Introduction. Risk prediction models, using either machine learning or statistical algorithms, can act as inputs of a cost-effectiveness model when predicting costs and effectiveness of an intervention. This systematic review has two objectives: to evaluate methodological quality of the published models to predict diabetic coronary heart disease (CHD) risk; to evaluate whether the models were sufficiently reported to judge their applicability to the cost-effectiveness modelling.
Methods. A targeted review of journal articles published in English, Dutch, Chinese, or Spanish was undertaken in PubMed, Embase, Scopus, Web of Science, and IEEE Explore from 1 January, 2016 to 31 May, 2021. To assess the methodological quality and reporting of the models, we used PROBAST (Prediction model Risk Of Bias Assessment Tool), CHARMS (a Checklist for critical Appraisal and data extraction for systematic Reviews of prediction Modelling Studies), and a checklist (Betts 2019) summarizing the application of cardiovascular risk prediction models to health technology assessment.
Results. Our search retrieved 6,579 hits, of which 18 models were eligible for inclusion. Among them, four studies developed machine learning models (2 recurrent neural networks, 1 random forest models, and 1 multi-task learning model) while 14 studies developed statistical models (8 Cox models, 5 logistic models, and 1 microsimulation model). More than 70 percent of models were of high methodological quality in aspects of participants (89%), predictors (72%), and outcomes (72%), while only five models (28%) in aspects of statistical analysis. For the reporting, only two models provided sufficient evidence in all aspects (i.e., participants, predictors, and outcomes) for judging their applicability to the cost-effectiveness modelling. Most models were reported sufficiently regarding participants (78%) and outcomes (72%), but only three models regarding predictors (17%).
Conclusions. To apply the CHD risk prediction models to cost-effectiveness modelling, concerns remain regarding the potential risk of bias due to inappropriate use of analysis methods, and regarding insufficient reporting on how to measure and assess the predictors.

PP39 Evidence Generation For Reimbursement Of Digital Health Applications (DiGAs) In Germany

Naomi Fujita-Rohwerder (naomi.fujita-rohwerder@iqwig.de) and Stefan Sauerland

Introduction. In 2019, the German government established a new evaluation procedure for digital health applications (DiGAs) to facilitate their reimbursement by statutory health insurance. The procedure involves the assessment of a DiGA's "positive healthcare effect", which is defined as a medical benefit and/or "a patient-relevant improvement of structure and processes". If the available clinical evidence is insufficient to prove the manufacturer's claim on the positive healthcare effect, but the claim seems plausible, the DiGA is provisionally reimbursed, and further clinical evidence within twelve months must be generated. DiGAs eligible for provisional or permanent reimbursement are publicly listed in the DiGA directory. In contrast to the usual pathways for reimbursement of healthcare technologies which involve IQWiG as the national HTA agency and the G-BA (Federal Joint Committee) as the decision-making body, the DiGA procedure is currently carried out by the national competent authority (BfArM) and thus outside the joint self-government. Furthermore, legal evidence requirements for DiGAs are comparatively low.
Methods. This work analyzed the suitability of clinical studies that intended to prove a DiGA's medical benefit. For this purpose, the key elements for clinical studies published in the DiGA directory and clinical trial registries were extracted and compared with the usual evidence requirements in the reimbursement context.
Results. As of October 2020, 20 DiGAs have successfully undergone the application procedure. Fourteen DiGAs (70%) were provisionally accepted. A randomized controlled study (RCT) design was chosen for all clinical studies to be conducted for further evidence generation. However, in four cases (28%), it is questionable whether the clinical study is suitable to demonstrate a medical benefit mainly due to the choice or operationalization of the primary endpoint (n=2), the timing of the endpoint survey (n=2) and/or the choice of the control intervention (n=1).
Conclusions. Even though all currently ongoing or planned clinical studies with DiGAs are RCTs, not all of them are adequate to demonstrate a medical benefit according to the usual evidence requirements.

PP40 Health Apps To Manage Depression: Can We Separate The Grain From The Chaff? EvalDepApps Project

Carme Carrion (mcarrionr@uoc.edu), Ariadna Sales-Masnou, Sophie Eis, Noemí Robles, Elisa Puigdomènech, Andrea Duarte-Díaz, Josep Vidal-Alaball, Lilisbeth Perestelo, Meritxell Davins and Oriol Solà-Morales

Introduction. The use of mobile applications in the treatment of health issues is more frequently becoming common practice. Apps are fast, versatile, and manageable tools that allow the empowerment of patients and professionals, and can reduce the possible stigmatization suffered by some patients, mainly in mental health. There are more than 325,000 health apps on the market, but their impact remains unclear. There are several initiatives to define how health applications should be assessed, however, all of them address only partial aspects of the evaluation. The theoretical frameworks existing to date highlight the need to develop new tools and methodologies to assess mobile applications whose objective is the management of specific pathologies.
Methods. The primary goal of the EvalDepApps project is to develop and pilot an assessment tool for mobile applications whose main objectives are the treatment, monitoring or social support of people suffering from depression. The project is inspired by the results and

https://doi.org/10.1017/S0266462322001805 Published online by Cambridge University Press