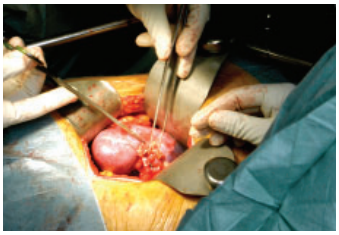
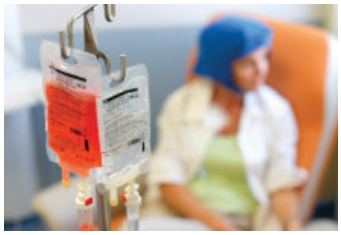
research



Predicting failure of kidney transplants p 353



Bias in trials supporting new cancer drugs p 354

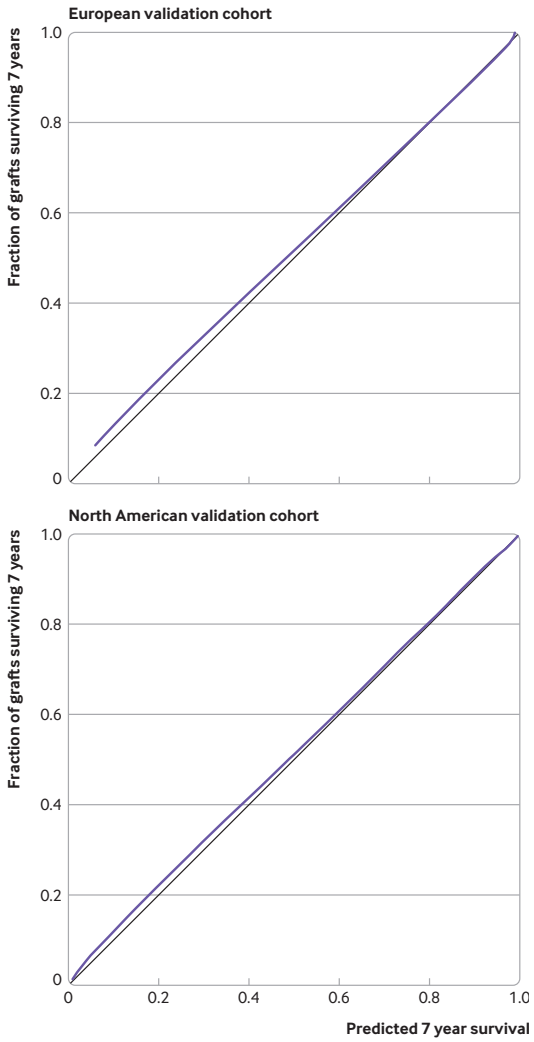
Aerobic fitness training versus relaxation in subacute phase of stroke p 356

ORIGINAL RESEARCH International derivation and validation study

Prediction system for risk of allograft loss in patients receiving kidney transplants

Loupy A, Aubert O, Orandi BJ, et al

Cite this as: *BMJ* 2019;366:l4923
Find this at: http://dx.doi.org/10.1136/bmj.l4923



Calibration plots at seven years of iBox risk scores for European and US validation cohorts. Black line represents perfectly calibrated model, and blue line represents optimism corrected iBox model

Study question Can a prediction system that assesses kidney transplant patients' individual risk of long term allograft failure be developed and validated?

Methods This was an international derivation and validation study involving 10 academic transplant centres. The risk score was derived within the Paris Transplant Group's deep phenotyped cohort. The performance of the score was then assessed in two independent cohorts from Europe and the US. The score was further validated in different clinical scenarios including type of immunosuppressive regimen used and response to rejection therapy and in three randomised therapeutic clinical trials.

Study answer and limitations Among the 7557 kidney transplant recipients included, 1067 (14.1%) allografts failed after a median post-transplant follow-up time of 7.12 (interquartile range 3.51-8.77) years. Eight functional, histological, and immunological prognostic factors were independently associated with allograft failure and were then combined into a risk prediction score (iBox). The iBox system was accurate in predicting allograft failure (C index 0.81, 95% confidence interval 0.79 to 0.83) and showed generalisability across centres worldwide and in common clinical scenarios, as well as in randomised clinical trials. Despite the validation of the iBox risk prediction score in an interventional setting, future trials need to determine whether a strategy based on a systematic risk evaluation compared with an empirical approach might improve clinical management.

What this study adds This study shows the value of a readily implementable risk prediction system that may help to guide monitoring of kidney transplant recipients. The iBox represents a valid and early surrogate endpoint for clinical trials and drug development in transplantation.

Funding, competing interests, and data sharing INSERM–Action thématique incitative sur programme Avenir (ATIP-Avenir) and Fondation Bettencourt Schueller provided financial support. A technical appendix is available from the corresponding author at alexandreloupy@gmail.com.

Study registration ClinicalTrials.gov NCT03474003.

Flawed evidence underpins approval of new cancer drugs

ORIGINAL RESEARCH Cross sectional analysis

Design characteristics, risk of bias, and reporting of randomised controlled trials supporting approvals of cancer drugs by European Medicines Agency, 2014-16

Naci H, Davis C, Savović J, et al

Cite this as: *BMJ* 2019;366:l5221
Find this at: http://dx.doi.org/10.1136/bmj.l5221

Study question What are the design characteristics, risk of bias, and reporting adequacy in pivotal randomised controlled trials that supported approvals of cancer drugs by the European Medicines Agency (EMA) from 2014 to 2016?

Methods This was a cross sectional analysis that focused on the primary endpoint of each pivotal randomised controlled trial. The main outcome measures were study design characteristics (randomisation, comparators, and endpoints); risk of bias using the revised Cochrane tool (bias arising from the randomisation process, deviations from intended interventions, missing outcome data, measurement of the outcome, and selection of the reported result); and reporting adequacy (completeness and consistency of information in trial protocols, publications, supplementary appendices, clinical trial registry records, and regulatory documents).

Study answer and limitations Between 2014 and 2016, the EMA approved 32 new cancer drugs on the basis of 54 pivotal studies. Of these studies, 41 (76%) were randomised controlled trials and 13 (24%) were either non-randomised studies or single arm studies. 39/41 randomised controlled trials had available publications and were included in our study. Overall, 19 randomised controlled trials (49%) were judged to be at high risk of bias for their primary outcome. Concerns about missing outcome data (n=10) and measurement of the outcome (n=7) were the most common domains leading to high risk of bias judgments. Fewer randomised controlled trials that evaluated overall survival as the primary endpoint were at high risk of bias than those that

COMMENTARY We must raise the bar to ensure real benefits for patients

In this issue, Naci and colleagues examine the risk of bias in randomised controlled trials that support European approvals of cancer drugs from 2014 to 2016. Their study confirms and extends the existing body of research that raises serious concerns about low standards of evidence supporting new cancer drug approvals.[2] Only 10 of the 39 randomised controlled trials (26%) that supported these approvals assessed overall survival as a primary outcome. Quality of life was reported in 17 of the randomised controlled trials (44%), mainly as a secondary outcome.

These authors also assessed how rigorously these trials were designed and conducted through a standardised risk of bias tool developed by the Cochrane Collaboration, which focuses on factors that exaggerate measured outcomes,[6] and through a review of European Medicines Agency regulatory documents.[7] Nearly half of the trials, 19/39 (49%), were judged to be at high risk of bias, which indicates that treatment effects might have been exaggerated. Trials that evaluated surrogate outcomes such as progression-free survival were at high risk of bias more often than those that evaluated overall survival. For 10 of the 32 new drugs (31%) approved over this period, regulators identified additional problems such as unplanned early termination, questionable clinical benefits, or use of inappropriate comparators. These concerns rarely surfaced in published reports.

Barbara Mintzes
barbara.mintzes@sydney.edu.au
Agnes Vitry
See bmj.com for author details

Patient access

In countries with public health insurance systems, patient access to cancer drugs largely depends on Health Technology Assessment agencies' funding recommendations.[8] Similar to regulatory agencies, Health Technology Assessment agencies are under pressure to fund new cancer medicines quickly.[9] Poor quality clinical trials and uncertainty about patient relevant outcomes create even greater challenges when a treatment's cost effectiveness is being compared with existing care. In Australia, the Pharmaceutical Benefits Advisory Committee, which makes national funding recommendations, frequently rejects cancer medicines because of uncertain clinical evidence.[10] Germany's Health Technology Assessment agency (Institute for Quality and Efficiency in Healthcare) assessed 82 new cancer drugs and indications from 2011 to 2017, and found 54% to have considerable, major, or minor benefits, but 39% had no proof of added benefit, 2% did worse than comparators, and for 5%, evidence was inadequate to quantify benefits[9] (B Wieseler, personal communication, 2019).

Uncertainty and exaggeration of the evidence that supports approval of cancer drugs causes direct harm

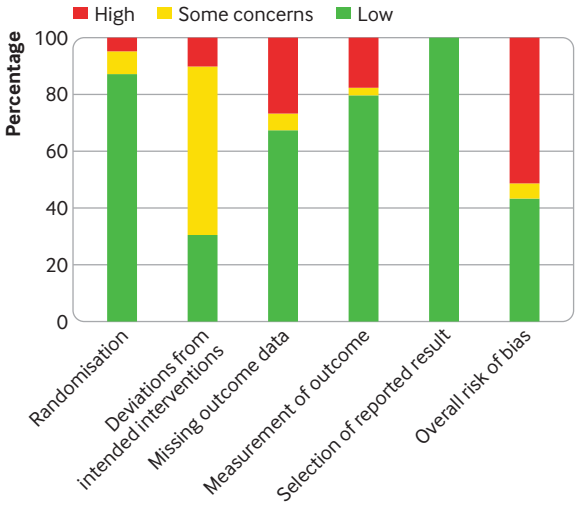
Uncertainty and exaggeration of the evidence that supports approval of cancer drugs causes direct harm if patients risk severe or fatal adverse effects without likely benefit, or forgo more effective and safer treatments. Inaccurate evidence also leads to intangible harms if it encourages false hope and creates a distraction from needed palliative care.[11] The average price of a course of cancer treatment in the US routinely exceeds $100 000 (£82 100; €90 800), with little correlation between extent of established health benefits and pricing.[12]

Professional bodies, including the European Society for Medical Oncology and the American Society of Clinical Oncology, have called for the bar to be raised. These bodies have defined clinically meaningful outcomes, including overall survival, quality of life, and adverse effect profile,[13 14] and

evaluated surrogate efficacy endpoints (2/10 (20%) *v* 16/29 (55%), respectively). This study examined the risk of bias, rather than bias itself. Therefore, it remains a possibility that trial results are unbiased despite the methodological flaws identified in our assessments.

What this study adds Around half of randomised controlled trials that supported European cancer drug approvals from 2014 to 2016 were assessed to be at high risk of bias based on characteristics of their design, conduct, or analysis. Trials that evaluated overall survival were at lower risk of bias than those that evaluated surrogate measures of clinical benefit.

Funding, competing interests, and data sharing This study was funded by the Commonwealth Fund, Higher Education Funding Council in England, Health Action International (HAI), National Institute for Health Research (NIHR) Biomedical Research Centre at University Hospitals Bristol NHS Foundation Trust and the University of Bristol, NIHR Collaboration for Leadership in Applied Health Research and Care West (CLAHRC West) at University Hospitals Bristol NHS Foundation Trust, and Arnold Ventures. See competing interests statement in full paper on bmj.com. No additional data available.



Risk of bias assessment that used combined information from scientific literature and regulatory documents and was based on primary efficacy endpoints

developed tools to assess clinical value.[15 16] On average, premarket testing for overall survival takes one extra year compared with use of surrogate outcomes.[17] A year may seem a long wait for a patient with advanced life threatening disease. However, several policy options have been proposed to facilitate earlier access to experimental treatments, such as clinical trial participation, compassionate access programmes, or managed entry agreements with ongoing funding conditional on more solid evidence.[18 19]

With increasing reliance on single armed trials to support cancer drug approvals, Naci and colleagues' study shows that trial evidence alone is not enough. Quality assessment of that evidence is also needed to ensure that these trials accurately estimate treatment effects.

Cite this as: *BMJ* 2019;366:l5399

Find the full version with references at http://dx.doi.org/10.1136/bmj.l5399

BURGER/PHANIE/SPL

Physical Fitness Training in Patients with Subacute Stroke (PHYS-STROKE)

Nave AH, Rackoll T, Grittner U, et al

Cite this as: *BMJ* 2019;366:l5101
Find this at: http://dx.doi.org/10.1136/bmj.l5101

Study question Is an aerobic physical fitness intervention in the subacute phase after stroke superior to relaxation plus standard care in terms of maximal walking speed and activities of daily living?

Methods This multicentre, randomised controlled, endpoint blinded trial included 200 participants with moderate to severe subacute stroke (days 5-45 after stroke). Participants were assigned to receive either aerobic, bodyweight supported, treadmill based physical fitness training (n=105) or relaxation sessions (n=95), each for 25 minutes, five times weekly for four weeks, in addition to standard rehabilitation therapy. The primary outcomes were change in maximal walking speed and change in Barthel index score (range 0-100 points, higher scores indicating less disability) at three months after stroke compared with baseline. Safety endpoints were recurrent fatal or non-fatal cardiovascular events, including stroke, readmission to hospital, and death.

Study answer and limitations Compared with relaxation sessions, aerobic physical fitness training did not result in a significantly higher change in mean maximal walking speed (treatment effect 0.1 m/s (95% confidence interval 0.0 to 0.2 m/s), P=0.23) or change in mean Barthel index score (0.0 (−5.0 to 5.0), P=0.99) at three months after stroke. A higher rate of serious adverse events was observed in the aerobic physical fitness training group compared with relaxation group (incidence rate ratio 1.81, 95% confidence interval 0.97 to 3.36). Findings are, however, only applicable to moderately to severely affected people with subacute stroke and should not be generalised to the stroke population at large.

What this study adds These results do not appear to support the use of aerobic, bodyweight supported, fitness training in people with subacute stroke to improve activities of daily living or maximal walking speed, and these findings should be considered in future guideline recommendations.

Funding, competing interests, and data sharing This study was funded by the German Ministry for Health and Education. See competing interests statement in full paper on bmj.com. The raw trial data are provided on a secure online repository (http://doi.org/10.5281/zenodo.3341240) and will be available three months after publication of the article.

Study registration ClinicalTrials.gov NCT01953549.

Results for primary efficacy outcome of change in maximal walking speed and Barthel index score from baseline to three months after stroke by aerobic physical fitness training or relaxation sessions (control group)

Primary outcomes	Aerobic physical fitness training (n=105)	Relaxation sessions (n=95)	Adjusted treatment effect*	P value
Mean (95% CI) maximal walking speed (m/s)	0.4 (0.3 to 0.4)	0.3 (0.2 to 0.4)	0.1 (0.0 to 0.2)	0.23
Mean (95% CI) Barthel index score	30 (24 to 36)	30 (23 to 36)	0 (−5 to 5)	0.99

Analyses based on multiple imputation.
*Treatment effects were analysed using analysis of covariance mixed models with three month outcome as dependent variable adjusted for baseline and additionally adjusted for sex, study centre, and functional ambulation category.