried out by surgeons who had done </= 35 such procedure in the previous year were successfully matched to patients whose surgeon had carried out more than 35 procedures. Patients in the former group had higher rates of dislocation (1.9% v 1.3%, P=0.006; NNH 172) and revision (1.5% v 1.0%, P=0.03; NNH 204). CONCLUSIONS: In a cohort of first time recipients of total hip arthroplasty, patients whose operation was carried by surgeons who had performed 35 or fewer such procedures in the year before the index procedure were at increased risk for dislocation and early revision. Surgeons should consider performing 35 cases or more a year to minimize the risk for complications. Furthermore, the methods used to visualize the relationship between surgeon volume and the occurrence of complications can be easily applied in any jurisdiction, to help inform and optimize local healthcare delivery.

82. **ARTÍCULO Nº: 4368**

De la Torre Barbero MJ, Estepa Luna MJ, Lopez-Pardo MM, Leon MM, Sanchez LF, Toledano RS. ***[Evaluation of the Andalusia Public Health System hospital websites in the period 2010-2012]***. Rev.Calid.Asist. 2014; 29(3): 127-134.

OBJECTIVES: Evaluate the quality, accessibility and presence of Web 2.0 tools in the Andalusia Public Health System hospitals websites METHODS: Observational, descriptive study carried out between 2010 and 2012. The variables analyzed were: quality, accessibility and innovation. The quality was evaluated using a Bermudez-Tamayo questionnaire. Accessibility was measured using the Web Accessibility Test (TAW) tool. Web 2.0 tools were identified by direct observation. RESULTS: A total of 31 of the 45 hospitals (68.9%) had a website in the year 2010, increasing to 34 (75.5%) in 2012. The average score+standard deviation (SD) of the Bermudez-Tamayo quality questionnaire was 11.1+3.8 points in 2010, and 12.3+3.9 points in 2012, observing a statistically significant difference of 0.25 being observed between the means (P=.007), 95% CI; 0.00 to 0.50) In the accessibility evaluation only 7 websites (n=31) in 2010, and 10 (n=34) in 2012, fulfilled the legal criteria for accessibility. The use of Web 2.0 tools has increased throughout the study. In 2010, 19.4% (n=6) of the hospital websites had this type of tool, in comparison to 58.8% (n=20) in 2012. CONCLUSIONS: In general, the quality of the websites studied is good. However, current legislation regarding accessibility is not fulfilled and must be revised and adapted to the current legal rules. There is an incipient use of Web 2.0 resources as education and communication strategies with regard to health.

83. **ARTÍCULO Nº: 4369**

Pinedo S, Zaldibar B, Sanmartin V, Tejada P, Erazo P, Miranda M et al. ***[Subacute care of stroke-affected patients. Satisfaction and results]***. Rev.Calid.Asist. 2014; 29(3): 150-157.

OBJECTIVE: To determine the satisfaction of the stroke inpatients and their caregivers in Rehabilitation Service and to analyze the effectiveness, social risk, and discharge destination. MATERIAL AND METHOD: Prospective longitudinal cohort multicenter study. An analysis was made of the social risk (Gijon Scale), co-morbidity (Charlson Index), disability (Barthel Index), effectiveness of the rehabilitation treatment, satisfaction (Pound Questionnaire) and discharge destination of 241 patients. An evaluation was also made on 119 caregivers 6 months post-stroke, recording age, family relationship, time care-giving, satisfaction with the information/training, and accessibility to the rehabilitation team. RESULTS: The patient profile is a 71 year-old male, with low/intermediate social risk, high co-morbidity and total/severe dependence, with 27.1% living alone. Almost all (96.6%) of the patients claimed to be satisfied/very satisfied with the treatment, with satisfaction with the recovery being lower (80.3%). The effectiveness was 32.5 +/- 20.4. Home was the discharge destination of 81.7% of the patients.The average age of the caregivers was 58.8 +/- 12.3 years, and 73.9% were women. The time dedicated to care-giving was over 6 hours per day in the 62% of the cases. Being satisfied/very satisfied with the received information was recorded by 89.9% of the caregivers. CONCLUSIONS: Patients admitted for stroke rehabilitation achieve significant functional gain during hospitalization and return to their homes in most cases. The satisfaction with the rehabilitation treatment and received information is high. The training of the caregiver is an aspect that needs improving.

84. **ARTÍCULO Nº: 4370**

Valverde-Bilbao E, Mendizabal-Olaizola A, Idoiaga-Hoyos I, rriaga-Goirizelai L, Carracedo-Arrastio JD, rranz-Lazaro C. ***[Medication reconciliation in primary care after hospital discharge]***. Rev.Calid.Asist. 2014; 29(3): 158-164.

OBJECTIVES: The primary objective of this study was to determine if changes prescribed in the usual treatment of patients at discharge from the hospital were updated in their active treatment sheet when they came to the Primary Care clinic. The secondary objectives included, determining whether the drug average varies between the admission and discharge, as well as, identifying other factors related to the modification of treatment during hospital admission including, among others, patient age or the number of drugs previously indicated. Finally, the relationship between the Primary Care Unit to which the patient belonged and the probability that the medication was reconciled was also examined. MATERIAL AND METHODS: This is an observational cross-sectional study conducted in the Bidasoa Integrated Healthcare Organization. The study included every patient over 65 years old with multiple medication (taking 5 or more drugs) belonging to this organization, and discharged from Bidasoa Hospital between 15th October and 11th November 2012. The information on hospital discharges during this period was sent from the hospital to those responsible for patient safety in the Primary Health Care Centers. Each patient clinical history was reviewed in order to confirm if a visit (at least once in the first two weeks after discharge) had been made to their Primary Care Unit, and whether there had been a change in their active treatment sheet. RESULTS: Two hundred sixty-one patients (n=261) were discharged from Bidasoa Hospital in the study period, and 80 met the inclusion criteria. The discharge report proposed a change in the active treatment in 39 of them (49%). Of these, 35 (90%) attended a Primary Care clinic, and the changes were included in their active treatment sheet in 24 patients, representing 68% of those who contacted Primary Care, and 61% of those who would have required changes. CONCLUSIONS: The results demonstrate the need to

establish a reconciliation medication program for patients on multiple medications after hospital discharge. Moreover, further studies are needed to investigate what may be the reasons why the changes to active treatment sheets are not taking place for some patients, despite these having visited Primary Care after having been discharged from hospital.

85. **ARTÍCULO Nº: 4371**

Martinez RF, Codina GJ, Deulofeu QP, Garrido CJ, Blasco CF, Gibanel G, X et al. ***[Indicators of healthcare quality in day surgery (2010-2012)]***. Rev.Calid.Asist. 2014; 29(3): 172-179.

INTRODUCTION: Monitoring quality indicators in Ambulatory Surgery centers is fundamental in order to identify problems, correct them and prevent them. Given their large number, it is essential to select the most valid ones. OBJECTIVES: The objectives of the study are the continuous improvement in the quality of healthcare of day-case surgery in our center, by monitoring selective quality parameters, having periodic information on the results and taking corrective measures, as well as achieving a percentage of unplanned transfer and cancellations within quality standards. MATERIAL AND METHOD: Prospective, observational and descriptive study of the day-case surgery carried out from January 2010 to December 2012. Unplanned hospital admissions and cancellations on the same day of the operation were selected and monitored, along with their reasons. Hospital admissions were classified as: inappropriate selection, medical-surgical complications, and others. The results were evaluated each year and statistically analysed using chi(2) tests. RESULTS: A total of 8,300 patients underwent day surgery during the 3 years studied. The day-case surgery and outpatient index increased by 5.4 and 6.4%, respectively (P<.01). Unexpected hospital admissions gradually decreased due to the lower number of complications (P<.01). Hospital admissions, due to an extended period of time in locoregional anaesthesia recovery, also decreased (P<.01). There was improved prevention of nausea and vomiting, and of poorly controlled pain. The proportion of afternoon admissions was significantly reduced (P<.01). The cancellations increased in 2011 (P<.01). CONCLUSIONS: The monitoring of quality parameters in day-case surgery has been a useful tool in our clinical and quality management. Globally, the unplanned transfer and cancellations have been within the quality standards and many of the indicators analysed have improved.