11:00 Monday August 20th - Niels K. Jerne - De-implementation and the Challenge of Tackling Overdiagnosis at the Level of the Consultation

DIAGNOSIS : INDUSTRY'S ROYAL ROAD TO MANIPULATE MODERN MEDICINE

JACQUES THIVIERGE
LAVAL UNIVERSITY, QUÉBEC, Canada

Abstract: HAVING BEEN A PRACTICING CHILD PSYCHIATRIST FOR THE LAST FORTY YEARS, I WANT TO TALK ABOUT THE PERVASIVE USE OF THE TERM DIAGNOSIS IN MODERN MEDICINE AND HOW THIS TERM IS BEING USED ANS MISUSED BY THE INDUSTRY, AND BY US, IN A WAY THAT HARMS OUR VERY PATIENTS.
THE ARGUMENT WILL SIT ON THE DETAILS OF TWO EXAMPLES: ONE FROM CLASSICAL MEDICINE, THE HYPERACTIVE BLADDER, THE OTHER FROM CHILD PSYCHIATRY, THE ATTENTION DEFICIT DISORDER WITH HYPERACTIVITY DIAGNOSIS.
IT WILL BE PROPOSED THAT MANY A DEFINITION OF THE TERM DIAGNOSIS ARE ACCEPTABLE, AND THAT WE HAVE NO TIME TO LOOSE FIGHTING FOR DEFINITIONS. WHAT WE ARE INTERESTED IN ARE THE CLINICAL REALITIES THEMSELVES WE WANT TO TALK ABOUT USING THE TERMS WE USE.
IN ORDER TO FIGHT BACK THE UNFORTUNATE AND HARMFUL CONSEQUENCES OF THE USE OF THE TERM DIAGNOSIS IN MODERN MEDICINE, THE AUTHOR WILL EXPLAIN HOW, SINCE A NUMBER OF YEARS, HE USED A DEFINITION OF THIS TERM WHICH CLARIFIES UP TO A USEFUL POINT FOR A GREAT NUMBER OF PATIENTS AND THEIR PARENTS, WHAT A VALID MEDICAL DIAGNOSIS SHOULD BE.
ACKNOWLEDGING THE MAJOR DIFFICULTIES OF OUR FIGHTING BACK FOR THIS TERM TO MEAN VALID MEDICAL PROBLEMS INSTEAD, AS TOO OFTEN NOWDAYS, DIAGNOSTIC-LIKE ENTITIES CREATED BY BUSINESSMEN AT THE SERVICE OF THE INDUSTRY, THE AUTHOR WOULD WELCOME THE OPPORTUNITY TO DISCUSS WITH OTHER CLINICIANS ABOUT THEIR WAYS TO DEAL WITH THIS PROBLEM, HOPING THEREBY TO IMPROVE HIS OWN WAY OF DOING IT.

Objectives: THE OBJECTIVE OF THE PRESENTATION IS TO STIMULATE A CRITICAL APPROACH TO THE DAILY USED TERM DIAGNOSIS IN THE ACTIVITIES OF OUR OCCIDENTAL HEALTHCARE SYSTEMS. FOLLOWING THIS PRESENTATION IT IS HOPED THE PARTICIPANTS WILL ASK THEMSELVES A NUMBER OF QUESTIONS BEFORE ACCEPTING A PROPOSED DIAGNOSIS AS A VALID MEDICAL DIAGNOSIS.

Method: THIS IS NO RESEARCH BUT THE REFLEXIONS OF A CLINICIAN ON HOW TO DEAL WITH THE CLINICAL PROBLEM CREATED BY HOW THE TERM DIAGNOSIS IS CURRENTLY USED IN OUR CORPORATE WORLD.

Results: NO QUANTIFIABLE RESULTS. THE RESULTS ARE THE ONES OF A CLINICAL EXPERIENCE DEALING WITH THE PROBLEMS CREATED FOR OUR PATIENTS WITH WHAT WE CALL VALID MEDICAL DIAGNOSIS IN OUR SOCIETIES N/A

Conclusions: THE AUTHOR WILL DISCUSS THE BENEFITS OF HIS DEALING WITH THE TERM DIAGNOSIS IN HIS CLINICAL PRATICE

Shared decision making in goals-of-care conversations with elderly patients:concerns and limitations

Aiane Plaisance[1,2], Annie LeBlanc[1,3], Patrick Michel Archambault[1,2]
[1]Université Laval, Québec, Canada. [2]CISSS-CA, Ste-Marie, Canada. [3]Mayo Clinic, Rochester, USA

Background: In goals-of-care conversations, patients' prognosis, level of functional autonomy, values and life goals are discussed in order to inform decisions regarding the use of life-sustaining interventions. Without such discussions, interventions that prolong life at the cost of decreasing its quality may be used without appropriate guidance from patients. Shared decision making (SDM) is recommended to support these goals-of-care conversations. Decision aids (DA), providing unbiased, evidence-based information to patients, can help clinicians engage in SDM. We developed a DA adapted to the context of an Intensive Care Unit (ICU) and a training program that could support these goals-of-care conversations.

Objectives: We aimed to: (i) determine to which extent DA use and training increase intensivists' SDM related skills, (ii) identify elderly patients' concerns regarding goals-of-care and whether and how they are addressed during the conversation, and (iii) identify opportunities for intensivists to improve their SDM skills.

Methods: We conducted a three-phase study using mixed-methods analysis, in a single ICU (Lévis, Canada), recruiting intensivists to participate in the training program and use the DA during real life goals-of-care conversations. We recruited elderly patients (>65 yrs) with whom intensivists intended to engage in a goals-of-care conversation. We videotaped goals-of-care conversations in three phases: (i) prior to the training session and DA availability, (ii) with the DA available for use, (iii) after the training session (and DA availability). We conducted a videographic analysis to assessed the degree of SDM related behaviour displayed (12-item OPTION scale, min-max scores: 0-48) and a retrospective qualitative content analysis to address goals-of-care elements addressed during conversations.

Results: We recruited 7 dyads per phase for a total of 21 patients (71% male; mean age 76 years) and 5 intensivists (80% male). None of the 21 conversations were supported by the DA. Median OPTION score were 12 (interquartile range [IQR]: 10-14), 10 (IQR: 7-12) and 9 (IQR:8-14) for the three phases respectively. Content analysis showed that intensivists tended to focus on medical interventions rather than talking about death/dying and avoided addressing options of not attempting cardiopulmonary/mechanical ventilation. When intensivists talked about death and dying, they used euphemisms and metaphors referring to the human body as a machine.

Conclusion: Our results show that the intensivists never used the DA and avoided discussing death and dying and the option of not attempting cardiopulmonary/mechanical ventilation in the case of a cardiac arrest/ respiratory failure during the three study phases.

The effect of a test ordering software intervention on the prescription of unnecessary laboratory tests - a randomized controlled trial

Carlos Martins
Family Medicine, Department of Community Medicine, Information and Health Decision Science, Faculty of Medicine of the University of Porto, Porto, Portugal, CINTESIS—Centre for Health Technology and Services Research, University of Porto, Porto, Portugal, Porto, Portugal

Objectives: The use of unnecessary laboratory tests is one of the factors that may contribute to overdiagnosis. The way software for electronic health records and laboratory tests ordering systems are designed may influence physicians' prescription. A randomised controlled trial was performed to measure the impact of a diagnostic and laboratory tests ordering system software modification.

Method: Participants were family physicians working and prescribing diagnostic and laboratory tests. The intervention group had a modified software with a basic shortcut menu changes, where some tests were withdrawn or added, and with the implementation of an evidence-based decision support based on the United States Preventive Services Task Force (USPSTF) recommendations. This intervention group was compared with usual software (control group). The outcomes were the number of tests prescribed from those: withdrawn from the basic menu; added to the basic menu; marked with green dots (USPSTF's grade A and B); and marked with red dots (USPSTF's grade D).

Results: Comparing the monthly average number of tests prescribed before and after the software modification, from those tests that were withdrawn from the basic menu, the control group prescribed 33.8 tests per 100 consultations before and 30.8 after ($p = 0.075$); the intervention group prescribed 31.3 before and 13.9 after ($p < 0.001$). Comparing the tests prescribed between both groups during the intervention, from those tests that were withdrawn from the basic menu, the intervention group prescribed a monthly average of 14.0 vs. 29.3 tests per 100 consultations in the control group ($p < 0.001$). From those tests categorised as USPSTF grade D, the intervention group prescribed an average of 9.8 vs. 11.8 tests per 100 consultations in the control group ($p = 0.003$).

Conclusions: Removing unnecessary tests from a quick shortcut menu of the diagnosis and laboratory tests ordering system had a significant impact and reduced unnecessary prescription of tests. The fact that it was not possible to perform the randomization at the family physicians' level, but only of the computer servers is a limitation of our study. Future research should assess the impact of different tests ordering systems during longer periods.

Barriers and facilitators to adopting Choosing Wisely recommendations in physiotherapy

Joshua Zadro[1,2], Christopher Maher[1,2], Rachael Dodd[1]
[1]School of Public Health, Sydney Medical School, The University of Sydney, Sydney, Australia.
[2]Institute for Musculoskeletal Health, Sydney Local Health District, Sydney, Australia

Objectives: The aim of this study is to explore physiotherapist's attitudes, views, and beliefs towards adopting the Australian Physiotherapy Association (APA) Choosing Wisely recommendations.

Method: We will conduct a qualitative study using semi-structured interviews; recruiting 25 registered physiotherapist members of the APA (unless data saturation occurs earlier). The APA will help us purposively recruit participants with different clinical backgrounds and levels of experience as they have access to all members' email addresses and data on demographics, years of clinical experience, and practice setting. An interview guide will be developed following a content analysis of ~300 surveys, previously completed by APA members, regarding their views on the Choosing Wisely recommendations. Interview questions will likely explore the

barriers and facilitators to adopting Choosing Wisely recommendations. Interviews will be conducted face-to-face or via telephone, lasting ~60 minutes. Interview data will be analysed using Framework analysis and will involve transcribing the interview from audio-recordings, identifying codes that could be linked by related concepts, grouping concepts into themes, and charting and mapping until the final themes emerge.

Results: This study will be designed and reported according to the Consolidated Criteria for Reporting Qualitative Research (COREQ). We anticipate to be in the middle of data collection at the time of the conference but are confident we will have some preliminary data to present.

Conclusions: Increasing adoption of the APA Choosing Wisely recommendations is important as the provision of low-value physiotherapy – physiotherapy that provides little-to-no benefit or causes harm, and diverts resources from high-value care – is high. Less than half of physiotherapists use guidelines in practice, and nearly 75% are willing to perform an unnecessary test if requested by a patient. This study will inform on the barriers and facilitators to reducing low-value physiotherapy and will have important implications for refining Choosing Wisely recommendations. Physiotherapists commonly treat people with musculoskeletal conditions (the leading cause of global disability) so replacing low-value physiotherapy with high-value physiotherapy could substantially reduce the enormous global burden of these conditions. The findings from this study will also be relevant to the hundreds of health disciplines worldwide that seek a better understanding of the barriers and facilitators to adopting Choosing Wisely recommendations.

Thyroid hormone treatment and hormone levels among fibromyalgia patients in a Finnish health center -are there indications of overuse?

Aleksi Varinen[1], Elise Kosunen[1], Tuomas Koskela[2]
[1]University of Tampere, Tampere, Finland. [2]Duodecim Medical Publications ltd, Helsinki, Finland

Objectives: From previous studies, it is known that the association between fibromyalgia and thyroid autoimmunity diseases exists. On the other hand, it was recently suggested that in many cases thyroid hormone treatment might be unnecessary. The aim of our study is to explore the thyroid hormone treatment among fibromyalgia patients in cross-sectional study based on data from fibromyalgia patients in the city of Nokia Health Center Finland.

Method; Fibromyalgia patients were searched from the electronic patient records. Patients filled five questionnaires and information from electronic patient records was gathered. Information from thyroid hormone treatment and thyroid-stimulating hormone (TSH) and free thyroxine (T4-V) levels were gained from patient records. We used the Finnish guidelines for hypothyroidism: TSH level over 4.2 mU/l was defined as subclinical hypothyroidism if the T4-V level was normal. The T4-V level under 11.0 pmol/l was defined as central hypothyroidism if TSH level was normal or low. TSH level over 4.2 mU/l and T4-V level under 11.0 pmol/l was defined as overt hypothyroidism. TSH levels between 0.27- 4.2 mU/l and T4-V levels between 11.0 – 22.0 pmol/l were defined as normal thyroid function. Cross-tabulation and Chi-Square test were used when categorical variables were present and two-sample t-test was performed with variables following a normal distribution.

Results: Altogether 208 patients were identified. 103 patients returned mailed questionnaires and were included in the study. 96 patients had fibromyalgia according to ACR 2010 criteria. From

that group 33 (34 %) had thyroid hormone treatment and 63 (66 %) had not. Statistical significance was not found between those taking thyroid hormone replacement and those not taking it in functional ability (p=0.36) or depression (p=0.71). From those 33 patients with thyroxine treatment, 16 had information regarding the initial TSH and T4-V levels before thyroid hormone treatment: Ten (63 %) patients had hypothyroidism based on the laboratory tests. Subclinical hypothyroidism was present in six cases, central hypothyroidism in three cases and as overt hypothyroidism in one case. Of 16 patients with thyroid hormone treatment, six patients (37 %) had normal thyroid function at the beginning of the treatment.

Conclusions: The occurrence of thyroid hormone treatment was much higher in our study population (34%) than in the previous study in Japan (8 %). In our study, over one third of the fibromyalgia patient using thyroid hormone treatment - whose initial thyroid hormone levels were available - did have normal thyroid function. There are likely several explanations to this. One might be that one of the main symptoms of fibromyalgia is fatigue, which also is common with hypothyroidism. Patients are aware of this and some may want to try out the thyroid hormone treatment even though their thyroid function is normal. Further studies are needed to confirm the potential association between functional syndromes like fibromyalgia and inappropriate thyroid hormone treatment.

Barriers and facilitators for de-implementation of unnecessary MRI's and arthroscopic surgeries in patients aged 50+ years with degenerative knee complaints among orthopaedic surgeons

Tessa Rietbergen[1], Ron Diercks[2], Ieke van der Wel[1], Elske van den Akker[1], Rob Nelissen[3], Enrike van der Linden[3], Perla Marang-van de Mheen[1], Leti van Bodegom-Vos[1]
[1]Leiden University Medical Center, Department of Biomedical Data Sciences, Leiden, Netherlands. [2]University Medical Center Groningen, department of orthopaedics, Groningen, Netherlands. [3]Leiden University Medical Center, Department of Orthopaedics, Leiden, Netherlands

Objectives: Approximately 25% of patients aged 50 years and over experience knee pain from degenerative knee disease. Despite the Dutch – evidence based - Choosing Wisely recommendation 'No arthroscopic surgery and no MRI for patients of 50 years and older with knee degenerative knee complaints' (in short: CW-recommendation), orthopaedic surgeons still frequently prescribe an MRI and/ or arthroscopic surgery to these patients. To de-implement the use of unnecessary MRI's and arthroscopic surgeries among patients with degenerative knee complaints, more insight is needed in the factors influencing the uptake of the CW-recommendation by orthopaedic surgeons. Therefore, this study aimed to assess the barriers and facilitators associated with the uptake of the CW-recommendation among Dutch orthopaedic surgeons.

Method: We performed an internet-based survey among 421 Dutch orthopaedic surgeons, of whom 243 (58%) responded. This survey included questions about background characteristics, uptake of the CW-recommendation (4-point scale), and 40 factors possibly influencing the uptake of the CW-recommendation (4-point scale). Factors were based on literature and interviews with orthopaedic surgeons, and classified according to the framework of Grol. Spearman rank correlations were used to identify factors associated with the uptake of the CW-recommendation. Next, as individual factors may be related to each other, we included factors significantly associated with the uptake of the CW-recommendation into a multivariable logistic

regression model ($p<0.05$). For this analysis, we dichotomized the answers on the uptake of the CW-recommendation and the individual factors because of few observations in some cells. Factors increasing the uptake of a CW-recommendation were classified as facilitators, factors decreasing the uptake as barriers.

Results: 200(82%) orthopaedic surgeons reported to adhere to the CW-recommendation for arthroscopic surgery and 197(81%) for MRI. De-implementation of arthroscopic surgery was independently associated with awareness (β=5.00 (95%CI 1.76till8.26), facilitator) and agreement with CW-recommendation (β=4.66 (95%CI 2.30till7.02), facilitator), belief in value of arthroscopic surgery (β=-1.88 (95%CI -3.10till-0.66), barrier) and clinical experience (β=-1.83 (95%CI-2.86till-0.81), barrier), knowledge about latest evidence (β=1.34 (95%CI 0.20till2.49), facilitator) and colleagues following the CW-recommendation (β=1.36 (95%CI 0.22till2.52), facilitator). De-implementation of MRI was independently associated with agreement with CW-recommendation (β=2.69 (95%CI 1.30till4.69), facilitator), belief in value of MRI (β=-0.91 (95%CI -1.76till-0.08), barrier) and clinical experience (β=-0.87 (95%CI -1.67till-0.08), barrier).

Conclusions: Barriers and facilitators were mostly found on the individual professional level, related to the awareness of, belief in and attitude towards the CW-recommendation. For the de-implementation of unnecessary MRI's and arthroscopies in patients aged 50 years and over with degenerative knee complaints, these barriers and facilitators should at least be included in any strategy to be likely to be effective.

11:00 Monday August 20th - Nielsine Nielsen Auditorium - Other

Incidental detection of thyroid abnormalities on low-dose computed tomography in the National Lung Screening Trial

Holli Loomans, Paul Pinsky, Barbara Dunn, Barnett Kramer
NIH/NCI, Bethesda, USA

Abstract in Full: Overdiagnosis, diagnosis of a cancer that would not ultimately cause symptoms or death, has become relatively common in a variety of cancers, including thyroid cancer (ThCa). The trend of ThCa incidence, which has increased substantially without a corresponding change in mortality, typifies overdiagnosis. An analysis of Surveillance, Epidemiology, and End Results (SEER) program data has suggested that overdiagnosis may be responsible for up to 60% of diagnosed papillary ThCa cases, the most common histotype of ThCa. The National Lung Screening Trial (NLST), a randomized trial of almost 54,000 current and former smokers that compared low-dose computed tomography (LDCT) to chest radiography (CXR) for the early detection of lung cancer, is a resource to investigate the potential impact of incidental overdiagnosis of ThCa even when screening for an unrelated cancer in anatomic proximity to the thyroid gland. Previously, Pinsky and colleagues used NLST data to examine renal-related abnormalities outside the intended lung field and found that renal tumors could be incidentally detected by LDCT. A similar phenomenon can be investigated in ThCa, examining the upper portion of the LDCT field. Preliminary analyses show more diagnosed thyroid cancers in the LDCT as compared to CXR arm during the screening phase (first 3 years) of the trial (23 versus 11) and overall (35 versus 25). We will examine the relationship of reported thyroid abnormalities on LDCT

and subsequent diagnoses of thyroid cancer. Specifically, we will assess the proportion of thyroid cancers diagnosed within one year of an LDCT screen that had a reported thyroid abnormality, as well as the overall rate of thyroid abnormalities seen on LDCT screens. We will also examine survival of thyroid cancers by trial arm and mode of diagnosis (observed on LDCT scan or not).

Objectives: Overdiagnosis, diagnosis of a cancer that would not ultimately cause symptoms or death, has become relatively common in a variety of cancers, including thyroid cancer (ThCa). The trend of ThCa incidence, which has increased substantially without a corresponding change in mortality, typifies overdiagnosis. An analysis of Surveillance, Epidemiology, and End Results (SEER) program data has suggested that overdiagnosis may be responsible for up to 60% of diagnosed papillary ThCa cases, the most common histotype of ThCa.
Our main objective is to assess whether thyroid abnormalities discovered incidentally on imaging studies carried out in the National Lung Screening Trial (NLST) are associated with overdiagnosis.

Method; The National Lung Screening Trial (NLST), a randomized trial of almost 54,000 current and former smokers that compared low-dose computed tomography (LDCT) to chest radiography (CXR) for the early detection of lung cancer, is a resource to investigate the potential impact of incidental overdiagnosis of ThCa even when screening for an unrelated cancer in anatomic proximity to the thyroid gland. Previously, Pinsky and colleagues used NLST data to examine renal-related abnormalities outside the intended lung field and found that renal tumors could be incidentally detected by LDCT. A similar phenomenon can be investigated in ThCa, examining the upper portion of the LDCT field.

Results: Preliminary analyses show more diagnosed thyroid cancers in the LDCT as compared to CXR arm during the screening phase (first 3 years) of the trial (23 versus 11) and overall (35 versus 25). We will examine the relationship of reported thyroid abnormalities on LDCT and subsequent diagnoses of thyroid cancer. Specifically, we will assess the proportion of thyroid cancers diagnosed within one year of an LDCT screen that had a reported thyroid abnormality, as well as the overall rate of thyroid abnormalities seen on LDCT screens. We will also examine survival of thyroid cancers by trial arm and mode of diagnosis (observed on LDCT scan or not).

Conclusions: Preliminary data suggest that more thyroid cancers were diagnosed in the LDCT than in the CXR arm. This suggests a possible association of overdiagnosis of ThCa with the more sensitive screening technology (LDCT).

What does “medicine” mean and why does it matter?

Astridur Stefansdottir
University of Iceland, Reykjavik, Iceland

Abstract in Full: We have by now realized that medicine is constantly evolving and spreading into new areas of people’s lives – not always to the benefit of those influenced by it. To be able to evaluate whether this expansion is justifiable it is necessary to have a common understanding on *what medicine is* and *what it should be*. To shed light on those questions I will draw attention to a tension between the goals and ends of medicine, on the one hand, and the societal forces influencing those goals, on the other. Callahan and Hanson (1999) state in their report on the goals of medicine that: “medicine should find its direction by means of a continuing dialogue with society in which each seeks its legitimate sphere, duties, and rights”. I will argue that in order for a “dialogue” of this kind to be fruitful it is necessary to underline the importance of self-awareness of the medical profession. Physicians need to turn their attention for a while away

from their daily work and reflect upon their own position and foundation. The questions they should pose themselves are: *Who am I?* and *What do I stand for?* These questions should be a central practice in the work of every doctor. If this reflection and awareness is not exercised, medical doctors are no longer reminded of their vocation and they lose sight of how to heed this vocation in their work. This neglect can cause the profession to lose her calling and autonomy in the delicate dialogue with society where medical boundaries are spelled out.

Objectives: Sharpen our understanding of the goals of medicine. Such an understanding is necessary to realize what lies within the spheres of medicine and what not.

Method: Philosophical investigation

Conclusions: If medical doctors are to be able to participate in a societal dialogue on the goals and hence limits of medicine, they need to reflect upon their own identity and form an understanding of what they stand for.

MEMORABLE: MEdication Management in Older people: Realist Approaches Based on Literature and Evaluation

Ian Maidment[1], Andrew Booth[2], Sally Lawson[1], Sylvia Bailey[3], Jane McKeown[2], Hadar Zaman[4], Judy Mullan[5], Geoff Wong[6]
[1]Aston, University, United Kingdom. [2]University of Sheffield, Sheffield, United Kingdom. [3]PPI Lead, Birmingham, United Kingdom. [4]Bradford University, Bradford, United Kingdom. [5]University of Wollongong, Wollongong, Australia. [6]Oxford University, Oxford, United Kingdom

Objectives: Older people are major users of medication. Their use of medication is increasing, over the last 20 years the number older people taking at least 5 medicines has quadrupled to nearly 50%. They should be taking medication that gives them more benefits than harms. As people get older or their health changes, the balance between benefits and harms for the medication they take will likely change. Medication management by health care practitioners (e.g. pharmacists, doctors and nurses) if done well is a process that can help older people decide if they need to be on certain medications – thus potentially helping to address the issue of over-treatment. However current evidence indicates that medication management is a complex process that is not easy to do well. The objective of the MEMORABLE project, which is funded by the NIHR, is to understand how medication management works and how it might be implemented effectively.

Method: This project uses a realist approaches to understand how, why, for whom and in what circumstances medicines management works. These are theory-driven ways to make sense of primary and secondary data, and are suited to making sense of complex interventions, such as medicines management, where outcomes are context sensitive.
MEMORABLE uses secondary data from the literature and primary data from up to 60 interviews with older people, carers and practitioners. Data will be synthesised within and across both datasets to set out and refine programme theories, about how medication management works. The findings will be developed into a framework, including guidance and recommendations, to improve practice.

Results: We have identified from the literature and interviews: - 3 candidate programme theories, one each relating to older people, their informal carers and care practitioners, which are being drafted and refined. These programme theories will be explored against a six stage process,

conceived as a framework for analysis of the literature and interview data. The six stage process is set out below: 1. Identifying the problem. 2. Obtaining medication. 3. Starting medication. 4. Continuing to take medication. 5. Reviewing medication. 6. De-prescribing (stopping inappropriate medication). - Emerging mechanisms of interest, trust, needs/concerns, dissonance are being explored.

Conclusions: MEMORABLE, which started in May 2017 and ends in December 2018, is using an innovative methodology, applied collaboratively, to develop an outcome-focused, evidence /experience-based framework to support and enhance medication management. This framework will be meaningful and relevant to those involved in the care of older people, as well as the wider community of interest.

Will the use of high sensitivity troponin result in overdiagnosis of myocardial infarction?

Jenny Doust, Paul Glasziou
Bond University, Robina, Australia

Objectives: The introduction of high sensitivity troponin (hs-Tn) redefines who is diagnosed with myocardial infarction (MI) and may increase the incidence of MI. The Preventing Overdiagnosis Working Group of the Guidelines International Network recently published a checklist for groups seeking to modify a disease definition in 2017. We used the checklist to determine if this clarifies the harms and benefits of the new test in the diagnosis of MI.

Method: We assessed the evidence for hs-Tn against each of the questions in the 8-item checklist:

1. Differences between the previous and the new definition
2. Changes to the incidence and prevalence of the disease
3. Trigger for considering the modification
4. Does it predict clinically important outcomes compared with the previous definition?
5. What is the repeatability, reproducibility, and accuracy of the new disease definition?
6. Benefit: What is the incremental benefit for patients?
7. Harm: What is the incremental harm for patients?
8. What is the net benefit and harm for patients?

Results: hs-Tn has higher analytical sensitivity, which allows myocardial infarction to be diagnosed earlier and treatment commenced. Patients with low levels at 4 hours can be safely discharged. hs-Tn also is predictive of clinically important outcomes and has higher precision than earlier forms of the troponin. Maintaining the 99th centile as the threshold for diagnosis of myocardial infarction and the methods used in studies to determine the reference limit leads to an increase in the numbers of people diagnosed with myocardial infarction, with significant differences between studies depending on the types of patients being tested. The evidence for assessing the benefits and harms of hs-Tn is limited, limited to a before-after study showing an improvement in health outcomes. The introduction of hs-Tn in Australia earlier than in the US may explain the divergence in the incidence of non-ST elevation myocardial infarction in the two countries.

Conclusions: The example of hs-Tn illustrates that rigorous evaluations of disease definitions cannot be isolated from the tests used to diagnose that disease. Despite the significant

consequences from the introduction of the test and the potential for overdiagnosis, it has been introduced because of improvements in analytical performance rather than on a thorough evaluation of potential harms and benefits.

Back to basics: overdiagnosis is about wrongful diagnosis

Bjørn Hofmann
NTNU, Gjøvik, Norway. University of Oslo, Oslo, Norway

Objectives: Clear definitions are crucial for measuring and handling the challenges of overdiagnosis. So far no unified definitions or measures of overdiagnosis exist. This spurs great controversies. Accordingly, the **objective** of this study is to analyze overdiagnosis in terms of *diagnosis* and thereby make it easier to comprehend and measure.

Method: Literature search (for conceptions, definitions and measures of overdiagnosis), qualitative content analysis, and philosophical concept analysis.

Results: Etymologically, overdiagnosis means too much diagnosis. However, etymology also points to another challenge with overdiagnosis: the difficulty to differentiate between what is important and not. One practical reason for this is that new indicators and increased precision de-couples diagnosis from manifest disease resulting in inflation in diagnosis. Thus, overdiagnosis is a result of (conceptually) removing diagnosis from disease. As such, diagnosis deflects medicine from its original goal: to help people who are suffering by identifying what is bothering them and by treating it to alleviate their pain.

Conclusions: We must revive medicine's original goal and re-connect diagnosis to what matters to professionals (knowledge) and to patients (avoid suffering). Only then can we measure and handle overdiagnosis.

Volume and variation of low-value care practices in the Netherlands using health insurance claim data

Tijn Kool, Eva Verkerk, Simone van Dulmen
Scientific Center for Quality of Healthcare (IQ healthcare), Radboud Institute for Health Sciences, Radboud University Medical Center, Nijmegen, Netherlands

Objectives: It is known for several years that a substantial part of healthcare, so called 'low-value care' is unlikely to benefit patients.Low-value care is health care for which there is evidence that it has no or little benefit for the patient, considering the costs, available alternatives, and/or patient preferences. In recent years, there have been several important contributions to quantify the volume of low-value care in the USA and Canada, based on health insurance claims data. As far as we know, there have been no quantifications in Europe in this important debate. This project aims at quantifying low-value care in the Netherlands as a base for attempts to reduce it. The purpose of this study was to measure the volume and variation of low-value care interventions based on claims data.

Method: We conducted a retrospective cohort study with data from one of the four major Dutch healthcare insurance companies. In order to select low-value care, we studied international papers about measuring low-value care and Choosing Wisely recommendations. We selected four recommendations based on their potential societal impact and potential for measuring it in

the available claims data:
Don't order an MRI during natural recovery nor after an operation for a hernia nuclei pulposi (HNP). Don't screen patients over 75 years for colorectal cancer. Don't use a Doppler or plethysmography investigation for the diagnosis of varices. Don't do imaging for non-specific low back pain unless red flags are present.
We developed an algorithm to identify the low-value care in the database with claims data from 2016.

Results

• The population of this health insurance company was comparable with the Dutch population on sex, age and social-economic status, representing a quarter of the Dutch population. • For 78% of the patients with an HNP, doctors ordered an MRI without an operation for HNP afterwards or they ordered the MRI after an HNP-operation. • 8% of the patients were diagnosed with varices with a Doppler or plethysmography investigation. • Less than 1% of the patients that were screened for colorectal cancer were older than 75 years. • The GP's ordered more than 3000 MRI's and more than 45000 X-rays for low back-pain. This was done by a small part of the GP's: more than 80% of the GP's did not order any MRI for low back pain and almost 80% of the GP's ordered 5 or less X-rays for low back pain.

Conclusions: We used Dutch health insurance data to quantify four low-value care practices. One of them showed a significant amount of low-value MRI's for HNP. This recommendation should get priority in de-implementation programmes. There is also overuse of imaging for low back pain. However, it can be attributed to a small number of GP's. The same conclusion can be drawn about not using Doppler or plethysmography investigation for diagnosing varices. For these low-value practices, there is room for improvement by targeting a selected group of healthcare providers. Most hospitals has successfully implemented the recommendation of not screening patients older than 75 years for colorectal cancer. Our research shows that it is crucial to quantify the amount of low-value care because there are major differences between healthcare providers. By quantifying, the low-value care practices that should be de-implemented with priority can be identified.

13:30 Monday August 20th - Niels K. Jerne - De-implementation and the Challenge of Tackling Overdiagnosis at the Level of the Consultation

A Cochrane review of strategies to increase adoption of the Ottawa Ankle Rules and reduce unnecessary imaging

Joshua Zadro[1,2], Amabile Dario[1,2], Christopher Maher[1,2]
[1]School of Public Health, Sydney Medical School, The University of Sydney, Sydney, Australia.
[2]Institute for Musculoskeletal Health, Sydney Local Health District, Sydney, Australia

Objectives: The aim of this review is to establish the effectiveness of existing strategies to increase adoption of the Ottawa Ankle Rules (OARs) and reduce ankle/foot imaging.

Method: We will conduct a Cochrane systematic review according to the Methodological Expectations of Cochrane Intervention Reviews standards, within the Cochrane Musculoskeletal Group. A comprehensive keyword search (combining terms synonymous with "implementation" and "Ottawa Ankle Rules") will be performed in MEDLINE, EMBASE, CINAHL, Cochrane CENTRAL, Scopus and Web of Science from the earliest record to the time of search. Additional articles will be identified by hand-searching references lists and forward searching of included articles. We will include randomised controlled trials, uncontrolled trials, and interrupted time-series investigating strategies to increase adoption of the OARs. The primary outcome will be documented adherence to the OARs. The proportion of unnecessary ankle/foot imaging requests and the total number of ankle/foot imaging requests will be secondary outcomes. Two reviewers will independently perform the selection of studies, extract key data (e.g. trial characteristics, intervention parameters, outcomes), and assess the risk of bias of included studies.

Results: We anticipate to have extracted all study data by the conference and are confident we will be able to present preliminary results.

Conclusions: Nearly 10% of people suffer an ankle injury in their life; but although less than 20% have a fracture, 70-95% receive imaging. In the absence of a fracture, imaging does not inform management and exposes patients to unnecessary/potentially harmful radiation. The OARs are a clinical decision tool with nearly 100% sensitivity for ruling out ankle/foot fractures, thereby indicating those who don't require imaging. These rules have been validated in numerous countries, endorsed in practice guidelines for over two decades, and more recently included in Choosing Wisely lists. Successful implementation of the OARs could reduce unnecessary ankle/foot imaging and time spent in emergency departments. However, the OARs aren't commonly used in practice. Identifying effective strategies to increase adoption of the OARs could reduce unnecessary ankle/foot imaging among various healthcare professionals and guide implementation activities to reduce low-value care across health disciplines.

The Risk Work of Overdiagnosis and Overtreatment

Natalie Armstrong
University of Leicester, Leicester, United Kingdom

Abstract in Full: Overdiagnosis and overtreatment are increasingly being recognised as a significant problem in contemporary healthcare but are yet to receive any significant sociological attention. Overdiagnosis and overtreatment occur when a diagnosis is 'correct' according to current standards but the diagnosis or associated treatment has a low probability of benefitting the patient, and may instead be harmful. There are now numerous resources and tools available that seek to facilitate decision making and support patients and professionals to become comfortable with doing less rather than more. As an example, the Choosing Wisely campaign provides lists of interventions that may be unnecessary and seeks to support patients to ask questions of professionals about the necessity of these. As yet though scant attention has been paid to the 'risk work' that such initiatives create for health professionals. Taking decision making about population-based screening as an example, this paper will explore the work professionals do in order to attempt to navigate the risks and uncertainties of screening for both themselves and their patients. While participation in screening is increasingly being framed in terms of informed choice and emphasis is being placed on decision aids and shared decision making, evidence shows that professionals lack confidence and guidance on how to support

patients to make decisions, feel uncertain of the evidence about harms and benefits of screening, lack external triggers to prompt meaningful engagement with patients, and, importantly, are concerned about being held accountable for decisions. Examination of the interactional work at the 'sharp end' of screening offers important insights into the way in which this form of work gets done by professionals and is experienced by patients, including problematizing ideals of non-directive counselling and the challenges of working with ambiguous or contested knowledges about risk and uncertainty.

Changes in prescription routines for treating hypothyroidism between 2001 and 2015 - a population-based study of 929,684 primary care patients in Copenhagen

Bjarke Medici[1,2], Birte Nygaard[2], Jeppe la Cour[2], Mia Grand[3], Volkert Siersma[3], Dagny Nicolaisdottir[3], Bent Lind[4], Niels Olivarius[3], Christen Andersen[3]
[1]University of Copenhagen, Copenhagen, Denmark. [2]Herlev and Gentofte Hospital, Copenhagen, Denmark. [3]Department of Public Health, Copenhagen, Denmark. [4]Department of Clinical Biochemistry, Amager Hvidovre Hospital, Copenhagen, Denmark

Objectives: Primary objective - To investigate changes in s-TSH threshold used by the general practitioners to initiate L-T4 therapy over the period 2001-2015 in Copenhagen.

Method: Retrospective analysis of all s-TSH measurements between 2001 and 2015 performed at the general practitioners' joint laboratory merged with The Danish Register of Medicinal Products Statistics and The Danish National Patient Registry.
For each year in the period the estimated treatment threshold was calculated from the s-TSH measurements performed in that year as the level of s-TSH where the estimated probability of starting L-T4 therapy is 50%. Also the median level s-TSH at therapy initiation was recorded for each year.

Results: 929,684 individuals with 2,975,277 s-TSH measurements were included for the calculations. While the size and composition of the population serviced by the general practitioners remained virtually unchanged throughout the 15-year period, the number of performed s-TSH measurements increased from 110,886 in 2001 to 292,911 in 2015 and the number of patients initiating L-T4 therapy increased from 786 in 2001 to 1,825 in 2015. The median s-TSH at L-T4 therapy initiation decreased from 10 mU/l (IQR 5.2-29.7) in 2001 to 6.8 mU/l (IQR 5.1-11) in 2015, while the estimated treatment threshold decreased considerably from 28.3 (95% CI 21.0-40.2) mU/l in 2001 to 14.2 (95% CI 12.0-18.0) in 2007 where it remained relatively unchanged for the rest of the study period.

Conclusions: This study performed on a sizeable primary care population demonstrates a considerable fall in the s-TSH threshold when initiating L-T4 therapy in hypothyroid patients, while the number of patients with s-TSH of 5-10 mU/l starting therapy increased considerably. Measuring TPO-antibodies appears to have influenced the s-TSH threshold considerably. The study also reveals that intensified TSH measuring from the 2010 level to the 2015 level did not lead to any further discoveries of hypothyroidism cases suggesting that very few hypothyroid persons remain undiagnosed.

Barriers and facilitators associated with uptake of the Dutch Choosing Wisely recommendation regarding MRI and knee arthroscopy among orthopaedic patients with degenerative knee complaints

Tessa Rietbergen[1], Ron Diercks[2], Ieke van der Wel[1], Elske van den Akker-van Marle[1], Rob Nelissen[3], Enrike van der Linden[3], Perla Marang-van de Mheen[1], Leti van Bodegom-Vos[1]
[1]Department of Biomedical Data Sciences, section Medical Decision Making, Leiden University Medical Center, Leiden, Netherlands. [2]Department of Orthopaedics, University Medical Center Groningen, Groningen, Netherlands. [3]Department of Orthopaedics, Leiden University Medical Center, Leiden, Netherlands

Objectives: About 25% of the patients who are fifty years and older suffer from degenerative knee pain and functional loss. According to Dutch orthopaedic guideline recommendations these patients are not required to be diagnosed and treated routinely with an MRI and/or an arthroscopy. The importance of these recommendations are underscored in the Choosing Wisely recommendation 'No arthroscopic surgery and no MRI for patients of 50 years and older with degenerative knee complaints' (in short: CW-recommendation) issued by the Dutch orthopaedic association. However, despite this recommendation, many patients with degenerative knee complaints still expect their orthopaedic surgeon to order an MRI and/or perform an arthroscopy for diagnosis and/ or treatment. To improve the uptake of the CW-recommendation, this study aimed to investigate which barriers and facilitators were associated with the unnecessary referral for an MRI or an arthroscopy from a patient point of view.

Method: We performed an online questionnaire among 138 Dutch orthopaedic patients ≥50 years with degenerative knee complaints, of whom 116 (84%)responded. This survey included questions about background characteristics, received care and factors influencing unnecessary referral for an MRI/ arthroscopy from a patient perspective (4-point Likert scale). These factors were based on literature and interviews among patients with degenerative knee complaints, and classified according to the framework of Grol. Spearman rank correlations were used to investigate which factors were associated with unnecessary referral for an MRI/arthroscopy. Next, as individual factors may be related to each other, we included factors significantly associated with unnecessary referral for an MRI/arthroscopy into a multivariable logistic regression model ($p<0.05$). For this analysis, we dichotomized the answers on received care and individual factors because of few observations in some cells. Factors increasing unnecessary referral were classified as facilitators, factors decreasing unnecessary referral as barriers.

Results: 74 (64%) patients received an MRI and 58 (50%) underwent a knee arthroscopy. Two factors were independently associated with the unnecessary referral for an MRI: whether the orthopaedic surgeon discussed the added value of an MRI in the care trajectory with the patient (β=3.72 (CI 95% 0.88 till 6.56, barrier) and the patient's own preference for an MRI (β=2.04 (CI 95% 0.19 till 3.90, barrier). Three factors independently increased the likelihood of referral for arthroscopic surgery and were thus barriers for the recommendation uptake: strong preference of orthopaedic surgeon for arthroscopy (β=3.64 (CI 95% 2.19 till 5.08), positive experiences of relatives and friends of the patient regarding an arthroscopy (β=1.68 (CI 95% 0.43 till 2.92) and preference of the patient for an arthroscopy (β=1.39 (CI 95% 0.07 till 2.71).

Conclusions: Barriers and facilitators among patients for unnecessary referral for an MRI or arthroscopy were found at the level of the individual patient (preference of patient), the social context of the patient (positive experiences of relatives and friends), and the individual professional (discussion of added value, preference of orthopaedic surgeon). Strategies to improve the uptake of the CW-recommendation among patients, by reducing the pressure of

patients toward orthopaedic surgeons, should at least address these barriers and facilitators to be likely to be effective in reducing unnecessary MRI and arthroscopy.

The Overdiagnosis of Chest Sepsis in Children: A Quality Improvement Project.

Rebecca Bradford-Duarte, Jillian McKenna, Valerie Milton, Ali Bokhari
Dartford and Gravesham NHS Foundation Trust, Dartford, United Kingdom

Objectives: The aim of this project was to test our clinical impression that a significant number of children and young people are diagnosed with and treated for "chest sepsis" unnecessarily. We also wanted to investigate the hypothesis that children received unecessary investigations, including Chest X-rays (CXRs) and measurement of acute phase reactants such as C-Reactive Protein (CRP), which are not recommended by British Thoracic Society (BTS) guidelines and would have been treated inappropriately with intravenous (IV) antibiotics.

Our main objectives were thus:
1) To design a Quality Improvement Project to investigate whether we overdiagnose chest sepsis, over investigate children presenting with possible signs and symptoms of chest sepsis and over use IV antibiotics
2) To evaluate adherence to BTS and NICE sepsis guidelines
3) To suggest measures to implement to prevent over investigation and unnecessary use of antibiotics in children with chest sepsis

Method: An audit was undertaken within the Paediatric Department in an NHS District General Hospital setting in the United Kingdom. Data was collected retrospectively from May 2017 to January 2018. The inclusion criteria were any child >10 days old admitted to the ward and treated with antibiotics for suspected lower respiratory tract infection (LRTI). Patients with underlying chronic lung disorders (except asthma/wheeze), cardiac conditions, immunodeficiency or those who are immunosuppressed were excluded. Patients were categorised by age (<5, 5-12, >12) and then stratified according to sepsis risk, as per NICE and BTS guidelines. The data were then analysed assessing adherence to these guidelines during clinical assessment and management following presentation with signs and symptoms of a LRTI.

Results

• Data from 156 patients were collected: 132 aged under 5, 21 aged 5-11 and 3 aged 12-17. • 10% were low, 21% moderate and 69% high risk for sepsis on presentation. • 36 patients were noted to have low oxygen saturations on admission, of which 64% had co-existing wheeze. • 66% of patients, of which 73% were high risk and 27% were low or moderate risk for sepsis were treated with IV antibiotics, most commonly Ceftriaxone. • 33% were treated with oral antibiotics, most commonly Co-amoxiclav. Only 9.8% of this group were treated with oral Amoxicillin as per BTS guidelines. • The majority (88%) had a CXR. Half were reported as normal, of which 71% were treated with IV antibiotics. 78% of those treated with oral antibiotics had a CXR. • 78% of patients had blood tests, which all included a CRP, which is not recommended to be routinely measured.

Conclusions: First phase audit results support the hypothesis that children presenting with signs and symptoms of a possible LRTI are overdiagnosed and over investigated. 61% of children were under the age of two and treated with antibiotics even though BTS guidelines suggest this age group is much less likely to have a bacterial infection. Over half of those with low oxygen saturations had co-existing wheeze, suggesting over-treatment of patients with wheeze for LRTI.

88% of our patients had CXRs, which is contrary to BTS guidelines, that suggest routine CXRs are unnecessary. 71% of children reported to have a normal CXR received IV antibiotics. We plan to implement a change of practice within the Paediatric department to prevent unnecessary investigations, notably CXRs and blood tests and to reduce unnecessary antibiotic use that comes with long term risks for the wider population. Following this we plan to re-audit.

Evaluating the five wise choices in wound care: are nurses and physicians choosing wisely?

Eva W Verkerk, Getty Huisman-de Waal, Lydia Overtoom, Rudolf B Kool, Simone A Van Dulmen
Radboud university medical center, Radboud Institute for Health Sciences, IQ healthcare, Nijmegen, Netherlands

Objectives: In 2014, the five wise choices for acute wounds were released in the Netherlands. These five evidence-based recommendations are meant to engage caregivers and patients in conversations about unnecessary procedures. Following these wise choices increases patient safety and reduces waste of resources. The wise choices are:

Do not clean the wound with saline
Do not soak the wound in cleanser
Do not use bandages on a primary closed wound
Do not use expensive bandages when gauze suffices
Do not discharge a patient without giving instructions

We evaluated nurses' and physicians' awareness of the five wise choices for acute wounds and their perceived barriers and facilitators for implementation in the Netherlands.

Method: An online survey was sent to nurses and physicians working in the emergency department and surgical wards in 13 hospitals in the Netherlands, and to members of the Dutch Association for Emergency physicians. In addition, we interviewed patients and professionals on their perceived barriers and facilitators to following the wise choices.

Results: The survey was completed by 160 nurses and 60 physicians. The majority indicates that they are aware of the wise choices. However, in practice they do not always choose wisely, in particular for recommendations 1 and 3 and in emergency care. We interviewed 17 nurses and 6 physicians. In addition to the lack of knowledge on wound care, important barriers to choosing wisely are their environment and the habits in their teams. Another important barrier is that nurses and physicians assume that patients want their wound cleaned with saline and covered with a bandage. In our 20 interviews with patients we found that while some patients indeed prefer this, the majority trusts the expertise of their nurse or physician and accepts their choice.

Conclusions: Although the majority of nurses and physicians are aware of the wise choices for acute wounds, they do not always choose wisely in practice. The factors that influence this most are the lack of knowledge, environment, habits, and perceptions of the patients' preferences. There is room for improvement and we recommend to change the environment to facilitate the wise choices, educate nurses and physicians about wound care and the actual preferences of the patient and specialize a nurse in wound care on every department.

13:30 Monday August 20th - Nielsine Nielsen Auditorium - Turning Citizens In to Patients Unnecessarily

Use of regular feedback of ranked performance data to Family Medicine trainees in the Sultanate of Oman, to support a reduction in the overuse of resources.

Robin Davidson
Sultan Qaboos University, Muscat, Oman

Objectives: A previous audit comparing the use of resources between final year trainees in Family Medicine (FM) in Oman and General Practice (GP) in the United Kingdom (UK), showed that, in patients attending for a same day/walk-in appointment, trainees in Oman prescribed significantly more items than those in the UK. The number of investigations ordered was the same although the average age of the patients in Oman was much lower.
The purpose of this project was to reduce what was considered to be over use of these resources.

Method: One cohort of 18 FM trainees (2015-2016) in Oman was given anonymised, ranked, performance data, for all the prescription items and investigations they had ordered for walk-in patients the previous month. They were encouraged to think about rational use of these resources and to compare their performance with their peers. The mean prescription items/patient and the mean investigations/patient across the group of residents was calculated with 95% confidence intervals for each outcome. As the number of trainees in the cohort was small and they all worked closely together it was not possible to split them and have a contemporaneous control group. As an alternative control group, the same analysis of prescription and investigations use was carried out for the cohort of FM trainees working the previous year (2014-2015). This group also contained 18 trainees, matched for age, sex and experience.

Results: Comparing the five months before and after this intervention, the total number of patients seen was 4228 and 2893 respectively. The total number of investigations fell from 1.27 items/patient(CI 95% 0.95-1.59)to 1.16 items/patient(CI 95% 0.96-1.36), a reduction of 8.6% and the number of prescriptions fell from 1.79 items/patient (CI 95% 1.66-1.92) to 1.57 items/patient (CI 95% 1.48-1.66), a reduction of 12.3%. In the 2014-15 cohort, there was no such reduction. Indeed, there was a tendency toward an increase in use of resources. The total number of investigations increased from 1.01 to 1.22 items/patient and the number of prescriptions increased from 1.82 to 1.91 items/patient. (CI 95% available but no space to include them here)

Conclusions: Discussion of the rational use of the resources is an important part of the education of FM trainees in Oman. This study has demonstrated that giving regular (monthly) feedback of anonymised, ranked performance data to these trainees supports a significant reduction in their use of resources. These data were gathered manually. A system to automatically generate and feedback this information should be developed.

A randomised on-line survey to explore how disease labels, psychological traits and illness risk perceptions affect behavioural intentions.

Rae Thomas[1], Mark Spence[2], Rajat Roy[2], Elaine Beller[1]
[1]Centre for Research in Evidence-Based Practice, Bond University, Gold Coast, Australia. [2]Bond Business School, Bond University, Gold Coast, Australia

Objectives: Negative consequences of medical labelling have been reported in research literature[1] and differences in an individual's intention to undertake further testing have been shown in studies that randomly assigned participants to labelled and unlabeled hypothetical medical scenarios[2]. When given information about overdiagnosis of polycystic ovary syndrome after medical scenarios, all groups (irrespective of whether the medical label was used) reduced their intention to have follow-up tests[3]. What is unknown, is how an individual's psychological traits such the predisposition to seek medical care, emotional stability, extraversion, and locus of control and their perceptions of risk and stigma toward the health condition might impact a person's decision to undertake further tests when exposed to either a labelled or unlabeled medical scenario.

Method: A randomised controlled online survey was distributed to 256 participants aged 45-70 years in three countries (Australia, Ireland and Canada). Participants completed trait-based measures including health locus of control, regulatory focus (promotion/prevention), self-perceptions of medical usage, and health risk orientation. Participants were then randomised to receive two scenarios (stratified for age, gender and country). Scenarios described the outcome of a recent health test using either medical terms ("labelled") or condition descriptions ("descriptive"). There were "labelled" and "descriptive" scenarios for four health conditions known for controversies over threshold changes (pre-diabetes, mild hypertension, mild hyperlipidaemia, and chronic kidney disease stage 3a). Each scenario informed participants they were close to the threshold and gave participants information about overdiagnosis. Post-scenario, participants rated their perception of illness risk and stigma. Between group differences for intentions to pursue a follow-up test was the primary outcome. We also assess what traits may have impacted their decision.

Results: Preliminary analyses suggest that after adjusting for two scenarios per person, there was no significant difference between the "labelled" (n=129) and "descriptive" (n=127) groups in their intention to have follow-up tests (95% CI -0.77 to 0.33 points). In a multivariable regression model, there was a significant increase in intentions to pursue further tests when participants were: high users of medical interventions (p

Conclusions: Previous research has consistently found a labelling effect, but the cause of the effect is unclear. Our findings both contrast and expand upon previous research. We analyzed four different health conditions with controversies around the threshold. All scenarios were "close to the cut-off". It is unclear why our "labelled" and "description" scenarios did not produce significant differences in intentions to undertake further tests, as has been found in previous studies. It may be that by first eliciting psychological trait measures related to health we cued participants to think about their health, which counteracted labelling effects. Future studies might reverse the data collection order (respond to illness scenarios prior to answering trait-based measures) to explore whether the labelling effect reappears. If this were the case, it would suggest that how we communicate to people about their health is more challenging than whether we label the health condition or not.

Women's acceptance of overdetection in breast cancer screening: can we assess harm-benefit trade-offs?

Anne Stiggelbout[1], Jolyn Hersch[2], Tessa Copp[2], Gert-Jan Liefers[1], Gemma Kacklyn[3], Jesse Jansen[3], Kirsten McCaffery[3]
[1]Leiden University Medical Center, Leiden, Netherlands. [2]Th University of Sydney, Sydney, Australia. [3]University of Sydney, Sydney, Australia

Abstract in Full: Aim was to assess 1) acceptance by women in the general public of overdetection (OD) in breast cancer screening, for different scenarios describing treatment after OD (mastectomy; lumpectomy; lumpectomy followed by radiotherapy; lumpectomy followed by radiotherapy and hormonal therapy), and 2) correlates of acceptance.
We recruited a random sample of Dutch and Australian women, stratified by age (45-75), through an online survey company. We assessed women's preferences for either *no screening* or *screening*, for each treatment scenario (randomized order), using pictographs of important screening outcomes (breast cancer deaths averted, overdetected, cured irrespective of screening) for 1000 women screened over 25 years. For each scenario, we presented five pairs of numbers: deaths avoided (remained constant at 5) + number of women overdetected, varied to 0, 5, 30 (Australian model), 15 (Independent Panel's estimate), and 2 (Dutch model). We assessed screening history, breast cancer in family/friends, prior attitudes, social norms, risk perception, worry, anticipated regret, perceived seriousness of diagnosis, comprehension of OD, health literacy, and numeracy. Using MANOVAs we compared the acceptance between treatment scenarios, and assessed associations and interactions with the other variables.
We received 854 responses, and deleted 51 due to inconsistent replies, leaving 400 Australian and 403 Dutch respondents. The majority (49-57%) strongly preferred screening, even at the 5:30 ratio of deaths-avoided:OD; 10-15% would never screen regardless of the ratio; around 15% would screen at 5:5. Only a slight effect (p=0.08) was seen for the effect of scenario, with lumpectomy having the highest acceptance of screening, then lumpectomy-radiotherapy, mastectomy, and lumpectomy-radiotherapy-hormones. Preference for screening was associated with having children and screening history. No effects were seen for health literacy and numeracy, but better comprehension of OD correlated with a less strong preference for screening (r= -0 .35 to -0.40). The strongest psychological predictors were prior screening attitude (.41-.45), subjective norm (.35-.39), and anticipated regret (0.37-0.39). In a MANOVA, only comprehension, attitude and subjective norm remained significant.
We found very strong preferences for screening, even with 6 cancers overdetected for every death averted. To a large extent this was explained by poor comprehension of OD.
This finding, and the insensitivity to treatment burden point to a general difficulty of questioning women about screening. Women have an overwhelming and uncritical feeling that screening is always beneficial. Better information and education are needed. Only once we create a good understanding of the drawbacks of screening can we assess a true harm-benefit trade-off.

09:00 Tuesday August 21st - Niels K. Jerne
De-implementation and the Challenge of Tackling Overdiagnosis at the Level of the Consultation

A collective approach to recommendations for de-implementation or reform of a national screening program. A case study from the French civic and scientific

inquiry into breast cancer screening

Alexandra Barratt[1], Karsten Juhl Jørgensen[2], Philippe Autier[3]
[1]Wiser Healthcare, University of Sydney, Australia. [2]Nordic Cochrane Centre, Copenhagen, Denmark. [3]International Prevention Research Centre, Lyon, France

Objectives: In October 2016 the French Minister of Health released the report of an independent inquiry into mammography screening, presenting two options for the future: end the national breast screening program, or end the current program and put in place a radically reformed program. Our objective was to critically analyse the report and the process that led to it.

Method: In September 2015, the Minister of Health ordered a civil and scientific inquiry, overseen by an independent committee of leading health (oncology, general medicine, epidemiology, public health) and social science professionals (anthropology, law, economics, history of science and bioethics), all free of ties to breast screening. As well as evidence reviews on specific questions, two substantial consultations were held: a civil consultation with a group of 27 women from different regions and diverse socio-economic groups, and a parallel consultation with a group of 19 independent health professionals. Each of these in-depth consultations took place over 5 days.

Results: The citizen consultation recommended against keeping the program as currently implemented. The health professionals consultation recommended continuing the program, but with major reforms. The steering committee itself found an unexpectedly intense scientific controversy about mammography screening, noting limited evidence on key breast screening outcomes, and incomplete knowledge of the natural history of breast cancer. In its final report, the steering committee recommended that if the program were reformed rather than ended, there was a need for neutral, complete information; a research program into the natural history of breast cancer(s); improved program evaluation; an end to screening average risk women under 50 years of age.

Conclusions: The French inquiry is strikingly different from other recommendation panels (eg USPSTF, American Cancer Society, International Agency for Research on Cancer) in process and result. Panels rarely seek the values and preferences of citizens in formulating recommendations yet screening impacts the lives of asymptomatic citizens. That the French inquiry included perspectives of citizens and unconflicted professionals may explain the difference in outcome. Seeking informed citizens' views and preferences through in-depth, community deliberations can enable meaningful information sharing and exchange of views between citizens of diverse background. Such approaches may be useful to inform and support screening de-implementation processes.

Clinician, patient and general public beliefs about diagnostic imaging for low back pain: A qualitative evidence synthesis

sweekriti Sharma[1], Adrian C Traeger[1], Benjamin J Reed[2,3], Denise A O'Connor[2,3], Tammy C Hoffmann[4], Carissa Bonner[1], Chris G Maher[1,5], Rachelle Buchbinder[2,3]
[1]School of Public Health, University of Sydney, Sydney, Australia. [2]Monash Department of Clinical Epidemiology, Cabrini Institute, Malvern, Victoria, Australia. [3]Department of Epidemiology and Preventive Medicine, School of Public Health and Preventive Medicine, Monash University,, Melbourne, Victoria, Australia. [4]Faculty of Health Sciences and Medicine, Centre for Research in

Evidence-Based Practice, Bond University, Gold Coast, Queensland, Australia. [5]Institute for Musculoskeletal Health,

Objectives: Low back pain (LBP) accounts for an estimated 83 million years lived with disability every year, making it the number one cause of disability worldwide. Australian Medicare data suggests that $243 million was spent on spinal radiographs in 2016/17, despite it being unnecessary for the majority of patients. Unnecessary diagnostic imaging places patients at risk of overdiagnosis. This can occur when diagnostic imaging detects normal age-related changes, such as intervertebral disc degeneration, that are common in asymptomatic people. Despite harms of unnecessary imaging, it is unclear what patients, clinicians and general public believe the role of diagnostic imaging is in the management of low back pain. The objective of this study was to systematically review the qualitative research that has explored clinician, patient and general public beliefs about diagnostic imaging for LBP.

Method: Studies that used qualitative methods were included if they interviewed the general public about LBP management, clinicians who treat LBP, and/or patients with LBP and explored beliefs about diagnostic imaging for LBP. We defined beliefs as propositions or set of propositions held true, and imaging as any diagnostic imaging (e.g. X-ray, CT scan, MRI and so on). Studies were ineligible if they were not published in English. We searched 5 databases (MEDLINE, EMBASE, CINAHL, AMED, PsycINFO). Our search strategy was comprehensive rather than purposive, so that we could locate all relevant studies representing the phenomena of interest. Two reviewers independently screened articles and extracted data. Initial synthesis of the results was done by open coding results into key themes and subthemes. The protocol was registered on PROSPERO (CRD42017076047).

Results: Of 6157 studies from the search, we screened 429 full texts, and 63 studies met our inclusion criteria. Of those included, 32 studies were with patients, 22 were with clinicians, four were with community members with LBP, five were with a mixed sample and none were with general public. We identified five key themes from our preliminary synthesis: (1) clinical presentation justifies scanning eg, severe, worsening, long lasting and relapsing pain warrants scanning; (2) perception that scans have benefits to clinicians eg, help make correct diagnosis, locate the source of pain, reduce risk of litigation if they did not use imaging; (3) perception that scans have benefits to patients eg, provide reassurance that there is no serious problem, provide evidence that they are in pain; (4) scans have potential harms eg, unnecessary disease labelling, radiation exposure; (5) health system drivers eg, ordering tests because health insurance/social security require it.

Conclusions: Our review identified that clinicians, patients and community members with LBP have misconceptions about the value of imaging. The belief that imaging is an important diagnostic tool for LBP was commonly expressed among clinicians, patients and community members with LBP. These beliefs are at odds with evidence that diagnostic imaging often adds little value to clinical decision-making or patient outcomes. Therefore, public health interventions such as education campaigns should target mistaken beliefs that imaging is part of standard procedure for the assessment of low back pain and are more informative than clinical evaluation.

Evaluating two decision aids for Australian men to support informed choice about prostate cancer screening

Kristen Pickles, Alex Barratt, Jolyn Hersch, Kevin McGeechan, Kirsten McCaffery
University of Sydney, Sydney, Australia

Objectives: Prostate-specific antigen (PSA) screening is highly contested, with well-documented harms. Public and professional communication has historically been highly varied and potentially confusing. Government-endorsed Australian guidelines make no recommendation to routinely offer PSA screening, but instead advocate informed decision-making, including advising clinicians to 'offer evidence-based decisional support to men considering whether or not to have a PSA test'. There is little evidence of widespread adoption of decision support in primary care. This study will assess and compare the acceptability and comprehensibility of two evidence-based decision aids (long and abbreviated) among a community sample of Australian men with varying educational backgrounds.

Method: A community sample of 3000 men aged 45-60 years will be recruited by a company experienced in panel survey sampling to participate in this study. Quota sampling will ensure strong representation of men with lower levels of educational attainment. Men will be randomised to view either the full decision aid (adapted from a mammography decision aid previously evaluated by RCT (Hersch et al, Lancet 2015)) or a new abbreviated version of the decision aid, and then answer standardised questions in an online survey. Analyses will compare the two aids with respect to ease of understanding, acceptability and confidence in decision-making, and men's attitudes and intentions to screen.

Results: At this time, results of the survey are pending. We will report the percentages of men who find each decision aid understandable and acceptable to use, stratified by level of educational attainment. We will test to see whether understanding and acceptability differ according to whether men have had a PSA test. We will assess, as secondary outcomes, measures of decision quality including decisional conflict (uncertainty in decision making) and informed choice. A well-established measure of informed choice will be used; it combines measures of men's knowledge, attitudes to PSA testing and intentions to screen to determine the proportion of men who were able to make an informed choice.

Conclusions: A thorough understanding of men's needs relating to decisions about PSA testing, and assessment of the comprehensibility and usability of materials, are critical elements to the development of a high quality decision aid. This study will provide important evidence regarding the effectiveness of a short summary version compared with a longer, more comprehensive decision aid. Findings of this study will inform the implementation of the decision aid in clinical practice nationally and add to global knowledge on decision aid design and performance.

"Practicing Wisely": a hands-on workshop to decrease overuse at the level of the consultation in primary care.

Guylene Theriault[1,2], Genevieve Bois[3], Rene Wittmer[3,4], Geraldine Lachance-Fortin[3]
[1]Quebec Medical Association, Montreal, Canada. [2]McGill University, Montreal, Canada. [3]Quebec College of family physicians, Montreal, Canada. [4]Montreal University, Montreal, Canada

Abstract in Full: In 2011, the *Ontario College of family physicians* (OCFP) developed a workshop entitled "Don't Just Do Something, Stand There". This workshop for primary care physicians quickly became very popular. It focused on helping them negotiate age-appropriate screening

options and address different health topics without medicalizing patients or turning healthy people into sick people.
The workshop was revamped and further elaborated in 2016 and is now given in most provinces in Canada. The focus is the same as the previous workshop but it aims to be even more practical and to give more tools to family doctors so they feel able to apply what they learned in daily practice. It aims to give them the ability to have what are sometimes difficult or complex conversations with their patients, and to practice shared-decision making.
The workshop consists of four 1h30 modules introducing a reflection on overimaging, overprescribing, cancer screening and overmonitoring of disease. The themes covered are as follows: imaging for low back pain and ovarian cysts, prescribing PPI and statins in primary care, screening for prostate cancer, breast cancer and colon cancer, the evidence-based periodic health visit, monitoring diabetes, deprescribing in the elderly and communication with the patient on issues of overuse.
The workshop begins with a reflection on the impact of "too much medicine" and continues through the topic via different clinical scenarios. Participants are guided through the cases and given different web sites and resources to explore in each section. Tutors make participants work in small and large groups to reflect on their own practice.
We surveyed the participants before the workshop and at 6 and 12 weeks after the workshop specifically asking them if they had changed certain practices and if so which one. If they didn't we tried to identify the barriers they encountered. If they did we ask them to describe their experience and also their perception of the patients' reaction to these changes.
The goal of our presentation is to inform, share and discuss our results while reflecting on our experience. Quebec is the Canadian province that has given the most workshops in 2017 and the demand for the workshops is sustained.

Objectives
-Give an outline of the workshop
-Share the results of the feedback given by participants as well as the impacts described in the participants answers
-Review which general themes seem to have greater retention in our participants, and which themes remain more challenging to address.

Method: For each workshop, the participants filled a questionnaire before and immediately after the workshop. They also had to choose two practices they would change and after 6 and 12 weeks, they were asked to provide us with feedback using structured and semi-structured questions on their successes or challenges.
We analyzed the responses to these questionnaires. We used quantitative measures when appropriate and analysed recurrent themes on answers asking them to describe how they were successful and/or the difficulties encountered in changing practices.

Results: Feedback was in general very favourable to this workshop, and most participants reported having used the tools and concepts that were discussed during the workshop. Many themes that can lead to overuse are discussed through the workshop. We have seen that deprescribing PPIs, shared decision making for statins in primary prevention and cancer screening counselling seem to be the practice change chosen the most often by participants. Some logistical issues are viewed as barriers but also, contradicting opinions by specialist colleagues can sometimes make applying what was learned difficult. It seems though, as per the participant's feedback, that patients appreciate these changes.

Conclusions: We present Québec's experience with "Practicing Wisely" a hands-on workshop aimed at decreasing overuse. We opted, which is different than in other provinces, to ask about practice changes in a descriptive manner, trying to better qualify successes and identify barriers. Our future goal is to prepare tools and material to address issues raised, so we can make the next version of the workshop even more relevant and useful. We believe such an interactive workshop, based on small and large group discussions, simple but common clinical scenarios, and use while in session of different resources is one possible (albeit one of many) way to help change in practice at the consultation level.

Preventing overtreatment in older age by prioritising medicines

Kristie Rebecca Weir[1], Carissa Bonner[1], Vasi Naganathan[2], Kirsten McCaffery[1], Andrew McLachlan[3], Stacy Carter[1], Lyndal Trevena[4], Jim Colvin[5], Debbie Rigby[6], Jesse Jansen[1]
[1]Wiser Healthcare, Sydney School of Public Health, The University of Sydney, Sydney, Australia. [2]Centre for Education and Research on Ageing (CERA), Ageing and Alzheimer's Institute, Concord Hospital, The University of Sydney, Sydney, Australia. [3]Faculty of Pharmacy, The University of Sydney, Sydney, Australia. [4]Sydney School of Public Health, The University of Sydney, Sydney, Australia. [5]Health Consumers New South Wales, Sydney, Australia. [6]Consultant clinical pharmacist, Brisbane, Australia

Objectives: Reducing inappropriate and harmful medicines in older people with comorbidity is a major public health issue. GPs and pharmacists play a pivotal role in optimising medicines and reducing treatment burden. This includes reviewing medicines to confirm they are (still) needed/safe and align with patient goals. Medication reviews conducted by pharmacists aim to resolve problems associated with polypharmacy and involve recommendations to the patient's GP on how to do this. However, there are barriers to uptake and limited evidence underpinning medication reviews. Our study aimed to explore GP and pharmacists' views on the role of patient goals and preferences in medication reviews and deprescribing.

Method: Semi-structured interviews were conducted with GPs (n=25), pharmacists (n=11) who refer to and conduct medication reviews; with varying background characteristics and experience levels. Transcribed audio-recordings of the interviews were coded using Framework Analysis.

Results: Most participants recognised the importance of getting a sense of patient's goals/preferences in relation to medicines. However, often these discussions were initiated by the patient, some participants tended to interpret goals only in terms of clinical outcomes (ie therapeutic index) and goals were not routinely discussed. Overall, participants considered medication reviews useful for preventing medication errors, reducing treatment burden and to support deprescribing. However, some GPs expressed doubts about the effectiveness of reviews, suggesting they may be limited in what they can achieve as only one aspect of patient care. Important barriers were highlighted related to patients (resistance to the review, misunderstanding about the aim); pharmacists (importance of patient goals/preferences being reflected in recommendations, reviews targeted to high-need patients with recommendations actioned by GPs for patients to benefit) and GPs (limited information from the GP at referral, limited follow-up afterwards and recommendations not being actioned).

Conclusions: Participants reported a range of benefits for medication reviews in terms of optimising care and reducing medication burden for older people. However, we identified limitations importantly goals and preferences may not be routinely incorporated in decision-

making about medicines and some GPs are doubtful as to the clinical benefit of medication reviews. This may influence barriers reported by participants such as patient resistance to reviews, misunderstanding of the purpose and limited follow up afterwards. This study highlights the challenges GPs and pharmacists face in collaboratively managing polypharmacy and providing care that aligns with patient goals and preferences.

Design and rationale for an implementation trial to improve care for low back pain in emergency departments

Gustavo Machado[1], Bethan Richards[2], Chris Needs[3], Rachelle Buchbinder[4], Ian Harris[2], Kirsten Howard[1], Kirsten McCaffery[1], Laurent Billott[5], Matthew Oliver[2], Eileen Rogan[2], Rochelle Facer[2], David Lord Cowell[6], Chris Maher[1]
[1]University of Sydney, Sydney, Australia. [2]Sydney Local Health District, Sydney, Australia. [3]Sydney Local Health District, Sydney, Australia. [4]Monash University, Melbourne, Australia. [5]University of New South Wales, Sydney, Australia. [6]Dubbo Base Hospital, Dubbo, Australia

Objectives: When low back pain is managed in the emergency department overdiagnosis and overtreatment are common. Key problems are overuse of imaging, opioid medicines and hospital admission. The overall aim of the Sydney Health Partners Emergency Department (SHaPED) trial is to implement and evaluate an evidence-based model of care for acute low back pain. The outcomes of the trial reflect the key messages in the model: (1) patients with non-specific low back pain do not require imaging; (2) where medicines are used, simple analgesics should be the first option; (3) patients with non-specific low back pain should be managed as outpatients.

Method: A stepped-wedge cluster randomised controlled trial will be conducted to implement and evaluate the model of care for acute low back pain at four emergency departments in New South Wales, Australia. Clinician participants will be emergency physicians, nurses and physiotherapists. Codes from the Systematised Nomenclature of Medicine—Clinical Terms—Australian version will be used to identify low back pain presentations. We anticipate ~2000 patient participants. The intervention, targeting emergency clinicians, will comprise educational materials and seminars and an audit and feedback approach. Health service delivery outcomes are routinely collected measures of imaging (primary outcome), opioid use and inpatient admission A random subsample of 200 patient participants from each trial period will be included to measure patient outcomes (pain intensity, physical function, quality of life and experience with emergency service). An economic evaluation will be undertaken from the health system perspective

Results: The SHaPED trial received ethical approval from the RPAH HREC (reference: X17-0043). The trial is registered with the Australia New Zealand Clinical Trials Registry: ACTRN 12617001160325.The conference presentation will discuss the design and rationale for the trial.

Conclusions: We hypothesised that active implementation of an evidence-based model of care for low back pain will improve emergency care by reducing inappropriate overuse of tests and treatments (i.e. imaging, opioids, admission to hospital) and improving patient outcomes.

09:00 Tuesday August 21st - Nielsine Nielsen Auditorium - Turning Citizens In to Patients Unnecessarily

Opportunity to reduce overdiagnosis and overtreatment of cervical abnormalities: de-intensifying the Australian National Cervical Screening Program

Rachael Dodd, Brooke Nickel, Jolyn Hersch, Carissa Bonner, Sally Wortley, Kirsten McCaffery
University of Sydney, Sydney, Australia

Objectives: Reductions to the Australian Cervical Screening program from two yearly to five yearly primary HPV screening, from the increased age of 25 (previously 18), offers an opportunity to reduce overdiagnosis and overtreatment of cervical abnormalities and its associated harms. The changes, implemented in December 2017, will likely reduce overtreatment particularly in women under 25 years, for whom HPV is usually transient and cervical abnormalities are more likely to regress. Strong resistance has been demonstrated previously in both the US and UK, when efforts have been made to de-intensify screening programs. We explored the views and attitudes of screening aged women to the changes, and their understanding of the rationale behind the changes with the aim to develop public communication that mitigates concerns about de-intensifying cancer screening programs.

Method; Six focus groups were conducted in November 2017 in the Sydney area, with 49 women aged 18-74, recruited through random digit dialling. Eligible women had not previously been diagnosed with cervical cancer or had a hysterectomy. Focus groups were structured around a presentation of information about the changes to the cervical screening program, with discussions of the information facilitated throughout. The focus groups were analysed thematically.

Results: Only a third of women had heard something about the changes, mainly either the increased interval between tests or the increased starting age. Questions were raised about the test, with awareness of human papillomavirus (HPV) evidently limited. Explaining clearly the difference between the two tests (Pap smear vs HPV test), and that the procedure is exactly the same for both tests, was important to women. Understanding of the new test was key to alleviate concerns about the extended screening interval. Communicating the rationale of the changes to women, in a clear and coherent way, was paramount for acceptance of the new program. Information presented about the potential of over-detection and unnecessary tests did not appear to concern women greatly, but one group showed some surprise shown of the harms that could result from treatment of cervical abnormalities.

Conclusions: De-intensifying screening programs should be accompanied by clear and coherent communication of the changes, to limit negative concerns from the public. Communicating about over-treatment in this cohort of women did not appear to spark concern, but perhaps further focus on communication of this would create a greater understanding in women about the benefits and harms of screening. The findings of this study contribute to an understanding of what information women seek about changes to the Australian Cervical Screening Program.

“People say it is dangerous”. Psychosocial Effects of Labelling People with Mild Hypertension: a Qualitative Study.

János Valery Gyuricza[1,2], John Brodersen[2], Lucas Bastos[1], Ana Flávia D'Oliveira[1]
[1]Departamento de Medicina Preventiva da Faculdade de Medicina da Universidade de São Paulo, São Paulo, Brazil. [2]Forskningsenheden for Almen Praksis Center for Sundhed og Samfund Københavns Universitet, Copenhagen, Denmark

Abstract in Full: Mild hypertension is a very common asymptomatic condition present in people at low cardiovascular risk. These people represent the greatest amount of those labelled with hypertension. Best available evidence does not support pharmacologic treatment for mild hypertension in order to reduce cardiovascular mortality. Moreover, overdetection of hypertension is also taking place supported by public awareness campaigns, screening programmes, easy access to testing and poor clinical practice, enhancing the overdiagnosis potential. Evidence suggests that this diagnosis is harmful in many ways: from side effects of treatment to psychosocial consequences of labelling. However, the psychosocial effects of labelling are potential harms that still need better understanding.

The aim of this study was to assess the potential psychosocial labelling effects in people diagnosed with mild hypertension.

10 Single and 5 focus group interviews were conducted in São Paulo, Brazil among persons labelled with mild hypertension without comorbidities. Volunteers were selected among general population either from a list of patients from a primary healthcare clinic or social media and social network, with a broad range of characteristics including sex, age, level of education, ethnic origin and time from diagnosis. Data was subjected to qualitative content analysis by three of the authors independently followed by discussions to generate themes and categories.

Preliminary results confirm that the label has impacts in a broad range of psychosocial dimensions, e.g. trust in own body, psychological reactions, social stigma and overmedicalization. Although informants had a broad range of characteristics, they shared similar stories, understandings and effects of labelling. The empirical material is still being analysed and final results will be presented at the conference.

The diagnosis of mild hypertension is a significant milestone and has impact on daily life. Most of the impact is regarded as negative psychosocial consequences or harm; however, sometimes it might be ambiguous.

Objectives: The aim of this study was to assess the potential psychosocial labelling effects in people diagnosed with mild hypertension.

Method: 10 Single and 5 focus group interviews were conducted in São Paulo, Brazil among persons labelled with mild hypertension without comorbidities. Volunteers were selected among general population either from a list of patients from a primary healthcare clinic or social media and social network, with a broad range of characteristics including sex, age, level of education, ethnic origin and time from diagnosis. Data was subjected to qualitative content analysis by three of the authors independently followed by discussions to generate themes and categories.

Results: Preliminary results confirm that the label has impacts in a broad range of psychosocial dimensions, e.g. trust in own body, psychological reactions, social stigma and overmedicalization. Although informants had a broad range of characteristics, they shared

similar stories, understandings and effects of labelling. The empirical material is still being analysed and final results will be presented at the conference.

Conclusions: The diagnosis of mild hypertension is a relevant milestone and has impact on daily life. Most of the impact is regarded as a negative consequence or harm; however, sometimes it might be ambiguous.

Increasing prescription of opioid analgesics and neuropathic pain medicines for spinal pain in Australia.

Stephanie Mathieson[1], Lisa Valenti[2], Christopher Maher[3], Helena Britt[2], Qiang Li[4], Andrew McLachlan[5], Christine Lin[3]
[1]Musculoskeletal Health Sydney, University of Sydney, Sydney, Australia. [2]Family Medicine Research Centre, University of Sydney, Parramatta, Australia. [3]Musculoskeletal Health Sydney, Sydney, Australia. [4]The George Institute for Global Health, Newtown, Australia. [5]Faculty of Pharmacy and Centre for Education and Research on Ageing, The University of Sydney, Concord, Australia

Objectives: Limited evidence exists on secular trends of analgesic medicines for spinal pain. We investigated general practitioner's (GP) recommendations of analgesic medicines for spinal pain and investigated characteristics associated with their recommendation.

Method: We accessed data on spinal pain consultations from the Bettering the Evaluation and Care of Health (BEACH) database, a nationally representative database on GP activity in Australia. Data extracted included consultation details and management provided. Medicines recommended were grouped as simple analgesics, non-steroidal anti-inflammatory drugs (NSAIDs), opioid analgesics or neuropathic pain medicines. Multivariate logistic regression determined if patient characteristics and GP characteristics were associated with medication recommendations.

Results: We analysed BEACH data for 9100 GPs who managed 39,303 patients with spinal pain between 2004 and 2014. Over the decade, analgesic recommendations increased. After accounting for patient and GP characteristics, there was a significant increase in the rate single-ingredient opioid analgesics (e.g. oxycodone) [annual relative increase of 6% (Rate Ratio (RR) 1.06 (95% CI 1.05–1.07)] and neuropathic pain medicines (e.g. pregabalin) [annual relative increase of 19% (RR 1.19 (95% CI 1.16 to 1.22)] were recommended; and a significant decrease in the rate NSAIDs were recommended [annual relative decrease of 4% (RR 0.96 (95% CI 0.95 to 0.97)]. Logistic regression identified several patient and GP characteristics associated with medicine recommendations, e.g. stronger opioids were less likely recommended for Indigenous patients [Odds Ratio 0.15 (95% CI 0.04 to 0.56)].

Conclusions: GP's analgesic recommendations for spinal pain have become increasingly divergent from guideline recommendations over time.

Human papillomavirus (HPV) infection of the anogenital tract detected through molecular tests leads to overtreatments and waste of resources.

Carlo Liverani, Ermelinda Monti, Giada Libutti

Reference Center for the prevention, diagnosis and treatment of HPV related genital pathology. Department of Mother and Infant Sciences, University of Milan. Foundation IRCCS Ca' Granda Ospedale Maggiore Policlinico, Milano, Italy

Abstract in Full: Primary cervical cancer screening through HPV DNA testing is currently being recommended by major guidelines worldwide for all women aged 30 to 65 years.
While management of high grade cervical lesions histologically confirmed is well standardized, there is still considerable confusion regarding the management of an HPV positive result. Cervical cytology should be performed in all these cases, but some Authors recommend genotyping for HPV 16/18 first, while others recommend dual-staining with p16/Ki-67 first, in an effort to reduce referral to colposcopy.
The main problem employing HPV testing based strategies in cervical cancer screening is the fact that both positive and negative HPV results are often misinterpreted or overestimated. An HPV positive, cytology negative woman, should repeat both tests after one year interval. Actually, patients too often tend to undergo immediate colposcopic examination, increasing health care costs and patients' anxiety, without benefit and potentially resulting in overtreatments.
Not rarely clinicians start treating HPV infections detected with molecular tests with surgery, laser, cryotherapy, interferon, 5-florouracil. Then multiple preventive, diagnostic, and therapeutic activities are initiated, both in women and their partners, with strict follow up programs, more tests and more interventions. What many health professionals actually do is test women under 30 years of age; re-screen every 1-2 years; test for low risk HPV types; test anal, vulvar, penile, oral sites; test male partners; test to screen for sexually transmitted infections. All these indications are not recommended and may lead to wrong decisions, with well documented but poorly recognized ill effects.
Screening policies designed to achieve an optimal cost-effectiveness ratio are being misinterpreted and have been declined into daily clinical practice, with unmeasurable downstream consequences.
Concerns on sexual relationships are frequently reported, even after having provided detailed explanations. Besides, there is no urge to detect too early lesions that have a very slow progression rate, and might have been detected with repeat cytology a couple of years later. Both doctors' attitudes and women's expectations are difficult to meet with the widespread utilization of different molecular tests not applied consistently according to shared recommendations. The economic, social, and psychological impact of HPV screening seems to have outweighed presumed benefits: the risks are a waste of resources, raise in costs and anxiety, and underrecognition of true disease.

Objectives: To assess emotional, relational, and sexual aspects in women screened for cervical cancer through HPV DNA testing, and to document whether clinicians are willing to respect guidelines recommendations on HPV testing.

Method: We analyzed data from a survey regarding more than 400 gynaecologists in Italy to assess knowledge, attitudes and practice among gynecologists. To test women's reactions we also conducted an observational, prospective, quantitative, case-control study: a group of 90 HPV DNA positive women were interviewed and compared with 61 controls (women with unknown HPV status), through specific questionnaires provided to each patient.

Results: After a negative HPV result, 26.5% of gynecologists recommend repeating HPV testing within 1-3 years instead of 5. Testing the male sexual partner is being recommended by 45.1% of specialists dealing with patients with low-grade cervical cytology and a positive HPV result.

Women who tested HPV positive experienced negative psychosocial responses even in the presence of a normal cytological smear and a normal colposcopy. Questionnaires revealed various degrees of anxiety, fear, anger, shame, regret, overestimation of cancer risk, concerns about loss of reproductive functions, concerns about negative reactions from friends, family, or sexual partners, concerns about partner infidelity, as well as changes in physical intimacy activities or sexual refusal. Many women felt urgent need for treatment. Answers may be grouped into four categories: - Grief of the discovery: why is this happening to me? - Couple: symbol of marital infidelity - Family: weight of prejudice - Everyday life: abiding memory

Conclusions: A high proportion of gynaecologists participating in the survey, reported inappropriate uses of HPV testing. This may lead to unnecessary follow-up and potential overtreatment. As a consequence, in Italy many women are undergoing unnecessary tests, raising medical costs and anxiety both in women and their families. Not rarely cervical excision procedures are being performed for low grade lesions in young women, instead of waiting for spontaneous regression. Women who tested HPV positive experienced different psychological reactions, ranging from anxiety and fear to severe depression, isolation and sexual refusal. HPV testing can have important social and psychological effects on women and can potentially damage their wellbeing particularly when used outside an organized prevention program. HPV positivity may also cause a prolonged psychosocial burden on women even after having had the necessary follow-up for their cervical abnormalities and having provided adequate explanations.

Psychosocial consequences of participating in a national colorectal cancer screening programme

Jessica Malmqvist[1,2], Volkert Siersma[1], Mie Hestbech[1], John Brodersen[1,2]
[1]Center for Research and Education in General Practice, Copenhagen, Denmark. [2]The Research Unit for General Practice in Region Zealand, Copenhagen, Denmark

Abstract in Full
Background: Screening for life-threatening diseases has potential benefits and harms. In screening for colorectal cancer (CRC) some potential harms are the negative psychosocial consequences of being diagnosed with polyps or receiving a false-positive result.
In 2014 a national CRC screening programme was implemented in Denmark, using an immunochemical faecal occult blood test (iFOBT) as the screening method and targeting all citizens aged 50-74. In the implementation period of the screening programme (2014-2017) citizens were invited to participate in a randomised design. According to the latest Danish CRC screening report from 2016, the positive rate has been approximately 7%. All citizens with positive results were urged to have a colonoscopy where subsequently 63.5% had a clean colon or benign polyps, 31% had adenomatous polyps and 5.5% were diagnosed with CRC. Hence, thousands of healthy people have undergone a colonoscopy to confirm that they do not have CRC. Moreover, thousands of people were diagnosed with polyps: benign and adenomatous. Previous studies investigating psychosocial consequences of CRC screening have shown diverging results. However, all these studies were performed using generic questionnaires, which have shown not to be adequate in measuring psychosocial consequences in a medical screening setting.

Objectives: The objective of this study was to investigate psychosocial consequences of participating in a national colorectal cancer screening programme, with a condition-specific

questionnaire specifically developed for measuring psychosocial consequences in a CRC screening setting.

Method: The study was a longitudinal study with a randomised design. We performed data collection in the Region Zealand county of Denmark in the period of April 2017 – March 2018. We enrolled participants with a positive iFOBT consecutively and matched them on gender, municipality and year of birth in a 2:1 design with citizens with negative results, non-attenders and a control group randomised to screening in the last part of 2017. We sent a condition-specific questionnaire four days after the participants received their iFOBT result and again eight weeks after the final diagnosis. The primary outcome was psychosocial consequences measured by a condition-specific questionnaire.

Results: Preliminary results and conclusions will be presented at the conference.

09:00 Tuesday August 21st - Holst Auditorium - The Impact of Power Driven Overdiagnosis: The Role of Regulators & Health Authorities

The introduction and demise of full body computed tomography (CT) scanning in Australia: implications for preventing overdiagnosis

Chris Degeling[1,2], Eliza J McEwin[3,2], Ray Moynihan[4,2], Stacy M Carter[1,2]
[1]University of Wollongong, Wollongong, Australia. [2]Wiser Healthcare, Sydney, Australia. [3]University of Sydney, Sydney, Australia. [4]Bond University, Gold Coast, Australia

Objectives: Full body computed tomography (CT) scans for screening came to Australia in 2002, accompanied by a wave of advertising promising life-changing benefits without disclosing equally life-changing potential harms. Within a year, public access to self-referred full body CT screening was withdrawn as a consequence of sustained professional and political pressure and legislative changes. We examine how self-referred full body CT scans were made available to Australian consumers, and then, just as rapidly, how they were withdrawn. We evaluate and discuss the social, cultural and system factors that contributed to their de-implementation. Our analyses are intended to inform horizon scanning processes with early identification of new screening technologies that have a high risk of overdiagnosis. The aim is to encourage discussion and debate as how best to intervene to enable early identification of new testing technologies that are likely to produce unacceptable levels of overdiagnosis.

Method: To identify all available materials relevant to the roll out and withdrawal of full-body CT screening in Australia, systematic searches of PubMed, Factiva, ANZ Proquest, Informit, AustLii, and Pandora/Trove were undertaken for the period 1995-2014, resulting in the aggregation of 48 unique items relevant to the topic. Taken together, these databases archive media reports, advertising materials, website content, legal judgements, parliamentary proceedings, peer-reviewed journals, and government and corporate grey literatures. Sample materials were qualitatively reviewed through an iterative process of testing, revising and refining our insights and theoretical generalisations against an emerging conceptual map, and feedback from the research team. This cycle of searching, mapping and critical inquiry continued until new textual materials were not providing substantive new insights. Using the withdrawal of full body CT scans

in Australia as a case example, our analyses focused on the politics and mechanisms through which a new screening technology was de-implemented.

Results: The rapid de-implementation of full-body CT scans in Australia arose from interconnected sociocultural and system factors, an unusual harm profile for the technology, and key contributions from people and groups within and outside of medicine. Media, legislation, health professionals, governments and regulators all played a part in a well-coordinated and joint effort. Crucially, legislation enacted for environmental protection rather than health care quality was amended to provide a mechanism for effective action. These findings can help map the challenges, pathways and mechanisms for caution about and de-implementation of future, potentially harmful, testing technologies. While the relative value of each stakeholder's role is unclear, their collective collaboration and interaction achieved this unusual example of successful, early de-implementation. Each stakeholder exercised their influence and authority within their respective remits: it is possible that without the actions of each, the final outcome may not have been achieved.

Conclusions: In Australia, restricting public access to privately operated self-referred full-body CT scans is an oft-cited example of a screening technology that caused harm and which was successfully withdrawn. Within this national context, the case of full-body CT scans suggests that neither single actors nor robust health technology assessments are likely to be sufficient to protect people from the unwanted consequences of too much medicine. These findings are both salient and timely, given growing concerns about overdiagnosis and the emergence of new forms of testing technologies, such as whole genome screening of healthy people. Against a background of technological optimism and commercial opportunism, a commitment to vigilance and cross-sectoral coordination across a range of decision making roles and institutions will be key to efforts to prevent the harms of overdiagnosis.

Variation in diagnostic test requests and outcomes: a preliminary metric for OpenPathology.net

Jack O'Sullivan[1], Carl Heneghan[1], Rafael Perera[1], Jason Oke[1], Brian Shine[2], Ben Goldacre[1]
[1]Centre for Evidence-Based Medicine, University of Oxford, Oxford, United Kingdom.
[2]Department of Clinical Biochemistry, University of Oxford, Oxford, United Kingdom

Abstract in Full: Efforts to reduce healthcare costs have led to the development of metrics to identify unwarranted variation in care. Previous work assessing diagnostic tests is limited, despite their substantial contribution to expenditure. We explored C-reactive Protein (CRP) and Erythrocyte Sedimentation Rate (ESR) tests ordered across Oxfordshire NHS General Practices, the proportion of tests that yielded an abnormal result, and identified practices that had a proportion of abnormal CRP and ESR results 3 standard deviations below the mean. We estimated the adjusted average proportion of abnormal CRP and ESR tests that yielded abnormal results from each practice, after adjusting for differences in practice populations. These proportions were plotted against the total CRP and ESR requests per practice. We constructed funnel plots to identify practices 3 standard deviations below the mean proportion of abnormal CRP and ESR tests. We analysed 143,745 CRP and 30,758 ESR requests from 69 practices. Twelve (17%) and 7 (10%) practices were more than 3 standard deviations below the mean for CRP and ESR testing respectively. Two practices (3%) were below the 99.8% limit for both CRP and ESR ordering. Variation in the proportion of tests with an abnormal result shows promise for auditing variation in care.

Objectives: To identify General Practices that order significantly more normal ESR and CRP tests compared with their peers.

Method: We explored C-reactive Protein (CRP) and Erythrocyte Sedimentation Rate (ESR) tests ordered across Oxfordshire NHS General Practices and generated the proportion of tests that yielded an abnormal result for each practice.
We estimated the adjusted average proportion of abnormal CRP and ESR tests that yielded abnormal results from each practice, after adjusting for differences in practice populations. These proportions were plotted against the total CRP and ESR requests per practice. We constructed funnel plots to identify practices 3 standard deviations below the mean proportion of abnormal CRP and ESR tests.

Results: We analysed 143,745 CRP and 30,758 ESR requests from 69 practices. Twelve (17%) and 7 (10%) practices were more than 3 standard deviations below the mean for CRP and ESR testing respectively. Two practices (3%) were below the 99.8% limit for both CRP and ESR ordering.

Conclusions: We used robust and conservative methods to identify two practices (3%) that requested a significantly lower proportion of CRP and ESR tests yielding abnormal results. We also identified 12 (17%) and seven (10%) General Practices that requested a significantly lower proportion of CRP and ESR tests yielding abnormal results, respectively. Variation in the proportion of tests with an abnormal result shows promise as an automated tool for auditing variation in care, and may contribute to improving quality and cost effectiveness.

Five years of EMA-approved systemic cancer therapies for solid tumours – a comparison of two thresholds for meaningful clinical benefit

Nicole Grössmann[1], Joseph Del Paggio[2], Sarah Wolf[1], Richard Sullivan[3], Christopher M. Booth[4], Katharina Rosian[1], Robert Emprechtinger[1], Claudia Wild[1]
[1]Ludwig Boltzmann Institute for Health Technology Assessment (LBI-HTA), Vienna, Austria. [2]Department of Medicine, Division of Medical Oncology, University of Toronto, Toronto, Canada. [3]Institute of Cancer Policy, King's College London, King's Health Partners Comprehensive Cancer Centre, London, United Kingdom. [4]Department of Oncology and Public Health Sciences, Queen's University, Kingston, Ontario, Canada

Objectives: Approximately, 800 drugs and vaccines are currently under investigation in clinical trials for the treatment of cancer. Roughly, 80% of those are first-in-class therapies, and around 73% are intended as personalised and, therefore, targeted medicines.
Therefore, several societies have proposed frameworks that attempt to support the optimal use of limited health care resources, while offering a standardised and transparent tool to evaluate the benefit of novel cancer therapies. One prominent tool is the European Society for Medical Oncology Magnitude of Clinical Benefit Scale (ESMO-MCBS). Our objectives were to investigate the extent of European Medicines Agency (EMA)-approved cancer drugs that meet the threshold for 'meaningful clinical benefit' (MCB), defined by the framework, and determine the change in the distribution of grades when an adapted version that addresses the scale's limitations is applied.

Method: We identified all approval studies of cancer drugs indicated for solid tumours that received marketing authorisation by the EMA between 1st January 2011 and 31st December 2016. We previously proposed adaptations to the ESMO-MCBS addressing its main limitations, including the use of the lower limit of the 95% confidence interval in assessing the hazard ratio.

To assess the MCB, both the original and adapted ESMO-MCBS were applied to the respective approval studies.

Results: In total, we identified 70 approval studies for 38 solid cancer drugs. 21% of therapies met the MCB threshold by the original ESMO-MCBS criteria. In contrast, only 11% of therapies met the threshold for MCB when the adapted ESMO-MCBS was applied. Thus 89% and 79% of therapies did not meet the MCB threshold in the adapted and original ESMO-MCBS, respectively.

Conclusions: In most of the cancer drugs, the MCB threshold is not met at the time of approval when measured using both ESMO-MCBS scales. Since approval status does not translate into a MCB, stakeholders and decision makers need to continually assess the benefitrisk ratio of new cancer drugs to ensure a balanced and an equitable distribution of resources in our health care systems.

One year mortality from selected cancers: a dubious success of screening, Russia

Julia Ledovskikh, Vasiliy Vlassov
National research University Higher School of Economics, Moscow, Russian Federation

Background: Early diagnosis is a preferred action to reduce the cancer mortality – second leading cause of death in Russia. Since 2006 the programs of the screening for different cancers were introduced. Since 2015 the reduction of the one year mortality (1YM) is reported as a major success. Objective of this study is to analyze the size of the effect and potential side effects.

Methods: We analyzed national cancer statistics for years 2007-2016, calculating the additional cases found, mortality and the change of the 1YM in cancers stage III-IV.

Results: Four selected cancers were reported as a success story of screening with reduction of 1YM by 21-49% during 10 years. The size of the 1YM reduction is influenced by the increased detection of cancer stages I-II. Recalculated to the number of cancers stage III-IV the 1YM reduction is 2-3 times lower. The side effect of the inflated 1YM reduction is the increase of the early stage cancers by 37-137%
In the population the mortality from breast cancer decreased 14%, for prostate cancer increased 15% and not changed significantly for cervix and colon cancers.

Discussion and conclusion; The screening for four prevalent cancers led to the increased incidence mostly due to early I-II stage cancers. During 10 years the number late stage cancers did not reduced. Reduction of the 1YM significantly depends from the inflated denominator – the cases found. In Russia the cancer care is in short supply. By the expert opinions 2/3 of needed does not have access to adequate chemotherapy and ¾ to the radiotherapy. Expansion of the screening does not led to the reduction of the incidence of the late stage cases but exaggerate the demand for the scarce treatment.

Overdiagnosis of low back pain

Adrian Traeger[1], Sweekriti Sharma[1], Rachelle Buchbinder[2,3], Ian Harris[4,1], Chris Maher[1]
[1]University of Sydney, Sydney, Australia. [2]Cabrini Institute, Malvern, Australia. [3]Monash University, Melbourne, Australia. [4]Whitlam Orthopaedic Research Centre, Liverpool, Australia

Objectives: Low back pain cannot be overdiagnosed, at least not in the narrow sense of the word. However, it is a common symptom, and one that is often given a diagnostic label (slipped disc, pinched nerve, instability, arthritis, degeneration, and so on), despite there being no reliable way of determining the pain source in most cases. Once the symptom is labelled, however, the consequences begin to resemble those of overdiagnosis: many will experience no clinical benefit from receiving a diagnostic label, but will feel less well because of it, and are more likely to undergo costly, invasive treatments with questionable efficacy. We will present our current understanding of overdiagnosis of low back pain, discuss how low back pain might differ from other well-known examples of overdiagnosis, and detail efforts from within our research group and beyond to develop and evaluate solutions.

Method: If one were to consider the broader definitions of overdiagnosis proposed by Carter et al. (*BMJ* 2015;350:h869), healthcare for low back pain would have examples abound: disease mongering ("Pain as the 5thvital sign" campaign by US Veteran's Affairs), overutilisation (spinal injections, opioids), overdetection (diagnostic imaging), overtreatment (spinal fusion surgery, early physiotherapy), and false positives (red flags for serious pathology). In 2013 Americans spent US$81.6 billion on care for low back pain. How did we end up here?

Results: Unlike other well-known examples, overdiagnosis of low back pain appears to have little to do with altering disease definitions or thresholds, or providing screening programs for the healthy. Some people with low back pain may receive no diagnosis but are overtreated.

Conclusions: We argue that many of the problems with overdiagnosis and overtreatment of low back pain arise because people enter a health system that is set up to encourage inappropriate care and discourage appropriate care.

PHYSICIAN STRESS AND BURNOUT- CAUSE OR CONSEQUENCE OF OVERDIAGNOSIS?

Ananta Dave
National Health Service (NHS)., Birmingham, United Kingdom

Abstract in Full: There is emerging literature about possibly inceasing rates of burnout and stress among doctors from all specialties (Gerada 2017), and in various regions of the world, though it has been studied more in the developed world. For instance burnout has been linked to higher rates of prescribing (Shanafelt and Noseworthy, 2016). Burnout can adversely affect quality of clinical practice and decision making and therefore there is a moral and ethical imperative to address this issue (Imison, 2018).
In this session, the various reasons why burnout may lead to overdiagnosis and subsequent overtreatment are discussed, with examples, and with relevant literature references. I also suggest that overdiagnosis (if associated with large volume and speed of working) can in turn lead to burnout in doctors. It is hoped that this will generate debate among the audience and we can collectively discuss strategies to address this.
A case is made for a systems approach involving an open and kind culture, early identificaton of doctors with problems, easy access to help and support for those affected including mentoring, coaching, supervision and re-training if required. It is also important to work closely with patient and carer representatives and collect data linking physician and organisational health to clinical practice.
Nurturing those who care can ensure that they can then care in a compassionate and efficient manner, thus ensuring good outcomes for patients, staff and organisations.

Objectives
to understand the correlation between physican burnout and overdiagnosis and over treatment
to understand the mechanisms by which it can occur
to discuss ways of addressing this issue and redressing the balance in diagnosis

Method
Literature review
Anonimised and hypothetical case scenarios
Reflection on a system approach
Group discussion

Results: At the end of the session I hope there will be increased understanding of the link between physician burnout and overdiagnosis and how we can individually and collectively play a role in resolving this.

Conclusions: Burnout among caregivers is an important mechanism by which overdiagnosis can take place

14:00 Tuesday August 21st - Niels K. Jerne - Other

Evaluating the content of Choosing Wisely recommendations and the prevalence of interdisciplinary finger pointing

Joshua Zadro[1,2], Ian Harris[1,2,3], Christopher Maher[1,2]
[1]School of Public Health, Sydney Medical School, The University of Sydney, Sydney, Australia. [2]Institute for Musculoskeletal Health, Sydney Local Health District, Sydney, Australia. [3]South Western Sydney Clinical School, University of New South Wales, Sydney, Australia

Objectives: The aim of this study is to evaluate the content of existing Choosing Wisely lists and determine the proportion of Choosing Wisely recommendations that discuss tests or treatments; whether recommendations are worded appropriately (i.e. target a reduction in low-value care and provide recommendations that are actionable and direct); and whether recommendations targeting low-value income generating treatments are directed towards society members or non-members ('finger pointing').

Method: We will perform a content analysis of all Choosing Wisely recommendations, extracting recommendations from Choosing Wisely websites and The Diana (Dissemination of Initiatives to ANalyse Appropriateness) in Healthcare website (including data on the professional society, year and country of the recommendation). Two researchers will independently apply the following checklist to determine the frequency of recommendations that are tests or treatments; recommended 'for' or 'against' a procedure ('do' vs. 'don't'); include an alternative when advising against a procedure; are direct ('don't') or indirect ('consider avoiding', 'don't routinely', etc.); and advise against an income generating treatment (and whether the treatment is typically performed by members, non-members, or both). Disagreements will be resolved by discussion or consultation with a third reviewer if necessary. The frequency of coded data will be presented using descriptive statistics (counts and percentages).

Results: We found 1,265 Choosing Wisely recommendations across the United States (n=535, 42.3%), Canada (n=297, 23.5%), Italy (n=175, 13.8%), Australia (n=172, 13.6%), the United Kingdom (n=56, 4.4%) and the Netherlands (n=30, 2.4%). There were 200 (15.8%) recommendations from surgical societies, 155 (12.3%) from allied health or nursing societies, 131 (10.4%) from general medical societies (e.g. palliative care, geriatrics, internal medicine), and 50 (4.0%) from emergency medicine or critical care societies. Coding of the recommendations is ongoing and we will present the full results of the analyses at the conference.

Conclusions: There are currently 1,265 Choosing Wisely recommendations published worldwide; but there is yet to be an evaluation of their content. The wording of Choosing Wisely recommendations could support or prevent adoption, so our findings are the first step towards improving recommendations to facilitate their use in practice. Our study will also determine whether recommendations are adequately covering low-value treatments (particularly income generating treatments), and whether these recommendations are directed to members or non-members. Healthcare professionals are 10 times more likely to discuss low-value tests with their patients compared to low-value treatments; possibly reflecting concerns about losing income. Our findings could have strong implications for creating transparency in Choosing Wisely and ensuring that societies aren't using recommendations directed at non-members ('finger pointing') to avoid publishing recommendations against low-value treatments that could affect their members' income.

Overdiagnosis: a multi level problem

Bjørn Hofmann[1,2], Wendy Rogers[3], Stacy Carter[4], Lynette Reid[5]
[1]University of Oslo, Oslo, Norway. [2]NTNU, Gjøvik, Norway. [3]Macquarie University, Sydney, Australia. [4]University of Wollongong, Wollongong, Australia. [5]Dalhousie University, Halifax, Canada

Objectives: Overdiagnosis is a complex problem that is challenging to conceptualise and define. This challenge contributes to problems with measuring it and developing ways to decrease rates of overdiagnosis. The objective of this study is to identify the different levels at which overdiagnosis can be conceptualised and defined in order to improve information for patients' decision making, estimates by health services researchers, and decisions of health policy makers.

Method: Literature search (for definitions, conceptualisations, and estimates of overdiagnosis), qualitative content analysis, and philosophical concept analysis.

Results: Overdiagnosis can be defined on three levels: personal, professional, and population. On the personal level overdiagnosis occurs when a person is given a diagnosis of a condition that otherwise would not have progressed to manifest disease. From a professional perspective overdiagnosis is linked to the definition of disease and can be identified in terms of dysfunction that is not at significant risk of causing harm. On a population level overdiagnosis is defined in terms of correct diagnoses that carry an unfavourable balance between benefits and harms. To make sense for individuals, overdiagnosis has to be explained (prognostically) in terms of the individual's chance of being diagnosed (and treated) for what would otherwise have been manifest disease. Arriving at this information requires attention to professionals' conception of disease, as well as epidemiologists' estimates of the proportion of individuals who are diagnosed without overall benefits.

Conclusions: On a personal level overdiagnosis involves being unnecessarily diagnosed with (and treated for) a disease. To better estimate the chance of this occurring for individuals we need agreement regarding professional conceptions of disease and population-based estimates.

Systematic review of overdiagnosis in cervical cancer screening: How should we define overdiagnosis in cervical cancer screening?

Chisato Hamashima[1], Teruhiko Hearasawa[2], Takafumi Katayama[3], Seijyu Sasaaki[4], Satoyo Hosono[5], Keika Hoshi[6]
[1]Teikyo University, Tokyo, Japan. [2]Fujita Health University, Nagoya, Japan. [3]Hyogo Prefecture University, Akashi, Japan. [4]St Luka's International Hospital, Tokyo, Japan. [5]Tokai clinic, Nagoya, Japan. [6]Kitazato University, Sagamihara, Japan

Objectives: Cervical cancer screening is a common strategy for cancer control worldwide. Although its real target is invasive cervical cancer, the incidence has not been high in developed countries, and precancerous lesions have now become the actual target of cervical cancer screening. Therefore, *cervical* intraepithelial neoplasia (CIN) 3 has now been generally identified as the actual target for early detection and treatment, while, in some countries, CIN2 has become the treatment target. The definition of overdiagnosis in cervical cancer screening has been unclear. Although most cases of CIN have a high possibility of disappearing, CIN2 and CIN3 lesions have been routinely resected when detected by cervical cancer screening. To clarify the traditional concept of overdiagnosis in cervical cancer screening, a systematic review was performed.

Method: Medline, Cochrane Central, Embase, and Igaku-Cyuo Zasshi were searched until January 2018. The articles were original articles limited to English-language or Japanese-language publications. Search terms such as 'cervical cancer', 'cancer screening', 'cytology', 'Pap smear', 'HPV testing', and 'overdiagnosis' were used. A modeling approach was also included. Additional references cited in candidate articles were included as needed. To select appropriate articles regarding the concept and frequency of overdiagnosis, a two-stage review process was used: the title and abstract were initially checked and then potential papers were subsequently reviewed. Finally, studies of overdiagnosis in cervical cancer screening were selected.

Results; One modeling reported from the Netherlands and two articles from a Finnish study which included in a randomized controlled trial for HPV testing were selected. In the modelling approach, the frequencies of overdiagnosis in the screening period were estimated to be 74.8% for CIN1+, 68.0% for CIN2+, and 55.4% for CIN3+. Over the subjects' lifetime, the frequencies of overdiagnosis were 70.6% for CIN1+, 63.2% for CIN2+, and 50.0% CIN3+. In the first report in the Finnish study, the gap in the cumulative incidence of detected invasive cancers between the Pap smear group and the HPV testing group suggested overdiagnosis of HPV testing. Based on a 4.5-year follow-up from the first screening of this study, the frequency of overdiagnosis was 20.3 (/100,000) for Pap smear and 39.6 (/100,000) for HPV testing.

Conclusions; In cervical cancer screening, precancerous lesions have been identified as the target of cancer screening. These lesions have been resected, and the adoption of this approach has expanded despite the high possibility of the disappearance of these lesions. Overdiagnosis of cervical cancer screening has not been investigated until recently and the studies regarding overdiagnosis have been few. However, its frequency was high in recent

reports. Until recently, overdiagnosis has been ignored in cervical cancer screening and has led to overtreatment of precancerous lesions. In developed countries, the incidence of cervical cancer has decreased and has not become a serious burden. In addition, the natural history of the development of cervical cancer has also been clarified. Although cervical cancer screening has high impact of reeducation from cervical cancer, the balance of benefits and harms including overdiagnosis should be reconsidered.

Primary health care education and Antibiotics overuse.

Nawras Azzam
MOH, Huraymila, Saudi Arabia

Abstract in Full: My name is Nawras Azzam, I am an ENT specialist currently working in Saudi Arabia for the past 3 years as the head of emergency department which has a capacity of 100 beds. I am originally from Syria in the Middle East.
Over the counter Antibiotic and its overuse has been discussed before in different parts of the world and some countries were able to control its use by making it illegal to give antibiotic to patients without a doctor's prescription. However, Saudi Arabic is still one of many countries that still faces the issue of over the counter abuse of antibiotic.
The case study that was contacted regarding URTI and febrile patients showed that 50-60% of these patients who were surveyed received anti-microbial therapy. The source of medicine was from PHCs, private hospital and pharmacies."The medication was provided by PHCs, private hospitals and pharmacies"
This uncontrolled prescribing of antibiotics has led to wrong diagnosis, wrong use, over use, and unsuspected bacterial resistance of antibiotics.
When faced with a patient of medical history, I found myself in a dilemma. Questioning myself if I should start from the beginning or continue with more investigation to confirm the diagnosis?
According to a study that was conducted in Saudi Arabic, the conclusion was as follows:
What is required is correct identification of the illness and good diagnosis methods if needed. In addition, primary health care education is the cornerstone of avoiding over diagnosis and over use of antibiotics.
MOH in Saudi Arabia introduced a new intuitive to stop the use and purchase of antibiotics without a prescription and has required a swab and culture for some specific antibiotics related to 3rd and 4th generations.
Furthermore, we have started educational programs in the high schools and Social institutions by giving lectures regarding antibiotics use, the differences between viral and bacterial infection and when to use anti microbial therapy.
Around 40-50% of patients have received our new guidelines. The feedback was very positive. They were very satisfied with the outcome. They were pleased to understand their kids who get infected with URTI or only suffering of fever.
Thank you

On the relevance of definitions: three conceptually challenging issues in overdiagnosis

Lynette Reid[1], Stacy Carter[2], Bjørn Hofmann[3], Wendy Rogers[4]
[1]Dalhousie University, Halifax, Canada. [2]University of Wollongong, Wollongong, Australia.

[3]Norwegian University of Science and Technology, Trondheim, Norway. [4]Macquarie University, Sydney, Australia

Objectives: The working definition of overdiagnosis is that it is the diagnosis of a condition that would not otherwise have become clinically significant. This definition highlights prognostic questions. To inform policy and clinical decision-making, epidemiologists have focused on estimating the extent of overdiagnosis, reaching widely diverging conclusions. Our objectives in this conceptual research are to:

- analyze existing conceptual definitions of overdiagnosis, in order to
- identify additional factors that help to put methodological disagreements into perspective.

Method: Conceptual analysis based on philosophical methods for identifying phenomena, analysing definitions, challenging assumptions and refining concepts.

Results: Conceptual definitions differ about what is overdiagnosed (harmless disease or "indicative phenomena" not themselves disease) and how treatment-related harms and benefits matter for identifying overdiagnosis. The disagreement usefully highlights three factors that are important in describing and communicating overdiagnosis. 1. The wide variation in overdiagnosis estimates arises in a circular fashion from assumptions about disease incidence and dynamics. Escaping the circle requires addressing the plausibility of basic causal disease mechanisms and criteria for setting disease boundaries. 2. There is a focus on variations in estimates of overdiagnosis. However, these contested estimates are an order of magnitude greater than the potential benefits against which they are weighed. 3. The focus on potential benefits from early intervention leads to neglect of the nature of overdiagnosis harms. While information about the rates of overdiagnosis is important, the seriousness of overdiagnosis harms is also critical for all levels of decision making.

Conclusions: Existing efforts to refine epidemiological estimates should continue and will help to inform patients; this analysis of conceptual definitions of overdiagnosis highlights additional issues that will complement advances in epidemiology by placing the estimates in ethical and practical context. We need to consider how estimates depend on assumptions about underlying disease incidence and dynamics, whether the range of disagreement matters when placed in perspective with corresponding benefits and harms, and how to weigh and act on the harms.

14:00 Tuesday August 21st - Nielsine Nielsen Auditorium - Other

Trends in stage-specific breast cancer incidence in New South Wales, Australia: insights from 25 years of screening mammography

Gemma Jacklyn[1], Kevin McGeechan[1], Les Irwig[1], Nehmat Houssami[1], Stephen Morrell[2], Katy Bell[1], Alexandra Barratt[1]
[1]The University of Sydney, Sydney, Australia. [2]The University of NSW, Sydney, Australia

Objectives: Screening mammography aims to improve breast cancer (BC) prognosis by increasing the incidence of early-stage tumours in order to decrease the incidence of late-stage cancer, but no reports have investigated these potential effects in an Australian population.

Therefore, we aim to describe temporal trends in the incidence of stage-specific breast cancer in New South Wales, Australia, between 1972 and 2012.

Method: An observational study of all women who received a diagnosis of BC from 1972-2012 as recorded in the NSW Cancer Registry, a population-based registry with almost complete coverage and high rates of histological verification. We analysed trends in stage-specific incidence before screening and compared them to periods after screening began. Our primary group of interest was women in the target age range of 50-69 years, though trends in women outside the target age were also assessed.

Results: Screening was not associated with lower incidence of late-stage BC at diagnosis. Incidence for all stages remained higher than prescreening levels. In women aged 50-69 years, the incidence of carcinoma in situ (CIS), localised and regional BC has more than doubled compared to the prescreening era, with incidence rate ratios ranging from 2.0 for regional (95% CI 1.95-2.13) to 121.8 for CIS (95% CI 82.58-179.72). Before the introduction of screening there was a downward trend in distant metastatic BC incidence, and after the introduction of screening there was an increase (IRR 1.8; 95% CI 1.62-2.00). In women too young to screen the incidence of late-stage BC at diagnosis also increased, whereas localised disease was stable.

Conclusions: The incidence of all stages of BC has increased over the past forty years, with the greatest rise seen during the established screening period for women aged 50-69 years. Our findings suggest that some of the expected benefits of screening may not have been realised and are consistent with overdiagnosis.

To PET or not to PET?

Andreea Badea[1], Sonia Vanderby[2,1], Paul Babyn[3,1]
[1]University of Saskatchewan, Saskatoon, Canada. [2]Health Quality Council, Saskatoon, Canada. [3]Saskatchewan Health Authority, Saskatoon, Canada

Objectives: Concerns about unnecessary medical imaging have led to the development of guidelines for a variety of indications and modalities by radiology and specialty groups. This study evaluates the utility of four national imaging guidelines, which assess the value of positron emitting tomography/computed tomography (PET/CT). As PET/CT is an emerging technology, the evidence base for the development of guidelines is not as established as other modalities. Of the indications commonly explored by PET/CT, the evaluation of lung cancer and lung nodules has the greatest evidence base and is therefore the most evaluable sector for guideline comparison. The primary goal of this study is to determine which of the four guidelines has the greatest coverage in a real-world clinical setting. The secondary goal of this study was to assess the convenience of the utilization of the selected guidelines for medical imaging specialists in a high-volume practice setting.

Method: PET/CT exam requests for suspected or confirmed lung cancer and lung nodules within a specified common seven day period were retrospectively accessed from four centers in Canada. Requests were de-identified and combined to create a single data set. Requests were then evaluated using four sets of recognized radiology guidelines: (a) the Canadian Association of Radiologists Referral Guidelines (CAR-RG), (b) the American Association of Radiologists Appropriateness Criteria (ACR-AC), (c) National Comprehensive Cancer Network (NCCN) Practice Guidelines, and (d) the Institut National d'Excellence en Santé et en Services Sociaux (INESSS) Interactive Tool. Results were compared in terms of (i) condition/symptom coverage

and (ii) recommendations obtained with particular attention to conflicting or differing results. Rates of high-value, low-value and unscorable requests were tabulated and rates were compared among guidelines. As a secondary assessment of the coverage, the usability of each set was subjectively measured.

Results: In total, 74 requests qualified for the inclusion criteria. Of these, 51% were requested for confirmed or highly suspicious cases of lung cancer and 49% were for the evaluation of lung nodules. 36-57% of all exams evaluated were considered high-value and 3-7% were considered low-value. Requests that could not be assessed totaled 31-50% for lack of information and 1-9% for the lack of an applicable guideline. Guideline coverage ranged from 64-72%. Of the guidelines used, the INESSS Interactive Tool was the most user-friendly in terms of organization and information provided with the value assessment. While the ACR-AC are very well developed for the radiologic management of cancers, the guidelines for their imaging was more sparse and difficult to navigate. The NCCN guidelines provide an easy to use summary document. The CAR-RG while well organized by body system, provided vague guidelines.

Conclusions: While the choice of guidelines used by institutions is generally determined based on the country they are in, our research has demonstrated that in some instances, it is prudent to use the guidelines that are most developed for that indication. For lung cancer cases and lung nodule evaluation, the INESSS guidelines performed the best with only 28% of total requisitions unable to be scored, whereas the CAR-RG guidelines cover the least indications with 36% of requests unable to be assessed. These results also indicate areas where national guidelines would benefit from further development. It is important to note that the INESSS guidelines also resulted in the most high-value assessments at 42% of requests. This might suggest that the issue does not always lie in low-value ordering, but perhaps the lack of guideline coverage and clarity in some guidelines resulting in higher proportions of low-value scores.

Research misconduct can promote overtreatment. A multi-institutional case study from Japan, with implications for osteoporosis management with vitamin K.

Alison Avenell[1], Andrew Grey[2], Mark Bolland[2], Greg Gamble[2]
[1]University of Aberdeen, Aberdeen, United Kingdom. [2]University of Auckland, Auckland, New Zealand

Objective: To examine the impact of delays in investigating and retracting randomised controlled trial (RCT) reports informing guidelines for osteoporosis treatment with vitamin K.

Methods: In 2007, concerns were raised with a leading medical journal about implausible results of 17 RCTs from a Japanese group from three universities, including 3 RCTs in a 2006 systematic review of vitamin K to improve bone mineral density and prevent fractures. No action was taken by the journal. In 2013, our systematic review of 33 RCT reports by this group indicated widespread concerns, but was not published until 3.5y later after several journal submissions. By April 2018, 21 of the 33 RCTs had been retracted, for reasons including data fabrication, concerns with data integrity, honorary authorship and self-plagiarism. To investigate the impact of the RCTs on systematic reviews and guidelines, we undertook citation searching in Scopus and Web of Science. Here we focus on the three RCT reports of vitamin K supplementation, and their subsequent impact.

Results: The 2006 systematic review of vitamin K included only 7 small RCTs with fracture outcomes, including these 3 RCT reports, all published by one journal. To date 2 of the 3 RCTs have been retracted. The 3 RCTs had been cited 157 times by April 2018. The systematic review has been cited 190 times. Particularly of note is that the review's meta-analysis plot for fracture outcomes was reproduced as the main evidence to support vitamin K for the prevention of fractures for the 2011 Japanese osteoporosis guidelines (replaced in 2015). The systematic review originally presented an odds ratio of 0.23 (95% CI, 0.12-0.47) for hip fracture, but a recent correction, omitting these 3 trial reports, amends this to 0.30 (95% CI 0.05-1.74). This is still mentioned as a 'large effect' even though based on only 3 events. Two subsequent trials in Japan with 6361 participants have failed to demonstrate that vitamin K prevents fractures.

Conclusion: 11 years have passed since concerns were raised about these RCT reports, which strongly influenced 2011 Japanese guidelines. Much prompter investigation of concerns about research misconduct is needed. We have found that journals and publishers are averse to flagging up expressions of concern, slow to investigate concerns, unwilling to accept publication of detailed statistical investigations, or retract articles. All of which spread the pernicious influence of research misconduct leading to inappropriate or ineffective treatment.

The Impact of Prognostic Estimates on Surgical Decision Making in the Setting of Severe Traumatic Brain Injury: A Survey of Neurosurgeons

Theresa Williamson[1], Marc D Ryser[2,3], Jihad Abdelgadir[1], Monica E Lemmon[4], Mary Carol Barks[5], Rasheedat Zakare[6], Peter A Ubel[5,6]
[1]Department of Neurosurgery, Duke University Medical Center, Durham, USA. [2]Department of Surgery, Division of Advanced Oncologic and GI Surgery, Duke University Medical Center, Durham, USA. [3]Department of Mathematics, Duke University, Durham, USA. [4]Department of Pediatrics, Duke University Medical Center, Durham, USA. [5]The Fuqua School of Business, Duke University, Durham, USA. [6]Duke School of Medicine, Duke University, Durham, USA

Objectives: Surgical decision making in the setting of severe traumatic brain injury is challenging because in a very short time window, neurosurgeons must weigh benefits and risks of invasive interventions that can both enhance the chance of recovery and prolong suffering. Although evidence-based prognostic tools are available, little is known about their impact on the decision-making process of surgeons. The objective of this study was to evaluate the effect of an evidence-based clinical prediction model, the CRASH calculator, on neurosurgeon prognostic beliefs and treatment recommendations.

Method: We performed a survey of neurosurgeons based on a convenience sample at the Congress of Neurological Surgeons International Meeting. All participants received information about a hypothetical patient with severe traumatic brain injury and subdural hematoma. Participants were further randomized to receive no additional information (control group) or receive CRASH-based clinical outcome predictions (CRASH group). In a first part, participants were asked to state their prognostic beliefs about the hypothetical patient's 30-day mortality and functional recovery at six months (continuous probabilities, range: 0%-100%). In a second part, participants were asked whether they would recommend craniotomy or medical management. Continuous outcomes were analyzed using linear regression models. Binary outcomes were analyzed using log-binomial regression models and relative risks (RR) for CRASH vs. control were estimated.

Results: A total of 139 neurosurgeons with median experience of 13 years (IQR: 5-27) completed the survey. Neurosurgeon prognostic predictions were highly variable and ranged from 0% to 100% in both 30-day mortality (median: 57.5% in control, 42.5% in CRASH) and 6-month functional recovery (median: 30% in control, 20% in CRASH). Providing CRASH estimates only modestly decreased prognostic estimates (p=0.14 for mortality; p=0.06 for functional recovery). Compared to the control arm, participants who received CRASH estimates were more likely to recommend non-surgical management (RR: 2.3, 95% CI: 1.16-4.9, p=0.02). Overall, 6-month functional recovery prognosis was a strong predictor of the decision to recommend craniotomy (p = 0.003). An exploratory mediation analysis suggested that prognostic beliefs may play a mediating role for the effect of CRASH on the treatment recommendation.

Conclusions: Neurosurgeon prognostic beliefs in severe traumatic brain injury are highly variable and correlate with treatment recommendation. Those who received evidence-based prognostic estimates were more likely to recommend non-surgical management to those who didn't receive such information. To reduce recommendation variability among neurosurgeons and facilitate informed decision making, there is a need to synthesize available evidence and effectively disseminate it to practitioners.

A critical interpretive synthesis of recommendations for De-intensification and de-IMPLEmentation from population Screening (DIMPLES)

Tammy Clifford[1,2], Stuart Nicholls[3], Pearl Atwere[2], Lindsey Sikora[2], Zahra Montazori[2], Richard Ashcroft[4], Jeff Botkin[5], Jamie Brehaut[3], Doug Coyle[2], Angus Dawson[6], Lesley Dunfield[1], Ian Graham[3], Jeremy Grimshaw[3], John Lavis[7], Beth Potter[2], Marcel Verweij[8]
[1]Canadian Agency for Drugs and Technologies in Health (CADTH), Ottawa, Canada. [2]University of Ottawa, Ottawa, Canada. [3]Ottawa Hospital Research Institute, Ottawa, Canada. [4]Queen Mary University of London, London, United Kingdom. [5]University of Utah, Salt Lake City, USA. [6]University of Sydney, Sydney, Australia. [7]McMaste University, Hamilton, Canada. [8]Wageningen Universit & Reseach, Wageningen, Netherlands

OBJECTIVES: Screening has generally been met with enthusiasm, and, in some cases, has been associated with demonstrable reductions in morbidity and mortality. In other cases, the balance of benefits and harms may be less clear and may evolve over time. Recommendations to reduce (de-intensify) or stop screening altogether (de-implementation) have occurred, but have proved to be controversial.
The goal of this study was to review existing recommendations to better understand the stated rationales for de-intensification or de-implementation of established population screening programs. In doing so we:

1. Identify examples of population screening programs recommended for de-intensification or de-implementation and describe the characteristics of these program changes
2. Describe the stated reasons for de-intensification or de-implementation
3. Explore the information cited for justification of recommendations

METHOD: We conducted a Critical Interpretive Synthesis (CIS) of published literature. This included an extensive search (websites of major screening organizations, reference searches, and content expert input) in addition to an electronic search of standard data bases. Data were analysed, coded and labeled in an inductive manner, and thus allowed for further

searches that explicitly sought out contrasting or conflicting evidence. The review has been registered with International Prospective Register of Systematic Reviews PROSPERO (CRD42016035279).

RESULTS: Of 9570 titles total of 66 documents, were included for analyses. Of these, 55 covered adult screening conditions and were largely cancer-related. Infant conditions included congenital toxoplasmosis (1), Down syndrome (1), fetal movement count (2), hearing test (1), urinalysis (3) and neuroblastoma (3). Of 107 recommendations, 50 related to de-intensification (e.g. increase of start age, extension of screening interval), 33 called for de-implementation, and 24 called for changes to the screening modality (often due to reduction in invasiveness or in association with increased periodicity). Explicit reasons were provided for 49 recommendations, while only 41 recommendations cited specific information in the justification.

CONCLUSION: We identified examples of programs across jurisdictions, age ranges, and clinical areas. Cancer screening dominates the examples identified. Recommendations to de-intensify or de-implement programs varied in terms of the level of evidence cited, as well as the stated rationales. Only around half of the identified recommendations provided an explicit rationale, and less than half of all recommendations were supported by a specific citation of evidence. Given the contested nature of de-implementation decisions, there is a greater need for transparency regarding the rationale behind recommendations and clearer articulation of the evidence used to support specific recommendations.

Quality Indicators aiming to help Primary Care in Sweden to balance between what to do and what not to doQuality Indicators aiming to help Primary Care in Sweden to balance between what to do and what not to do

Malin André[1], Ulrika Elmroth[2], Eva Arvidsson[3], Jörgen Månsson[4]
[1]Dep of Public Health and Caring Sciences, Uppsala University, Uppsala, Sweden. [2]Swedish Association of Local Authorities and Regions, Stockholm, Sweden. [3]Academy for Health and Care, Jönköping County, Sweden. [4]Department of Public Health and Community Medicine, The Sahlgrenska Academy at Gothenburg University, Gothenburg, Sweden

Objective - To construct evidence based indicators, automatically retrieved from electronic medical records (EMR), for quality improvement including indicators to prevent overtreatment and overdiagnosis

Method - "Primary Care Quality" is a national system consisting of nationally specified quality indicators. The data for the indicators is automatically retrieved from EMRs. The indicators are presented at the HC centres for benchmarking but also to identify individual patients.

The indicators reflect quality of care for patients with chronic diseases, infections as well as qualities like priority setting and multimorbidity. The project was initiated by the professional organisations in Primary Care and most of the indicators are built on evidence based national guidelines. In line with similar systems in many countries most indicators assess certain measures being taken, but in addition this system includes several indicators aiming to point out possible overdiagnosis and overtreatment

Result - The indicators aiming to prevent overtreatment concern use of antibiotics, bensodiazepines, hypnotics, proton pump inhibitors and medication among the elderly.

A few indicators concern overdiagnosis (laboratory test in common infections). However, comparing prevalence of different diagnoses at different HCs may a way to inspire discussion on possible overdiagnosis, especially prevalence of different infectious diseases as well as depression and anxiety. Moreover there are indicators for multimorbidity, which aim to balance disease specific indicators since guidelines usually have to be adjusted to the individual patient in patients with more than one chronic disease.

Conclusion - Although most indicators concern what should be done, quite a few concern overtreatment and overdiagnosis. Indicators for multimorbidity are constructed to balance disease specific indicators.

It is possible to construct indicators from automatically retrieved data from the EMR for use in discussions on overtreatment and overdiagnosis in primary care.

How effective the indicators are for reducing overtreatment and overdiagnosis is not known yet.

14:00 Tuesday August 21st - Holst Auditorium - The Impact of Power Driven Overdiagnosis: The Role of Regulators & Health Authorities

Drivers of general health checks in China and the risk of overdiagnosisDrivers of general health checks in China and the risk of overdiagnosis

Yu Wang[1], Mengxuan Zou[2], Hui Miao[1]
[1]Peking University, Beijing, China. [2]University of Bristol, Bristol, United Kingdom

Objectives: In the last decade, the industry of general health checks has been growing rapidly in China, increasingly people are keen on physical examination. However, there is very few study focusing on general health checks and overdiagnosis compared with the great amount of such studies conducted in many western countries. Therefore, the aim of this study is to explore and analyse the drivers of the prevalence of general health checks in China and its consequence especially overdiagnosis.

Method: literature review and in-depth interview with stakeholders were used in this study.

Results: The drivers of the prevalence of general health checks in China are various and complicated, including policy support and government guidance, health system and medical insurance system need to be improved, occurrence of cardiovascular disease (CVD) in China grows faster than expectation and the improvement of public health literacy and attitude to general health checks. Last but not least is the guide of Media and the cultural differences between East and West countries. The general health checks in China is at high risk to develop to overdiagnosis, yet it is not enough for the government or researchers to pay attention to it.

Conclusions: The current popularity of general health checks and its potential effect on overdiagnosis are inevitable in the development of social economy and health system in China. All stakeholders need to pay attention to this big issue of the situation of general health checks in China and its serious consequence like overdiagnosis. Meanwhile, the government,

practitioners, policy makers and scientific researchers should pay attention and deal with them as soon as possible.

The diagnostic "bubble" - overdiagnosis revisited from economic perspective using osteoporosis as an example

Christoffer Bjerre Haase, John Brodersen
Centre of Research & Education in General Practice, Department of Public Health, University of Copenhagen, Faculty of Health Sciences, Copenhagen, Denmark

Objectives: Many individual drivers for overdiagnosis have been identified. Out of those, the economical one has been considered the strongest with the global market of pharmaceuticals now valued at 1.000 billion U.S. dollars and is expected to grow to 1.300 billion in 2020.
Despite the common awareness of the above, studies so far lack a more coherent approach explaining the structures and dynamics in society that facilitate overdiagnosis rather than examining its single driver alone.
The following study offers a more comprehensive explanation to overdiagnosis, by approaching the concept with the economical point of view as the major tool. Osteoporosis, which is widely accepted as an overdiagnostic condition, is used as an example. The suggested approach opens a new possibility to identify regulations that are needed in order to reduce overdiagnosis and its undesired consequences.

Method: The analysis is based on abductive methodology, in which evidence about osteoporosis is contextualised into the interdisciplinary bubble theory.
This economic theory explains how economical assets are being traded at prices significantly departing from their fundamental value, which constitutes a potential risk of creating a bubble. Latest well-known bubble was the financial crunch in 2008. Recently, the theory has been extrapolated to explain situations in also other fields than finance, such as scientific bubbles, political bubbles and information bubbles. The inflated value of a specific entity, due to various facilitating elements, is what these non-financial scenarios have in common with the financial bubbles.
The presented study applies the facilitating elements known from bubble theory to the empirical evidence from osteoporosis. Furthermore, it evaluates the consequences of overdiagnosis in the selected case and its relevance to the other bubbles.

Results: Doctors, patients, pharmaceutical industry, patient associations etc. all have a stake in the use of the osteoporosis diagnosis. Therefore, they may be considered as distinct actors. The concept of speculation - profiting from trade of the asset instead of its use - correlates with the three main reasons for overdiagnosis (disease mongering, lowering thresholds and over-detection) that are all evident in the case of osteoporosis. Similarly to the financial market, the medical market for osteoporosis seems to be configured in ways to boost the use of the diagnosis. Actors in financial markets are susceptible to social influence and so it seems, in general, in the medical field. By the very definition of a bubble, it can be argued that the development of osteoporosis diagnosis can be seen as one.

Conclusions: Although there are many similarities, the analogy between bubbles in finance and the situation of osteoporosis as an overdiagnostic condition is not perfect. That is mainly because of insufficient solid and consistent theory of bubbles. However, the bubble theory does offer a new approach to explain overdiagnosis with a more coherent explanation than just study of

individual drivers. Furthermore, this new approach shows the value of interdisciplinary research in order to understand complex phenomena such as overdiagnosis, by complementing the empirical evidence with theory from other scientific spheres.

Systematic influences effect on participation rate in a fictional medical screening programme

Christian Patrick Jauernik[1], Or Joseph Rahbek[1], Thomas Ploug[2], Volkert Siersma[1], John Brodersen[1]
[1]The Research Unit for General Practice and Section of General Practice, Department of Public Health, University of Copenhagen, Copenhagen, Denmark. [2]Center for Applied Ethics and Philosophy of Science, Department of Communication, Aalborg University, Copenhagen, Denmark

Objectives: The objective of screening programmes for cancer is to reduce mortality and morbidity of the concerned cancer disease. However, cancer screening programmes may also cause overdiagnosis, false-positive results, overtreatment as well as psychological or physical harms. National health agencies and cancer organizations prioritize a high screening participation rate in favour of a comprehensive, understandable and evidence-based informed choice. In pursuit of this goal, these organisers make use of different systematic influences: Use of relative risk reductions; Emission of harms/exaggeration of benefits; Pre-booked appointments; Explicit recommendation of participation; Inducing fear - often in combinations.

The objective of this study was to asses if the different categories of systematic influences had a significant effect on the intention to participate in a screening programme for a non-communicable life-threatening disease, when applied in an invitational folder, and whether the applied systematic influences are recognized by the participants.

Method: We created seven different pamphlets, inviting the reader to a screening programme for a fictional non-communicable life-threatening disease. Five pamphlets encompassed one of the five abovementioned influences, one encompassed none, and one encompassed all five influences combined.

In different public places in Denmark, 600 random passer-by were randomised to one of the seven pamphlets. After reading the pamphlet participants were asked: 1) whether they intended to participate in the fictional screening programme, and 2) whether they were able to correctly identify the applied influences.

Statistical analysis (Chi^2) was used, to measure the effect the systematic influences had on intention to participate. A descriptive analysis of the participant's ability to identify the influences was also conducted.

Results: Ad 1) A statistically significantly ($p < 0.05$) higher proportion intended to participate in the groups receiving pamphlets containing relative risk reductions (OR 2.09 (95% CI: 1.16-3.75)), omission of harms/exaggeration of benefits (OR 4.30 (95% CI: 2.38-7.77)) explicit recommendations (OR 1.83 (95% CI: 1.03-4.25)), fear appeals (OR 2.39 (95% CI: 1.35-4.25)), and all combined (OR 8.60 (95% CI: 4.43-16.69)). Ad 2) Only a minority of the participants were able to correctly identify systematic influences in the pamphlets, varying from 4.1% to 29.9%. Participants who correctly identified a systematic influence in a pamphlet had a decreased

intention to participate compared to those who found the pamphlet informational or those who indicated the pamphlet was influencing their choice but failed to identify where.

Conclusions: In this study, four of five categories of systematic influence proved to increase intention to participate statistically significantly, except for the category pre-booked appointments. The latter category also increased the intention to participate but not in a statistically significant manner. Only a minority were able to identify the systematic influences, and since not identifying an influence is associated with an increased tendency to intend to participate, our results suggests, that some participants intended to participate in the screening programme because they were not able to recognize that they were being influenced. This effect seems to increase with increasing potency of the influences.

Antibiotics-impregnated calcium sulfate in surgery: A case of broader use of innovative medical practice

Renée Drolet, Geneviève Asselin, Alice Nourissat, Martin Coulombe, Marc Rhainds
CHU de Québec - Université Laval, Quebec, Canada

Objectives: Surgical site infections (SSI) are a major source of mortality and morbidity, especially with prosthesis or medical implants. Prevention and control of SSI refer to numerous prophylactic measures, including administration of systemic antibiotic therapy. Local antibiotherapy mediated by bioresorbable bone substitute impregnated with antibiotics has been suggested as a complementary strategy. Beads of synthetic calcium sulfate ($CaSO_4$) impregnated with antibiotics are prepared in the operating room and implemented at the surgical site to ensure local diffusion of antibiotics. A health technology assessment (HTA) was carried out in our university hospital to assess the effectiveness and safety of antibiotics-impregnated $CaSO_4$ (Al-$CaSO_4$) for the prevention and treatment of SSI in orthopedic, vascular and neuromodulation surgeries.

Method: A literature review was conducted in several indexed databases and gray literature. Systematic review, primary studies, and guidelines on the efficacy and safety of Al-$CaSO_4$ used for prevention or treatment of SSI in orthopedic, vascular or neuromodulation (neurostimulator implant) surgeries were retrieved. Two reviewers independently performed selection, quality assessment, and data extraction.The primary outcome was SSI. Secondary outcomes included infection eradication, time to eradicate infections and additional interventions related to infections. Data extraction from Electronic Patient Record (EPR) was performed to review local use. A cohort of 99 patients for whom Al-$CaSO_4$ was used during orthopedic, vascular or neuromodulation surgeries between August 18th 2015 and July 31th 2018 was analyzed. Evidence-based review and local perspective were shared with an interdisciplinary group including orthopedic and vascular surgeons, pharmacists, infectiologists, and hospital managers.

Results: Twenty-one studies on the efficacy of Al-CaSO4 were included. Available evidence suggests a beneficial effect of Al-CaSO4 in the surgical treatment of osteomyelitis whereas any conclusion can be drawn for other indications. There is insufficient evidence to support the use of Al-CaSO4 in the prevention of SSI. Few cases of hypercalcemia and renal insufficiency have been reported with Al-CaSO4 use. Data from 99 EPRs totaling 113 surgical procedures (orthopedic n = 54, vascular n = 37 and neuromodulation n = 22) showed rapid adoption and broader use of with Al-CaSO4 since its introduction in 2015. Osteomyelitis treatment represents 2.7% of cases in our hospital. Al-CaSO4 was used mainly in the context of revision surgeries (74 %)

and prevention of SSI (65 %). A lack of standardization for the preparation and documentation of the use of Al-CaSO4 was observed.

Conclusions; Limited evidence supports the use of Al-CaSO4 in osteomyelitis treatment whereas safety profile remains to be proven. Broader indications with Al-CaSO4 use than those assessed in the literature, advocate for a better control and surveillance when a new medical practice is introduce in healthcare.

Evaluation of Strategies to Prevent Overdiagnosis of Melanocytic Skin Lesion Biopsies: A Decision Analysis

Anna Tosteson[1], Stephanie Tapp[1], Linda Titus[1], Heidi Nelson[2], Gary Longton[3], Tracy Onega[1], Lisa Reisch[4], Patricia Carney[2], Raymond Barnhill[5], David Elder[6], Martin Weinstock[7], Michael Piepkorn[8], Joann Elmore[4,9]
[1]Dartmouth College, Lebanon, USA. [2]Oregon Health & Science University, Portland, USA. [3]Fred Hutchinson Cancer Research Center, Seattle, USA. [4]University of Washington School of Medicine, Seattle, USA. [5]University of Paris Descartes Faculty of Medicine University, Paris, France. [6]University of Pennsylvania, Philadelphia, USA. [7]Brown University, Providence, USA. [8]Dermatopathology Northwest, Bellvue, USA. [9]University of California Los Angeles, Los Angeles, USA

Objectives: The Melanocytic Pathology Study (MPath) reported variation in community pathologists' interpretations of melanocytic skin lesions relative to reference diagnoses developed through consensus by a panel of three experts. Little is known about the impact of second (2nd) opinion strategies on false positives (FP- overdiagnosis relative to consensus diagnosis) or false negatives (FN-underdiagnosis relative to consensus diagnosis) in melanocytic lesion diagnosis, or on patient care costs incurred within one year of biopsy.

Method: Lesion severity was classified into five classes based on the nature of clinical follow-up care required. Relative to the reference diagnoses, community pathologists overcalled (FP) or undercalled (FN) as follows: Class I (FP:7.8%), Class II (FN:62.8%, FP:12.5%), Class III (FN:54.1%, FP:5.5%), Class IV (FN:48.1%, FP:9.1%), Class V (FN:27.9%). We assessed second opinion strategies on (1) concordance between community pathologists' diagnoses and diagnoses rendered by the reference panel, and (2) patient care costs incurred during the first year following biopsy. Second opinion strategies assessed included: no 2nd opinion; 2nd opinion obtained for all lesions; 2nd opinion required for some lesions by institutional policyor based on pathologists' preference. For each second opinion strategy, decision analysis was used to estimate the expected percent of concordant diagnoses, FN, and FP. Standardized care pathways were used to estimate care costs in the year following biopsy.

Results: Without a 2nd opinion, 83.2% of biopsies received a concordant diagnosis with 8.0% FP and 8.8% FN. Concordance increased under all 2nd opinion strategies and was highest (87.4%) with universally obtained 2nd opinions, resulting in 3.6% FP and 9.1% FN While the proportion of FN cases was fairly consistent across 2nd opinion strategies (range: 8.8 to 9.2%) the proportion FP cases ranged from 3.6% to 7.6%. Per 100,000 biopsies, the costs were estimated as $118.6 million with no 2nd opinions, and 127.6 million with 2nd opinions obtained for all lesions. Second opinion strategies based on institutional policy and/or pathologist preference reduced FP cases without appreciable change in FN cases, and led to lower costs in the year following diagnosis (approximately $117 million/100,000).

Conclusions: While 2nd opinion strategies did not appreciably alter the proportion of FN cases, they did result in fewer FP cases. If selectively implemented, 2nd opinion strategies have the potential to save resources and improve care in the year following biopsy. Such strategies could be mandated through regulatory channels.

The impact of orthodontic treatment regulation in the German public health sector on the overuse of orthodontic services

Alexander Spassov[1], Bernard Braun[2], Dragan Pavlovic[3], Hartmut Bettin[1]
[1]Institute of the History of Medicine, Greifswald, Germany. [2]SOCIUM, Bremen, Germany.
[3]Department of Anestesiology and Intensive health care, Halifax, Canada

Objectives: The provision of orthodontic (dental) care is part of the German Public Health sector comprising 90% of German population. The current regulations for the provision of orthodontic care in the public sector have been adopted in 2004 as a negotiation-process between providers and public health insurance companies without patient participation.
To date patient-related data on the impact of the 2004 regulation on the quality and quantity of orthodontic services are lacking. In addition, patients (mostly adolescents) preferences and perceptions toward orthodontic treatment are almost unknown.
Therefore, we first examined the content of the 2004 regulations in terms of formal criteria for treatment access, diagnostic and treatment guidelines as well as treatment duration. Secondly, we analyzed medical health record data to obtain information on real utilization of orthodontic services as diagnostic and treatment procedures. Finally we examined patient's preferences and their role in the decision-making process.

Method: In the first stage we examined the formal criteria of the regulation of orthodontic services regarding treatment access, appropriateness criteria for diagnostic and treatment procedures as well as quality control.
In the second stage we surveyed adolescents aged 10 to 14 years (n=2.991, 29% response rate) insured by a public health care insurance company before or at the beginning of treatment related to their perceptions of toward orthodontic treatment need using validated and standardized questionnaires.
Finally, in the third stage we analyzed medical record data of patients undergoing orthodontic treatment from 2012-2017 (n=5.514) insured by a second, independent public health care insurance company to obtain quantitative data on the real utilization of detailed orthodontic services including diagnostic records, treatment procedures, costs as well as treatment duration time.

Results: Treatment guidelines adopted 2004 are mandatory for both providers and insurance companies. They regulate access through objective criteria and define criteria for the use of diagnostic and treatment procedures. Parallel guidelines adopted simultaneously regulate payment system and quality control. The described regulation has following impact on utilization of orthodontic services: Diagnostic procedures as panoramic x-ray and cephalograms were performed routinely (85-90%) although their use is limited to defined diagnoses according to radiation guidelines and should not exceed 30%.. About 64% of the patients received removable appliances despite their inappropriateness and inefficiency in comparison to fixed appliances. The mean treatment duration time was 36 months given that the published standard for a mean duration should be no longer than 18-20 months. Dentists were reported (81% of respondents) to

be the primary driver for patients to start treatment. 93% of the patients reported to no complaints before starting treatment.

Conclusions; The use of inappropriate diagnostic and treatment procedures as well as prolonged treatment duration indicate overuse of orthodontic services in German public health sector. As diagnostic and treatment procedures are strictly regulated by treatment guidelines adopted by provides and public health insurance companies the overuse of orthodontic services seems to be driven by regulators, health authorities and providers.

11:30 Wednesday August 22nd - Niels K. Jerne - De-implementation and the Challenge of Tackling Overdiagnosis at the Level of the Consultation

The STARS Back Pain App - using real time emergency department data to address overdiagnosis

Macedo Gustavo[1], Mauricio Oliveira[2], Noel Baidya[2], Hannah Storey[2], Bethan Richards[2], Chris Maher[1]
[1]University of Sydney, Sydney, Australia. [2]Sydney Local Health District, Sydney, Australia

Objectives: When low back pain is managed in the emergency department overdiagnosis and overtreatment are common. Measuring this is usually cumbersome. An online data analytics and visualisation tool was designed and developed to capture, store, analyse and visually present ED care data of patients presenting with low back pain.

Method: This project was conducted in collaboration with the Performance Monitoring, System Improvement & Innovation Unit of the Sydney Local Health District (SLHD). An online data analytics and visualisation tool was designed and created using Qlik Sense® by a multidisciplinary team of researchers, clinicians, and information technology experts

Results: The online data analytics and visualisation tool (STARS Back Pain App) was developed within the SLHD Targeted Activity & Reporting System (STARS). It displays the total number of presentations for low back pain at the three SLHD's EDs, as well as subsequent admissions to hospital. Data displayed in the app reflect ED practice for low back pain management, such as proportion of patients receiving: i) laboratory tests, ii) imaging, and iii) pain medications. The app also displays demographics and characteristics of patients, including age, gender, days and hours presenting, mode of arrival, and emergency triage category. The app allows interactive analysis using innovative visualisation techniques.

Conclusions: The STARS Back Pain App will provide emergency clinicians with a summary of their clinical performance. It will also allow us to efficiently measure unwarranted clinical variation and drive practice change using and audit and feedback approach to avoid inappropriate use of tests and treatments for low back pain.

Impact of Full-Field Digital Mammography versus Film-Screen Mammography: Systematic Review

Rachel Farber, Katy Bell, Nehmat Houssami, Kevin McGeechan, Sally Wortley, Michael Marinovich, Alexandra Barratt
Wiser Healthcare, University of Sydney, Australia

Objectives: Most breast screening programs worldwide have replaced screen-film mammography (SFM) with full-field digital mammography (FFDM) in expectation of technical, clinical and economic advantages. However, we are only just now able to begin to measure the effects of this practice shift in population screening on health outcomes among asymptomatic women eligible for population screening. This systematic review aims to assess the impact of digital mammography on breast cancer detection rates at screening and on interval cancer rates, as indicators of additional net benefit through early detection, or additional net harm from overdiagnosis.

Method: We searched Medline, Premedline, PubMed, Embase, NHSEED, DARE and Cochrane databases and identified 2139 potentially eligible papers. 31 papers were included after exclusions for relevance, duplication and other exclusion criteria. Primary outcomes are detection rates and interval cancer rates. Secondary outcomes include recall rates, false positive rates, and positive predictive values. Results are stratified by first and subsequent screening rounds.

Results: Preliminary results for primary outcomes are available at the present time and reveal a small increase in screen detected cancers across all studies. However, in 7 studies with data on interval cancer rates, we observed no statistically significant increase in detection rate, nor a reduction in interval cancer rates. Final data for these primary outcomes, and for secondary outcomes, are being prepared and will be presented at the conference.

Conclusions: Overall there has been a small increase in screen-detected cancers with the transition from film to digital mammography screening. However we observed no reduction in interval cancers, and the effect, if any, remains unclear. This observed pattern of results is consistent with a possible, small increment in cancer detection which may result in future benefit for screened women, but is also consistent with a net increase in overdiagnosis. These data reinforce the need to carefully evaluate effects of future changes in technology such as 3D mammography to ensure incremental changes to screening programs do not lead to a poorer ratio of benefit to harm from screening.

Innovative approaches to proactively identify members with special medical needs

Milan Mrekaj[1], Alvertos Fiorantis[2], Janne Kaariainen[2], Michelle Carolan[1], Epaminondas Sourlas[1]
[1]BUPA UK Insurance, London, United Kingdom. [2]BUPA UK Insurance, Manchester, United Kingdom

Objectives: Bupa's purpose is longer, healthier, happier lives. We do this by providing a broad range of healthcare services, support and advice to people throughout their lives. Bupa is commited to becoming the most customer centered health & wellbeing organisation in the world. Meeting a patient's individual care needs is right at the heart of this comittment.
The objective is to develop a predictive model which accurately identifies patients whose claiming behaviour is likely to escalate in the near future. This allows for timely referral to specialist support nurses, medical directors/forums for discussion and input and case

coordination to help the most vulnerable patients at their time of need, potentially avoiding unnecessary treatment at the same time.

Method: The modelling dataset is a sample of a half a million patients with Bupa PMI cover who claimed or were due to claim on their policy in last 15 months and more than 650 indicators. The indictors included member demographics and claims-based variables with severity (claimed amount), frequency (number of care episodes), and timing (months since last treatment) aspects.
Taking into consideration the business needs, we wanted to create a model that generates both accurate predictions and meaningful "insights", which could be converted into triggers for patients' case management. We considered several traditional statistical methods (logistic regression) and more innovative machine-learning techniques (mainly tree based models). The latter can capture very complex relationships and therefore be more accurate but often lack insights.

Results: Compared to the traditional method we ran, tree based algorithms, in particular xgboost, provided the highest accuracy, with 2 out of 3 patients correctly classified. Despite the general belief that machine-learning models are considered "black boxes", we were able to generate 3 levels of insights: • A list of the most important factors at a population level (age, previous cancer claim, etc.). • Insights at individual indicator level. For example, we found that once over 55, a patients' likelihood of their care escalating increases dramatically. • The contribution each indicator has on patient level to their individual probability.

Conclusions: This talk demonstrates the value and potential applications of predictive modelling in the UK private medical settings. Such an application enables us to create triggers for case management, pathways tailored for an individual patient, and potentially avoiding unnecessary treatment.

Do nudge-type interventions change clinician treatment, screening, and testing behaviours?: a systematic review

Mary O'Keeffe[1,2], Adrian Traeger[2,1], Tammy Hoffmann[3], Carissa Bonner[2], Jason Soon[4,5], Chris Maher[2,1]
[1]Institute for Musculoskeletal Health, Sydney, Australia. [2]Sydney School of Public Health, University of Sydney, Sydney, Australia. [3]Faculty of Health Sciences and Medicine, Bond University, Gold Coast, Australia. [4]Royal Australasian College of Physicians, Sydney, Australia. [5]Menzies Centre for Health Policy, Sydney School of Public Health, University of Sydney, Sydney, Australia

Objectives: To review all randomised controlled trials that have assessed the effect of nudge-type interventions on clinician treatment, screening, and testing behaviours across all health conditions.

Method: A systematic review of randomised controlled trials. Two reviewers are conducting the electronic searches of six databases independently.
We are defining a nudge as any simple, brief, low-cost, individual level, easy to avoid intervention intentionally designed to alter the presentation or framing of choices or information within the clinician's immediate environment (the clinical setting) to target and change a specific clinician behaviour in a predictable way, while preserving the clinician's freedom of choice, and not forbidding any clinician options/behaviours.

Studies comparing nudge type interventions to usual care, non-nudge interventions or sham nudges will be included.

Studies measuring changes in treatment delivery, testing, and screening utilisation will be included. These will include any procedure, test or treatment.
The GRADE approach as recommended in the *Cochrane Handbook for Systematic Reviews of Interventions* will be used to assess to overall quality of evidence.

Results: The review is still in progress, and so the results are unavailable at this time.

Conclusions: Finding effective ways to change clinician behaviour to ensure the provision of high-value care, and the removal of low-value or harmful care, is a priority across healthcare systems. The concept of nudging has gained traction in policy circles and as a way to improve health care delivery, but empirical evidence to support this idea is limited, especially its application to changing clinician behaviour. We do not know if there is robust evidence to show that nudging is better at changing clinician behaviours over standard usual care practices. Furthermore, depending on the targeted health condition and the specific nudge being used, it is unclear if the nudge is aiming to change clinician behaviour, to promote low value or high value healthcare (e.g. nudges to increase or decrease screening). This will be the first ever systematic review to provide results on the effect of nudging on clinician behaviours.

11:30 Wednesday August 22nd - Nielsine Nielsen Auditorium - Turning Citizens In to Patients Unnecessarily

Mono-criterial thresholds are a likely source of over-testing and over-treatment

Mette Kjer Kaltoft[1,2], Jesper Bo Nielsen[2], Jack Dowie[3,2]
[1]Odense University Hospital Svendborg Sygehus, Svendborg, Denmark. [2]University of Southern Denmark, Odense, Denmark. [3]London School of Hygiene and Tropical Medicine, London, United Kingdom

Objectives: To establish that the optimal decision for the individual person, which is preference-sensitive and multi-criterial, can be identified without any mono-criterial diagnostic classification, such as one based on threshold segmentation of a risk or disease scale. These segmented subgroups may be needed in *research* to produce the probabilistic prognoses for each option required in person-centered *decision making.* However, their use in the decisional context will often lead to suboptimal decisions for the individual, leading to over-testing and over-treatment from their point of view. Such classifications also jeopardise the obtaining of fully informed and preference-based consent to subsequent actions.

Method: Personalised decision support tools based on the technique of Multi-Criteria Decision Analysis (MCDA) have been produced for many conditions, including various cancers and cardiovascular conditions, and for the screening, testing and treatment contexts. Some are original, most are translations of published studies into MCDA format. A person-centred MCDA-based decision support tool requires the performance of all options on all criteria to be established on a continuous 0-1 scale. (Network Meta-Analyses provide the preferred evidential

basis for these performance rates, but expert opinion is almost always required for option-criteria cells that lack robust evidence. The person is themselves the expert on a criterion such as Treatment Burden). Any threshold placed on any one of the multiple criteria in the analysis interferes with the cross-criteria validity of a value-based, compensatory MCDA and hence jeopardises its overall integrity.

Results: A full set of option evaluation scores is produced by the personalised decision support tool without any reference to mono-criterial risk classes or disease stages, such as those based on PSA and Gleason score thresholds for prostate cancer, on euroSCORE thresholds in cardiovascular disease, and Frax® fracture risk thresholds in bone health. The significant weights assigned to the other criteria, most of which are harms, ensure that the optimal decision for many individuals is not that based on the focal criterion, implicitly assumed to be solely or very highly weighted. As one specific example, in the piloting of a tool with patients having a first DXA bone scan (n=32), the mean weight to Avoiding a Fracture was 43%, with 37% assigned to Avoiding Side-effects, and 20% to Avoiding Treatment Burden.

Conclusions: Testing and/or treating based on segmentation of a continuously-scaled criterion ignores the other criteria analytically, leaving them to be 'taken into account'. It also undervalues the other criteria important to the person, most of which are harms. The segmented mono-criterion is almost always benefit-focused, so the excessive weight given to it, relative to these harms, logically leads to over-testing and over-treatment for that person. The multi-criterial decision support tools confirming this pattern have been evaluated by an embedded multi-criterial Decision Quality PRO, but here we make empirical claims only at the proof-of-method level. (If the person asks to be provided with a label (e.g. 'high', 'moderate' or 'low' risk) this request should be addressed ex post engagement with the tool and they should be offered support in a personalised assessment of the benefits and harms of labeling.

Decision Support and Knowledge Translation Tools to Highlight the Benefits and Harms of Screening: An Analysis of Online Access and Dissemination of the Canadian Task Force for Preventive Healthcare Resources

Eddy Lang[1], Danielle Kasperavicius[2], Danica Buckland[2], Rossella Scoleri[2], Julia E. Moore[2], Brett Thombs[3], Sharon Straus[2]
[1]University of Calgary, Calgary, Canada. [2]St. Michael's Hospital, Toronto, Canada. [3]McGill University, Montreal, Canada

Objectives: The Canadian Task Force for Preventive Healthcare (CTFPHC) has developed a bilingual library of fifteen tools to support clinical and shared decision-making for screening in primary care. https://canadiantaskforce.ca/tools-resources/. These include clinical algorithms and frequently asked questions as well as infographics that are designed to facilitate Knowledge Translation (KT). We report the proportion of guidelines in the CTFPHC library which recommend against screening based on harms such as overdiagnosis. We also set out to describe the access and dissemination metrics for these resources from both language and geographical perspectives as well as in terms of inter-resource comparisons.

Method: Resources include clinical and patient algorithms, clinician and patient FAQs and infographics that are designed to communicate harms and benefits including over-diagnosis and false positive estimates. Resources were developed with input from clinicians and patients.

We examined website access and report on dissemination strategies for CTFPHC KT tools using website analytics as well as data on from the Canadian Medical Association Journal (CMAJ). Results are reported descriptively for the years 2016 and 2017 inclusively.

Results: Most CTFPHC guidelines (11/15) recommend against screening. The most widely viewed were the clinical algorithm for Hypertension and the "1000 person tool" for prostate cancer which highlights harms and benefits. English versions were viewed approximately 7000 times each in 2016 and 2017 with the French version accessed at 10% of that. Childhood obesity, developmental delay and cognitive impairment screening were among the least viewed with approximately 700 total views in English (2016/2017). In 2017, 70 000 hard copies of KT tools for abdominal aortic aneurysm, Hepatitis C screening and tobacco smoking in children and adolescents were distributed with the CMAJ. City-based analysis of access in Canada is in keeping with population levels with the exception of Ottawa where the CTFPHC is based and was ranked third. Both in 2016 and 2017, Brazil logged the third most sessions with the CTFPHC website after Canada and the US.

Conclusions: The CTFPHC guidelines highlight the harms of screening by demonstrating rates of overdiagnosis and have recommended against screening in controversial areas such as breast cancer screening for women aged 40-49 and colon cancer screening for men aged greater than 75; colonoscopy is also recommended against as a screening tool . The CTFPHC resources to support decision-making receive modest online traffic. The screening scenario for which the harms likely outweigh the benefits i.e. prostate cancer screening are among the most widely seen. These resources are accessed beyond the intended Canadian audience.

Loss of insurability: an inadequately elucidated harm of screening and overdiagnosis.

James Dickinson[1], John Webb[2]
[1]Family Medicine, University of Calgary, Calgary, Canada. [2]Department of Community Health Sciences, U of Calgary, Calgary, Canada

Abstract in Full: Life Insurance assists people to participate in modern society: it is often used to provide security for debts such as mortgages and business overheads, as well as for families in the event of major illness or death. After a serious diagnosis is made, it is well known that health insurance may be unobtainable: it is less known that life insurance also may be either unobtainable or prohibitively priced. We have anecdotal experience of patients who have been harmed through being unable to obtain insurance for home mortgages, after medically trivial diagnoses such as abnormal cervical smear tests.
Therfore we sought to understand the process of underwriting and how risks are assessed. We found limited public discussion of the effect of positive screening tests on insurabilty, but higher cost, postponement of contracts, or refusal are consequences.
Physicians and the public need to be aware that this is a harm of screning, and therfore insurance should be obtained before undertaking screening for disease.

Objectives; We sought to understand the process of underwriting life insurance and assessing risks, particularly about how positive screening tests and their follow-up might affect insurabilty and premium costs.

Method: We searched for evidence of company policies about how risks are assessed and the thresholds used to decide to increase the cost of policies, or decide not to insure a proposal. We also searched the insurance business journals for articles discussing this issue.

Results: Only limited information is available because much is regarded as commercial in-confidence. We found one article that explicitly described reduced life insurance eligibility as a potential harm of cancer screening. We focussed on cancer screening results and their implications. Some companies have rigid rules, such that almost any cancer diagnosis will not be accepted. Others will postpone acceptance for varied lengths of time. Some have nuanced understanding of the limited risk after diagnosis of cervical pre-cancers, and papillary thyroid cancers. However, since overdiagnosis that occurs in many cancers is indistinguishable from fatal disease, it is largely treated the same.

Conclusions: The implication is that people should be warned that having a positive screening test may cause inappropriate loss of insurability, so that they should be sure to obtain any needed life insurance before having such tests. Physicians need to be aware of this hazard to financial health caused by well-meaning screening.

Challenges and uncertainties regarding Polycystic Ovary Syndrome (PCOS) and the potential for overdiagnosis: Clinicians' views and experiences

Tessa Copp[1], Jolyn Hersch[1], Jenny Doust[2], Kirsten McCaffery[1], Anuja Dokras[3], Ben Mol[4], Jesse Jansen[1]

[1]Wiser Healthcare, University of Sydney, Sydney, Australia. [2]Wiser Healthcare, Bond University, Gold Coast, Australia. [3]University of Pennsylvania, Philadelphia, USA. [4]Monash University, Melbourne, Australia

Objectives: The diagnostic criteria for polycystic ovary syndrome (PCOS) were broadened by experts in 2003, resulting in the inclusion of women with milder forms. This has contributed to a substantial increase in the number of women diagnosed. PCOS is associated with adverse reproductive, metabolic and cardiovascular outcomes, however not all women fitting the expanded criteria may be at risk. Although signs of PCOS often first appear during adolescence, symptoms of PCOS overlap with normal signs of pubertal development, making diagnosis challenging. Currently there are no clear criteria to differentiate normal variability from the abnormality of PCOS, especially in young women. Although the diagnosis is appropriate and beneficial for some women, such as those with severe forms, the benefit is uncertain for women with milder variants of the syndrome. We aimed to understand clinicians' views about PCOS, and how they manage and communicate these uncertainties with their patients in practice.

Method: Doctors who diagnose and manage adult and/or adolescent women with PCOS (general practitioners, endocrinologists and gynaecologists) around Australia were recruited via professional organisations, specialist groups and through snowballing. We conducted 30 semi-structured telephone interviews. Topics included challenges and uncertainties with diagnosis and management of PCOS, patient communication, benefits and harms of a diagnosis, and the potential for overdiagnosis. Interviews were audio-recorded, transcribed and analysed thematically.

Results: Clinicians expressed a range of views regarding the uncertainties in the diagnosis and management of PCOS. Many clinicians reported difficulties with dispelling women's preconceptions about PCOS (e.g. never able to conceive) as a result of misinformation online,

and discussed the challenges with communicating the unpredictability of long-term consequences. Many clinicians expressed the importance of early diagnosis as it provided the opportunity to make lifestyle changes and initiate family planning. On the other hand, others were cautious about labelling women prematurely or inaccurately, and hesitant to discuss long-term consequences in the first consultation out of concern it might cause unnecessary anxiety, especially for adolescents and young women.

Conclusions / Discussion: Clinicians who diagnose and manage PCOS provide valuable perspectives regarding the current issues surrounding diagnostic criteria and uncertainties regarding diagnosis and management. The key findings and their implications for future research and clinical practice will be discussed, such as the importance of effective communication and tailored care in minimising the potential harmful impact of the diagnosis and improving patient-centred outcomes.

Pharmacovigilance and participatory medicine through social media – we are still not there

Fran Quattri
Jiangsu University, Zhenjiang, China

Abstract in Full: Turning citizens into patients may not be just unnecessary, but also premeditated. Disease branding is a marketing technique that prompts existing or potential patients to get checked and consult their doctors for diseases that may or do not even exist. It is hazardous, pervasive and poorly regulated misinformation.
We conducted a systematic literature review of scientific publications, searching for the most branded conditions over the last 15 years, and found hundreds of symptoms and diseases that we should be aware and that are daily promoted, on- and offline. The mostly branded conditions retrieved through our study are the chronic ones, followed by assumed chemical imbalances in our brain, and by pituitary disorders.
According to participatory medicine, patients in and out hospital wards have the most significant responsibility when it comes to their health. They are the ones who shall educate themselves and decide what is best for them and their dear ones.
The good news is that this is already happening.
In our digitalized era, there are patients or those speaking on their behalf who comply with the suggested treatments, observe the reactions, and share information and experiences online, through social media. These social debates should trigger changes in the clinical trial experience as well as in the post-marketing experience – but do they?
In this research, we decided to focus on two of the previously retrieved most branded conditions and followed the tweets about them posted by both those in favor and those against, by those who underwent treatment and took drugs, and those who promote both. Our outcomes confirm that we have to think deeper and broader when it comes to participatory medicine, to allow it to be genuinely participative.

Objectives: Identify the mostly branded conditions and symptoms in the medical literature in the last fifteen years (2002-2017). Understand and acknowledge the social media dialogue of users of suggested drugs and treatments and compare it with the mainstream information on the same as provided by producers and supporters.

Method: A systematic literature review of mongered conditions in the last fifteen years, followed by a comprehensive social media data analysis of tweets on two of the most branded conditions.

Results: We followed the post-marketing experience of users, the active online participation of patients and those speaking on their behalf. We collected their narratives on side effects and adverse reactions to a particular drug or treatment. In light of these inputs, we could not find significant changes in the marketing communication, production or distribution of the same drug and treatment, thus leading us to think that some users' voices were left unheard.

Conclusions: Undoubtedly, a participatory medicine that takes into account the social media discourse could lead to improvements and upgrades in drugs and treatments, a win-win situation for producers and researchers. Nevertheless, our outcomes mirror a reality that shows the opposite. Even by assuming that a patient is merely a client for the drug producer, there still exists an ethical obligation to customer care. In the medical field, customer care should represent an empathic approach to the customers' experiences with the drug or the treatment, followed by action triggered by their complains and suggestions. Participatory medicine can include online participation as an active form of social engagement to real-time health issues. In this way, even if treated as a client, the patient will at least be a happy customer.

Reducing inappropriate antibiotic use among infants through an educational intervention targeting new parents

Jette Nygaard Jensen, Ida Scheel Rasmussen, Tina Marloth, Magnus Arpi
Department of Clinical Microbiology, Herlev and Gentofte Hospital, University of Copenhagen, Herlev, Denmark, Herlev, Denmark

Background: Prescribing antibiotics for common non-serious or inaccurately diagnosed respiratory tract infections in childhood is common, especially among infants. Most infections are self-limiting. However, symptoms of infections often lead to treatment with antibiotics which increases the risk for unbeneficial treatment with no effect other than side effect(s). Studies have shown that parental knowledge about infections and antibiotics is limited and that general practitioners experience requirements for antibiotic from the parents. Maternal and child health nurses (MCHN) in Denmark visit all infants 4-5 times during the first year of life and provide counseling and guidance for new parents. Information about infections and antibiotics is not included in the visits, so far.

Objective: This study investigated whether education of new parents through MCHN could increase parents' knowledge about infections and antibiotics and reduce inappropriate use of antibiotics among infants.

Method: Parents of infants born in 2017 in three municipals in the Capital Region participated in the intervention. A booklet with information about bacteria, viruses and antibiotics was developed for new parents as target audience. Furthermore, MCHN in the three municipalities participated in a tailored seminar - organized by microbiological experts - to achieve the latest knowledge about infections and antibiotics among children. The parents received a baseline questionnaire at six months of age and a follow-up questionnaire at ten months of age measuring their knowledge about infections and antibiotics. In between the two questionnaires–

at eight month of age – the MCHN informed/educated the parents about bacterial infections and viral infections and the appropriate use of antibiotics.

Results: In April 2018, 450 parents had received a baseline questionnaire (response rate 70%) and 123 parents had received a follow-up questionnaire (response rate 60%). The study is ongoing, but the first preliminary results indicate that knowledge about infections and antibiotics among parents has increased significantly at some indicators. However, detailed results will be available at the time of the conference.

Conclusions: This study will elucidate whether education of parents is a strategy to increase the understanding of infections and antibiotics. An increased understanding of infections and antibiotics among parent is expected to decrease the unnecessary requirements for antibiotics for children with self-limiting infections.

11:30 Wednesday August 22nd - Holst Auditorium - The Role of Risk Factors in Overdiagnosis

Incremental benefits and harms of the 2017 American College of Cardiology/American Heart Association High Blood Pressure Guideline

Katy Bell[1], Jenny Doust[2], Paul Glasziou[2]
[1]University of Sydney, Camperdown, Australia. [2]Bond University, Gold Coast, Australia

Objectives: The recent recommendations from the American College of Cardiology and the American Heart Association (ACC/AHA) to lower the thresholds for defining hypertension follows a general pattern across medical specialties, whereby disease definitions are more frequently widened than narrowed. Such widened definitions usually label people as unwell, even if they are at low risk of a disease, and thus have the potential to cause harm.
We aimed to assess the incremental benefits and harms of the definition used by the ACC/AHA guideline as compared to that used by the Seventh Report of the Joint National Committee on Prevention, Detection, Evaluation, and Treatment of High Blood Pressure (JNC7) using the checklist.

Method: We compared the previous and new definitions for Hypertension using a newly developed 8-item checklist:

1. What are the differences between the previous and the new definition?
2. How will the new disease definition change the incidence and prevalence of the disease?
3. What is the trigger for considering the modification of the disease definition?
4. How well does the new definition of disease predict clinically important outcomes compared with the previous definition?
5. What is the repeatability, reproducibility, and accuracy of the new disease definition?
6. Benefit: What is the incremental benefit for patients?
7. Harm: What is the incremental harm for patients?
8. What is the net benefit and harm for patients?

Results: For the majority of adults newly classified as having high blood pressure under the ACC/AHA guideline (80% of those newly diagnosed who have 20% 10-year risk or history of CVD – 2.8 million Americans), the incremental benefits may often outweigh the incremental harms, leading to net benefit. For some people with high risk, (such as the elderly, diabetics, and people with renal disease), and for the 11% of newly diagnosed who have 10-20% 10-year risk (3.4 million Americans), the benefits and harms may often be in rough balance.

Conclusions: The 2017 ACC/AHA guideline would classify 31 million additional people in the United States as having hypertension. For the majority of these people, who are at low risk and not recommended for drug treatment, physicians should not label them as having hypertension. Physicians should continue to support healthy choices with regard to diet and physical activity regardless of whether the patient's systolic blood pressure is above or below 130 mmHg. When there is a question of starting blood pressure medication, the risk of cardiovascular disease should be estimated using a reliable risk calculator and the potential benefits and harms discussed with the patient. Some people are willing to accept a moderate increase risk of a cardiovascular event to avoid taking daily medications, increased doses or more medications, and others are not In this situation, informed and shared decision-making is essential.

Cause of death among lung cancer patients in the National Lung Screening Trial: Competing causes of death as a hallmark of overdiagnosis

Danielle Durham, Pamela Marcus
The National Cancer Institute, Bethesda, USA

Objectives: Much of the literature on cancer overdiagnosis due to screen detection focuses on the assessment of trends in mortality, incidence, and indolent cancers. Overdiagnosed cases after cancer screening also occur because screen-detected patients die of a non-cancer cause of death (competing cause of death) prior to their hypothetical symptomatic date of cancer diagnosis. The purpose of this study is to characterize cause of death after lung cancer diagnosis for participants in the National Lung Screening Trial (NLST).

Method: Participants were 55-74 years at randomization to chest x-ray or low dose computed tomography (LDCT) and were asymptomatic current or former smokers with a 30 pack-year history. The primary endpoint was lung cancer specific mortality verified through death certificate and medical records. Participants lost to follow-up were submitted to the National Death Index. Median follow-up duration was 6.5 years. We focus on LDCT as it is the standard screening modality for lung cancer detection.
We examined cause of death by cancer characteristics and detection method, calculating time from date of diagnosis to date of death (or censor if alive at the end of the trial) by plotting Kaplan-Meier curves. We enumerated the experience (health history, tumor characteristics, treatment and complications) of those who died of a cause other than lung cancer. In a sensitivity analysis we examined cause of death as reported by death certificate compared with endpoint verification.

Results: Of the 53,248 participants randomized, 2,403 were diagnosed with lung cancer during the trial, 1,079 of which were in the LDCT arm. At the end of trial follow-up, 54% did not have a reported death, 41% had a reported death due to lung cancer, and approximately 4% had a death due to a cause other than lung cancer. Among those with death from non-lung cancer causes, 60% were diagnosed with stage I lung cancer and 78% had a positive LDCT screening

test. Competing causes of death were primarily other cancers and chronic diseases such as cardiovascular and respiratory illnesses. Overall, those with a cause of death other than lung cancer appear to live longer (median 957 days (IQR 315-1307)) than those with a lung cancer cause of death (median 300 days, (IQR 109-691)).

Conclusions: Of lung cancer patients who were deceased at that end of the trial, the majority of deaths were due to lung cancer. Among those who were diagnosed but died of competing causes, it may be of public health importance to characterize their experience to elucidate the harms due to screening as well as possible health interventions for co-existing conditions.

Overinvestigation in the elderly? The role of guidelines on emergency intracranial imaging: Experience from a DGH.

Roshini Kulanthaivelu[1,2], George McInnes[3]
[1]Poole Hospital NHS Foundation Trust, Poole, United Kingdom. [2]University Hospital Southampton, Southampton, United Kingdom. [3]Poole NHS Foundation Trust, Poole, United Kingdom

Objectives: The aim of our project was to illustrate the effects that recent NICE head injury and stroke guidelines have had on the role of emergency intracranial imaging in patients older than 75 years and identify the reasons for this.

Method: We looked at the number of CT head scans performed in patients over the age of 75 years in the emergency department during the course of a month in 2013 (before the NICE head injury guidelines were updated) and compared this with those performed during a comparable time period in 2018. We analysed the number, the indications and compared these with those outlined in the NICE head injury, NICE and RCP stroke guidelines

Results: During June 2013, 22 patients over 75yrs were scanned compared with 139 in March 2018. Head injury (45% in 2013, 57% in 2018) followed by stroke (36% vs 29%) and other (headache, collapse, confusion) (14% vs 18%) were listed as indications, however the proportion of positive findings reduced in 2018 from 41% to 11% (9 to 11 patients). 54% of those performed for head injury in 2018 strictly complied with head injury guidelines. These guidelines are however open to interpretation, for example, although imaging for patients on NOACs is not explicitly advised, it is for bleeding disorders. On analysis of clinical information provided, multiple risk factors identified in each of these guidelines, were often listed as indications for intracranial imaging, rather than differential diagnoses. Incidental findings increased in 2018 (4%) and a number of patients underwent multiple CT scans in the preceeding 3-6 months for the same indication (2%).

Conclusions: Guidelines have reduced the threshold for imaging elderly patients and encouraged a shift towards investigation based on risk factors rather than individual circumstances. This unsurpisingly, has resulted in over investigation of elderly patients in particular, who often have multiple comorbidities and are subsequently at higher risk. Additionally, the increasing porportion of elderly patients in the population and seeking urgent care are factors to be considered. The only way investigations can be rationalised is through better assessment of elderly patients (a view that is shared by the Royal College of Emergency Medicine). This is difficult, especially against a background of defensive medical practice. However, even the smallest improvement can have significant cost saving implications, and so

by increasing awareness we hope to shift the focus back to choosing the right investigations at the right time for the right patient, as surely this is ultimately in the patient's best interests?

Is there really an epidemic of vitamin D deficiency? An investigation of the evidence base for vitamin D supplementation.

Mark Bolland[1], Andrew Grey[1], Alison Avenell[2]
[1]University of Auckland, Auckland, New Zealand. [2]University of Aberdeen, Aberdeen, United Kingdom

Objective: Rickets and osteomalacia are rare diseases caused by vitamin D deficiency. The US Institute of Medicine considers that there is a risk of vitamin D deficiency with serum 25(OH)vitamin D <30nmol/L, and potential risk between 30-50 nmol/L, relying on evidence from surrogate markers, e.g. parathyroid hormone suppression, unmineralised bone on biopsy. 25(OH)D shows considerable analytical variability, is not the bioactive metabolite, and is expensive to measure. Academic enthusiasm for, media coverage of, and public interest in, vitamin D supplementation to prevent a wide range of diseases, with some claims that serum 25(OH)D ≥75nmol/L is optimal, have led to marked increases in requests to measure 25(OH)D (119% increase in NE Scotland over 4y). We investigated whether clinical endpoints from randomised controlled trials (RCTs) supported supplementation for diseases other than rickets and osteomalacia.

Method: In December 2015, we searched Pubmed, recent systematic reviews, and trial registries for RCTs of vitamin D with surrogate or clinical endpoints in adults. We examined whether vitamin D affected a range of clinical endpoints according to baseline 25(OH)D status, whether mean/median baseline 25(OH)D in RCTs changed over time, and whether ongoing trials will evaluate clinical endpoints from vitamin D supplementation in populations at risk of deficiency.

Results: We found 547 RCTs of vitamin D supplementation, with 137 reporting clinical endpoints, and 118 reporting baseline 25(OH)D. Mean/median baseline 25(OH)D was <25, 25-49, ≥50nmol/L in 12(10%), 62(53%), and 44(38%) of RCTs, respectively. No effect was evident for nonskeletal outcomes in the full dataset. Of 12, mostly small, RCTs in populations with 25(OH)D <25nmol 8 had neutral results and 4 showed benefit (3 primary endpoints, 1 secondary endpoint). Trials reporting subgroup analyses for 25(OH)D <20-32nmol/L recapitulated the main analyses in all 7 trials, with only 1 trial showing benefit. Baseline 25(OH)D increased over time, and large (≥1000) ongoing trials are likely to have baseline 25(OH)D >40-50nmol/L.

Conclusions: The evidence supporting supplementation with vitamin D in populations with 25(OH)D <25nmol/L is weak for outcomes other than rickets and osteomalacia, and does not suggest benefit above this threshold. Assessing vitamin D status for healthy populations, with little risk of rickets or osteomalacia from very low sunlight exposure, is not clinically useful.

The Impact of a Regulatory Nudge on the Overuse of Low Dose Codeine in Manitoba, Canada

Shawn Bugden[1], Kevin Friesen[2], Jamie Falk[2]
[1]Memorial University of Newfoundland, St. John's, Canada. [2]University of Manitoba, Winnipeg, Canada

Objectives: Low-dose codeine products (≤ 8mg codeine/tablet) have been available over-the-counter from Canadian pharmacies for many years. In the era of the North American opioid

crisis there has been concern about the volume of low-dose sales and their potential to contribute to the problem. This concern extended to the potential for acetaminophen/paracetamol toxicity. There has been widespread criticism about the lack of assessment by pharmacists in the provision of these products. In 2016, Manitoba became one of the few provinces to impose further regulation on the provision of low-dose codeine products. After February 2016 these products could only be dispensed with a prescription. Pharmacists were able to provide these prescriptions but only after a complete and documented assessment (sign and symptoms, medical history, length and severity of condition). In this study we explore the impact and potential unintended consequences of this regulatory nudge on low-dose codeine use.

Method: A drug utilization analysis was conducted for the year before and after the policy change. Sales information was reported by the College of Pharmacists of Manitoba from the Quintiles IMS Canadian Drug Store and Hospital Purchases Audit. Prescription codeine use was assessed using the Manitoba Population Research Data Repository using data released in January 2018. The effects of this policy change on the use of other prescribed opioids were assessed by piecewise Poisson regression analysis.

Results: In the year before the regulatory nudge sales data reported 52.5 million low-dose codeine tablets sold in Manitoba. This fell by 94% to 3.3 million tablets in the year after the change and this was confirmed in the research database which also showed dispensations for 3.3 million tablets. The prescribing of low-dose codeine was split evenly between pharmacists (49%) and physicians (48%). In its first year as a prescribed product, low-dose codeine use declined from 0.30 to 0.26 million tablets/month. Poisson regression analysis found no significant effect of this changeover on the trends in the utilization of higher strength codeine products, tramadol, and oxycodone with acetaminophen products.

Conclusions: Manitoba's regulatory nudge reduced low-dose codeine usage by 94% without obvious negative implications. Other provincial or national policy changes should be considered to implement this simple change on a nationwide basis.

The use of theories and frameworks to understand and address the reduction of low-value healthcare practices: a scoping review

Gillian Parker, Nida Shahid, Whitney Berta
University of Toronto, Toronto, Canada

Background: There is recognition that the overuse of procedures, testing and medications strains the healthcare system financially and can cause unnecessary stress and harm for patients. In recent years, several initiatives—such as Choosing Wisely—have targeted the reduction or elimination of low-value practices in healthcare. Researchers have begun to use theory or develop frameworks to elucidate the dynamics of de-implementation and support efforts to reduce low-value practices. The purpose of this scoping review was to identify and characterize the use of theories and frameworks to understand and address the reduction of low-value healthcare practices.

Methods: We conducted a systematic review of MEDLINE, EMBASE, CINAHL and Scopus databases from inception to January 2018. Building on previous research, 44 key terms were used to search the literature. To be included, papers had to present an explicit theoretical

approach or framework. The database searches identified 725 unique articles for which titles and abstracts were screened for inclusion; 74 items were selected for full text review.

Results: Forty studies met the inclusion criteria. Over 50% of included articles were published in the last two years. Of studies which used a theoretical approach, the majority used psychological theories, such as the Theory of Planned Behaviour or applied behavioural science concepts to develop interventions. The included studies demonstrated a maturating of the use of theory in this field, with a progression from the use of classic theories to multi-theory frameworks to dual processing models. Theories or frameworks were used primarily to identify barriers and facilitators or develop conceptual clarity. Articles used theory to understand how provider decision-making, knowledge or perceptions of social pressure play a role in overuse. The majority of studies addressed low-value care at the provider level. Antibiotic overuse, polypharmacy and appropriate prescribing were the practices targeted most frequently in the included studies.

Conclusions: De-implementation is an emerging field of research. This scoping review was the first to review the use of theory in efforts to reduce low-value practice. The results of this review can provide direction and insight for future primary research in the use of theory to support de-implementation and reduction of low-value healthcare practices.

11:30 Wednesday August 22nd - Henrik Dam Auditorium - Psychiatry and Overdiagnosis

MEDREV: feasibility study of a pharmacy de-prescribing and health psychology intervention to improve care for people with dementia with BPSD in care homes

Dr Ian Maidment on behalf of the MEDREV Team
Aston University, Birmingham, United Kingdom

Objectives: "Behaviour that Challenges" is common in older people with dementia in care homes and traditionally treated with anti-psychotics. Guidelines across the globe have aimed to reduce the use of anti-psychotics in people with dementia and therefore reduce harm. These initiatives have only been partially successful.
This submission reports results on a feasibility study (MEDREV) funded by the NIHR. The study combined a medication review by a specialist dementia care pharmacist combined with staff training with the overall objective of limiting the inappropriate use of psychotropics.

Method: Care homes were recruited. People meeting the inclusion (dementia diagnosis; medication for behaviour that challenges), or their personal consultee, were approached. A specialist dementia care pharmacist reviewed medication and recommended the de-prescribing of inappropriate psychotropics. Care staff received a 3-hour training session promoting person-centred care and GPs brief training.
The primary outcome measure was the Neuropsychiatric Inventory-Nursing Home version (NPI-NH) at 3 months. Other outcomes included quality of life (EQ-5D/DEMQoL), cognition (sMMSE), health economic (CSRI) and prescribed medication. Qualitative interviews explored expectations and experiences.

Results: Five care homes and 34 of 108 eligible residents (31.5%) were recruited. Medication reviews were conducted in 29 of the 34 residents recruited and the pharmacist recommended de-prescribing or reviewing a medication in 21 of these. Fifteen (71.4%) of these 21 medications were antidepressants. 57.1% (12 of 21) of recommendations were implemented; although implementation took a mean of 98.4 days. Themes for non-implementation of recommendations will be presented. In total, 164 care staff received training; 142 from the care homes and 22 GP staff. Twenty-one participants (Care Home Managers=5; General Practitioners=3; Nursing Staff=2; Care Staff=11) were interviewed.

Conclusions: People with living with dementia continue to be prescribed inappropriate psychotropics for beaviour that challenges and the current policy may simply have shifted prescribing from anti-psychotics to other equally inappropriate psychotropics. We found the study feasible although our approach may need to be modified to improve the uptake of our reviews and reduce the delay in implementation.

Psychiatrization of Society: A Wake-up Call for Debate

Timo Beeker, Samuel Thoma, Martin Heinze, Sebastian von Peter
Medical School Brandenburg Theodor Fontane, Rüdersdorf, Germany

Abstract in Full: Over the last decades, there has been a marked increase in diagnosed psychiatric disorders worldwide causing a growing burden for health-care systems and society. In the USA, a full half of the population is claimed to meet the criteria for a DSM-IV disorder over the course of their lives. Typical explanations are that advancing destigmatization of mental illness and psychiatric treatment may be uncovering the real number of mental disorders for the first time. Also, contemporary working and living conditions are held responsible for causing a massive increase in psychiatric morbidity. On the other hand, changes in epidemiology, inflation of psychiatric diagnoses and widespread overdiagnosis are criticized as having substantially contributed to this increase. Although all of these factors seem relevant, the question remains, if there could be a more general, higher-order process behind these developments, both connecting and explaining them. We identify this process as a progressing psychiatrization of society, which causes as well as reflects the rising diagnoses of mental disorders.

Psychiatrization and its various sub-processes are increasingly relevant in the light of ongoing and profound social changes in countries of the Global North, which coincide with a reorientation of psychiatric services (e.g. through digitalization and gradual substitution of inpatient-care with outpatient services). Simultaneously, Western psychiatric concepts and mental health services are exported to the Global South widely unchallenged. In this context, there is an urgent need for a broad debate about psychiatrization and its numerous negative consequences: On an individual level, these are mainly related to overdiagnosis and overtreatment, e.g. medication adverse effects and harms from long term use. Through pathologization of minor disturbances of psychological wellbeing and various life issues, psychiatrization can also promote disempowering changes to subjectivity and sense of the self, co-produce avoidable patient careers and create dependencies on mental health services. On a societal level, psychiatrization predetermines preferences for ineffective and short-term medical interventions which require individuals to cope with social problems, instead of finding collective long-term solutions. These would be located rather in the political than the medical field. From a public health point of view, psychiatrization might stimulate a further adaptation of mental health services to the needs of the "worried well" and borderline cases while the provision of care for the severely and chronically ill is reduced.

Objectives: Our aim is to stimulate debate and research about psychiatrization as a highly complex, diverse, globally effective process of great importance for society as a whole. Its effects can be harmful to individuals in many ways and detrimental to society and public healthcare systems. Empirical preliminary studies on psychiatrization processes and their effects are largely missing. Theoretical work on single aspects can be found essentially outside of psychiatric research itself, i.e. in the canon of antipsychiatric literature of the 1960s and 70s or sociological studies on medicalization. Further studies will need to update theoretical understanding of psychiatrization and fill conceptual gaps in order to initiate a theory-led, transdisciplinary research program, which can empirically establish its various manifestations and consequences. Additionally, further research will have to acknowledge the vast body of experience of (anti-)psychiatric researchers and organizations through collaborative projects between professionals and service users.

Overestimation of depression prevalence in meta-analyses via the inclusion of primary studies that assessed depression using screening tools or rating scales rather than validated diagnostic interviews

Brooke Levis[1,2], Xin Wei Yan[2], Chen He[1,2], Andrea Bendetti[1], Brett D. Thombs[1,2]
[1]McGill University, Montreal, Canada. [2]Lady Davis Institute for Medical Research, Montreal, Canada

Objectives: Estimates of the prevalence of depression should be based on validated diagnostic interviews to determine case status and not depression screening tools or symptom rating scales, which are not intended for this purpose. Using cutoff scores from screening or rating scales to estimate prevalence tends to over-estimate prevalence substantially, particularly in low-prevalence groups (Thombs et al, CMAJ, 2018). Authors of meta-analyses, however, sometimes base estimates of depression prevalence on depression screening tools or rating scales. The objectives of this study were to (1) determine what depression ascertainment methods are commonly used to classify cases of depression in primary studies included in meta-analyses of depression prevalence; (2) determine what terminology is used in meta-analyses to describe prevalence rates based on different methods; and (3) examine the extent to which including studies that count positive screens as cases results in the overestimation of depression prevalence in recent meta-analyses of depression prevalence.

Method: We searched PubMed from 2008-2017 for meta-analyses that reported pooled prevalence of depression in the abstract. For each included meta-analysis, we recorded whether the abstract reported a pooled prevalence based on primary studies that used (1) diagnostic interviews only, (2) screening tools or rating scales only, and (3) a combination of diagnostic interviews, screening tools or rating scales, and other methods (e.g., chart diagnoses, self-report). If multiple prevalence estimates were reported for a category (e.g., for different subgroups), we extracted data only for the first prevalence estimate in the category. For each meta-analysis, for each prevalence estimate, we recorded whether the abstract indicated the types of ascertainment methods included, the terminology used to describe the prevalence value, and the number of studies pooled, the pooled sample size, and the pooled prevalence. For the combination category, we determined how many pooled primary studies used a validated diagnostic interview.

Results: 69 eligible articles were included, and 81 pooled prevalence estimates were extracted (9 for diagnostic interviews only, 36 for screening tools or rating scales only, 36 for combinations).

Mean pooled prevalence was 17% for interviews only, 31% for screening tools or rating scales only, and 22% for combinations. Among 11 articles that reported prevalence for interviews or for combinations and also for screening tools or rating scales, prevalence was always higher based on screening tools or rating scales. Only 10 of 36 meta-analyses that combined studies that used screening tools or rating scales indicated this in the abstract; 22 of 36 referred to the prevalence as for "depression" or a "depressive disorder", despite using screening or rating tools. 5 studies that did not report a prevalence based on diagnostic interviews in the abstract provided one in the text; on average, it was half the value reported in the abstract.

Conclusions: Most meta-analyses of depression prevalence combine prevalence estimates from primary studies that used methods other than validated diagnostic interviews to assess depression. Some meta-analyses describe the resulting pooled prevalence as prevalence of "depressive symptoms", but most describe prevalence of "depression" or "depressive disorders" despite not assessing disorders. Among 69 included meta-analyses, prevalence of depression was always higher when based on depression screening tools or rating scales than when based on diagnostic interviews or a combination of ascertainment methods. Including studies that assess depression based on methods other than validated diagnostic interviews exaggerates the true prevalence of depression. These are preliminary results. Final results will include, for the combination category, a comparison of published prevalence estimates and prevalence estimates based on re-analysis using only the studies that assessed depression using a validated diagnostic interview.

Applying the Checklist for Modifying the Definition of Disease to Attention Deficit/Hyperactivity Disorder (ADHD) age of onset criterion (AOC)

Rae Thomas, Sharon Sanders, Paul Glasziou, Jenny Doust
Centre for Research in Evidence-Based Practice, Bond University, Gold Coast, Australia

Objectives: Widening disease definitions is a major driver of overdiagnosis. In response, the Preventing Overdiagnosis working group of the Guidelines International Network developed a checklist to provide guidance on issues to consider when modifying disease definitions. This checklist recommends panels outline definition changes and the trigger for change, and examine research informing the potential changes to prevalence, the prognostic ability and precision of the disease definition, the potential benefits and harms of the disease definition and balance between them.

Using this checklist as a framework, we examined the documented considerations of the panel responsible for modifying the ADHD diagnostic criteria, focusing on the age of onset (AOC) criterion, which widened the definition by changing the requirement that symptoms causing impairment need to be present before the age of 7 (DSM-IV), to the presence of symptoms before age 12 (DSM-5).

Method: For the checklist items requiring panels consult research studies (e.g. prevalence, prognosis, precision, benefit and harm), we examined the research considered by the panel modifying the DSM-IV ADHD AOC. We recorded the research studies identified by the panel and described how these informed their conclusions. We appraised these studies for risk of bias (ROB), and on their ability to address the checklist item (e.g. limitations in design and generalisability). We conducted searches to identify research that would have been available to the panel at the time of the modification and compared it to the research considered by the

panel. We also identified studies related to checklist items published since the modification, assessed ROB, extracted data and compared the consistency of findings of these studies to the conclusions reached by the panel.

Results: DSM-5 panel documentation cited two studies: a "systematic literature review" of studies related to the AOC conducted by panel members (the methods of this study were not documented), and a longitudinal cohort study assessing prevalence. Cited within the review, were studies we assessed as related to some checklist items (precision, prognosis and benefit) however no reference to these constructs were reported in the document and no systematic appraisal of this research or comment on the strength of the evidence was provided. The cohort study reported 'negligible' change in prevalence. We appraised the ROB in the cohort study to be low, however, study design precluded confidence in the prevalence estimate, and subsequent research reported larger prevalence increases with AOC changes. We found overlap in the studies identified by the panel and the studies we identified as being available at the time which we assessed for ROB and strength of evidence.

Conclusions: Minimal documentation of the considerations and decisions of the panel limits transparency and makes it impossible to judge the rigor of the process behind the modifications to ADHD diagnostic criteria. The information available suggests that rigorous consideration of important issues identified by the checklist did not occur, although this may be a problem of reporting. Panels modifying DSM ADHD diagnostic criteria comprise clinical and research leaders. Critical thinking and rigorous methods are their forte. Future changes to DSM diagnostic criteria should ensure all process are documented clearly and rigorous appraisal of research used to support any further changes. Use of the checklist for modifying disease definitions would ensure a more thorough and transparent assessment of important issues prior to recommending changes, and that these changes can be more robustly supported.

Drug treatment of ADHD in children and adolecents – tenuous scientific basis

Charlotte Lunde[1], Henrik Vogt[2], Ole Jacob Storebø[3]
[1]Lovisenberg Sykehus, Nic Waals Institute, Oslo, Norway. [2]Allmennmedisinsk forskningsenhet, Norwegian University of Science and Technology, Trondheim, Norway. [3]Syddansk Universitet, Odense, Denmark

Abstract in Full: This presentation highlights recent publications showing that the evidence for the efficacy of pharmaceutical treatment in children with ADHD diagnosis is weak,especially over the long term and that it comes with previously unrecognized risks of side-effects.Despite widespread pharmacological treatment of ADHD, the first comprehensive systematic review on the use of methylphenidate in children and adolescents was not published until 2016. A Cochrane group examined 185 randomised controlled trials with more than 12 000 children and adolescents and concluded that existing studies were of such low quality due- to high risk of bias – that it was not possible to say for certain whether methylphenidate is beneficial for children and adolecents with an ADHD diagnosis. An additional review from Cochrane concludes that the evidencebase for the use of amphetamines in ADHD is similarily weak. Spring 2017, the latest results form the Multimodal Treatment Study of Children with ADHD (MTA) concluded, 16 years after the start of the study, tht longterm use of stimulants is associated with suppression of adult height, but no reduction in symptoms.

A new Cochrane review (in print 2018) included 260 relevant non-randomized studies with 2,283,509 participants worldwide. The occurrence of serious adverse events in the comparative cohort and patient-controlled studies was 1.36 times more frequent in participants that used methylphenidate when compared to controls. The proportion of withdrawals from methylphenidate due to serious adverse events in cohort studies without a control group was 1.50%, and the proportion of withdrawals due to adverse events of unknown severity was 7.30%. In addition, the proportion of withdrawals due to non-serious adverse events was 6.2%, and 16.2% for unknown reasons.

These findings are discussed in light of overdiagnosis. Children with ADHD have genuine and serious problems. However, we cannot ignore the fact that several recent studies have yielded only weak evidence to support the extensive use of medication that occurs today. This state of affairs should trigger renewed public and expert discussion on the pharmacological treatment of ADHD in children and adolecents.

Objectives: A presentation of recent reviews on the pharmacological treatment of ADHD in children and adolecents in light of overdiagnosis.

Method: discussant review

Results: Recent studies have yielded only weak evidence to support todays extensive use of ADHD-medication to childen and adolecents.

Conclusions: The state of affairs should trigger renewed public and expert discussion

Study doping: Prevalence of psychostimulants use among university students in the Netherlands.

Huib van Dis[1,2], Samuel Rasche[1], Isabelle van Bennekom[1], Jop van Kempen[1]
[1]University of Amsterdam, Amsterdam, Netherlands. [2]

Objectives: In the Netherlands, a considerable increase is seen in number of prescriptions for psychostimulants (e.g. methylphenidate, dexamphetamine, atomoxetine) in the national health database (GIP-ZIN, 2018). A more then 5-fold increase in users is reported, from 42.000 users in 2002 to 221.000 users in 2016. The official indications given by the regulatory authority CBG (Medicine Evaluation Board) are ADHD in children and adolescents and narcolepsy. Off-label use for the diagnoses adult ADHD is also increasing. Regular use of prescribed and non-prescribed psychostimulants is seen among university students with the aim of improving study-results. This pilot is aimed at clarifying the prevalence psychostimulant use on a university campus.

Method: 493 Dutch students of the campus van de University of Amsterdam participated November/December 2017 in digital survey concerning the use of stimulants, the use of other medications and recreational drugs. Detailed information was gathered on the reasons for the use of stimulants: medical use, recreational use, to improve study-results during preparation and of exams, during exams, and as a help in writing bachelor and master-theses

Results: Surveys of 491 students were available for analysis: mean age was 21 years. 44 students (9 %) reported a diagnosis ADHD or ADD. Use of psychostimulants (ever) was reported by 37% of the students; 19 % used stimulants recently (past few months). Methylphenidate (Ritalin, Concerta) was used most frequently (85%). 34% of the 184 users of psychostimulants had been

diagnosed with having ADHD of ADD, 66% had no diagnosis. 78% of all stimulants users reported they had been using psychostimulants for studying purposes.

Conclusions: Many users of psychostimulants are not well informed about cognitive enhancement effects that could be expected, or about somatic and psychological adverse side effects. Some students indeed report increase alertness, cognitive enhancement effects as increased concentration and staying awake. Most frequently reported adverse side effects were unexpected sleep problems, loss of appetite, palpitations, dry mouth, anxiety, feelings of disorientation. There seem to be a need among students to be informed more adequately about possible adverse side effects of study doping.

14:00 Wednesday August 22nd - Niels K. Jerne - De-implementation and the Challenge of Tackling Overdiagnosis at the Level of the Consultation

De-Implementing Unconsidered Cancer Screening: *Primum non Nocere*

Ronald Adler
University of Massachusetts Medical School, Worcester, MA, USA

Abstract in Full: People – clinicians, patients, and policy-makers – overestimate the benefits and underestimate the harms of cancer screening. Though its net benefits are typically over-estimated, people find the rationale for it – namely, that early detection saves lives – easy to understand. This concept has been widely promoted and accepted even though it is rarely true.

The primary benefits of cancer screening are reassuring healthy people that they don't have cancer and reducing cancer-specific mortality. However, the cancer-specific mortality benefits are surprisingly small – undercut in part by improved cancer treatments. Meanwhile the harms of cancer screening are common and substantial. They include overdiagnosis of indolent cancers, false positive test results and the resulting cascade of additional interventions, opportunity costs, and the extraordinary associated financial expense. Such factors are particularly important because the population participating in cancer screening is – by definition – healthy.

The dramatic success of cervical cancer screening proved the concept that early detection can save lives. The obvious – though erroneous – assumption was that "early detection saves lives" was a universal principle that could be successfully applied to other cancers. Unfortunately, that promise has not been fulfilled. This is primarily due to the biology of other cancers and the fact that they do not share the features that make cervical cancer so amenable to screening: it grows slowly, is frequently progressive, and has easily-detected precursor lesions that are readily treated.

Unfortunately, the "conventional wisdom" that cancer screening is fundamentally beneficial is now deeply entrenched in the minds of clinicians and patients and in policies and guidelines. This is driven by multiple factors, including advocacy campaigns, misguided "quality" metrics, and the fact that while most people know cancer survivors, they confuse improved survival rates with improved mortality rates.

With the probable exceptions of cervical and colorectal cancer screening, widespread, uncritical promotion of cancer screening is not justified by the evidence. As cancer screening efforts have intensified, and cancer treatments have improved, the net effects tend towards harms over benefits. In this context, it becomes appropriate to ask, "Is it ethical for organizations to promote cancer screening (through guidelines, media campaigns, etc.), and for clinicians to offer patients cancer screening *without accurately describing the anticipated potential benefits and harms?*"
In this seminar, these and other questions will be explored. The evidence will be reviewed, and the knowledge gaps – especially among clinicians – identified. Strategies for addressing the challenge of de-implementing unconsidered cancer screening will be developed and critiqued.

Objectives: After participating in this session, participants will:

1. more clearly understand the balance of benefits and risks associated with various cancer screening programs.
2. have developed strategies for more effectively communicating the benefits and risks of cancer screening.
3. have grappled with some of the most challenging questions regarding cancer screening.
4. have considered various strategies for de-implementing unconsidered cancer screening.

Method: Pictograms, which can be effective with all stakeholders, will be used to communicate the likelihood of the most significant cancer screening outcomes: cancer-specific mortality, false positives, and overdiagnosis.
Per participant interest, the following questions will be addressed:

1. Is it ethical for clinicians to offer patients cancer screening or for organizations to promote cancer screening (through guidelines, media campaigns, etc.) *without accurately describing the potential benefits **and** harms?*
2. Assuming it is time to de-implement unconsidered cancer screening, what strategies will be most effective?
3. How best to communicate that cancer screening decisions should be approached as an evaluation of trade-offs, balancing the benefits against the harms in the context of individual characteristics, values, and preferences?
4. How should we prioritize efforts to improve cancer screening?
    A. Promote shared decision-making.
    B. Match intensity to sub-population risk.
    C. Modify intervals, techniques, etc.
    D. Modify thresholds for action.
    E. Develop better tests.

Results: Clinicians are easily misled by cancer screening statistics, mistaking survival for mortality benefit, failing to appreciate the importance of lead-time bias, and not recognizing the significance of overdiagnosis. Evidence that demonstrates the benefits and harms of cervical, colorectal, lung, breast, and prostate cancer screening, including variation in the performance characteristics and value of different screening programs will be presented. Except for cervical cancer, this will include a lack of all-cause mortality benefits. The magnitude of harms, including false positives and overdiagnosis will be compared. Evidence that clinicians and patients overestimate the benefits and underestimate the harms will be shared. While shared decision-making has been shown to lead to "better" decisions, it is undermined by inaccurate perceptions of benefits and risks. Decision aids can be helpful in multiple ways: addressing

knowledge gaps, promoting communication, and illustrating the trade-offs that are pertinent in choosing whether to be screened.

Conclusions: It is time to acknowledge that unconsidered, widespread cancer screening may be doing more harm than good – and at substantial human, financial, and opportunity costs. Careful examination of cancer screening programs reveals some variation in outcomes, but they generally fall far short of their perceived value. Though there are individual cases in which early detection saves lives, this does not accurately describe cancer screening programs, because in a population, there is no all-cause mortality benefit. Meanwhile, as cancer screening has been more widely implemented and its downstream effects have become more invasive, the harms of false positive results and overdiagnosis have grown in significance. It is time to de-implement widespread cancer screening in ways that retain the potential value of cancer screening while mitigating the harms. This seminar aspires to identify the best paths forward as we address some of the most pressing challenges regarding cancer screening.

Overdiagnosis and error in general practice: the need of a new approach

SERGIO MINUÉ LORENZO[1], ALBERTO FERNANDEZ AJURIA[1], CARMEN FERNANDEZ AGUILAR[1], JOSE JESUS MARTIN MARTIN[2] [1]ANDALUSIAN SCHOOL OF PUBLIC HEALTH, GRANADA, Spain. [2]UNIVERSIDAD DE GRANADA, GRANADA, Spain

Objectives: Diagnostic error (DE) is defined as a diagnosis unintentionally delayed (sufficient information was available earlier), wrong (another diagnosis was made before the correct one), or missed (no diagnosis was ever made) (Graber,2012). More recently, it has been proposed to include overdiagnosis in DE definition (Zwaan, 2015).

Usually, the standard to analyze diagnostic process is the appropriateness of code assigned to a problem using the International Classification of Diseases (ICD). However at least one third of common symptoms do not have a disease-based explanation (Kroenke 2014). This approach mean all expressions of illness must be associated to a disease label, against the context of general practice (Heath,2011).

Objectives:

1. To assess diagnostic codes allocated to new cases of dyspnea in General Practice
2. To study diagnostic error in patient attended by dyspnea in Primary Care
3. To determine the relationship between the occurrence of diagnostic error and patient harm.

Method: Cohort study of new episodes of dyspnea in patients receiving care from general practitioners (GPs) and GPs trainees at Primary Care Practices in Granada (Spain).

In addition to filling out the electronic medical record of the patients attended, each physician fills out two specially designed questionnaires about the diagnostic process performed in each case of dyspnea: the first questionnaire includes questions on the physician's initial diagnostic impression, the three most likely diagnoses (in order of likelihood), and the diagnosis reached after the initial medical history and physical examination. It also includes items on the physicians' perceived overwork and fatigue during patient care. The second questionnaire records the final diagnosis once it is reached. The complete diagnostic process is peer-reviewed to identify and classify the diagnostic errors.

Our methods is based on the Zwaan methodology on Diagnostic errors in dyspnea patients treated in Dutch hospitals (Zwaan,2009).

Results: Nine general practitioners and five GP trainees participated in this study. Initially, 373 records were collected for a year. 217 new cases of dyspnea were finally registered. Each of them received a diagnostic code by their general practitioner. The average number of cases recorded by each physician was 21.Mean time at until diagnosis was 32.89 days Final diagnosis included acute asthma exacerbation(37),acute bronchitis (34), asthma (30), COPD (27), other disease of respiratory system (26), anxiety (24) and congestive heart failure (20). In peer review process, Diagnostic testing requested were considered appropriate in 81.56% of the cases, Diagnostic process performed was considered not appropriate in 12.44%, and Diagnosis was wrong in 13.2% (32 of 217). Diagnostic error-related harm was not found. However, 4,15 % of the cases in which a diagnostic tests were requested, adverse effects were found.

Conclusions: Although 1/3 of common symptoms don't have a disease-based explanation,all new cases of dyspnea were categorized with a diagnostic code.It could can produce overdiagnosis, unnecessary tests and adverse effect. In our study, DE (13,2%) was similar to the only akin study (Zwaan,2012), performed in five Dutch hospitals(13,8%). While in our study, carry out in PC, harm was not found as result of DE, patient was harmed in 11,3% cases in the Dutch study. The use of disease classifications as standard to define DE in PC may not reflect the general practice complexity: broad spectrum of symptoms, highest frequency of illness over disease, time as strategy to diagnosis (Irving, 2013). According to this, diagnosis-related harm (produced by missing, delayed or wrong diagnosis as well as early diagnostic and unnecessary procedures) could be an appropriate way to study DE in PC

The Imaging Learning Network (ILN): a healthcare provider - industry collaboration to reduce inappropriate imaging through clinical decision support.

William Baughman, MD[1], Robert Ferguson, MD[1], Eric Gross, MD[2], Jonathan Darer[3]
[1]MetroHealth, Cleveland, USA. [2]UC Davis, Sacramento, USA. [3]Siemens Healthineers, Malvern, USA

Abstract in Full: Computerized physician order entry of imaging in the US can be characterized by poor communication and frequent inappropriate or incorrect orders. The ILN was launched in June 2016 as a collaboration between eight major health systems and a clinical decision (CDS) provider to: a) support the development of high-quality clinical guidelines, b) enhance end-user interactions with CDS and c) accelerate learning through multi-site evaluation. ILN members have developed a logical framework and clinical indication vocabulary that enables imaging appropriate use criteria to be translated into digitally executable interventions. Ordering providers including emergency medicine, primary care, oncology, radiologists and vendor meet weekly and use an online portal to review clinical guidelines and optimize CDS informatics. Substantial attention is paid to the quality of submitted indications to assess the impact of imaging CDS upon provider communication. Current CDS interventions include suspected pulmonary embolism (PE), headache, minor head trauma, neck pain, low back pain, hip pain, shoulder pain and lung cancer.

The ILN members are designing imaging management programs to accompany imaging CDS implementation including establishing performance benchmarks (e.g., rates of imaging, imaging yield, adherence to guidelines), feedback reports, identification of outliers and the use of peer-to-peer interactions.

Conclusions: The Imaging Learning Network is a provider – industry collaboration well positioned to design, implement and evaluate interventions to enhance the value of imaging.

Implementing tools for reduction of overdiagnosis and overtreatment in clinical practice - position papers

Ronen Bareket[1,2,3], Anat Gaver[4,2,3], Eitan Lavon[4,2], Ian Miskin[2,5,6]
[1]Departement of family medicine, Sackler School of Medicine, Tel-Aviv University, Tel Aviv, Israel. [2]sraeli Society for the Reduction of Overdiagnosis, Tel Aviv, Israel. [3]Department of Family Medicine, Rabin Medical Center and Tel Aviv & Dan districts, clalit, Tel Aviv, Israel. [4]Departement of family medicine, Sackler School of Medicine, Tel Aviv, Israel. [5]Department of Family Medicine, Hadassah-Hebrew University Medical School, Jerusalem, Israel. [6]Department of Clinical Microbiology and Infectious Diseases, Hadassah Medical Center, Jerusalem, Israel

Objectives: The Institute for Quality in Medicine (IQM) is one of the arms of the IMA (Israeli Medical Association, representing professionally physicians in Israel). The institute publishes position papers and guidelines written by professional societies of the IMA. These documents are part of the curriculum for specialization qualification and considered as guidelines for common practice by the courts. The Israel Society for the Reduction of Overdiagnosis and Overtreatment (ISROD) was founded in 2016 under the auspices of the IMA.
A position paper, written by a multidisciplinary team, was recently published by ISROD and IQM. The paper addresses all professional societies and guideline writing panels; detailing rationale and need to address overdiagnosis and overtreatment (OdX), elaborating mechanisms which lead to OdX and giving recommendations aimed at implementation of methods for reducing OdX in new clinical guidelines.
We aimed to write a new position paper for the clinicians.

Method: A new multidisciplinary team was established……. In order to understand and relate to the different clinical contexts in which overdiagnosis happens. (as many types of clinical settings and dilemmas:) The team members come from various fields of medicine: internal medicine, emergency medicine, hematology, orthopedic, urology, general surgery, family medicine, pediatrics. ההמלצות גיבוש על משהו פה להוסיף צריך

Results: Key points that will be included in the position paper are: Use of time as a diagnostic tool and as a strategy to prevent overdiagnosis, engaging patients in shared decision making, asking themselves and encouraging patients to ask the" 4 questions" about offered test and treatments (natural history, benefits, harms, alternatives), next step consideration, keeping some healthy skepticism relating to benefits of medical interventions, looking for opportunities for Quaternary Prevention in each interaction with patients, offering deprescribing and "undiagnosing" when appropriate, coping with mistakes, CME, and role of the medical director.

Conclusions: The target of ISROD's first position paper was guideline's writing panels within the scientific societies. The target of the second position paper is actually each Israeli physician. We hope to provide individual doctors across medical specialties with tools to reduce overdiagnosis on a daily basis.

Screening and breast cancer mortality and in São Paulo State, Brazil: an ecological analysis

Simone Diniz, Francisco Chiaravalloti
School of Public Health - University of São Paulo, SÃO PAULO, Brazil

Objective: Identify the factors associated with the age-standardised breast cancer mortality rate in the municipalities of State of São Paulo (SSP), Brazil, from 2006 to 2012. In São Paulo, 79% of women 50-69 years old report a mammography in the last two years (based on the Ministry of Health recommendation); this rate is 90.5% for more educated women in Brazil; the recommendation of specialists' societies is annual mammography starting at age 40.

Methods: Ecological study of the breast cancer mortality rate standardised by age, as the dependent variable, having each of the 645 municipalities in the SSP as the unit of analysis. Setting: The female resident population aged 15 years or older, by age group and municipality, in 2009 (mid-term), obtained from public dataset (Informatics Department of the Unified Health System). Women 15 years or older who died of breast cancer in the SSP were selected for the calculation of the breast cancer mortality rate, according to the municipality and age group, from 2006 to 2012. Main outcome measures were mortality rates for each municipality, using the age structure of the population of SSP in 2009 as the standard.

Results: In the final linear regression model, breast cancer mortality, was directly associated with rates of mammography ($p<0.0001$), nulliparity ($p<0.0001$), and use of private healthcare ($p=0.006$), in the municipal level. The same paradoxical trend of increased mortality with increased screening, was found in recent individual-level studies. In the general female population in the SSP, compared with women with lower education and income, wealthier women generally used the private sector, were more exposed to carcinogens such as more radiation from mammography - higher exposure from more frequent (annual versus biannual), screening starting earlier (40 instead of 50), and more frequently using higher doses (more frequent use of digital equipment). Additional biopsies resulting from over-testing may contribute to further increase the risks. Wealthier women also have less children, and less breastfeeding opportunities, and are more exposed to oestrogens, in contraceptive and hormone replacement.

Conclusion: The association between mammography rate with increased breast cancer mortality adds to the evidence of a probable overestimation of benefits and underestimation of risks associated with this form of screening. Health services need to confront the problem of overdiagnosis and overtreatment, expanding health literacy and informed choice for patients, investing on primary prevention (safer options on contraceptive and menopause management, nutrition and physical activity resources, social support for breastfeeding etc). Additional studies and innovative methodologies should explore if there is a causality link between mammography rate, and increased all-causes and breast cancer mortality.

Reducing over investigation in suspected PE.

Iain Mactier, William Peagam, Gavin Leslie, Ailsa Doak, Paul Travnicek
Port Macquarie Base Hospital, Port Macquarie, Australia

Objectives: Computerised tomography pulmonary angiography (CTPA) studies are commonly requested in the diagnostic workup of suspected pulmonary emboli (PE). CTPA involve significant contrast and radiation exposure. There is also a risk of over or incidental diagnosis that may prompt further harmful investigations or treatment. Efforts should therefore be made to avoid use of CTPA when possible. Pretest probability scores - such as the Wells Score - and D-dimer can be used to reduce unnecessary CTPA requests. A low Wells Score and negative D-

dimer almost excludes PE - without additional diagnostic workup. A diagnosis of PE in approximately 15% of CTPA is suggested as an acceptable yield by the British Royal College of Radiologists. We conducted a quality improvement project in our emergency department (ED) at Port Macquarie Base Hospital, to encourage judicious use of CTPA and improve the PE diagnostic yield.

Method: We retrospectively collected data for patients that presented over a three month period. All patients that had a CTPA requested during their ED attendance were included. The patients were identified through a search of electronic records. After the initial audit posters were displayed and teaching sessions organised. It was envisaged this would improve use of the Wells Score and - ultimately - more judicious CTPA requesting. The audit cycle was repeated for the exact same time period one year later. Evidence of documented Wells Score use was sought in the medical records. We evaluated the formal radiologist report for each CTPA. All data were anonymised and input into a secure spreadsheet. As data were gathered retrospectively formal ethics approval was not deemed necessary.

Results: For the initial audit - 1st June to 31st August 2016 - 57 patients were identified. Only 7% - 4 - patient records had evidence of Wells Score use. The CTPA diagnostic yield was corresponding low with only 9% - 5 - studies identifying a PE. 87 patients were identified for the repeat audit. Wells Score use increased to 13% - 11 - and the CTPA PE yield to 16% - 14. No patient - in either audit - had a low Wells Score and PE on CTPA.

Conclusions: The interventions to encourage more judicious requesting of CTPA appear to have been successful. The absolute number of CTPA requested increased but so did the diagnostic yield for PE. This suggests a more appropriate use of CTPA, likely due to increase use of the Wells Score. In the repeat audit our ED exceeded British College of Radiologist standards on the expected PE yield from CTPA. The increase in CTPA requests may be explained by a larger overall number of ED attendances. We recognise some patients may have had a Wells Score applied but not documented. This is an inherent weakness of retrospective audit. The sample size was also relatively small. Nonetheless we recommend other ED evaluate their use of Wells Score in the context of suspected PE. It is a simple - and effective - way to minimise over investigation for this common presentation.

14:00 Wednesday August 22nd - Nielsine Nielsen Auditorium Turning Citizens In to Patients Unnecessarily / Psychiatry and Overdiagnosis

Under- versus overdiagnosis: Exploring the benefits and harms of a PCOS label and its impact on women's psychosocial wellbeing, lifestyle and behaviour

Tessa Copp[1], Jolyn Hersch[1], Kirsten McCaffery[1], Jenny Doust[2], Anuja Dokras[3], Ben Mol[4], Jesse Jansen[1]
[1]Wiser Healthcare, University of Sydney, Sydney, Australia. [2]Wiser Healthcare, Bond University, Gold Coast, Australia. [3]University of Pennsylvania, Philadelphia, USA. [4]Monash University, Melbourne, Australia

Background: Diagnostic criteria for polycystic ovary syndrome (PCOS) were expanded in 2003, resulting in the inclusion of women with milder forms and a substantial increase in the total number of women diagnosed. This has raised concerns about overdiagnosis and unnecessary disease labelling. PCOS is associated with adverse reproductive, metabolic and cardiovascular outcomes, however not all women fitting the expanded criteria may be at risk of these. Women with PCOS also have higher rates of psychological distress than their peers. However, it is unclear whether this is due to the condition, its symptoms, or the psychological impact of being labelled with PCOS. We aimed to explore the benefits and harms of a PCOS label from the perspective of women diagnosed with PCOS who vary across the spectrum of disease severity, and investigate how the diagnosis affected their psychosocial wellbeing, lifestyle choices and behaviour.

Method: Women diagnosed with PCOS aged 18-45 in Australia were recruited via Facebook. We conducted 25 semi-structured face-to-face and telephone interviews. Topics included experience with diagnosis and management, doctor-patient communication, information provision and diagnostic satisfaction, perceived benefits and harms of diagnosis, and impact of the diagnosis on psychological wellbeing, life decisions, behaviour and social environment. Interviews were audio-recorded, transcribed and analysed thematically.

Results (preliminary): Women expressed a range of experiences regarding the diagnosis and management of their condition. Perceived benefits included validation and explanation of bothersome symptoms, increased understanding about their body and why it behaves in certain ways, and better access to treatment. Perceived harms included increased worry and anxiety about the future, and misperceptions and confusion about fertility, which resulted in unplanned pregnancies among some women. Adverse impacts on self-esteem and relationships were also reported, as well as significant out of pocket costs for specialist appointments, screening tests and medication.

Discussion: Diagnostic criteria for PCOS have expanded without clear evidence of benefit. It is vital the benefits and harms of a diagnosis for women across the spectrum of symptom severity are investigated. Participants interviewed in this study described various positive and negative impacts of a diagnosis on their lives, which were influenced by factors such as symptom severity, communication at diagnosis and their relationships with their clinicians. The implications of these findings for future research and clinical practice will be discussed.

Sarcopenia: A case study in how the phenomenon of overdiagnosis is generated

Jacob Bülow[1], Lars Holm[2], John Brodersen[3]
[1]Institute of Sports Medicine Copenhagen (M81), Bispebjerg Hospital, Copenhagen, Denmark. [2]School of Sport, Exercise and Rehabilitation Sciences, University of Birmningham, Birmingham, United Kingdom. [3]Section of General Practice, University of Copenhagen, Copenhagen, Denmark

Objectives: In 1989 the term Sarcopenia was introduced to describe the phenomenon of age-related loss of muscle mass. From 2010 to 2014 six consensus definitions were presented based on the previous two decades of research within the field of sarcopenia. Despite the fact that loosing muscle mass originally was seen as a normal process of ageing, it was classified as a disease in 2016 and dedicated its own ICD-10 code.

The arguments for classifying age-related loss of muscle mass as a disease put forth by the consensus papers and leading scientist in the research field of sarcopenia are sparse and use tautological reasoning in its attempt to legitimate this transformation.
The purpose of the study is firstly to clarify whether or not Sarcopenia is another case of overdiagnosis due to overdefinition, and if so, secondly to give new insight in how the phenomena of overdiagnosis arise.

Method: The analysis in this study is based on an historical approach to the scientific literature on Sarcopenia. An understanding of the transformation of the phenomenon from normal to pathological will be approached by analysing the conceptualization of sarcopenia in cornerstones articles in the research field of sarcopenia chronological from the introduction of the concept through the different decades up until its admission in ICD-10. The current consensus definitions of Sarcopenia and the argumentation for its admission in ICD-10 will be compared to the general understanding of the concept of overdiagnosis, to clarify whether or not this new disease is a case of overdiagnosis.

Results: During the decade of 1990 the transformation of sarcopenia from a normal to a pathological condition started. The phenomenon was made guilty by association to both mortality and decreased physical function, with the problem that over 50% of the population over 80 years would be considered as sarcopenic. From 2000 a decreased physical function became a criterion, to narrow down the amount of patients categorized with sarcopenia, in the effort of legitimizing it as an independent pathological phenomenon. It culminated in admission of an ICD-10 code in 2016, which put Sarcopenia in line with other age-related diseases such as osteoporosis, diabetes and hypertension. Comparing the consensus definitions with the concept of overdiagnosis, Sarcopenia seems to be a case of overdefinition. The change of the conception of Sarcopenia from a normal to a pathological process is a lowering of the threshold of what we in general categorize as disease.

Conclusions: With the rewarding of an independent ICD-10 code and its acceptance as a disease in the geriatric research field, Sarcopenia has become another example of the increase in overdiagnosis in especially the modern western societies. Despite differences in historical development the definition of Sarcopenia shows great similarities with other phenomena such as osteoporosis and hypertension where overdiagnosis is present. The historical analysis of the transformation of Sarcopenia gives insights into how phenomena, otherwise considered as completely normal and compatible with life, can change its ontological status into a treatment demanding disease. Further investigating of its comparability to other age related phenomena could possibly elucidate further mechanism behind the appearance of overdiagnosis. This could contribute to the understanding of the phenomenon of overdiagnosis and how to prevent further expansion of this unfortunate development.

Widening Disease Definitions in Gestational Diabetes: an evaluation of changing guidelines

Christiana Naaktgeboren[1], Jenny Doust[2], Paul Glasziou[2], Julia Lowe[3], Mariska Leeflang[4]
[1]Julius Center, Utrecht, Netherlands. [2]Bond University, Brisbane, Australia. [3]University of Toronto, Toronto, Canada. [4]Amsterdam Medical Center, Amsterdam, Netherlands

Objectives: The incidence of gestational diabetes mellitus (GDM) is rapidly increasing worldwide, raising a concern of overdiagnosis. While population trends such as obesity, increased age at

motherhood, and ethnic changes play a role in this increase, another major cause is the widening of the diagnostic criteria. The primary aim of this study is to evaluate what factors were taken into consideration when new diagnostic criteria for GDM were made. The Guidelines International Network (G-I-N) Preventing Overdiagnosis workgroup recently developed guidance for modifying the definition of disease, including a checklist for items to consider when widening disease definitions (1). The secondary aim of this study was to pilot the use of this G-I-N checklist.
(1) Doust J, Vandvik PO, Qaseem A, et al. Guidance for Modifying the Definition of DiseasesA Checklist. JAMA Intern Med. 2017;177(7):1020–1025. doi:10.1001/jamainternmed.2017.1302

Method: Documents in which currently used criteria for gestational diabetes were proposed were the focus of this study. Changes to thresholds, timing of testing, and the combination of abnormal test results required were considered changes to definitions, but changes to screening strategies were considered outside the scope. These definition documents were found by backward reference searching from 5 recent reviews of gestational diabetes and through searching of websites of professional and guideline organizations. Documents containing new definitions were assessed against the 8-item G-I-N checklist to evaluate what domains were considered when proposing a new disease definition.

Results: We identified 14 documents which proposed modifying the diagnosis of GDM. Four types of definitions were observed: a percentile definition similar to laboratory reference ranges (n=4); harmonization with type two diabetes mellitus (n=6); a risk-based assessment examining the risk of maternal and fetal adverse outcomes (n=3); and one informed by a health technology assessment (n=1). None of the 14 documents considered all 8 criteria in the G-I-N checklist. All described the new definition in detail and all but one described the trigger. None estimated the impact on the prevalence of GDM. The prognostic ability of the definitions was only assessed by risk-based criteria (n=3) and little attention was given to accuracy, repeatability, or reproducibility(n=2). Potential benefits were mentioned by half (n=7) and harms by fewer (n=4). The balance between harms and benefits was only discussed by 3.

Conclusions: Our analysis of the changes to criteria for GDM reveals a complex history of definitions stemming from 4 conceptual bases. There appears to be a paucity of primary research data used in the development of definitions for GDM. While harms and benefits of changing the definition were sometimes mentioned, there was no explicit consideration or quantification of the benefits versus harms, making thresholds chosen appear arbitrary. Given the impact of seemingly modest changes to disease definitions have on the incidence of disease, we consider definitional changes to be a substantial task for guidelines, one that requires a separate panel. Such panels should use G-I-N's published 8-item checklist of elements to consider when modifying definitions. For key items, rapid systematic reviews should be considered. Finally, panels could be more cautious in applying dichotomous disease labels and instead use a stratified terminology that reflects a spectrum of risk.

The evaluation of a breast cancer screening decision aid in the community setting.

Kimberly Cuetara Gibbons, Gene Elizabeth Harkless
University of New Hampshire, Durham, USA

Abstract in Full: For over 40 years, breast cancer screening (BCS) has been recommended to women by healthcare providers as well as professional organizations and non-profit awareness

advocates. Recently, attention is being paid to quantification of the actual benefits and harms of BCS with mammography. The harms include overdiagnosis, overtreatment and mandate of mammography screening that may not be in alignment with the individual values and preferences of women ages 40-69. This evidence translation project developed a decision aid , "My Personal Decision" (MPD). The project piloted the decision aid (DA), directly to a community-based convenience sample of women ages 40-69 (N=66). This resource for practice was then evaluated by community participants for alignment with personal value and preference-based educational needs about BCS.

Objectives; The purpose of this quality improvement project was to pilot a clinical practice decision aid (DA) to provided evidence-based information about the harms and benefits of BCS directly to a community-based, convenience sample of women at average risk for breast cancer between the ages of 40 and 69. The specific aim of this project was to enhance preparedness for decision-making through the provision of a DA that is perceived as useful to the user. The online DA was evaluated by community participants for alignment with personal value and preference-based educational needs, usefulness and influence on personal beliefs about BCS.

Method: An online, informational DA, "My Personal Decision" (MPD), was constructed, with permission, from the Gotzsche, Hartling, Nielsen, & Brodersen (2012) leaflet, based on the findings of the systematic review by Gotzsche and Jorgensen (2013). The MPD was recorded using Powerpoint and was placed within a Qualtrics survey. This was piloted with a community-based convenience sample of 66 women in the spring of 2018. Information included in the MPD reviewed the definition of screening, benefit of lives saved from mammography, the harms of mammography screening including false alarms, more extensive surgery and after treatment, overdiagnosis/ overtreatment, pain at the examination, and false reassurance. The numeric risk data in the DA was presented using icon arrays in a theater format.

Results: Beliefs about harms vs. benefits of mammography screening, risk of over diagnosis, risk of overtreatment, and the usefulness of the DA in preparing for informed decision-making, along with structured and unstructured feedback questions about the DA were obtained. Participants (n=42) reported a change in response to the statement "There are more benefits than harms related to breast cancer screening (mammography). " Prior to the DA, 36 women responded with strongly agree/ agree that benefits of BCS outweighed harm. After completing the DA, 19 women continued to agree with this statement (P=0.0258, Fisher's exact test). Most participants (81%) reported they agree/ strongly agree with "This presentation helped me to make an informed decision about breast cancer screening." As well, 83% agreed/ strongly agreed that "After watching this presentation, feel better prepared to make a decision about breast cancer screening." Further analysis will be provided at the presentation.

Conclusions: Change in beliefs after evidence-based information about the harms and benefits of BCS was demonstrated in this small, community based sample. Evaluation feedback from the participants before and after use of the tool confirmed that MPD was helpful and enhanced preparedness for breast cancer screening decision-making. Overall, this project demonstrates that providing evidence-based information directly to women outside of a clinical consultation may be acceptable, informative, and perceived as valuable in the BCS decision-making process. These findings are important as understanding of evidence-based information is a fundamental element in the empowerment of the decision maker with the goal of informed decision making.

Defining Overdiagnosis of Mental Health Disorders: Secondary Analysis of an Overdiagnosis Scoping Review

Kimberly Turner[1], Ian Shrier[1,2], Brett Thombs[1,2]
[1]McGill University, Montreal, Canada. [2]Lady Davis Institute, Montreal, Canada

Objectives: The term "overdiagnosis" was first used as early as 1924, when J. D. Adamson used it to describe negative implications of attempts to achieve early diagnosis of pulmonary tuberculosis (CMAJ, 1924). Since then, the term has been used most frequently to describe negative outcomes from screening for early-stage asymptomatic cancers. Typical definitions describe overdiagnosis as detection of early-stage asymptomatic conditions that would never have led to morbidity or mortality. In other areas, including mental health, overdiagnosis can occur among people who experience symptoms but whose symptoms do not reflect disorders and may not be amenable to healthcare interventions. Thus, more recent alternative definitions have focused on diagnosis among people who would not be expected to experience net benefit. The degree to which such definitions have been adopted is not clear. Our objective was to describe how the term overdiagnosis has been defined explicitly or operationalized implicitly in mental health.

Method: A scoping review of overdiagnosis across medical disciplines searched PubMed in August 2017 for published articles that used keywords related to overdiagnosis. Articles from the scoping review were eligible for the present analysis if they were classified in the scoping review as related to mental health, excluding neurocognitive disorders, and if they used the term overdiagnosis in the text of the article and not just in the title. We extracted basic information about the article and whether it included an explicit or implicit definition of overdiagnosis. Explicit definitions were extracted. If the definition was implicit, the reviewer provided an explanation of how overdiagnosis was operationalized in the study or article. Data were extracted by one reviewer with validation by a second reviewer, and any disagreements resolved by consensus. Explicit and implicit definitions were grouped into categories by one investigator and verified by a second investigator.

Results: 148 articles were included. Of the 14 articles that explicitly defined overdiagnosis, 9 defined it as a false positive diagnosis, 2 as misdiagnosis (diagnosing people with one disorder rather than another), 1 as diagnosis of an individual who would not be expected to benefit from treatment, and 2 had vague descriptions. In the other 134 articles, implicit definitions fit into 4 categories; 68 articles implicitly defined overdiagnosis as diagnosis of people who do not meet diagnostic criteria, 59 as misdiagnosis, 13 as diagnosis resulting from overly broad or changed diagnostic criteria; and 2 as no net benefit from diagnosis. There were 13 with unclear or difficult to classify definitions. There was overlap of definitions with several articles fitting into more than one category. The most significant overlap involved 13 articles that were classified as both misdiagnosis and diagnosis of people who do not meet diagnostic criteria.

Conclusions: Definitions of overdiagnosis commonly used in the context of screening for asymptomatic early-stage disease are not generally applicable in mental health where diagnoses are not made in the absence of symptoms. There is not, however, an agreed upon definition of overdiagnosis in mental health. Results from the present review indicate that the term is used most commonly in the field to describe potential drivers of overdiagnosis, including diagnosing individuals who do not meet diagnostic criteria and overly broad diagnostic criteria, as well as misdiagnosis, which may not always reflect overdiagnosis. Some articles define overdiagnosis in mental health as occurring when there is no net benefit from diagnosis; that is, when individuals are diagnosed, but expected benefits from the diagnosis would not be

expected to exceed harms. Agreement on an approach to defining overdiagnosis in mental health is needed so that evidence of overdiagnosis can be more readily evaluated.

14:00 Wednesday August 22nd - Henrik Dam Auditorium - Other

Overdiagnosis, overtreatment and low-value care in physiotherapy: a scoping review

Joshua Zadro[1,2], Christopher Maher[1,2], Mary O'Keeffe[1,2]
[1]School of Public Health, Sydney Medical School, The University of Sydney, Sydney, Australia.
[2]Institute for Musculoskeletal Health, Sydney Local Health District, Sydney, Australia

Objectives: The aim of this scoping review was to define concepts and map the available research in the area of overdiagnosis, overtreatment and low-value care relevant to physiotherapy by answering the following questions:
i) When did the concept of 'low-value' physiotherapy start to be discussed in the literature?
ii) What terms are being used to describe overdiagnosis, overtreatment, and low-value care relevant to physiotherapy?
iii) What is the prevalence of physiotherapy care that is low-value, high-value, or unknown value, and does this differ across areas of physiotherapy practice?
iv) What proportion of physiotherapy care is delivered in accordance with evidence-based clinical practice guidelines?
v) Which strategies aimed at improving the quality of physiotherapy care are effective?

Method: A keyword search will be run in MEDLINE, EMBASE, CINAHL, Cochrane CENTRAL, AMED Scopus and Web of Science (earliest record to April 2018) combining terms synonymous with "low-value care" and "physiotherapy". Additional articles will be identified by hand-searching references lists and forward searching of included articles.
Articles discusses low-value physiotherapy (e.g. editorials, Choosing Wisely reports), investigating physiotherapy treatment practices and adherence to guidelines, and evaluating strategies to improve the quality of physiotherapy will be included. There will be no restriction on the physiotherapists clinical specialty (musculoskeletal, cardiopulmonary and neurological), work setting (e.g. private, public, community), experience, ethnicity, age or gender.
Two reviewers will independently perform the selection of studies and extracted key data (e.g. study characteristics, outcome data relevant to research questions). Disagreement will be resolved through discussion. Analyses will be qualitative for questions i-ii), and quantitative for questions iii-v) where data is considered sufficiently homogenous for meta-analysis.

Results: This study will be reported according to the "Preferred reporting items for systematic reviews and meta-analyses" (PRISMA) statement. At the time of submission we were extracting data from the included studies. Hence, I am confident to present the full findings of this review at the conference.

Conclusions: Overdiagnosis, overtreatment and low-value care – care that provides little-to-no benefit or causes harm, and that is an unwise use of scarce health resources – are concepts receiving substantial attention in medicine. However, it is unclear how much attention these concepts are receiving in physiotherapy. Understanding whether these concepts are receiving attention in physiotherapy and evaluating the prevalence of low-value physiotherapy is

important because the profession is rapidly expanding. Identify the areas of physiotherapy practice where the quality of care must improve and identifying effective strategies for improving guideline adherence in physiotherapy could have enormous implication for improving the health of millions of people worldwide who seek physiotherapy each year.

A Tale of Two Studies: Diagnostic algorithms and clinical practice guidelines minimize overdiagnosis and overtreatment and maximize survival in lung cancer screening.

Frederic Grannis
City of Hope Natinal Medical Center, Duarte CA, USA

Abstract in Full: At the beginning of the new millennium two major lung cancer screening (LCS) studies used different research approaches to study computerized tomographic (CT) LCS. The National Lung Cancer Screen study (NLST) used a randomized control design to compare with chest roentgenogram (CR). The study did not utilize a diagnostic algorithm. The International Early Lung Cancer Action Program (IELCAP) was a prospective, multi-institutional, single-arm registry study incorporating diagnostic algorithms for baseline and annual-repeat screens http://www.ielcap.org/protocols .
Although the NLST demonstrated substantial decrease in lung cancer (>20%) and all-cause (>7%) mortality, study subjects had lower 5-year survival (60%) and experienced higher levels of false positive test results and invasive diagnostic tests and surgical operations in patients with benign nodules than IELCAP.
The IELCAP study used a form of "learning healthcare system", analyzing accumulating registry data on a semi-annual basis with step-wise modification of the algorithms based upon analysis of results and consensus of assembled principal investigators. The updated IELCAP algorithms have been incorporated into the clinical practice guideline of the National Comprehensive Cancer Network and other professional organizations. Although the IELCAP design does not allow precise estimation of mortality reduction, it has established that application of the algorithms results in substantially higher 10-year actuarial lung cancer specific survival (>80%), lower (13%) rate of false positives and fewer invasive tests and operations on benign nodules. Modifications in the algorithm to date primarily relate to raising the threshold for a positive test result to solid nodules with an average diameter of 6 mm. or greater and changes in the management of non-solid nodules, regardless of size. The algorithms reflect current uncertainty as to how often and how rapidly in-situ adenocarcinomas will progress into invasive adenocarcinomas. Extant data suggest that careful observation using annual CT scans allows identification of progression to invasive adenocarcinoma with demonstration of transition from non-solid to part solid nodules containing a growing solid component, without allowing neoplasms to progress in stage.
Based upon retrospective analysis of IELCAP data suggesting that minimally invasive, sub-lobar resectons provide equivalant survival, with lower morbidity, an offshoot, prospective registry study,
'Initiative or Early Lung Cancer Research on Treatment
' (IELCART), is currently accruing participants to determine whether small screen-detected LC can be treated safely and effectively using minimally invasive, sub-lobar pulmonary resection or radiation therapy modalities.

Objectives: The objective of this presentation is to review current efforts to reduce potential

overdiagnosis and (more important) to avoid overtreatment of screen-detected lung cancer, and their effectiveness and safety.

Method: Review and analysis of NLST, IELCAP and NCCN diagnostic algorithms, clinical practice gudelines and published results on lung cancer screening.

Results: Application of diagnostic algorithms, updated after analysis of accumulating research data, yield improvements in survival and reduction in false-positive test results, further non-invasive and invasive testing and overtreatment in the form of resection of benign pulmonary nodules or potentially-overdiagnosed in-situ adenocarcinomas.

Conclusions: Application of 'learning healthcare system" principles to prospective registry research models in lung cancer screening offer potential to improve the safety and effectiveness of lung cancer screening while simultaneously avoiding over-treatment of potentially-overdiagnosed lung cancer. A substantial number of lung cancer deaths can be prevented by lung cancer screening.

Preferences for papillary thyroid cancer management and the impact of terminology: a discrete choice experiment

Brooke Nickel[1,2], Kirsten Howard[1], Juan P Brito[3], Alexandra Barratt[1], Ray Moynihan[1,4], Kirsten McCaffery[1,2]
[1]Wiser Healthcare, University of Sydney, Australia. [2]Sydney Health Literacy Lab, University of Sydney, Australia. [3]Knowledge and Evaluation Research Unit, Mayo Clinic, Rochester, MN, USA. [4]Centre for Research in Evidence-Based Practice, Bond University, Gold Coast, Australia

Objectives: Given recent evidence of overdiagnosis and overtreatment of small papillary thyroid cancers (PTCs) and other low risk cancers, strategies are needed to help patients consider less invasive management options. This study aims to: (1) determine which management-related factors influence preferences for management options for PTC; (2) determine the trade-offs in treatment factors people are willing to accept; and (3) understand how terminology (cancer terminology versus non-cancer terminology) influences preferences and benefit harm trade-offs.

Method: PTC management preferences were evaluated using a discrete choice experiment (DCE). Participants were presented with a series of questions, asking them to choose between three management options for PTC (thyroidectomy, hemi-thyroidectomy and active surveillance) described by varying combinations of attribute levels. Participants were randomised to receive one of two frames of information based on the terminology used to describe the condition: 'papillary thyroid cancer' or 'papillary thyroid lesion'. The outcome variable was management preference; attributes and sociodemographic characteristics were explanatory variables. Analyses were conducted using a mixed logit (ML) model in NLOGIT 5.0.

Results: The DCE was completed by 2054 men and women aged 18 years and above with no history of thyroid cancer. Participants preferred options with less frequent follow-up, lower out of pocket costs, lower chances of having voice and calcium level problems, a lower risk of developing invasive thyroid and of dying of thyroid cancer. The chance of experiencing fatigue did not influence preferences. Benefit harm trade-offs found respondents were willing to accept a higher likelihood of adverse effects to avoid invasive thyroid cancer diagnosis or thyroid cancer death when the condition was described as a cancer compared to a lesion. For both the 'cancer' and 'lesion' terminology, the health literacy consistently impacted preferences for

management options. Those with lower health literacy were significantly more likely to prefer more invasive management options (full surgery compared to partial surgery and monitoring).

Conclusions: This study makes an important contribution to our understanding of how attributes of management options, terminology and patient characteristics, in particular health literacy, influence management decision making for PTC. As guidelines supporting less aggressive management continue to emerge as a result of increasing evidence of the indolent nature of PTC and other low risk cancers, strategies to deal with potential overtreatment are critically needed. Furthermore, clinician understanding of what drives patient management preferences and decision making is fundamental in order to help patients make decisions that align with their preferences.

Can We Detect Overdiagnosis Early? *Exploring indicators of possible overdiagnosis outside cancer screening contexts.*

Katy Bell[1], Jenny Doust[2], Leon Smith[1], Ian Harris[3], Rachelle Buchbinder[4], Louise Cullen[5], Alexandra Barratt[1]

[1]University of Sydney, Sydney, Australia. [2]Bond University, Gold Coast, Australia. [3]University of New South Wales, Sydney, Australia. [4]Monash University, Melbourne, Australia. [5]University of Queensland, Brisbane, Australia

Objectives: Although the risk of overdiagnosis in cancer screening was identified almost 50 years ago, evidence of cancer overdiagnosis was slow to emerge, and has required development of new approaches and study designs. As a result, responses to mitigate overdiagnosis, and to inform the public and professionals, have been delayed. This meant cancer screening activities became well established in clinical practice and policy before the problem overdiagnosis was recognised, making reversal difficult.
It would be beneficial to avoid repeating the same delays outside the cancer setting - whether in the context of screening programs (eg gestational diabetes screening among pregnant women) or in managing already symptomatic patients (eg children with ADHD symptoms, adults with joint pain) - so that timely policy responses to overdiagnosis in these settings can be developed. We therefore aimed to identify early "indicators" of potential overdiagnosis.

Method: We used concepts described in key papers about methods for detecting and quantifying cancer overdiagnosis, the clinical utility of diagnostic tests, and for modifying the definition of diseases, to identify early indicators of potential overdiagnosis. We identified four indicators: 1. Expanded detection of condition/disease; 2. Subclinical forms of the condition/disease; 3. Increasing diagnosis and treatment of condition/disease; and 4. Balance of benefit vs harm of diagnosis/treatment likely to be unfavourable based on existing evidence. We provide an overview of the proposed approach, and illustrate it using two examples that are common in everyday, clinical investigation of non cancer conditions: MRI in the investigation of adult knee pain, and high sensitivity troponin testing in emergency department workup of patients with suspected myocardial infarction. We evaluate and discuss the application of the indicators to these examples, using analysis of routinely collected, local data, and current research evidence.

Results: Example One: Is there potential overdiagnosis of knee pathology in older Australians? Knee MRI is positive on indicators 1, 2 and 4 for possible overdiagnosis (expanded detection of condition (disease); subclinical forms of the condition (disease); balance of benefit vs harm of diagnosis/treatment likely to be unfavourable). Indicator 3 (increasing diagnosis and treatment

of the condition) is inconclusive. Example Two: Is there potential overdiagnosis of Myocardial Infarction in Australians? Non ST Elevation Myocardial Infarction (NSTEMI) meets indicators 1 and 3 (expanded detection of condition (disease); increasing diagnosis and treatment of condition (disease)). Indicators 2 (subclinical forms of the condition (disease)) and 4 (benefit vs harm of diagnosis/treatment likely to be unfavourable) are inconclusive.

Conclusions: The proposed indicators may have utility in helping to identify possible overdiagnosis sooner than would normally come about through standard processes of surveillance of changes in clinical testing and its impact. We found that in our two examples, more indicators were clearly positive for knee MRI than for hs Tn/NSTEMI, suggesting more evidence of potential overdiagnosis for the former than for the latter. The approach is necessarily different from methods that have been developed for cancer screening because, in non-cancer conditions, patients may be asymptomatic (as in the context of early detection of disease) OR symptomatic yet still at risk of overdiagnosis. These four indicators may help identify tests that warrant the investment of research resources to assess the impact of the new testing patterns and disease definitions on long term, patient-relevant outcomes to avoid harm through overdiagnosis and overtreatment.

Evaluation of a Choosing Wisely Canada initiative to reduce unnecessary radiology and laboratory testing in the emergency department

Janessa Griffith[1,2], Noah Ivers[2], Sacha Bhatia[2], Wendy Levinson[1]
[1]University of Toronto, Toronto, Canada. [2]Women's College Hospital Institute for Health System Solutions and Virtual Care, Toronto, Canada

Introduction: With an aging population placing pressure on Canada's healthcare system, reducing unnecessary ordering of tests is one way of improving safety without compromising quality of care. Unnecessary care is not only wasteful, it can have implications for patient safety, such as exposure to radiation from diagnostic imaging (DI) tests: 2-3 computed tomography (CT) scans involve amounts of radiation that have been associated with increased cancer risk. Choosing Wisely Canada (CWC), a physician driven campaign aimed to reduce unnecessary care, could help to promote more appropriate ordering practices. Through CWC, Canadian specialty societies have produced numerous recommendations to avoid common tests, treatments, and procedures that may be unnecessary and harmful.

Objectives: The recommendations have been implemented in many Canadian practice settings, and one large community hospital in Ontario, Canada has made it a strategic priority. Specifically, the hospital employed various interventions such as changing medical directives and electronic order sets, engaging leadership, and holding educational meetings. Through a pre-post study, this hospital found that laboratory testing in the ED was reduced by 31%, representing approximately £88,877 over one year. Although these preliminary results indicated decreases in laboratory test ordering, our research team is currently conducting a more rigorous study, as pre-post studies are problematic for causal attribution.
Therefore, the purpose of this research is to evaluate the implementation of DI- and laboratory-related CWC recommendations into the emergency department at a large community hospital, including ehealth interventions.

Methods: We will use an explanatory sequential mixed-methods design to explain and explore the impact of CWC upon this hospital's emergency department testing practices. We will use

interrupted time series analyses to assess changes in monthly testing rates one year before and three years after the implementation of interventions. We will also conduct key informant interviews with clinicians and hospital leadership to gain insight into how and why the effects were achieved in this context.

Results: Findings from these methods will be triangulated and used to develop an implementation framework that can be generalized for use in other settings.

Conclusions: In order to promote patient safety and the responsible use of health care resources, it is critical to determine ways to reduce unnecessary tests. Through this comprehensive program evaluation, we plan to gain insight into what interventions are most impactful.

Natural History of Ductal Carcinoma in Situ in the Absence of Locoregional Treatment

Marc D Ryser[1,2], Donald L Weaver[3], Fengmin Zhao[4,5], Mathias Worni[6], Lars Grimm[7], Ruth Etzioni[8], Terry Hyslop[9], Sandra J Lee[4,5], E Shelley Hwang[10]
[1]Department of Surgery, Division of Advanced Oncologic and GI Surgery, Duke University Medical Center, Durham, USA. [2]Department of Mathematics, Duke University, Durham, NC, USA. [3]Department of Pathology and Laboratory Medicine, University of Vermont and UVM Cancer Center, Burlington, VT, USA. [4]Department of Biostatistics and Computational Biology, Dana-Farber Cancer Institute, Harvard Medical School, Boston, MA, USA. [5]Department of Biostatistics, Harvard T.H. Chan School of Public Health, Boston, MA, USA. [6]Department of Visceral Surgery and Medicine, Inselspital, Bern University Hospital, University of Bern, Bern, Switzerland. [7]Department of Radiology, Duke University Medical Center, Durham, NC, USA. [8]Program in Biostatistics, Fred Hutchinson Cancer Research Center, Seattle, WA, USA. [9]Department of Biostatistics & Bioinformatics, Duke University Medical Center, Durham, NC, USA. [10]Department of Surgery, Division of Advanced Oncologic and GI Surgery, Duke University Medical Center, Durham, NC, USA

Objectives: Guideline-concordant care for women diagnosed with ductal carcinoma in situ (DCIS) consists of surgery in the form of a lumpectomy (with or without radiation therapy), mastectomy, or bilateral mastectomy. There is growing concern that a significant fraction of these women may be overtreated for indolent or slowly growing disease that would, in the absence of treatment, not develop into symptomatic or clinically significant breast cancer during their remaining lifetime. However, because the vast majority of women diagnosed with DCIS undergo treatment, the risks of invasive progression and competing death in the absence of definitive surgery and radiation therapy remain uncertain.

Method: Survival analysis of patient-level data from DCIS patients who did not receive definitive surgery or radiation therapy as recorded in the US National Cancer Institute's Surveillance, Epidemiology, and End Results (SEER) program (1992-2014). Kaplan-Meier curves were used to estimate the net risk of subsequent ipsilateral invasive breast cancer by tumor grade. Sensitivity analyses using information about cause of death were performed to account for potential ascertainment biases in the registry. The cumulative incidences of ipsilateral invasive breast cancer, contralateral breast cancer, and death were estimated using competing risk methods.

Results: A total 1,286 DCIS patients who did not undergo locoregional therapy were identified. Median age at diagnosis was 60 years (inter-quartile range [IQR]: 51-74), with median follow-up of 5.5 years (IQR: 2.3-10.6). Among patients with tumor grade I/II (n=547), the 10-year net risk of ipsilateral invasive breast cancer was 12.2% (95%CI: 8.6%-17.1%), compared to 17.6% (95% CI: 12.1%-25.2%) among patients with grade III (n=244). In sensitivity analyses, the estimates increased to 14.9% (95% CI: 10.4%-19.2%) and 27.9% (95% CI: 20.3%-34.7%), respectively. Among all patients who did not receive locoregional treatment, the 10-year cumulative incidences of the competing risks of ipsilateral invasive breast cancer, contralateral breast cancer and all-cause mortality were 10.5%, 3.9%, and 24.1%, respectively.

Conclusions: Despite limited available data, our findings provide an approximation of the risk of DCIS progression and competing risks in absence of definitive surgery and radiation therapy. Results indicate a limited propensity for invasive progression in patients with non-high-grade DCIS, and thus a substantial potential for overtreatment in these patients.

14:00 Wednesday August 22nd - Holst Auditorium - The Role of Risk Factors in Overdiagnosis

Statins for the primary prevention of cardiovascular disease: a simulation of eligibility, costs, patient preferences and number-needed-to-treat in the context of changing clinical guidelines

Paula Byrne[1], John Cullinan[1], Paddy Gillespie[1], Rafael Perera[2], Susan M Smith[3]
[1]National University of Ireland, Galway, Ireland. [2]Nuffield Department of Primary Care Health Sciences, University of Oxford, Oxford, United Kingdom. [3]Royal College of Surgeons in Ireland, Dublin, Ireland

Abstract in full
***Background*:** Changes in clinical guidelines for primary prevention of cardiovascular disease (CVD) have widened eligibility for statin therapy.
***Methods*:** We analysed a cohort of people aged over 50 without established CVD from The Irish Longitudinal Study on Ageing, determined their eligibility for statin therapy based on seven consecutive European clinical guidelines and modelled the associated potential cost. We calculated the number-needed-to-treat (NNT) to prevent one major vascular event in patients at the lowest baseline risk for which each guideline recommends treatment, as well as for those at *low*, *medium*, *high* and *very high* risk according to the most recent guideline. We compared these to the published NNTs that patients report is required to justify taking a daily medicine.
***Results*:** The proportion of statin-eligible patients increased from 7.95% in 1987 to 61.27% in 2016 with associated costs rising from €13.9m to €107.1m per annum in real terms. The NNT for those at the lowest risk for which each guideline recommends treatment rose from 40 to 400, with the proportion of statin-eligible patients achieving NNT levels needed to justify taking a daily medicine falling accordingly. At least 400 *low-risk* patients would have to be treated to prevent one major cardiovascular event compared to 25 or fewer *very high-risk* patients, according to 2016 guidelines.
Conclusion: Increased eligibility for statin therapy has implications for the medicalisation of large proportions of our population and healthcare budgets. Decisions to take and reimburse statins should be considered on the basis of expected cost effectiveness and acceptability to patients.

Objectives - The aim of this paper was four-fold:
Firstly, to estimate the increasing proportions of people considered eligible for statin treatment according to each of seven European Society of Cardiology/European Atherosclerosis Society (ESC/EAS) guidelines from 1987 to 2016.
Secondly, to estimate potential cost increases associated with each consecutive guideline recommendation.
Thirdly, to calculate the NNT to prevent one major vascular event in patients at the lowest baseline risk for which each guideline recommends treatment, as well as for those at *low*, *medium*, *high* and *very high* risk according to the most recent guideline.
Finally, to compare these NNTs with those reported by patients as being the minimum benefit they would need to justify taking a daily medicine, by calculating the proportion of those eligible for statins who would be on or below an 'acceptable' NNT as reported in a recent systematic review of patient preferences for taking cardiovascular medicines.

Method

1.Times trends in eligibility for statins - We used treatment thresholds recommended by seven consecutive guidelines to calculate the proportion of a sample of participants aged over-50 years from The Irish Longitudinal Study on Ageing eligible for statin therapy according to each guideline.
2. Cost increases due to widening eligibility - We estimated a weighted average annual cost per patient taking statins and the total annual cost.
3.Changes in Numbers-Needed-to-Treat due to treatment threshold changes over time -We estimated the NNT to prevent one major vascular event for those at the lowest risk for which each guideline recommends treatment, as well as for those considered *low, medium, high and very high risk according to the most recent ESC/EAS guidelines.*
4.Patient perspective - We considered these NNTs in the context of the minimum acceptable risk reduction that patients say is necessary to justify a daily intake of medication.

Results: The proportion of statin-eligible patients increased from 7.95% in 1987 to 61.27% in 2016 with associated potential costs rising from €13.9m to €107.1m per annum. The NNT to prevent one major vascular event in those at the lowest levels of risk for which statins could be recommended was 40 according to the 1994 and 1998 guidelines; 73 according to the 2004 and 2007 guidelines; and 400 according to the 2012 and 2016 guidelines. According to the 2016 ESC/EAS guidelines, 400 low-risk people would have to be treated to prevent one major vascular event, between 53 and 400 moderate-risk people, between 25 and 53 high-risk people and 25 or fewer very high-risk people. Only some of those classified as high and very high-risk, according to the 2016 ESC/EAS guidelines, would reach an acceptable NNT of less than 30.

Conclusions: Changes in recommendations for the use of statins would result in almost two thirds of over-50s in Ireland and similar countries being considered eligible for statin therapy. This has implications for the medicalisation of large proportions of our population, as well as for already resource constrained healthcare budgets. The value for money of the widening use of statins should be considered from both a societal and individual perspective. The decision to take and reimburse statins could be informed by NNTs, which are large in some risk categories. As we have seen from our analysis, a proportion of our sample would require significantly greater reductions in absolute risk to justify taking a daily medication. The patient's decision to take statins should be considered in the context of shared decision-making and the relevant NNT so that informed choices can be made relevant to their individual baseline risk.

Physicians' perceptions of inappropriate laboratory testing in clinically admitted patients

Bram Vrijsen, Maarten ten Berg, Christiana Naaktgeboren, Laura Vos, Wouter van Solinge, Karin Kaasjager
UMC Utrecht, Utrecht, Netherlands

Background: Overtesting in laboratory medicine is common. We set up this study to ascertain the appropriateness of laboratory test ordering in admitted patients and to investigate staff's perceptions of, attitudes towards and reasons for the inappropriateness.

Methods: Information was collected on all laboratory tests ordered for patients admitted to the internal medicine ward of our hospital, a large university medical center, from January 2011 to December 2016. Several prespecified measures of inappropriate testing were assessed using these data including the number of fixed combinations of tests, the rate of repetition of laboratory test results within normal range, the number of laboratory test orders per week, and the time to repetition of CRP measurement and its clinical consequences.
A survey was performed to investigate perceptions of and attitudes towards the appropriateness. The results of this survey were used as starting point in a series of focus group discussions.

Results: In 97% of orders that included an ALT, AST was also measured. For sodium and potassium, thus occurred in 95% of orders. Laboratory test results within the normal range are repeated in up to 85% of cases. Admitted patients on average have 5.7 laboratory test orders per week. In 59% of patients in whom a CRP was measured in the Emergency Department directly prior to admittance, CRP was repeated within 24 hours. In 88% of cases, this did not lead to changes in patient management.
In the survey, respondents' estimates of the number of laboratory test orders per week are substantially lower than the actual number (3.9 vs. 5.7). Nevertheless, the majority of respondents think the number of laboratory test orders is too high. Most point to the residents as being most responsible for ordering laboratory tests.
In focus groups the causes of excessive laboratory test ordering were explored. Most prominently, the residents felt they did not receive enough feedback. At the same time they considered laboratory test ordering to be too trivial to ask for more feedback from their supervisors.

Conclusion: Inappropriate laboratory test ordering is common and underestimated.. Residents working on the wards are considered to be most responsible for inappropriate laboratory test orders. They feel they need more feedback to change their behaviour, yet think the subject is too trivial to ask for feedback themselves. This implies that interventions to reduce inappropriate laboratory test ordering should focus on facilitating feedback.

Risk factors associated with potentially serious incidental findings and with serious final diagnoses on multimodal imaging in the UK Biobank Imaging Study

Lorna Gibson[1], John Nolan[1], Thomas Littlejohns[2], Edouard Mathieu[2], Steve Garratt[3], Nicola Doherty[3], Naomi Allen[2], Joanna Wardlaw[1], Caroline Jackson[1], Cathie Sudlow[1]

[1]University of Edinburgh, Edinburgh, United Kingdom. [2]University of Oxford, Oxford, United Kingdom. [3]UK Biobank Co-ordinating Centre, Stockport, United Kingdom

Objectives: Some form of clinical assessment, such as imaging, blood tests, primary or secondary care appointments and invasive procedures, is usually generated when healthy research participants receive feedback of potentially serious incidental findings (PSIFs, defined here as findings that indicate the possibility of conditions which, if confirmed, would carry a real prospect of seriously threatening life span, or of having a substantial impact on major body functions or quality of life). However, only around 1/5 PSIFs turn out to represent serious disease. Understanding the factors associated with increased risks of PSIFs and of serious final diagnoses may influence individuals' decisions to participate in imaging research, and will inform researchers' designs of appropriate PSIFs policies for future research studies in order to minimise the risks of overdiagnosis in healthy research volunteers.

Methods: We included all UK Biobank participants who undertook an imaging assessment up to December 2015 (n=7334, median age 63 [interquartile range 56-68], 3,804 [51.9%] women). Brain, cardiac and body magnetic resonance, and dual energy x-ray absorptiometry images from the first 1000 participants were reviewed systematically by radiologists for PSIFs. Subsequently, radiographers flagged potentially concerning images for review by radiologists. We classified final diagnoses as serious or not based on survey data from participants and their GPs, and clinical correspondence collected up to six months following feedback of a PSIF. We used binomial logistic regression models to investigate associations between PSIFs protocol, age, sex, ethnicity, socio-economic deprivation, private healthcare use, alcohol intake, diet, physical activity, smoking, body mass index and morbidity with both PSIFs and serious final diagnoses.

Results: Systematic radiologist review generated 13 times more PSIFs than radiographer flagging (179/1000 [17.9%] versus 104/6334 [1.6%], OR 13.3, 95% CI 10.3-17.1, p<0.001) and proportionally fewer serious final diagnoses (21/179 [11.7%] versus 33/104 [31.7%]). Older age was associated with increased odds of PSIFs and of serious final diagnoses under both protocols (sex-adjusted ORs [95% CI] for oldest [67-79 years] versus youngest [44-58 years] group for PSIFs and serious final diagnoses, respectively: 1.59 [1.07-2.38] and 2.79 [0.86 to 9.0] for systematic radiologist review; 1.88 [1.14-3.09] and 2.99, 95% CI [1.09-8.19] for radiographer flagging). No other investigated factor was convincingly associated with either PSIFs or serious final diagnoses.

Conclusion: Systematic radiologist review generates many more PSIFs, and proportionally fewer serious final diagnoses, compared to radiographer flagging. The risks of PSIFs and serious final diagnosis are most greatly affected by PSIFs protocol, and to a lesser extent by age. Only around 1/5 PSIFs are finally diagnosed as serious disease, and appropriate PSIFs protocol design is therefore paramount to minimise the risks of overdiagnosis in healthy research participants. Direct comparison of different PSIFs protocols are essential to inform the design of PSIFs protocols for future imaging studies.

Are we overestimating or underestimating cardiovascular events risk? The impact of using the American, British and European guidelines on cardiovascular risk assessment

Carlos Brotons[1], Irene Moral[1], Eva Maria Calvo-Bonacho[2], Mireia Puig[1], Luis Javier Quevedo-

Aguado[2], Diana Fernández[1], Carlos Catalina-Romero[2], Carlos Fernández-Labandera[2], Paloma Martínez-Muñoz[2]
[1]Sardenya Primary Health Care Center. Biomedical Research Institute Sant Pau (IIB Sant Pau), Barcelona, Spain. [2]Ibermutuamur, Madrid, Spain

Objectives: The guidelines of the American College of Cardiology/American Heart Association, the British National Institute for Health and Clinical Excellence and the European Guidelines on cardiovascular disease prevention are worldwide disseminated. They recommend the use of different tables to estimate cardiovascular risk: atherosclerotic cardiovascular disease risk-ASCVD, the QRISK2 and the SCORE. ASCVD and QRISK2 measure risk of cardiovascular morbidity and mortality due to atherosclerosis disease and due to coronary heart disease and stroke, respectively, and SCORE measures risk of cardiovascular death. According to ASCVD a high-risk person is considered when risk is $\geq$7.5%, according to QRISK2 $\geq$10% and according to SCORE $\geq$ 5%, in the next ten years. The objective was to evaluate and compare the impact of using the American, the British and the European guidelines on the identification of high risk patients and on the percentage of patients requiring statin therapy in a Spanish working population.

Method: Observational study conducted among Spanish workers whose companies had contracted health monitoring services from the *Sociedad de Prevencion de Ibermutuamur*, who underwent a medical examination between 2004-2007. Cardiovascular risk was calculated for each worker using the SCORE cardiovascular risk tables for low-risk countries, as well as the tables recommended by the American and British guidelines. Diabetic patients were excluded. Following the recommendations of the European Guidelines on Cardiovascular Prevention, treatment targets for patients at high (SCORE 5%-9%) or very high risk (SCORE $\geq$10%) are LDL-C concentrations of <100 mg/dL and <70 mg/dL, respectively. NICE and ACC/AHA recommendations do not stipulate therapeutic targets for LDL-C, therefore all patients at high risk are considered candidates for lipid lowering therapy. On top of that, ACC/AHA recommends treatment with lipid lowering therapy when LDL-C is >190mg/dl, regardless the cardiovascular risk.

Results: A total of 227,371 workers between 40 and 65 years were included (75.54% men; mean age, 47.96 years; 42.62% were smokers; 10.1% were hypertensives; 11.0% had dyslipidemia; 7.2% were treated with antihypertensive drugs; 3.7% were treated with lipid lowering drugs). Individuals at high risk was found in 4.42% of the population according to the SCORE tables and in 17.79% and 26.02% according to the British and American tables, respectively. Lipid lowering treatment would be recommended in these high risk patients, except for the American Guidelines that the percentage would increase up to 33.74% (after including non-high risk patients with LDL-C > 190 mg/dl).

Conclusions: We observed marked differences on the percentages of high risk patients when comparing the three different cardiovascular risk charts. The application of the American and British compare to the European guidelines would result in identifying more high risk patients and in treating a larger fraction of the population with lipid lowering drugs and with other intensive preventive pharmacotherapy such as use aspirin and anti-hypertensive agents, which would result in substantially increase costs. Clinicians may need to interpret cardiovascular risk estimates with caution in order to avoid overestimation of risk and overtreatment.