EDITORIALS

Editorials are usually commissioned. We are, however, happy to consider and peer review unsolicited editorials

Problems with the new junior doctor contract

Why junior doctors need to channel their anger

Jessamy Bagenal editorial registrar, *The BMJ* (NHS national medical director's clinical fellow scheme 2015-16), and general surgical registrar, North West Thames
jbagenal@bmj.com
Tom Moberly editor, BMJ Careers
Fiona Godlee editor in chief, *The BMJ*

Junior doctors are angry. In August 2016 the UK government plans to impose a new contract on doctors in England that threatens to extend their standard working hours while cutting their pay by up to 15%.[1]

More than 54 000 people have already signed an online petition calling on doctors to take strike action to oppose the imposition of the contract.[2] Some medical royal colleges have also taken the unusual step of commenting on contractual terms and conditions, an area usually reserved for trade unions. The Royal College of Paediatrics and Child Health, the Royal College of Physicians of Edinburgh, the Royal College of Psychiatrists, the Royal College of Surgeons, the Association of Surgeon in Training, and others have raised concerns about the potential effect of the changes on the safety of patients and on recruitment and retention.

Although many details about the proposed changes are missing, they may look fairly harmless at first sight. Changes are based on recommendations of the independent Review Body on Doctors' and Dentists' Remuneration. They include contractual safeguards to ensure that hours and rest periods are maintained; a separate hourly payment for unsocial hours, with an end or change to pay banding; an extension to the normal working day; and for pay to be based on actual progression to the next level when junior doctors take up a new post.[3] As always, the devil is in the detail. Under the proposals, unsocial hours are likely to be redefined so that working up to 21:59 on a Saturday night attracts the same rate of pay as working from 09:00 on a Tuesday morning.[4] The BMA says that it has struggled to find any junior doctor who would be financially better off in the new contractual system.[5] And indeed it has been estimated that even high intensity specialties such as anaesthetics and emergency care will see a pay cut.

Think of the public

Junior doctors are also concerned about the effect of the new reward structure on doctors who are not in a training scheme and those who take time out to do research or for illness or parental leave. The proposed changes mean that annual pay increments will be directly linked to appraisal and will be given only on progression to the next level of a recognised training programme. Many doctors work in roles that are not included in any such training programme, and so under the proposed changes they would be excluded from annual pay increments. The BMA is also concerned that the proposals could increase the pay gap between the sexes and deter trainees from training less than full time.[6]

Tired of doing more for less

Resentment over the proposed changes to the junior doctors contract has fed into wider frustration across the NHS, with staff generally reporting that they feel demoralised, disenfranchised, and undervalued.[7] As the NHS struggles to deliver the efficiency target it has been tasked with meeting, its workforce is tired of being asked to do more for less.

> Junior doctors need to turn their frustration and anger into meaningful engagement

It doesn't have to be like this. Enlightened employers such as Google and the consultancy firm Deloitte focus on their employees' happiness as a means to improve their business outcomes.[8] The NHS needs to do the same. Results of staff surveys correlate with mortality rates and patients' satisfaction.[9 10] And staff satisfaction is inversely related to staff sickness, which has been gradually increasing since 2009.[11]

NHS Employers has acknowledged that staff morale is low and that work-life balance and staff wellbeing need to be improved.[12] NHS England and NHS Employers have committed £5m (€6.9m; $7.8m) towards improving staff health, wellbeing, and support.[13] Welcome though this is, it works out at £3 per employee, and the proposed changes to the junior doctors contract are likely to undermine, rather than improve, doctors' work-life balance and their ability to work flexibly.

Most junior doctors' ignorance about negotiations over their pay and conditions has not helped. Negotiations between NHS Employers and the BMA ended abruptly 11 months ago when the junior doctors' negotiators walked out. Yet the great majority of junior doctors have only now begun, largely through social media, to participate in the debate. NHS Employers is holding meetings throughout the country to connect with doctors directly. Having got things badly wrong, it now has a clear responsibility to listen and respond to concerns over patient safety.

Now that junior doctors have turned their attention to the proposed changes to their pay and conditions, they need to understand the contractual process and engage with it constructively to ensure improvements in working conditions. Doctors need to be careful about taking strike action. Media coverage of a possible strike by doctors has been relatively favourable so far. After all, 72% of members of the public are proud of the NHS, and doctors are still ranked as the most trustworthy profession in opinion polls.[14] But public attitudes can quickly change, and focusing on pay may alienate the public, especially when politicians undermine doctors and their work.

The proposed changes seem to be harsh and lacking in clarity and planning. It's important that junior doctors dispute aspects that will affect patient safety and workforce retention. But they need to turn their frustration and anger with the current proposals into meaningful engagement so that their contract develops in a way that is beneficial to patient care, the health service, and doctors themselves. Junior doctors should focus their argument on the need to enhance the wellbeing and work-life balance of all NHS staff for the quality and safety of patient care to be improved. This is more likely to capture the public support that they otherwise risk losing.

Cite this as: *BMJ* 2015;351:h5077

thebmj.com

Letter: From the FDA, we still hear mostly thunderous silence (*BMJ* 2015;351:h3763)

These studies give cause for concern about whether most new drugs are any more effective than existing products or whether their safety has been adequately assessed

The FDA's new clothes

The FDA does not protect patients from harmful or ineffective drugs, but approves both

Donald W Light professor, School of Osteopathic Medicine, Rowan University, 2250 Chapel Avenue, Cherry Hill, NJ 08002, USA dlight@rowan.edu
Joel Lexchin professor, School of Health Policy and Management, York University, 4700 Keele Street, Toronto, ON M3J 1P3, Canada

The Vioxx disaster in the early 2000s triggered a crisis of mistrust in the US Food and Drug Administration (FDA), as evidence emerged that it had downplayed or ignored evidence of serious cardiovascular harm associated with Vioxx (rofecoxib), a cyclo-oxygenase-2 selective non-steroidal anti-inflammatory drug.

The result was a renewed emphasis on drug safety throughout a product's lifecycle. At the same time, drug companies, which provide most of the funds for the FDA's review of their drugs, kept pushing for faster approvals and new uses for old drugs, supposedly so that more patients could benefit. Any possible risks in getting new drugs to market more quickly would be offset by more intensive monitoring once they were being prescribed.

Two linked papers provide valuable accounts of how the FDA is using faster reviews for what it deems to be important new drugs and using supplemental approvals for existing drugs more widely.[1 2] This is just what patients and their doctors are said to want—more patients benefiting from taking more new drugs sooner, generating revenue for the companies to fund more breakthrough research.

Put in the context of the FDA's larger record, however, these studies give cause for concern about whether most new drugs are any more effective than existing products or whether their safety has been adequately assessed. The term "safe and effective" misleads patients and prescribers. Although the US Congress and the FDA require "substantial evidence of effectiveness" to approve new drugs, they require no evidence of substantial effectiveness.[3] Companies provide substantial evidence of effectiveness through trials that in most cases prove only that the product being tested has a non-zero level of effectiveness. The result is that independent reviews find that 85-90% of new drugs provide few or no advantages for patients.[4] The FDA's flexible criteria and low threshold for approval do not reward more research for breakthroughs but instead reward more research for minor variations that can clear this low threshold.

More drugs, weaker evidence

The growing number and widening application of expedited review programmes are accompanied by evidence that many of the clinical trials accepted by an industry compliant FDA have features that contribute to biased results and compromised science (see box on thebmj.com).[5 6] As a consequence, these trials are incapable of providing patients or doctors with valid information on what new clinical benefits a drug provides. The result is an ever larger number of drugs approved on the basis of weaker evidence and in shorter time periods. We documented this for cancer drugs,[5] and a much more comprehensive review comes to similar conclusions across many areas of medicine overseen by the FDA.[6] Yet both of the linked studies point out that Congress is poised to advocate for still more accelerated reviews based on even less evidence.

Do patients and doctors really want medicines for cancer and other life threatening conditions approved this way—quickly, with marginal evidence of real benefit? Do they know that faster reviews are associated with a significant increase in serious safety problems[7] and the risk of patients being admitted to hospital with or dying from adverse reactions?[8] Canadian data show that faster review increases the chances of harm serious enough to warrant a severe warning or market withdrawal from one in five to one in three.[9]

In most drug research, harm is called "safety" or "safety events," a fig leaf of pharmaceutical English covering up the real thing. The "risk-benefit ratio" can also obscure the real chance of serious harm. When the possibility of benefit declines, the chance of being harmed stays the same, so the ratio of harms to benefits increases.[10] Prescription drugs are the fourth leading cause of death in the United States and the third leading cause in Europe, according to one authority.[11 12]

These twin studies are part of a series drawing on impressive datasets assembled under Kesselheim's direction at Harvard University. However, these data are hard to abstract and collate and require searches through multiple FDA databases, along with Freedom of Information Act requests. Wang and Kesselheim could not locate the FDA medical reviews containing the clinical evidence for the basis of approval for 80% of the supplemental applications. Just one medical review was available among the 66 approvals in 2013-14. Only slightly more than 30% of supplemental approvals were supported by trials against active comparators, and more than 70% of approvals were based on trials using surrogate endpoints. Effectively, the FDA has been granting most supplemental approvals without evidence of meaningful clinical benefit.

FDA data on drug withdrawals are equally lacking. A recent review of safety warnings finally concludes that, "Remarkably, no comprehensive source of information on black-box warnings and withdrawals is available."[13]

The United States and other countries need an alternative paradigm—one in which research focuses on better medicines for patients rather than for profits, where clinical trials with low risk of bias look for real benefits and faithfully report harms. Such a paradigm of ethical, open, not for profit research already exists at research institutes such as the Mario Negri Institute for Pharmacological Research.[14] Although this institute accepts funding from drug companies, it operates under rules and practices for keeping drug research independent, transparent, and accountable. The institute's leaders have long advocated for publicly funded regulators whose deliberations are transparent and accountable. With so much misdirected investment, biased science, and harm resulting from industry directed research, with little offsetting benefit, perhaps it is time to consider the Mario Negri public health model for developing better medicines for patients.

Cite this as: *BMJ* 2015;351:h4897

EDITORIAL, p 9
RESEARCH, pp 11, 12

Clinical Review: Diagnosis and management of first trimester miscarriage (*BMJ* 2013;346:f3676)

Diagnosing a miscarriage

When is it safe to make the call?

Elizabeth Anne McCarthy senior lecturer
eamcca@unimelb.edu.au
Stephen Tong professor, University of Melbourne, Department of Obstetrics and Gynaecology at Mercy Hospital for Women, Heidelberg, Victoria, Australia

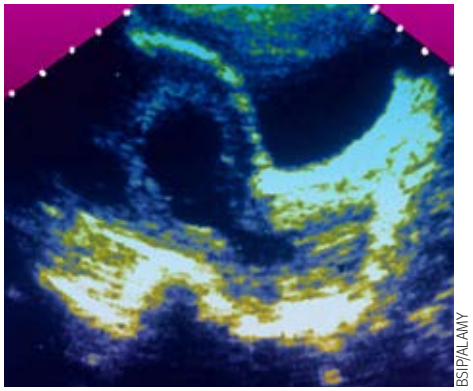
The diagnosis of a miscarriage will more than likely dash the hopes and dreams of couples anticipating the life changing possibilities of a new addition to their family—a new house, extended leave from work, relief in finally appeasing grandparents-in-waiting. It is therefore paramount that clinicians get it right. Always.

It was a touch unnerving when evidence emerged around 2011 suggesting that ultrasonography criteria advocated by official bodies may be too liberal, and that they risked some viable pregnancies being misdiagnosed as miscarriages.[1-4] In response, more conservative ultrasonography criteria were published as recently as 2013.[5]

Given these recent shifts in criteria to define something so important, it is timely that in this issue Preisler and colleagues present a large prospective multicentre study to inform about when it is safe to make the call that a pregnancy has failed.[6]

Over two and a half years the team recruited 2558 women across seven early pregnancy units in the United Kingdom.[6] They included women in whom it was unclear at the first ultrasound scan whether the pregnancy was ongoing, and plans were made for repeat scanning. By the follow-up scans performed at 11 to 14 weeks' gestation, about half of the pregnancies appeared healthy and the remainder had definitely miscarried. A decided strength of this study was the large sample size, which produced narrow confidence intervals to estimate positive predictive and specificity values.

Firstly, the authors established definitive criteria defining when it is safe to diagnose miscarriage at the first ultrasound examination. A single scan is diagnostic for miscarriage if there is an empty gestational sac with a mean diameter ≥25 mm, or there is an embryo present with a crown-rump length ≥7 mm but no cardiac activity. These measurements are in agreement with recent published guidelines[5] and supersede more liberal criteria for missed



BSIP/ALAMY

Tough call

miscarriage—smaller sac diameters (for example, 16 or 20 mm)[7 8] and shorter crown-rump length (for example, 5 or 6 mm).[7 8]

Furthermore, if the time between the first day of the last menstrual period is known with certainty to be ≥10 weeks, clinicians can safely diagnose a miscarriage if the mean gestational sac diameter is ≥18 mm without a visible embryo, or the embryonic crown-rump length is ≥3 mm without cardiac activity. This is pragmatically useful as many women present for their first ultrasound examination at 11-14 weeks' gestation to examine fetal anatomy and calculate the risk of aneuploidy.

It is paramount that clinicians get it right. Always

In a previous era, a tincture of time was the main diagnostic test for threatened miscarriage. Old textbooks recommended re-examining uterine size after three or four weeks to determine whether growth had arrested and thus made the diagnosis of a "dead ovum," "carneous mole," or missed abortion.[9] With the advent of ultrasonography in the late 1960s, early pregnancy development has become visible, matching embryology specimens seen in the anatomy museum.[10]

Watchful waiting: still a role

Despite ultrasound technology, we cannot get around the fact that often we still must resort to watching and waiting. In the era of in vitro fertilisation and home pregnancy tests, many women present early for their first ultrasound examination. Thus it is common at a viability ultrasound investigation that embryonic and gestational sac sizes are too small to be certain whether an early pregnancy is still viable or a miscarriage has occurred. Anxious women are told to return one or two weeks later for repeat scanning to see whether structures have grown.

However, the underlying evidence guiding practice in this scenario has been limited. Should repeat scanning be arranged 7, 10, or 14 days later? The work by Preisler and colleagues is an important advance that provides greater certainty. They conclude that if at initial ultrasound scanning an embryo with a crown-rump length ≤6 mm is seen with no cardiac activity, a diagnosis of miscarriage can be made if there is still no heartbeat at a repeat scan at least seven days later. If the mean sac diameter is <12 mm and an embryo is not seen, a miscarriage can be diagnosed if the sac diameter has not doubled after at least 14 days. Finally, if the mean sac diameter is ≥12 mm but there is no embryo, a miscarriage can be diagnosed if there is still no embryonic cardiac activity at least seven days later.

To parents experiencing even one miscarriage, the loss can be devastating. But to find answers to help we need data from more than the affected parents and more than one busy maternity service; we need evidence from thousands of women. Preisler and colleagues have gone a long way to achieving this by including seven centres in their study. Organisations such as the UK Miscarriage Association (www.miscarriageassociation.org.uk) can connect clinicians, researchers, women, and families with a threatened or confirmed miscarriage, helping continuous quality improvement and advancing research.

The Miscarriage Association website also provides woman centred support and information. The website advises: "Sometimes, it can take several scans before you know for sure what is happening. It can be very stressful dealing with this uncertainty—some women describe it as being 'in limbo.' You may need to find some support for yourself if this happens to you." These words remain true after the study by Preisler and colleagues. A tincture of time and a lot of human care remain important for women facing the possibility of a miscarriage.

Cite this as: *BMJ* 2015;351:h4769

RESEARCH, p 13

thebmj.com

Research: Frequency of discrepancies in retracted clinical trial reports versus unretracted reports: blinded case-control study (*BMJ* 2015;351:g4708)

The predictive power of an error count to identify rogue papers will be extremely low. You would have to kiss many princes to find a frog

Catalogue of errors in papers reporting clinical trials

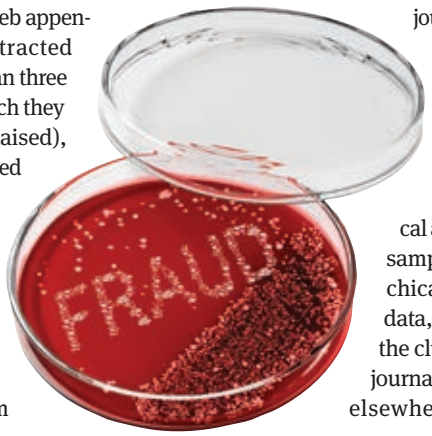
Errors are linked to retraction, but are an unreliable marker for fraudulent or harmful research

Nick Freemantle professor of clinical epidemiology and biostatistics
nicholas.freemantle@ucl.ac.uk
Greta Rait reader in primary care, Department of Primary Care and Population Health, PRIMENT Clinical Trials Unit, UCL Medical School (Royal Free Campus), London NW3 2PF, UK

The paper by Cole and colleagues[1] examines the association between discrepancies and retractions in clinical trial reports, concluding that discrepancies or errors could be an early signal of unreliability in clinical trials. It is another foray by these authors into the examination of error in reported research. Their previous paper[2 3] demonstrated a clear and concerning relation between error and inflated effect size in stem cell research. But is the identification of unreliable, misleading, or even fraudulent research really as simple as counting the number of errors or discrepancies in clinical trial reports? Cole and colleagues describe errors or discrepancies as red flags, but several limitations within their own study indicate that these red flags are an unreliable marker for deeper problems within a piece of research.

The authors undertake a case-control study and make recommendations based on sensitivity and specificity, even though they acknowledge that these metrics depend on prevalence and the research design is inappropriate.[4] There are far more unretracted reports of clinical trials (802 953 papers on PubMed, 6 July 2015) than retracted reports of clinical trials (379 papers on same date), so the predictive power of an error count to identify rogue papers will be extremely low. You would have to kiss many princes to find a frog.

This is illustrated by the authors' own table of included studies (web appendix 2). Fifteen (30%) retracted papers (frogs) had fewer than three errors each (the level at which they suggest that a red flag is raised), whereas 17 (34%) unretracted controls (princes) had more than three errors.



Cole and colleagues interpret their outcome (retraction) as if it is always homogeneous and harmful. But retracted papers do more or less harm depending on the reason for retraction. For example, a duplicate publication, while wrong, could paradoxically help a practitioner to identify an important piece of information that they might otherwise miss, while fabricated results lead more directly to patient harm if they wrongly suggest that an ineffective treatment works. Retraction cannot be considered as "the variable capable of providing the most clinically relevant and convincing evidence."[5] Instead, the authors could have chosen a more specific outcome such as retraction for misconduct.

The authors anonymise their selected papers, but provide National Library of Medicine identification numbers, making it possible—although tedious—for readers to identify the journals that published these papers and to scrutinise their study material. Basic demographics of participants are an essential part of good reporting practice. In this case, the authors could and arguably should have reported the journal of publication, and the specialty of each included paper. While reviewing the articles, we thought it was important to see that, for example, eight (16%) retracted papers came from the same anaesthetic journal, *Anesthesia and Analgesia*. On the basis of the journal titles alone, 15 (30%) are in the area of anaesthesia, an over-representation that requires further scrutiny. Indeed, when considering the retracted papers in anaesthesia, we quickly discovered evidence of systematic fraud[6] focused on the work of two authors, who between them have had 102 retracted articles.

Cole and colleagues analyse the association between errors and retraction in their case-control study. Of the study sample, five journals had multiple papers selected, which between them published 17 (34%) of the pairs of retracted and control papers. There was also a limited number of clinical areas covered by the study sample. Owing to this hierarchical nature of the authors' data, they failed to account for the clustering of papers within journals, which is curious when elsewhere they recognise that papers from the same journal cannot be considered independent.

So what should we make of the study's findings? The authors identify a relation between errors in clinical trial reports and retraction. Although this association may be real, it is unlikely to be helpful in practice if the main goal is to identify seriously misleading or fraudulent research that could harm patients. Identifying errors could increase our suspicion that a paper has a higher chance of being retracted, and that risk of retraction might be due to misleading or made up findings. But it could also be due to another perhaps less harmful problem such as duplication.

Mea culpa

A more fundamental question is, why were the many errors and inconsistencies identified by Cole and colleagues not picked up by journal editors or technical editors whose job it is to ensure the validity of the work they publish? Our experience is that many journals manage the publication process very well. It is surprising, therefore, that the so called "top scoring" retracted paper in Cole and colleagues' study (published by a BMJ group journal) contained 35 different errors. Almost all (n=34) the errors were numerical, which might have been identified by technical editors.

Schafer describes the considerable efforts that editors went to in uncovering the extent of research fraud in the work of authors in anaesthesia after errors were highlighted by readers,[6] and recognises the duty of journals to ensure high quality research. But he likened the growth of human knowledge to the weaving of a tapestry, and the sudden loss of retracted articles to the ripping of a thread. Many papers are published but fortunately few are fabricated or irrevocably flawed. We must expect both benign and malicious errors to occur in reports of research and devise ways to identify and correct errors of substance in the research publication process. However, Cole and colleagues' recommendation that we raise a red flag to clinical trial reports containing more than three errors is premature and not well supported by the evidence in their paper.

Cite this as: *BMJ* 2015;351:h4843

RESEARCH, p 14