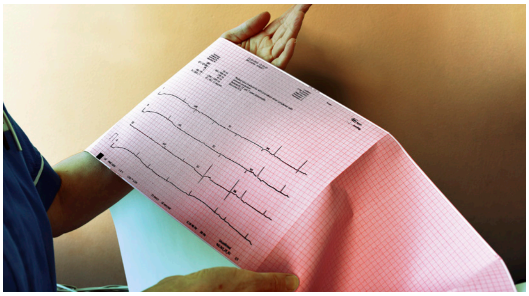
research



Lifetime risks of atrial fibrillation p 59

Predicting risk of kidney failure p 61

The growing burden of atrial fibrillation and its consequences

ORIGINAL RESEARCH Danish, nationwide, population based, cohort study

Temporal trends in lifetime risks of atrial fibrillation and its complications between 2000 and 2022

Vinter N, Cordsen P, Johnsen SP, et al

Cite this as: *BMJ* 2024;385:e077209

Find this at doi: 10.1136/bmj-2023-077209

Study question Have the lifetime risks of atrial fibrillation and subsequent complications changed over time in Denmark?

Methods In this nationwide population based cohort study, 3.5 million individuals aged 45-95 years without atrial fibrillation were followed up until incident atrial fibrillation. Individuals with incident atrial fibrillation but no prevalent complication were further followed up until incident heart failure, stroke, or myocardial infarction. The outcome measures were lifetime risks of atrial fibrillation and subsequent complications over two prespecified periods (2000-10 *v* 2011-22).

Study answer and limitations Over the past two decades, the lifetime risk of atrial fibrillation has increased from 24.2% in 2000-10 to 30.9% in 2011-22 (difference 6.7% (95% confidence interval 6.5% to 6.8%)). Among people with incident atrial fibrillation, heart failure was the most common complication, with a lifetime risk of two in five, which was double the lifetime risk of stroke after atrial fibrillation (42.9% in 2000-10 and 42.1% in 2011-22 (−0.8% (−3.8% to 2.2%))). From 2000-10 to 2011-22, the lifetime risk of heart failure showed almost no improvement and the lifetime risks of stroke and myocardial infarction after atrial fibrillation only slightly decreased, from 22.4% to 19.9% for stroke (−2.5% (−4.2% to −0.7%) and from 13.7% to 9.8% for myocardial infarction (−3.9% (−5.3% to −2.4%))). The Danish National Patient Registry provided information on atrial fibrillation diagnoses, therefore it was not possible to confirm the diagnosis through clinical examination.

What this study adds The lifetime risk of atrial fibrillation in Denmark increased between 2000 and 2022. The most common complications were heart failure, stroke, and myocardial infarction.

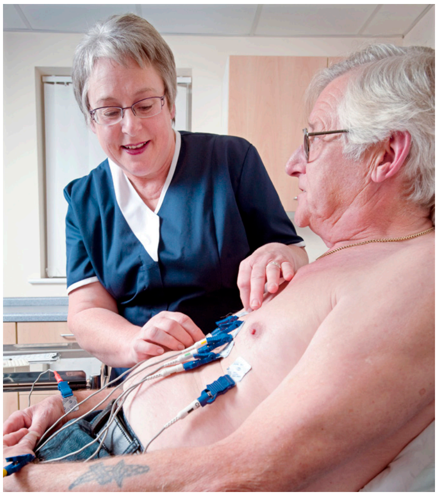
Funding, competing interests, and data sharing This study was supported by a research grant from the Danish Cardiovascular Academy (PD2Y-2022002-DCA). See full paper on bmj.com for competing interests. Access to the data used in this study can be granted on approval from the Danish Health Authority.

Atrial fibrillation is a major public health problem affecting 37 million people worldwide,[1] and conferring an increased risk of stroke, heart failure, myocardial infarction, and death, as well as quantifiable impairment in quality of life.[2] In the English NHS alone more new cases of atrial fibrillation are diagnosed each year than the four most common causes of cancer combined,[3] and direct expenditure on atrial fibrillation has reached £2.5bn.[4]

In their paper, Vinter and colleagues addressed important knowledge gaps relating to lifetime risk and sequelae of atrial fibrillation in a nationwide population based study using the population of Denmark from 2000 to 2022.[7]

Using administrative registry data from 3.5 million individuals, Vinter and colleagues estimated that the lifetime risk of atrial fibrillation for an individual aged 45 years or older increased from 24.2% to 30.9% between decades 2000-10 and 2011-22, a 28% relative increase. This risk was larger in men than in women and in individuals with prevalent heart failure, myocardial infarction, stroke, diabetes, or chronic kidney disease compared with people without these conditions. Among patients with an incident diagnosis of atrial fibrillation, heart failure was the most frequent complication, with a lifetime risk of 41.2%, double that of stroke (21.4%). Comparing the two prespecified periods, lifetime risk of heart failure after an atrial fibrillation diagnosis did not change, but absolute lifetime risks declined by 2.5% for stroke and by 3.9% for myocardial infarction.

Strengths of this observational study included the capture of data for a nationwide population of 3.5 million individuals, and use of sophisticated methods (the Aalen-Johansen estimator) to accurately calculate the cumulative incidence of atrial fibrillation and complications while accounting for left truncation and the competing risk of death. Limitations included the grouping of the population into two 10 year periods, which results in the loss of temporal resolution; the lack of reporting on ethnic group composition of the study population, which influences lifetime risk of atrial fibrillation;[8] and the absence of subgroup analysis by socioeconomic status, which affects incidence and outcomes of atrial fibrillation.[3,9]



The lifetime risk of heart failure outweighs the risk of stroke

Primary prevention needed

The finding that lifetime risk of atrial fibrillation has increased over the past two decades is not surprising because many other studies have shown increasing incidence of atrial fibrillation.[3,10] Nonetheless, routinely collected data show that contemporary lifetime risk of atrial fibrillation has increased to one in three because up to 35% of disease burden remains undiagnosed.[11] In contrast, the incidence of myocardial infarction has decreased over recent decades,[12] in association with national programmes of vascular checks to address key risk factors for ischaemic heart disease.[13] This new study reinforces the principle that analogous primary prevention programmes for atrial fibrillation are required to stem the apparent rise in incidence, associated disease burden, and cost.[2,14]

Unfortunately, the evidence base for primary prevention of atrial fibrillation predominantly relies on observational data and post hoc analyses of data from randomised clinical trials where atrial fibrillation was not prespecified as a primary or secondary endpoint, and occurrence was not systematically collected.[15] As a consequence, international guidelines do not provide specific recommendations for interventions to reduce the risk of newly onset atrial fibrillation.[2,16] While difficulties in identifying a group at sufficiently high risk for atrial fibrillation historically impeded primary prevention trials,[15] opportunities are now available to comprehensively estimate atrial fibrillation risk by considering multiple risk factors.[17,18] As such, Vinter and colleagues' findings should act as a call to prioritise prospective trials in this area.

Long term sequelae

The analysis is also noteworthy for quantifying long term risks of sequelae after an atrial fibrillation diagnosis. Atrial fibrillation care has improved considerably in recent decades, informed by randomised clinical trials showing that oral anticoagulation, and, more recently, catheter ablation, reduce the risk of stroke and death.[19,20] These interventions are being increasingly used worldwide.[21,22] International guidelines emphasise stroke prophylaxis in patients with atrial fibrillation[2]; yet, Vinter and colleagues' analysis shows that the lifetime risk of heart failure outweighs the risk of stroke.

The neglect of heart failure as a complication of atrial fibrillation in international guidelines is conspicuous because, similar to stroke, heart failure is associated with functional limitations, decreased quality of life, and poor prognosis,[23] and the subpopulation who have both atrial fibrillation and heart failure have a significantly increased risk of cardiovascular and all cause mortality.[24] Prospective cohort studies have established factors identifying people at high risk of heart failure after an atrial fibrillation diagnosis.[23,25] However, whether more intensive interventions directed towards modifiable cardiovascular risk factors could affect their long term incidence of heart failure has not been prospectively tested and requires further investigation.[25]

Interventions to prevent stroke have dominated atrial fibrillation research and guidelines, but no evidence suggests that these interventions can prevent incident heart failure. Alignment of both randomised clinical trials and guidelines to better reflect the needs of the real world population with atrial fibrillation is necessary.

Jianhua Wu Jianhua.wu@qmul.ac.uk
Ramesh Nadarajah
See bmj.com for author details

Cite this as: *BMJ* 2024;385:q826

Find the full version with references at http://dx.doi.org/10.1136/bmj.q826

Predicting the outcomes of chronic kidney disease in older adults

ORIGINAL RESEARCH Multinational, longitudinal, population based, cohort study

Predicting the risks of kidney failure and death in adults with moderate to severe chronic kidney disease

Liu P, Sawhney S, Heide-Jørgensen U, et al

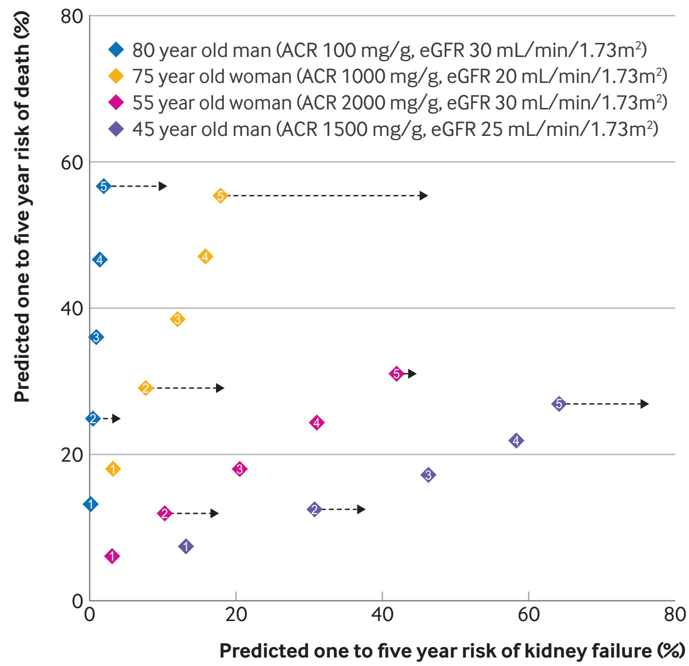
Cite this as: *BMJ* 2024;385:e078063
Find this at doi: 10.1136/bmj-2023-078063

Study question Can the risks of kidney failure and all cause death be accurately predicted in individuals with newly documented moderate to severe chronic kidney disease?

Methods A super learner meta-algorithm (KDpredict) was designed, trained, and tested to predict one to five year risks of kidney failure and all cause death in adults with newly recorded stage G3b to G4 chronic kidney disease (estimated glomerular filtration rate (eGFR) from 15 to 44 mL/min/1.73m^2) using routine population health data from Canada, Denmark, and Scotland. The super learner used cross validation to select the best performing learners among many prespecified regression models and machine learning algorithms, based on their ability to minimise prediction error. Predictors included age, sex, eGFR, albuminuria, without or with diabetes and cardiovascular disease. The index of prediction accuracy (a measure of calibration and discrimination, the higher the better) was used to compare KDpredict with the benchmark model, kidney failure risk equation.

Study answer and limitations KDpredict outperformed the kidney failure risk equation in predicting kidney failure risk: five year index of prediction accuracy of 27.8% (95% confidence interval 25.2% to 30.6%) *v* 18.1% (15.7% to 20.4%) in Denmark, and 30.5% (27.8% to 33.5%) *v* 14.2% (12.0% to 16.5%) in Scotland. Individual risk predictions from KDpredict, involving four or six variables, were accurate for both outcomes. The current version of KDpredict is a static tool to be used at the onset of the disease.

What this study adds KDpredict could be incorporated into electronic medical records or accessed online (http://kdpredict.com) to support holistic decision making in this patient population. The KDpredict learning strategy is designed to be adapted to local needs and regularly updated to accommodate changes in the underlying health system and care processes.



Two dimensional risk predictions in four hypothetical individuals. Two dimensional risk predictions at years 1 to 5 (kidney failure and death) from the four variable super learner and two and five year predictions of kidney failure only from kidney failure risk equation (black arrows). The kidney failure risk equation does not provide corresponding predictions at one, three, and four year horizons so these are not shown. Diamonds indicate point estimates (absolute risks) and 95% confidence intervals (width) and include numbers indicating prediction horizons from one to five years. Simultaneous predictions from KDpredict show how the predicted risks of kidney failure and mortality increase over sequential years. Most people with chronic kidney disease are older than 75 years and have a greater increase in the predicted risk of mortality over sequential years than kidney failure. The opposite happens to younger people. ACR=albumin to creatinine ratio; eGFR=estimated glomerular filtration rate

Funding, competing interests, and data sharing Canadian Institutes of Health Research funded this study. No competing interests declared. Regional health authorities may grant data access on request.

The BMJ is an Open Access journal. We set no word limits on *BMJ* research articles but they are abridged for print.
The full text of each *BMJ* research article is freely available on bmj.com.

The online version is published along with signed peer and patient reviews for the paper, and a statement about how the authors will share data from their study. It also includes a description of whether and how patients were included in the design or reporting of the research.

The linked commentaries in this section appear on bmj.com as editorials. Use the citation given at the end of commentaries to cite an article or find it online.

FAST FACTS Predicting the risks of kidney failure and death in adults with chronic kidney disease

The risk prediction model for kidney failure and death in people with chronic kidney disease presented in our study is a super learner. A super learner is an algorithm that repeatedly splits the data into training and test sets and then chooses the best performing model from a list of candidate prediction models. This article describes why and how the super learner was implemented in our study.

A medical risk prediction model reads the data of a patient and returns their predicted risks.[1 2] Generally speaking, the model makes predictions by referring to what happened to similar patients in the past, as recorded in a learning dataset. For example, a model could predict for a new patient with chronic kidney disease that within two years from now their risk of kidney failure is 8% and their risk of death is 13%. These predictions are interpretable as follows: out of 100 people who today are all like this patient, eight are expected to develop kidney failure and 13 are expected to die within the next two years. Notice that patients who first develop kidney failure and then die contribute to both outcomes. Whenever competing events prevent the outcome of interest, in our case death, medical decision making needs to account for the predicted risks of all events.

Risk prediction framework

Creating a medical risk prediction model based on electronic health records is challenging. A sound framework includes the definition of a clinically meaningful time zero, which is called prediction time origin, and one or multiple prediction time horizons.[2] Subsequently, the availability of predictor information at the time origin (patient age, sex, albuminuria, etc) should be verified. Inclusion of predictor variables that are accessible in a timely fashion and without great additional costs enhances model usability. For example, in our study the super learner could use four variables (sex, age, albuminuria, and estimated glomerular filtration rate, a marker of kidney function) or the same four variables plus history of diabetes or cardiovascular disease.[3] All these variables are routinely available during a clinical encounter.

Thomas Alexander Gerds
Pietro Ravani pravani@ucalgary.ca
See bmj.com for author details

Why use a super learner

The motivation for using a super learner is best explained by considering alternative strategies. A traditional statistical approach could be to prespecify a regression model and then perform a sequence of model goodness-of-fit tests based on the whole learning dataset. For example, the regression model could include additive and linear effects of all predictor variables (ie, all predictors without transformations or interactions). What should the modeller do if the model fit is rejected? A more complex model could be specified post hoc (eg, including interactions and non-linear relationships), followed by a new sequence of model goodness-of-fit tests. If the sample size is large, as in our study,[3] small deviations from model assumptions may be statistically significant even if they are not clinically relevant. Such a procedure would continue to reject the fit and eventually lead to a very flexible model. The challenge is that if the model goodness-of-fit tests are performed using the full learning dataset, overfitting is guaranteed. Overfitting means that the model has learnt too much about the learning dataset and hence will not work well in new patients. Instead, the super learner simultaneously considers many alternative models prespecified with variable degree of flexibility. By repeated cross validation, the super learner minimises overfitting hereby simulating the application of the candidate prediction models in new patients. The same applies to machine learning algorithms. A machine learning approach could be to include all predictor variables into a random forest. However, the forest requires the specification of hyperparameters (number of trees, terminal leaf size, etc) in a process called tuning. A tuning strategy works by selecting the constellation of hyperparameters with the highest cross validation prediction performance among different tuning parameters values. Hence, tuning of a random forest is a special case of super learning. These machine learning algorithms can be included together with the regression models in a super learner meta-algorithm.

Cross validation

A learner is an algorithm that takes in a learning dataset and returns a medical prediction model. The super learner uses cross validation to rank alternative learners based on their prediction performance. Cross validation randomly splits the learning data into separate sets and allows the learner to access only part of the data for training, while withholding the rest of the data for testing.[2] Different cross validation methods exist, including k-fold cross validation, bootstrapping with and without resampling. In its simplest form, cross validation splits the data into two parts, one time. Repeated random splitting removes the influence of the modeller.

The candidate learners

The modelling task consists of mapping the predictor variables to the predicted risk of the event of interest and the competing risks such that patients with similar values of the predictor variables have similar expected outcomes. To accomplish this task, the modeller can choose freely between traditional regression modelling strategies[4] and machine learning algorithms.[5] For both approaches, sensible models are only expected when the modelling team has advanced knowledge about the clinical setting and expertise on the chosen type of analysis.

How the super learner works

The ingredients of a super learner are the learning dataset, a series of candidate learners, a cross validation algorithm, and a measure of prediction performance. The super learner is expected to be as accurate as the best candidate learner that is tested.[1] An alternative super learner algorithm creates an ensemble by exploiting the discrepancies in risk prediction of a list of candidate learners and combining their risk predictions into a weighted average (ensemble super learner).[2] Methods for ensemble learning are currently limited in settings involving right censored data and competing risks.

Cite this as: *BMJ* 2024;385:q721

Find the full version with references at http://dx.doi.org/10.1136/bmj.q721

Key features of the super learner strategy

- Provides a systematic method for making a prediction model
- Combines traditional regression models with machine learning algorithms
- Expected to perform as well as the best candidate learner

COMMENTARY New prediction models developed with machine learning show promise

Clinical prediction models use mathematical functions to combine information from an individual's characteristics and use the resulting output to estimate the chances that a condition of interest is present (diagnostic models) or will occur in the future (prognostic models).[1] These estimates are in turn used to start or withhold treatments, order new tests, or educate individuals on their condition's outlook.

Prediction models are increasingly used in clinical medicine. In nephrology, estimators of kidney function are used to predict glomerular filtration rate because the direct measure of kidney function is not feasible in routine settings. In people with a diagnosis of chronic kidney disease (CKD) already, prediction models are used to estimate the risk of progression to renal replacement therapy and to predict other outcomes such as mortality or cardiovascular diseases.[2]

In their paper, Lui and colleagues reported the development and evaluation of KDpredict,[3] a suite of new models for predicting future kidney failure and all cause mortality in adults with moderate to severe CKD over one to five years. KDpredict was developed using the super learner strategy based on machine learning.[4] The super learner is a prediction method designed to find the best combination from a collection of prediction algorithms (also known as learners). Prediction algorithms are ranked according to prediction error (Brier score)—a measure of discrepancy between the expected risk of a particular outcome (from the model derived from a training dataset) and the risk observed in a validation dataset. A lower prediction error indicates a better model performance.

Liu and colleagues used the discrete super learner approach for evaluation, which selects the best prediction algorithm from a library of prespecified algorithms. They developed KDpredict using data from a cohort of patients in Alberta (Canada) and externally validated the model using patient cohorts from Denmark and Scotland. KDpredict showed that risk of death over five years was higher than the risk of progression to kidney failure in most subgroups with moderate to severe CKD


IMAGEBROKER/ALAMY

Further external validation of KDpredict in diverse settings by independent investigators is needed

at baseline. The model's performance was excellent when tested in the Danish and Scottish cohorts.

Differences between prediction models

Traditionally, the kidney failure risk equation, also developed in Canada,[5] has been the benchmark model for predicting outcomes among people with CKD, using the same core predictors as KDpredict. When both models were tested in patient cohorts from Denmark and Scotland, KDpredict performed better than the kidney failure risk equation at predicting end stage kidney failure.

Besides differences in model development (super learner for KDpredict *v* standard Cox regressions modelling for kidney failure risk equation), KDpredict and kidney failure risk equation differed in various ways. The population used to develop KDpredict was older (median age 77-80 *v* 69-70 years) and included people with more advanced CKD (estimated glomerular filtration rate <45 mL/min/1.73 m^2 *v* <59 mL/min/1.73 m^2). These dissimilarities may account for some of the difference in performance. While both models incorporated the same core variables, KDpredict included substantially more predictors overall. This broader approach makes more efficient use of predictive information, which translates into enhanced predictive performance. This approach can result in models that perform well in the derivation sample, but less well when tested in a different population. Furthermore, the increased complexity of models such as KDpredict means that routine uptake will depend on automated computation of risk. Widespread availability of smart devices and internet connectivity may help to overcome this issue.

More than meets the eye

Liu and colleagues cited various limitations of the kidney failure risk equation to justify the development of new and more complex models; however, some of these limitations are arguable. Firstly, although the current kidney failure risk equation does not account for competing risks, the investigators of the equation did develop an alternative model using competing risk approach that made no material difference to the original equation's performance.[5] Secondly, the need to predict all cause mortality for people with CKD (available with KDpredict but not with kidney failure risk equation) is debatable. Cardiovascular disease is the leading driver of mortality among adults with moderate to severe CKD so these individuals typically have multiple treatments to reduce cardiovascular disease, irrespective of their absolute cardiovascular disease risk. Therefore, estimating mortality risk is unlikely to substantially affect patients' clinical management, including preparing them for the possibility of end stage renal replacement treatment.

The proposed locally optimised decision support with KDpredict that moves away from the current one-size-fits-all approach would require a specific model for each setting. This customisation is not achievable because the data needed to optimise the model locally will not always be available. Furthermore, the opportunity to compare the models' performance within and across settings will be missed.

Additionally, the super learner approach to prediction models requires substantial input from experts with knowledge of the field of interest; knowledge that is generally derived from conventional approaches to prediction. As such, super learner approaches to prediction are better positioned as an extension of, or enhancement to, the conventional one-size-fit-all approach.

KDpredict is a potentially useful tool for risk stratification in people with CKD, to be integrated into routine care, alongside existing tools such as the kidney failure risk equation. But, further external validation of KDpredict in diverse settings by independent investigators is needed.[6,7]

Andre Pascal Kengne Andre.kengne@mrc.ac.za
Cindy George
Oluwatoyin Idaomeh Ameh
See bmj.com for author details

Cite this as: *BMJ* 2024;385:q749

Find the full version with references at http://dx.doi.org/10.1136/bmj.q749