EDITORIALS

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

thebmj.com Read BMJ Careers reporter Abi Rimmer's blog on her investigation into seven day working at bit.ly/1LeyAe6

Making sense of the evidence for the "weekend effect"

Sicker patients at the weekend, but even after adjustment for this their risk of death is higher

Paul Aylin professor of epidemiology and public health, Dr Foster Unit, School of Public Health, Imperial College London, London EC4Y 8EN, UK p.aylin@imperial.ac.uk

In their analysis of 2013-14 English hospital administrative data, Freemantle and colleagues again find an increased risk of death in patients admitted at a weekend compared with weekdays.[1] Their analysis combines both emergency and elective admissions and updates their earlier study.[2] They confirm a persistent "weekend effect" in England, which they claim is "not otherwise ignorable," and add to a substantial body of literature demonstrating this phenomenon both nationally[3-5] and internationally.[6-8]

The findings of such studies and the resultant media coverage are being used by politicians to bring about changes in working practices,[9] but against a strong background of criticism of these kinds of statistics.[10] At first glance, there is conflicting evidence about whether the weekend effect exists at all.[11 12] However, closer scrutiny shows that apparently "contradictory" studies tend to be smaller, carried out in single hospitals, and lack statistical power. Death after hospital admission, particularly for a planned surgical procedure, is relatively rare, and small studies simply don't have the numbers to be able to show an effect.

Risks and assumptions

Freemantle and colleagues find an increased risk of death of 10% for admissions on a Saturday and 15% for admissions on a Sunday compared with patients admitted on a Wednesday, which on the face of it, does not seem trivial. However, this represents a 10-15% increase on a relatively low risk of death overall (1.8%). The difference between relative risk and attributable risk is key. Researchers in this area have often focused on death as an outcome because it is important to patients and easily measured, but there is also evidence of a weekend increase in avertible errors in care leading to other complications.[13] Though the proportion of admissions ending in death is relatively low, there is an assumption that an increased risk of death may indicate a much greater burden of complications and error, causing unnecessary harm to patients.

An obvious criticism of some of these observational studies is that patients admitted at the weekend are simply sicker. Freemantle and colleagues do indeed find a higher proportion of sicker patients at the weekend but attempt to account for this by adjusting for case mix using a wide range of variables included in administrative data. They are not the first group to adjust for this,[3] yet the weekend effect seems to persist. No attempt to account for sicker patients in an analysis is perfect, but risk adjustment models based on administrative data have been shown to be as good, if not better, than models based on clinical data.[14] The possibility, however, of residual confounding can never be entirely ruled out.

In an attempt to get around this problem, one of my group's studies looked at planned surgical procedures and focused on the day of the week on which the procedure was performed. We found an increasing risk of death as the day of procedure approached the weekend, with 44% higher odds of death if an operation was carried out on a Friday (with the most critical postoperative period occurring over the weekend) compared with a Monday.[4] The case mix of patients varies much less in planned procedures on weekdays, and therefore confounding was less of a problem. In a follow-up paper, we also ruled out a lower level of experience in consultants carrying out procedures on a Friday as a contributory factor for this effect.[15]

Freemantle and colleagues look specifically at stroke as well as other conditions for which there is less reason to believe that patients would be sicker at the weekend. We have also examined emergency stroke admissions and, even after adjusting for case mix, found not only an increased risk of death for patients admitted at the weekend but also that patients were less likely to receive computed tomography on the day of admission, were less likely to receive thrombolysis, had higher rates of pneumonia, and were less likely to be discharged back to their usual place of residence.[16] This combination of both process and outcome measures strengthens the case for poorer quality of care at the weekend. If patients were sicker, we might expect higher rates of diagnostic tests and interventions.

Cause and effect

So if the "weekend effect" is real, what are the reasons behind poorer outcomes? Freemantle and colleagues point towards reduced services inside and outside the hospital as a possible explanation, but few studies have examined this important issue directly. Researchers in the US found a Friday effect in non-emergency major surgery but not in patients admitted directly to an intensive care unit, where there were more consistent levels of staffing throughout the week.[17] A recent UK study found that ward rounds by stroke specialists on seven days a week did not affect risk of stroke mortality but did find an association with the intensity of weekend staffing by registered nurses.[18] A recent, very large European study found a strong relation between overall hospital mortality and nurse staffing and education.[19]

Much evidence exists on the subject of weekend care, and there is a need for a comprehensive systematic review of findings. More research is needed to understand the complex relation between staffing levels and services and patient safety, using both process and outcome measures. With promised changes to how the NHS provides weekend and out of hours care, it will be an ideal opportunity to evaluate their impact on the "weekend effect."

Cite this as: *BMJ* 2015;351:h4652

FEATURE, p 14
ANALYSIS, p 16
OBSERVATIONS, p 22

Health professionals can model a humane response to the refugee crisis and speak out against the global inequity and injustice that is its root cause

Europe's refugee crisis: an urgent call for moral leadership

Offering asylum is a minimum standard of civilised society

Kamran Abbasi international editor, *The BMJ*
kabbasi@bmj.com
Kiran Patel medical director, NHS England (West Midlands)
Fiona Godlee editor in chief, *The BMJ*

Europe's refugee crisis is the greatest test of humanity faced by the world's rich countries this century. It isn't a new crisis. Nor was it difficult for politicians to anticipate. Refugees have fled to Europe since at least the premature optimism of the Arab Spring in 2011. Today, optimism is replaced by desperation, a promise of freedom overshadowed by death. Western nations rushed to support the democratic principles of the Arab Spring yet are reluctant to address the root causes and the consequences, which include civil war and state brutality, most notably in Syria. Oil rich Arab States have played their part by allowing political oppression and conflict to flourish in their region. A funding crisis in UN organisations is affecting the humanitarian effort in the Middle East, driving refugees to Europe in greater numbers.[1] Ignoring injustice and inequity in poorer countries and in areas of conflict has not prevented the consequences reaching the shores and borders of the rich world.

Over 300 000 refugees and migrants have crossed the Mediterranean this year, with an estimated 3000 dead or missing.[2] Tens of thousands are now entering over land. But as burdensome as the influx of refugees may seem to Europeans, the number of displaced people in Europe is a fraction of those in the Middle East, South Asia, and Africa.[3] Most refugees from Syria have been absorbed by its neighbours, Lebanon, Jordan, and Turkey. At the end of 2014, there were 19.5 million refugees worldwide, 86% of whom were in poorer countries, and 38.2 million people were internally displaced.[3] One in seven people across the world are now migrants, yet the international response to migrants is disappointing and unwelcoming.[4]

This evolving crisis is now presenting Europe's leaders and citizens with tough moral questions. The picture of a 3 year old child drowned on a Turkish beach has exposed the moral contradiction at the heart of Europe's approach, catalysing a change in public and political moods.[5]

Fragmented political response

While individuals are largely showing compassion and kindness, Europe's leaders continue to play to their own agendas. Germany's pragmatic form of compassion means processing asylum applications as quickly as possible, accommodating genuine refugees to help meet its labour needs but seeking to return economic migrants from the western Balkans.[6] Hungary's response panders to a domestic audience, focusing on the cultural dangers of accepting refugees while speeding their transit across Europe and out of Hungary.[7] The United Kingdom is populist, switching its tune in line with influential newspapers, from demonising refugees to sympathy.[8] New European member states, whose citizens were welcomed as migrants themselves after 1989, now reject immigrants.

Instead of fragmentation, we need a fair, humane, and unified approach. A meeting of European Union ministers of home affairs on 14 September to discuss the escalating refugee crisis offers Europe's major powers an opportunity to show strong commitment and moral leadership.

MICHAEL DEBETS/DEMOTIX/PA

Their welfare is our concern

Human rights are now, sometimes controversially, at the heart of modern Europe and might be its defining feature.[9] The right to health of all the world's inhabitants is enshrined in the Universal Declaration of Human Rights, as is the right to seek asylum from persecution.[10] Giving asylum to refugees, regardless of culture and religion, is a minimum standard expected of a civilised society. The World Health Organization and medical aid charities reinforce the importance of prioritising health services for refugees and asylum seekers.[11] In poorer countries, the health needs of refugees exhaust local healthcare resources to the detriment of both refugees and the host population.[12] The UN High Commission for Refugees has called for health services for refugees to be incorporated into health systems of host countries.[13]

The welfare of casualties of war and refugees should be of interest to every health professional. Health professionals can support by volunteering for humanitarian organisations[14]; lobbying politicians for action; treating patients regardless of race, religion, or refugee status; and taking a leadership role.

Medical journals have a part to play too by supporting health professionals in highlighting and addressing the social and political determinants of health in poorer countries and areas of conflict, as well as the plight of refugees. Health outcomes are improved more by tackling the social and political iniquities driving migration than by responding to health needs after migration. We must provide an adequate response to the refugees' immediate needs while we plan to tackle the underlying causes of their plight.[15]

Health professionals cannot save people from becoming refugees; that responsibility lies with politicians and civil society, but health professionals can model a humane response to the refugee crisis and speak out against the global inequity and injustice that is its root cause. Health professionals can remind society that seeking asylum is not a crime, that migrants are not necessarily a security risk, that refugees need and deserve our common humanity, and that migrants have not come to occupy Europe or hijack healthcare resources.

Without moral leadership from politicians, however, the death and suffering of vulnerable people will continue to shame us all.

Cite this as: *BMJ* 2015;351:h4833

thebmj.com

- Research News: Digoxin is linked to raised risk of death in patients with atrial fibrillation (*BMJ* 2015;350:h2387)
- Practice: Digoxin specific antibody fragments (Digibind) in digoxin toxicity (*BMJ* 2009;339:b2884)

Trials are best, ignore the rest: safety and efficacy of digoxin

Observational data might have misled doctors for years.

Graham D Cole research fellow
g.cole@imperial.ac.uk
Darrel P Francis professor of cardiology, International Centre for Circulatory Health, National Heart and Lung Institute, Imperial College London

Nearly 20 years after the digoxin in heart failure (DIG) trial showed that digoxin reduced hospital admissions for heart failure by 28% with no effect on mortality,[1] many clinicians still see this drug as a last resort—reflected in guidelines that reserve it only for those with severe or worsening heart failure when first and second line treatments have failed.[2] One reason might have been the numerous observational studies indicating the potentially worrying trade-off that digoxin is associated with an increased risk of death.

The linked article by Ziff and colleagues, a meta-analysis of digoxin treatment for heart failure or atrial fibrillation (or both), sheds light on the discordant findings between randomised and observational studies.[3] They show that the mechanism of the association with mortality in observational studies is that clinicians preferentially gave digoxin to sicker patients: when these patients died, there was therefore a true but misleading association between death and digoxin. The paper paints a devastating picture (see figure 3): the more vulnerable the observational design was to inadvertent bias, the stronger the observed association.

> The mechanism of the association with mortality in observational studies is that clinicians preferentially gave digoxin to sicker patients

Even more important is figure 2 in the paper. The unadjusted observational data give the most misleading message (risk ratio 1.76). Straightforward adjustment still gives a misleading message (ratios of 1.61 and 1.17). Even sophisticated propensity matching gives a misleading indication of harm (ratios of 1.18 and 1.07). If clinicians had only the observational data, they would likely form a strong and understandable opinion that digoxin increases mortality. Only the randomised dataset evaluating digoxin in patients with heart failure refutes this fallacy, decisively.

The penalty patients have paid for clinicians allowing observational data (pointing to increased mortality) to be stacked up against data from randomised controlled trials (pointing to reduced admissions and no increase in mortality) might have been underuse of an effective treatment. The 28% reduction in admissions for heart failure in the DIG trial could have provided worthwhile relief from the high burden of heart failure worldwide.[4]

Randomisation rules

Wrong again

Uncritical trust in observational associations in cardiology has, in recent years, been repeatedly exposed by randomised trials. For example, in myocardial infarction, randomised controlled trials overturned the observational logic of antiarrhythmics in patients with ventricular ectopy and showed a clear and consistent neutrality for intra-aortic balloon pump therapy where observational study results disagree with each other extensively (see table on thebmj.com). In each case, perusing the observational data was of no help in predicting the direction, let alone the magnitude, of the actual randomised findings

The demonstration by Ziff and colleagues of the uncorrectability of bias in observational studies of treatments is not a criticism of those studies. The problem is that experienced clinicians use a vast array of cues in their decision making, and many of these cues are difficult to describe, let alone quantify, so that they can't contribute to statistical adjustments in observational studies. For example, a patient might have one of many subtle and difficult to document signs of frailty that make the clinician hesitate to prescribe a particular drug. When that patient later dies, the statistics can adjust for documented risk factors (such as age), but this leaves all the undocumented cues used by the clinician lumped in with the (documented) treatment decision. Thus avoidance of the drug (and perhaps the use of an alternative) is correctly, but misleadingly, associated with mortality.

This is why when randomised data point in one direction, we should resist the temptation[9] to obscure the message by presenting observational data with an opposite association.

Curiously, guideline systems are still advising guideline writers[10] that observational datasets, if large enough, generate the same level of evidence (B) as a randomised trial. This unwisely encourages writers to balance results from the two sources and fosters a dangerous misconception that large study size is always good news for an observational dataset. In reality, large size increases the tendency for observational studies to identify false associations as statistically significant.[11]

A future for digoxin

Is it time to use digoxin more widely in heart failure? The DIG trial took place before widespread use of β blockers, cardiac resynchronisation therapy, and implantable defibrillators. It is conceivable that the benefits seen in the DIG trial are now being achieved by other means. The authors eloquently call for a randomised controlled trial (not more observational data) to answer this question.

An observational study cannot be mutated into a randomised trial by statistical adjustment or propensity matching: the randomisation is a crucial part of trial design. Specialisation matters. An airliner should always be flown by a pilot, if available, even if passengers are enthusiastic to contribute and happen to be more numerous. In the same way, we argue that clinical recommendations should always be guided by randomised trials if available, even if observational datasets are enthusiastically promoted and similarly more numerous.

Cite this as: *BMJ* 2015;351:h4662

RESEARCH, p 11

thebmj.com

- Analysis: When a test is too good: how CT pulmonary angiograms find pulmonary emboli that do not need to be found (*BMJ* 2013;347:f3368)
- Clinical Review: Diagnosis and management of pulmonary embolism (*BMJ* 2013;346:f757)
- Research: Safe exclusion of pulmonary embolism using the Wells rule and qualitative D-dimer testing in primary care (*BMJ* 2012;345:e6564)

Predicting pulmonary embolus in primary care

Decision rules supplement but don't supplant clinical judgment

Robert J Dachs vice-chairman, Department of Emergency Medicine, Ellis Hospital, Schenectady, NY, USA
dachsmd@aol.com

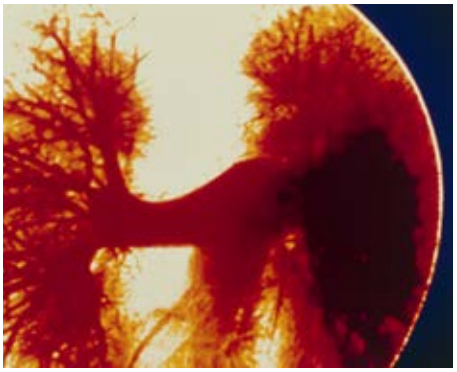
It is 4 30 pm, and your last patient of the day is a 42 year old woman in excellent health who awoke with pain located in the right infrascapular region that worsens with deep inspiration. She reports no dyspnea, cough, fever, or recent prolonged immobilization. Her vital signs and physical examination are normal. Your initial impression is that some type of musculoskeletal condition is causing her pain. Then you begin to wonder. Could this be a pulmonary embolus? Your gut says "no," but your brain continues to dwell on this possibility. You contemplate your dilemma and consider two competing options: stick with your impression and treat the patient for presumed muscle pain or send her for an advanced imaging study to rule out a pulmonary embolus.

In a linked paper, Hendrickson and colleagues attempt to provide a third option for the physician in a primary care office setting.[1] Their goal was to find a strategy using clinical decision rules that have been derived and validated in emergency department settings and that could be transported to the primary care setting and allow the safe discharge of a low risk patient while avoiding unnecessary radiology studies.

Famous five

To reach this goal, the authors did a systematic literature search that identified five validated clinical decision rules for pulmonary embolus: the Wells criteria, Geneva score, Charlotte rule, pulmonary embolus rule-out criteria (PERC), and Pisa rules. They determined that only the Wells and Geneva rules contained variables that they could retrospectively apply to the 598 primary care patients in the previously published AMUSE trial.[2] They concluded that these two scoring systems, when combined with a point of care D-dimer test, would be useful in the case scenario described above, with a nod towards the simplified Wells score being slightly "safer" than the simplified Geneva model.

How nice to have a third management option for patients whom you judge to be at low risk, but not at no risk, for a pulmonary embolus.



BSIP VEM/SPL

Unpredictable

But before incorporating this option into their clinical armamentarium, practicing clinicians must ask if Hendrickson and colleagues' results are generalizable to their own practice setting. Doing so quickly identifies an important limitation of this study: the specific D-dimer test that was used (Simplify D-dimer; Clearview, Inverness Medical, Bedford, UK) is not universally available. A local hospital or reference laboratory may use a qualitative D-dimer test such as an enzyme linked immunosorbent assay (ELISA) with higher sensitivity but lower specificity than bedside point of care qualitative tests.[3] Although a test with higher sensitivity (fewer "misses" or false negatives) may be appealing, such super-sensitive tests return more false positive results, exposing more patients to needless computed tomography scanning.[4]

Assuming that this particular bedside point of care D-dimer is available to the office practitioner, the next question is: "Do I agree with the authors' conclusions?" The authors conclude that the failure rate of their strategy is acceptable. However, the "failure rate" is not the same as the "sensitivity" of the test and may falsely reassure the clinician about the safety of this strategy.

In the study, an original Wells score of 4 or less with a negative point of care D-dimer test missed four out of 73 patients with a pulmonary embolus, whereas the simplified Wells score missed three of these same patients. The resulting sensitivity for each was approximately 95%. Dismissing at one's peril the associated 95% confidence intervals (which indicate possible sensitivities as low as 87-88%), can the practicing office clinician accept a strategy that misses one in 20 patients harboring a pulmonary embolus? I suspect that many clinicians, particularly those who are risk averse, would judge this sensitivity (the ability of this strategy to "rule out" a pulmonary embolus) to be inadequate and would likely avoid the proposed strategy.

Trust your judgement

Therefore, the search for an easily applicable clinical decision rule that helps the primary care doctor's clinical decision making in patients at low risk will continue. One such rule may already be available. The PERC rule had a sensitivity of 97.4% without any D-dimer test when studied in a large, diverse emergency department population.[5] The eight objective criteria (box, thebmj.com) are easily accessible in most office practices. Interestingly, if the PERC rule had been used in Hendrickson and colleagues' study, the three false negative cases would have been identified correctly. I suspect that if (and when) the PERC rule is tested in primary care, it will prove as useful for primary care doctors as it has for doctors in the emergency department.

For now, most primary care doctors will need to continue to rely on their clinical judgment or "gestalt" when managing low risk patients with suspected pulmonary embolus. When a clinical decision rule ultimately shows trustworthiness in the office setting, we still need to listen to our clinical judgment about a patient. Why? Because after years of knowledge acquisition during medical school, residency, and practice, our brains have incorporated many more than the five, six, or eight objective variables in any of these rules.

Clinical decision rules, like any diagnostic test, can only supplement, but never supplant, a clinician's judgment. Judgment or gestalt has recently been shown to trump both the Wells criteria and revised Geneva score in a head to head study.[6] Our years of training and experience have provided us with the skills and intuition to manage these cases. We should be comforted by this concept and trust our clinical acumen.

Cite this as: *BMJ* 2015;351:h4594

RESEARCH, p 13