Is quality of care improving in the UK?

Yes, but we do not know why

RESEARCH, pp 369, 370

Peter J Pronovost professor
Sean M Berenholtz associate professor, Department of Anesthesiology and Critical Care Medicine, Johns Hopkins University School of Medicine, Baltimore, MD, 21231, USA
Laura L Morlock professor and deputy chair, Department of Health Policy and Management, Johns Hopkins Bloomberg School of Public Health

Competing interests: The authors have completed the Unified Competing Interest form at www.icmje.org/coi_disclosure.pdf (available on request from the corresponding author) and declare: no support from any organisation for the submitted work; PJP and SMB have received grants or had contracts with the Agency for Healthcare Research and Quality, the World Health Organization, the Michigan Health and Hospital Association Keystone Center for Patient Safety and Quality, NIH/NHLBI, the Centers for Disease Control and Prevention, JHPIEGO, and the Robert Wood Johnson Foundation in the past three years; PJP has been a paid consultant for the Association for Professionals in Infection Control and Epidemiology, received speaking honorariums and travel expenses from various hospitals and the Leigh Speaker's Bureau, and received royalties from his book *Safe Patients, Smart Hospitals* in the previous three years; no other relationships or activities that could appear to have influenced the submitted work.

Provenance and peer review: Commissioned; not externally peer reviewed.

Cite this as: *BMJ* 2011;342:c6646
doi: 10.1136/bmj.c6646

The need to improve the quality of care is well recognised. Yet accomplishing this is complicated, messy, and uncertain, requiring that researchers tackle technical (science) and adaptive (emotional, social, cultural, and political) challenges.[1] Tension exists between those who say "just do something" to improve quality and those who say "science should be the guide."[2]

The two linked studies suggest that more science is needed. Benning and colleagues evaluated a large patient safety programme (the Safer Patients Initiative; SPI) in the United Kingdom, led by the Institute for Healthcare Improvement.[3 4] The Health Foundation initiated and supported the initiative and, laudably, an independent evaluation, grounded in theory and conducted by experts in epidemiology, biostatistics, medical sociology, health services research, and clinical medicine. They performed a quantitative and qualitative evaluation at organisational and patient levels. The evaluation included five substudies that looked at whether the interventions worked and why. In addition to using a rigorous research design, the authors conducted a state of the art analysis, which included using different approaches to evaluate changes over time in treatment and comparison hospitals. This evaluation will serve as a model for the field. It required, however, an interdisciplinary team of experts and appropriate research funding, both of which are rare.

The study's findings are partly encouraging and partly worrying. On the encouraging side, the study provides convincing evidence that safety and quality improved in NHS hospitals in England over the study period (about 18 months). This should provide comfort to UK citizens, the NHS, and parliament. Patients are less likely to be harmed from the care they receive. The NHS should try to understand why these improvements occurred and how they can be strengthened and replicated broadly across the UK.

For those who hoped SPI would transform care, the findings are disconcerting. The authors found that the initiative had no discernible additional effect on patient safety; care improved to the same extent in both treatment and comparison hospitals, highlighting the need for robust evaluation with concurrent controls. It is, of course, difficult to measure the impact of patient safety interventions, especially diffuse interventions like SPI. The initiative might possibly have provided benefits that were not measured or may emerge over time. It is also difficult in these types of large scale evaluations to find an appropriate comparison group. In areas where intervention hospitals were performing well at baseline, it would be difficult to show improvement.

The study should be a wakeup call to those implementing patient safety programmes. Too many patients in the UK, and the rest of the world, continue to experience preventable harm. The quality improvement field needs to embrace science, favour evidence over anecdote, and move beyond using only one generic framework for improvement (the plan, do, study, act cycle).[5] Different types of patient safety challenges exist, such as translating evidence into practice, improving teamwork and organisational culture, identifying and mitigating hazards, and reducing diagnostic errors. Each type of problem should be informed by specific theories, methods, and measures.[6]

Although well intentioned, it is not surprising that the SPI had less of an impact than the investigators anticipated. It was not grounded in a theory of organisational change.[7] It asked hospitals to implement 43 interventions, when most hospitals would find it difficult to implement three. Clinicians thought that many of the interventions were supported by weak evidence and that some measures were not valid. The initiative was largely top down, with limited input from local clinicians. Moreover, it did not target areas where teams performed poorly (in many of the areas, teams were performing nearly flawlessly before the initiative). The interventions and measures were not sufficiently pilot tested, and quality control over the quality improvement data collected by the local teams was virtually non-existent.

Clinicians who push back against patient safety interventions are often viewed as "knaves."[8] This study suggests that some of that resistance may be warranted. Some interventions focused on areas that were not problematic and used evidence and measures that doctors did not always perceive as valid, potentially souring clinicians' attitudes towards efforts to improve patient safety. We need clinicians to lead patient safety efforts. For this to happen, they must believe that interventions and measures are based on science and that their patients will benefit.

Yet when interventions deal with both technical and adaptive challenges, broad scale improvement in patient safety is possible. Several patient safety interventions have shown significant improvements in patient outcomes by having clinicians and researchers collaborate when developing and pilot testing the programme. Such innovations include centralised collection of performance measures and evidence summaries by the researchers, and local innovation of programme implementation by the clinicians.[9-12]

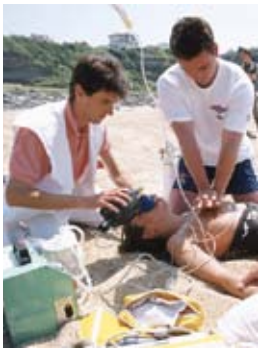
These studies provide three important lessons. Firstly, patient safety studies require robust design and evaluation.[13] Funding agencies need to support the development and implementation of patient safety programmes that include rigorous evaluation. These programmes should be grounded in change theory, include evidence based interventions, valid measures, and data quality control. Although theory and interventions evolve over time, patient safety programmes should be

developed in collaboration with clinicians and be pilot tested, and measures should be validated before broad implementation. Secondly, care in the UK is improving; we should understand how and why. Thirdly, quality improvement efforts must improve, embracing rather than running from science. The science for quality improvement differs from basic and clinical research. It requires input from clinicians, health service researchers, social scientists, and human factors and systems engineers; in addition, it uses change theory, mixed methods, and robust evaluation. These studies provide a model.

1 Heifetz RA. Leadership without easy answers. Belknap Press at Harvard Press, 1994.
2 Auerbach AD, Landefeld CS, Shojania KG. The tension between needing to improve care and knowing how to do it. *N Engl J Med* 2007;357:608-613.
3 Benning A, Dixon-Woods M, Ghaleb M, Suokas A, Dawson J, Barber N, et al. Large scale organisational intervention to improve patient safety in four UK hospitals: mixed method evaluation. *BMJ* 2011;342:d195. check
4 Benning A, Dixon-Woods M, Nwulu U, Ghaleb M, Dawson J, Barber N, et al. Multiple component patient safety intervention in English hospitals: controlled evaluation of second phase. *BMJ* 2011;342:d199.
5 Berwick DM. A primer on leading the improvement of systems. *BMJ* 1996;312:619-22.
6 Pronovost PJ, Goeschel CA, Marsteller JA, Sexton JB, Pham JC, Berenholtz SM. A framework for patient safety research and improvement. *Circulation* 2009;119:330-7.
7 Weiner BJ. A theory of organizational readiness for change. *Implement Sci* 2009;4:67.
8 Jain SH, Cassel CK. Societal perceptions of physicians: knights, knaves, or pawns? *JAMA* 2010;304:1009-10.
9 Pronovost PJ, Goeschel CA, Colantuoni E, Watson S, Lubomski LH, Berenholtz SM, et al. Sustaining reductions in catheter related bloodstream infections in Michigan intensive care units: Observational study. *BMJ* 2010;340:c309.
10 Miller MR, Griswold M, Harris JM, Yenokyan G, Huskins WC, Moss M, et al. Decreasing PICU catheter-associated bloodstream infections: NACHRI's quality transformation efforts. *Pediatrics* 2010;125:206-13.
11 Pronovost PJ, Freischlag JA. Improving teamwork to reduce surgical mortality. *JAMA* 2010;304:1721-2.
12 Neily J, Mills PD, Young-Xu Y, Carney BT, West P, Berger DH, et al. Association between implementation of a medical team training program and surgical mortality. *JAMA* 2010 (forthcoming).
13 Berenholtz SM, Needham DM, Lubomski LH, Goeschel CA, Pronovost PJ. Improving the quality of quality improvement projects. *Joint Commission Journal on Quality and Patient Safety* 2010;36:468-73.

Audiovisual feedback and quality of CPR

Evidence so far shows no improvement in clinical outcomes



BSIP, THEOBALD/SPL

RESEARCH, p 371

Peter Leman associate professor, Emergency Department, Royal Perth Hospital, Perth, WA, 6011, Australia
peter.leman@health.wa.gov.au

Competing interests: The author has completed the Unified Competing Interest form at www.icmje.org/coi_disclosure.pdf (available on request from the corresponding author) and declares: no support from any organisation for the submitted work; no financial relationships with any organisations that might have an interest in the submitted work in the previous three years; PL is the Australasian College for Emergency Medicine representative on the Australian Resuscitation Council, and as such is involved in the development of national resuscitation guidelines, including those for CPR.

Provenance and peer review: Commissioned; not externally peer reviewed.

Cite this as: *BMJ* 2011;342:c7108
doi: 10.1136/bmj.c7108

Effective cardiopulmonary resuscitation (CPR) is essential to maximise the chance of a successful outcome after cardiac arrest. In the linked cluster randomised trial, Hostler and colleagues assess the effect of real-time audio and visual feedback in people with an out of hospital cardiac arrest on whom resuscitation was attempted by emergency medical services.[1]

In October 2010, the International Liaison Committee on Resuscitation (ILCOR) published international consensus recommendations on CPR and emergency cardiovascular care.[2] This review concluded that a strong emphasis on delivering high quality chest compressions was essential. Specific recommendations included placing the hands in the centre of the chest; giving at least 100 compressions a minute; and pressing to a depth of at least 5 cm. Unfortunately, although the recommendations used the term "at least," no clear evidence exists that pushing faster than 100 compressions per minute or deeper than 5 cm is associated with better or worse outcomes.

The European Resuscitation Council (ERC) published its updated guidelines at the same time and stressed the importance of feedback to rescuers in maintaining high quality CPR.[3] The council acknowledged that the use of feedback devices "may be beneficial" but did not make it clear that evidence to support their routine use in any particular setting was lacking. The ILCOR review of feedback devices found no convincing evidence of improved long term outcomes with their use.[4] Feedback devices varied and included intra-arterial pressure lines, CPR sensing devices built into a standard monitor-defibrillator, transthoracic impedance devices, and simple external audio prompts.

ILCOR reviewed two large out of hospital cardiac arrest studies of feedback devices. The first, an observational before and after study, compared automated verbal feedback with no feedback; it showed an increase in 4 mm (95% confidence interval 2 to 6; P<0.001) of chest compression depth and a reduction in compression rate from 121 to 109 (P=0.001) with feedback switched on.[5] The number of people discharged from hospital alive did not differ significantly (odds ratio 1.5, 95% confidence interval 0.8 to 3.0), however. The second study used performance quality data feedback to teams after the introduction of a verbal feedback defibrillator for use during CPR.[6] It reported no significant improvement in compression depth or rate, or in the ratio of time performing compressions to no compressions.

Additional reviews of the quality of CPR that included smaller studies as well as studies on manikins drew similar conclusions. They also highlighted that more than feedback at the time of the CPR may be needed—for example, a system of ongoing debriefing, simulation, and training using recorded data might be useful.[7-10]

Hostler and colleagues report a multicentre cluster randomised controlled trial evaluating the effectiveness of real-time audiovisual feedback on the performance and outcome of CPR after out of hospital cardiac arrest in more than 1500 patients.[1] They found no significant benefit to the clinically relevant outcomes of return of spontaneous circulation or survival to hospital discharge. They measured how CPR was performed and found that feedback significantly reduced the compression rate (103.1/min *v* 108.0/min; P<0.001) towards the rates suggested by current guidelines. However, no convincing evidence exists that any rate between 100 and 120 is better than any other. Similarly, although the mean depth of compressions increased slightly (39.6 *v* 37.8 mm; P<0.001), it did not approach the currently recommended depth of at least 50 cm.

So why, in such a large study when primary evidence suggests that feedback in addition to ongoing debriefing and training should improve adherence to international CPR guidelines, was no clinical benefit seen? The study design may have contributed to the lack of an observed benefit. Hostler and colleagues mention the Hawthorne effect, whereby the prehospital personnel providing the CPR were aware that they were being assessed, so their performance may have improved independently of the intervention.

In addition, multiple crossovers occurred between clusters

Discuss community cardiopulmonary resuscitation on doc2doc, BMJ Group's global online clinical community, at http://bit.ly/eZYTfB

during the study period, and although it is unclear how this related to individual providers who used the same defibrillator and recording device in both arms of the study, some confounding is likely. It was interesting that in the early cluster rotations of the study, more patients arrived at hospital after an initial asystolic arrest with a return of spontaneous circulation in the feedback-off arm than in the feedback-on arm. This implies that the quality of CPR was perhaps better with no verbal feedback. This raises the question of how much the individual providers tried to raise the quality of their CPR in the "off" arm, knowing that the event was being recorded. So, in a well trained highly motivated system, feedback may be superfluous. The field rate of return of spontaneous circulation of more than 40% (the historical background rate was 20%) implies that either patient selection was biased or prehospital care was extremely good.

The results of this study suggest that audiovisual feedback during CPR may not be necessary in the field for well trained CPR providers in robust systems, although its role in the training environment is better supported by the evidence.[9 10] It can also reasonably be concluded that good CPR occurs over a narrow range of compression rates and depths, rather than at an absolute single value only.

References are in the version on bmj.com.

Measurement of blood pressure in primary care

Must be done carefully, or not at all

JIM VARNEY/SPL

RESEARCH, p 372

Jonathan Mant professor of primary care research, General Practice and Primary Care Research Unit, University of Cambridge, Cambridge CB2 OSR, UK
Richard McManus professor of primary care cardiovascular research, Primary Care Clinical Sciences, University of Birmingham, Birmingham, UK

Competing interests: All authors have completed the Unified Competing Interest form at www.icmje.org/coi_disclosure.pdf (available on request from the corresponding author) and declare: no support from any organisation for the submitted work; no financial relationships with any organisations that might have an interest in the submitted work in the previous three years; JM and RM have used automated blood pressure monitors in their research but have no financial or other link with the manufacturers.
Provenance and peer review: Commissioned; not externally peer reviewed.

Cite this as: *BMJ* 2011;342:d382
doi: 10.1136/bmj.d382

In the linked cluster randomised controlled trial, Myers and colleagues compare the quality and accuracy of blood pressure measurement of manual sphygmomanometers with automated blood pressure monitors in primary care.[1]

The value of measuring blood pressure in the clinic is debatable, especially when it is done by doctors.[2] Inherent variability means that true changes in blood pressure can be difficult to detect using clinic readings.[3] Furthermore, misclassification on the basis of clinic readings can occur in people with "normal" underlying blood pressure who are labelled as hypertensive because of the pressor effects of having blood pressure measured in the clinic (white coat hypertension) and those with hypertension who are labelled as normal because clinic readings are lower than the usual blood pressure (masked hypertension).[4] Out of clinic methods of measuring blood pressure, such as ambulatory blood pressure monitoring and home monitoring, are better predictors of cardiovascular outcome than clinic measurements.[5 6] Self monitoring and self management can lead to better control of blood pressure than monitoring blood pressure in the clinic.[7 8] However, the evidence base for the treatment of hypertension is based on clinic measurements, and it can be problematic to translate ambulatory and home readings to their clinic equivalent.[9 10] Therefore, methods to improve the accuracy of the measurement of blood pressure in the clinic remain relevant.

An evaluation of one such method is reported in Myers and colleagues' trial, which randomised practices to continue to use manual sphygmomanometers or to use an automated blood pressure monitor (BpTRU device).[1] This monitor takes six readings at set intervals (in this study, two minutes), and the first reading with the clinician present is ignored. Subsequent readings are performed automatically and do not require the clinician to be present: in this study, it was emphasised that the patient should be left alone. After enrolment, all participants had their blood pressure measured using ambulatory blood pressure monitoring and attended the primary care clinic where their blood pressure was measured using either the automated monitor or a manual device. Clinic systolic blood pressures were on average 5.4 mm Hg lower when measured using automated sphygmomanometers compared with manual ones. The readings from the BpTRU device were more similar to and correlated more closely with daytime ambulatory blood pressure monitoring than did the manual readings. This suggests that use of automated sphygmomanometers can improve the accuracy of blood pressure measurements performed in the clinic, and in particular, reduce the white coat effect.

Control practices received no training on how to measure blood pressure, so a better technique in the use of manual sphygmomanometers might have achieved similar results. Manual readings obtained by research nurses and research staff following a specific protocol correlate much better with ambulatory blood pressure monitoring than do doctors' readings.[10]

The randomised controlled design is a particular strength of this study, in that it enables an unbiased assessment of the potential difference between blood pressure measurement in routine clinical practice and under controlled conditions, in this case through the use of the automated blood pressure monitors. In this study, routine clinical practice overestimated systolic blood pressure by on average 5.4 mm Hg, and this could lead to overtreatment.

What are the clinical implications of the study? Firstly, it is a reminder of the importance of measuring blood pressure properly. Secondly, it provides empirical evidence that automated sphygmomanometers can improve the accuracy of blood pressure measurement in primary care.

Blood pressure is measured for two main reasons—to diagnose hypertension and monitor blood pressure in people with a label of hypertension. Given that a diagnosis of hypertension may lead to lifelong drug treatment, blood pressure must be measured accurately when making such a diagnosis. While this study highlights improvements in monitoring that can be made through using automated sphygmomanometers, it is not yet clear what the role of such devices is compared with ambulatory blood pressure monitoring and home monitoring. The study population was restricted to people with a label of hypertension, so those with masked hypertension will have been largely excluded.

Furthermore, although automated measurements reduced the possibility of a white coat effect, they did not eliminate it—some patients had substantially higher automated office readings than daytime ambulatory readings. In terms of monitoring blood pressure, automated sphygmomanometers will reduce the "noise" associated with operator error but will not eliminate it. For example, poor arm positioning and using the

wrong size of cuff may still occur and have significant effects on recorded blood pressure.[11]

Several questions remain. How long should the interval be between automated readings, and how many readings should be taken? The six readings taken two minutes apart as used in the study will be practically difficult to implement in many primary care settings. Are automated readings the equivalent of research measurements, or do adjustments need to be made to blood pressure target and treatment thresholds in the same way as for ambulatory measurements?[10]

Automated blood pressure measurement provides an alternative to manual and out of clinic monitoring. Its precise role depends on further evidence, some of which may come from the trial from which these data are drawn. In the meantime, doctors should remember that blood pressure thresholds and targets are based on measurements performed under controlled research conditions, and that poor technique can lead to substantial error.[2 11]

1 Myers MG, Godwin M, Dawes M, Kiss A, Tobe SW, Grant FC, et al. Conventional versus automated measurement of blood pressure in primary care patients with systolic hypertension: randomised parallel design controlled trial. *BMJ* 2011;342:d286.
2 Graves JW, Sheps SG. Does evidence based medicine suggest that physicians should not be measuring blood pressure in the hypertensive patient? *Am J Hypertens* 2004;17:354-60.
3 Keenan K, Hayen A, Neal BC, Irwig L. Long term monitoring in patients receiving treatment to lower blood pressure: analysis of data from placebo controlled randomised controlled trial. *BMJ* 2009;338:b1492.
4 Pickering TG, Gerin W, Schwartz JE, Spruill TM, Davidson KW. Franz Volhard lecture: should doctors still measure blood pressure? The missing patients with masked hypertension. *J Hypertens* 2008;26:2259-67.
5 Verdecchia P. Prognostic value of ambulatory blood pressure: current evidence and clinical implications. *Hypertension* 2000;35:844-51.
6 Bobrie G, Chatellier G, Genes N, Clerson P, Vaur L, Vaisse B, et al. Cardiovascular prognosis of "masked hypertension" detected by blood pressure self measurement in elderly treated hypertensive patients. *JAMA* 2004;291:1342-9.
7 Bray EP, Holder R, Mant J, McManus RJ. Does self-monitoring reduce blood pressure? Meta-analysis with meta-regression of randomized controlled trials. *Ann Med* 2010;42:371-86.
8 McManus RJ, Mant J, Bray EP, Holder R, Jones MI, Greenfield S, et al. Telemonitoring and self-management in the control of hypertension (TASMINH2): a randomised controlled trial. *Lancet* 2010;376:163-72.
9 Verberk WJ, Kroon AA, Kessels AG, de Leeuw PW. Home blood pressure measurement: a systematic review. *J Am Coll Cardiol* 2005;46:743-51.
10 Head GA, Mihailidou AS, Duggan KA, Beilin LJ, Berry N, Brown MA, et al. Definition of ambulatory blood pressure targets for diagnosis and treatment of hypertension in relation to clinic blood pressure: prospective cohort study. *BMJ* 2010;340:c1104.
11 McAlister FA, Straus SE. Measurement of blood pressure: an evidence based review; *BMJ* 2001;322:908-11.

Postmarketing studies of drug safety

A European initiative could help bring more transparency and rigour to pharmacoepidemiology

Sebastian Schneeweiss associate professor
schneeweiss@post.harvard.edu
Jerry Avorn professor, Division of Pharmacoepidemiology and Pharmacoeconomics, Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, MA 02120, USA

Competing interests: None declared.
Provenance and peer review: Commissioned; not externally peer reviewed.
Cite this as: *BMJ* 2011;342:d342
doi: 10.1136/bmj.d342

In the early days of randomised clinical trials, their results could be manipulated in several ways—protocols could be altered in light of early findings, sponsors could exert undue influence over what could be published, and some "unfavourable" results could be suppressed entirely. In the United States, the creation of the government clinical trials website (www.clinicaltrials.gov) greatly contributed to minimising these threats to honest science.[1] But requiring similar consistency, rigour, and transparency has been more difficult with observational studies, because any person or company with modest resources can purchase a large database of health insurance claims and perform a variety of epidemiological analyses with little or no accountability for the transparency, rigour, or visibility of such work.

In 2006, the European Medicines Agency took on this problem by creating the European Network of Centres for Pharmacoepidemiology and Pharmacovigilance (ENCePP) to provide registration, standardisation, and quality assurance for observational studies of the effects of drugs (www.encepp.eu/). To qualify for the "ENCePP seal," study organisers must agree to a code of conduct and transparency, meet a checklist of methodological standards, and agree to publicly post the study protocol as well as its results.[2]

"Best practices" for the conduct of epidemiological studies of the safety of drugs are less well standardised than those developed over the decades for clinical trials. Because of the multitude of designs and analytical choices available for observational studies, each with its own strengths and limitations, the validity of pharmacoepidemiological research findings is harder to appraise. People responsible for decisions about drug safety, who must often make difficult policy calls based on scanty non-randomised data, have long wished for a user-friendly way of determining the validity of a given study. The creation of the ENCePP process will do a great deal to make study designs more transparent and might provide some useful methodological guidance in the process.

Although the ENCePP approach will promote transparency in non-randomised drug safety studies, it cannot (and does not) claim to guarantee their quality. A methods checklist can remind researchers of areas that should be dealt with in any pharmacoepidemiological study protocol—an important development in the evolution of this still young field. But it does not ensure guidance or best practice beyond that. Other guidelines, including good pharmacoepidemiological practice,[3] purposely list only higher level principles of good science, but they stop short of making specific recommendations about methodological choices.

As with the registration of randomised trials, the idea of registering an observational study protocol is that predefined endpoint analyses are likely to be more valid and less easy to manipulate. But some have questioned this approach.[4] Is the association between a drug and a predefined end point (such as a given adverse event) necessarily more important or valid than a stronger and more precise effect estimate for an important end point (such as another unexpected adverse event) that was discovered during the analysis of the same study? Confining all analyses to a predefined protocol without allowing for additional considerations may reduce the potential for in depth discovery, because important new insights may emerge only during data analysis.[5] A more important advantage of registration may be that researchers are likely to work harder on a study design that they know will become public. In the end it is the quality of individual studies—that is, the control of biases—that most needs to be improved.

When biases are ruled out in a credible way, by the original or subsequent studies, epidemiological data can lead to sensible regulatory decisions, irrespective of what was predefined in a protocol.[6 7] Although the ENCePP seal will confer a degree of legitimacy regarding transparency and adherence to basic standards, studies that do not go through the ENCePP approval process should not be seen as necessarily yielding lower grade evidence.

The first ENCePP listed study will tackle the effectiveness and safety of long acting β agonists in patients with chronic obstructive pulmonary disease.[8] As required, the authors have posted their study design and pledged that they will adhere to ENCePP's code of conduct and research standards. So far, so good. But a review of the posted protocol raises several methodological concerns. Firstly, the study will combine prevalent and incident users of inhaled long acting β agonists. But ongoing users of these drugs are likely to represent "survivors" who have done well on that drug (and not had to stop taking it because of side effects).[9] Furthermore, unlike in clinical trials where participants' characteristics are assessed before treatment, the characteristics of these people will be assessed during treatment and therefore be subject to the effects of the study drugs being used.[10] In addition, the total duration of drug use may not be fully known from the proposed claims database, so that shorter term users might be compared with longer term users.[11]

Secondly, one comparison group will be "non-users" of any study drug. Yet untreated patients with chronic obstructive pulmonary disease are likely to have less severe disease than those who are treated.[12] If the investigators cannot determine the time of onset of the disease or of drug use, it may be difficult to be sure that the underlying differences between treatment and non-user comparison groups can be adequately adjusted for in the data analysis, because it can be difficult to measure the severity of chronic obstructive pulmonary disease with claims data.

An important advantage of the ENCePP process is that study designs will now be publicly available, providing an opportunity for feedback by peer researchers and revisions by the investigators, which will ultimately lead to better science. ENCePP deserves much credit for its successful implementation of this process to ensure improved transparency of study methods. But the quality of each study will still need to be judged on the basis of a detailed description of the final design and analytical approach, which ENCePP now makes possible.

An important next step for ENCePP could be to make study data, from source data to the final analytical datasets, available to peer reviewers. This would make it possible for findings to be reconsidered using different analytical strategies. Doing so would also encourage the original investigators to work even more thoughtfully, knowing that their analyses may be replicated and reassessed by their peers.

1 Zarin DA, Ide NC, Tse T, Harlan WR, West JC, Lindberg DA. Issues in the registration of clinical trials. *JAMA* 2007;297:2112-20.
2 European Network of Centres for Pharmacoepidemiology and Pharmacovigilance. ENCePP studies. 2010. www.encepp.eu/encepp_studies/index.html.
3 Andrews EA, Avorn J, Bortnichak EA, Chen R, Dai WS, Dieck GS, et al. Guidelines for good epidemiology practices for drug, device, and vaccine research in the United States. *Pharmacoepidemiol Drug Saf* 1996;5:333-8.
4 Vandenbroucke JP. Preregistration of epidemiologic studies: an ill-founded mix of ideas. *Epidemiology* 2010;21:619-20.
5 Vandenbroucke JP. Observational research, randomised trials, and two views of medical science. *PLoS Med* 2008;5:e67.
6 Schneeweiss S, Seeger JD, Landon J, Walker AM. Aprotinin during CABG surgery and risk of death. *N Engl J Med* 2008;358:840-2.
7 Wang PS, Schneeweiss S, Avorn J, Fischer MF, Mogun H, Solomon DH, et al. Risk of death in elderly users of conventional vs atypical antipsychotic medications. *N Engl J Med* 2005;353:2335-41.
8 Agabiti N. Long-term outcomes and adverse events of therapy with inhaled corticosteroids, long-acting beta-2-agonists and anticholinergic drugs in hospitalised patients with chronic obstructive pulmonary disease (COPD)—a cohort study based on health information systems in three Italian regions. 2010. www.encepp.eu/encepp/viewResource.htm?id=1669.
9 Moride Y, Abenhaim L. Evidence of the depletion of susceptibles effect in non-experimental pharmacoepidemiologic research. *J Clin Epidemiol* 1994;47:731-7.
10 Schneeweiss S. A basic study design for expedited safety signal evaluation based on electronic healthcare data. *Pharmacoepidemiol Drug Saf* 2010;19:858-68.
11 Ray WA: Evaluating medication effects outside of clinical trials: new-user designs. *Am J Epidemiol* 2003;158:915-20.
12 Glynn RJ, Schneeweiss S, Wang PS, Levin R, Avorn J. Selective prescribing led to overestimation of the benefits of lipid-lowering drugs. *J Clin Epidemiol* 2006;59:819-28

Use of skin lightening creams

Lack of recognition and regulation is having serious medical consequences

Yetunde M Olumide professor of medicine (retired), University of Lagos, Lagos, Nigeria
mercyolumide2004@yahoo.co.uk

Competing interests: None declared.

Provenance and peer review: Commissioned; not externally peer reviewed.

Cite this as: *BMJ* 2010;341:c6102
doi: 10.1136/bmj.c6102

Skin lightening (bleaching) cosmetics and toiletries are used to lighten the colour of darker skin. The practice, which is fuelled by racial prejudice, stems from the misconceptions that black skin is inferior and that someone with a fair skin is more attractive.

By definition, cosmetics are meant to improve the appearance of the skin or enhance the attractiveness of users, not to alter the basic structure of the skin. Skin lightening creams alter the chemical structure of the skin by inhibiting the synthesis of melanin and should therefore be regulated as drugs not cosmetics.

The active ingredients include hydroquinone, mercury, and highly potent fluorinated corticosteroid ointments and creams such as fluocinonide, betamethasone valerate, and clobetasol propionate. The list of ingredients has expanded because some manufacturers have introduced new chemicals of unknown safety—such as niacinamide, oxybenzone, and triethanolamine—to circumvent the efforts of government regulatory agencies that prohibited the use of the above chemicals in cosmetics and toiletries.[1] Some products do not have ingredient labelling or place of manufacture,[2 3] and inadequate regulation has provided users with easy access to cheap, substandard, and misbranded toxic products.

These products are associated with serious and life threatening complications because they are used for long periods on a large body surface area and often under hot and humid tropical conditions, which promote percutaneous absorption. Complications such as exogenous ochronosis and colloid milium were initially reported in people with coloured skin in North America and South Africa in 1975.[4] Exogenous ochronosis is related to the use of hydroquinone, and it presents as a dirty greyish brown waxy pigmentation on sun exposed

SEYLLOU/AFP/GETTY IMAGES

areas of the skin. The primary lesion of colloid milium is a translucent flesh or cream coloured papule of 1 mm to 5 mm diameter on sun exposed areas. Pigmented forms of colloid milium are associated with hydroquinone induced exogenous ochronosis. Lesions often coalesce into nodules or plaques, and may rarely be verrucous.

A recent literature review of the adverse effects of skin lightening cosmetics found that the complications have become more complex and serious because of the nature of the additives and the methods of application.[3] Exogenous ochronosis is a major complication. Topically applied phenolic intermediates such as hydroquinone may produce exogenous ochronosis. Hydroquinone specifically inhibits the enzyme homogentisic acid oxidase locally, resulting in accumulation of this substance on the collagen fibres at the site of application. Exogenous ochronosis shows identical histological changes on skin biopsy to those seen in alkaptonuria, an inherited disorder caused by lack of renal and hepatic homogentisic acid oxidase. Other complications are impaired wound healing and wound dehiscence; the fish odour syndrome[5]; nephropathy[6]; steroid addiction syndrome; predisposition to infection; a broad spectrum of cutaneous and endocrinological complications of corticosteroids, including suppression of the hypothalamic-pituitary-adrenal axis; and rarely skin cancer.[3 7]

The culture of bleaching has become common among black Africans. A study of market traders in Lagos, Nigeria, reported a prevalence of 27.6% in men and 49.7% in women.[8] Another study among students of a tertiary institution in Enugu, Nigeria, reported a prevalence of 71.9% in women and 53.3% in men.[2] Studies of women in Senegal and in Jordan reported a prevalence of 26% and 60.7%, respectively.[9 10]

This habit of bleaching has been driven by several factors. Women with lighter complexions are more often used to advertise a wide range of products including alcoholic beverages, toiletries, cosmetics (that are not even skin lightening creams), and clothing. Also, some highly visible women—for example, in the entertainment industry—bleach their skin. Therefore, the dominant signal being sent to undiscerning minds is that people with fair complexions are the beautiful and successful ones. Compared with non-users, people who use skin lightening products are more likely to believe that lighter skin tone has a positive role to play in self esteem, perception of beauty and youth, marriage, and employment opportunities compared with non-users.[10] Some claim they use the cream to treat skin blemishes such as acne and melasma.[2 3 9 10]

Outside Africa, these products are not sold in the regular cosmetic section of department stores or pharmacies. They are either manufactured by international companies for export only or are manufactured in Africa for local consumption and export (largely through smuggling) to countries where the demand for these products is high. They are usually found only in local shops where ethnic foodstuffs are sold or in ghettoes colonised by black Africans and other people with coloured skin.

When should clinicians suspect the use of bleaching agents? The clinical presentation depends largely on the duration of use and the active ingredient in the cosmetic. The clinical history is usually inaccurate because patients may deny using the products or pretend not to know what their function is. Suspicion should be aroused in any black person whose skin appears unnaturally light. Patients often exude a rotten fish body odour because of the excretion of trimethylamine in the breath, urine, sweat, saliva, and vaginal secretions. The most common complications are exogenous ochronosis and colloid milium as a result of hydroquinone. If mainly corticosteroid creams are used patients may have extensive superficial fungal infections of the skin; extensive eruptive filiform (digitate) viral warts; severe acne vulgaris on a background of bleached facial skin; folliculitis, particularly around the axillary folds, and stretch marks (striae). The skin bruises easily, and wound dehiscence after surgery and delayed healing of wounds may occur.

Some grossly inadequate efforts have attempted to combat the problem. In 1995, the Nigerian Food, Cosmetics and Drugs Regulatory Agency prohibited the manufacture and sale of cosmetics containing hydroquinone and mercury,[1] but it later endorsed a new cosmetic product—Venus Skin Toning Cream.

The international drug companies that manufacture corticosteroid creams are probably well aware of the misuse of their products as cosmetics. For financial reasons, some of them have chosen to turn a blind eye. Furthermore, the open trade between the borders of various African countries means that no country is immune, even if it tries to protect its own citizens through legislation.

Several efforts have been made by various people and corporate bodies—for example, the Medical Women's Association of Nigeria, Nigerian Association of Dermatologists, the Nigerian Medical Students Association, and some religious organisations—to educate the public about the inherent dangers of these products, but the impact, if any, is yet to be documented.

Complications can be prevented only if there is a positive change in attitude among all stakeholders. Efforts should be made to change people's unenlightened values and perceptions about beauty. The public needs to be told about the hazardous effects of these products. Skin lightening creams, especially potent fluorinated corticosteroid creams, should be available only as prescription drugs, as is the case in developed countries. Because of the open cross border trade between African countries, the various regional health communities in Africa must act in concert to enforce these regulations.

1 Cosmetic Product Regulations (Prohibition of bleaching agents etc). *Extraordinary Federal Republic of Nigeria Official Gazette* 1995;82(31c).
2 Ofondu EO. Cosmetic use of bleaching creams in Enugu [Nigerian FMCP dissertation]. National Postgraduate Medical College, Nigeria, 2003.
3 Olumide YM, Akinkugbe AO, Altraide D, Mohammed T, Ahamefule N, Ayanlowo S, et al. Complications of chronic use of skin lightening cosmetics. *Int J Dermatol* 2008;47:344-53.
4 Findlay GH, Morrison JGL, Simson IW. Exogenous ochronosis and colloid millium from hydroquinone bleaching creams. *Br J Dermatol* 1975;613-22.
5 Shelley ED, Shelley WB. The fish odour syndrome: trimethylaminuria. *JAMA* 1984;251:253-5.
6 Chakera A, Lasserson D, Beck LH Jr, Roberts ISD, Winearls CG. Membranous nephropathy after use of UK-manufactured skin creams containing mercury. *QJM* 2010; doi:10.1093/qjmed/hcq209.
7 Addo HA. Squamous cell carcinoma associated with prolonged bleaching. *Ghana Med J* 2000;34:3.
8 Adebajo SB. An epidemiologic survey of the use of skin lightening cosmetics among traders in Lagos, Nigeria. *West Afr J Med* 2002;21:51-5.
9 Del Giudice P, Yves P. The widespread use of skin lightening creams in Senegal: a persistent public health problem in West Africa. *Int J Dermatol* 2002;41:69-72.
10 Hamed SH, Tayyen R, Nimer N, Alkhatib HS. Skin-lightening practice among women living in Jordan: prevalence, determinants, and user's awareness. *Int J Dermatol* 2010;49:414-20.

Response on bmj.com

"In my country, Pakistan, the media plays a very negative role in promoting such cosmetic agents. Dozens of television adverts run day and night promoting creams and bleaches to make women fairer." Hunniya Waseem, emergency physician, Public Health Solutions, Pakistan

To submit a rapid response, go to any article on bmj.com and select "Respond to this article"

Discuss skin bleaching on doc2doc, BMJ Group's global online clinical community, at http://bit.ly/exknzk