ABSTRACT
Objective: To examine how leading clinical journals report research findings, aiming to assess how they frame their implications for medical practice and to compare that literature’s patterns with those of the management literature.
Data Source: Clinically relevant research articles from three leading clinical journals (N Engl J Med, JAMA, and Ann Intern Med).
Methods: Review of wording of a sequential sample from 2010, with categorisation, comparison among journals, and comparison with management literature.
Results: Clinical journals usually state that one approach did or did not differ from another approach (35 of 51 articles, 68.6%), but they recommended a specific course of action (‘therefore, x should be done’) in just 25.5%. One article gave instruction on how to implement the changes. Two-thirds of the reports called for further research. Half used tentative language. Management research articles nearly always specified who should use the information and drew from over 60 types of potential users, whereas the clinical literature named the audience in only 23.5% of clinicians.
Conclusions: Authors and editors of the clinical literature could test being more clear and direct in presenting implications of research findings for practice, including stating when the findings do not justify changes in practice.

INTRODUCTION
In March 2010, Bartunek and Rynes reviewed the leading professional literature in organisational management, showing that a minority of papers provide concrete and clear advice arising from the research and that most used tentative language in conveying their advice. We were curious as to the patterns and practices in the medical literature.
Clarke and Chalmers in 1998 reviewed 26 randomised clinical trials and found only two articles that deliberately integrated their new findings into existing literature. Lewis suggested that passive voice and unclear comments limited the usefulness of articles to healthcare decision makers. Lucas confirmed that medical journals usually give tentative result statements. Tunis and colleagues reported that clinical trial research regularly fails to give decision makers enough information to make well-informed decisions, perhaps because investigators have little incentive to reach out to practising clinicians or policy makers.
This paper examines recent research reports in leading medical journals to gain insight into how authors interpret the implications of their findings for medical practice.

METHODS
Journal and article selection
Since we were seeking to characterise broadly influential journals, we selected those with the highest ‘Journal Impact Factor (JIF)’, which tallies the citations of articles in one journal in all professional journals. The top three US medical journals by (2004–2009) JIF score are The New England Journal of Medicine (NEJM), The Journal of the American Medical Association (JAMA) and Annals of Internal Medicine (Annals). The subscribers are mostly clinicians; for example, JAMA reports 100% of their circulation goes to physicians, medical students, hospitals and firms associated with the medical profession.
Two authors (APO and JL) reviewed the table of contents for each issue of each journal, starting with the first edition of 2010, until we had identified more than 20 articles with titles that appeared to report primary research relevant to the medical care of human beings, excluding reviews, editorials and other items.
Review of selected articles
Each author independently reviewed each article. We excluded articles that did not report findings relevant to medical treatment (because they reported basic science, epidemiological description or methodological work). Our reviews focused upon identifying statements that articulated implications for practice in the Abstract or Discussion sections. In addition, we identified statements that the findings had implications for future research or for public health. We did not code guidelines or medical advice that did not arise from the current research. Annals publishes a short comment from the Editor, which we discuss separately.

After initial review, we developed categories to code and all three authors independently reviewed each article, identified the appropriate text and coded it. We resolved any coding disagreements by discussion. The Bartunek and Rynes article identified that managers prefer straightforward, clearly implementable prescriptions for practice. Thus, we examined the text that articulated implications for practice, being careful to identify tentative language (eg, words such as ‘may’, ‘speculate’, and ‘potentially’). We also tallied prescriptive language, such as ‘ought’ and ‘must’. We tabulated findings that applied to a subset of situations, which we called ‘contingencies’. Finally, we coded whether the article identified who (eg, the physician), if anyone, should take account of the findings.

We then compared the three journals, using overall $\chi^2$ tests; and, when appropriate, we compared pairs of journals using Bonferroni’s correction for significance level. We also compared the medical literature with the management literature using $\chi^2$ tests.

We tabulated text findings on Excel and managed coding and analyses with SPSS (PASW V.18).

RESULTS

The cohort and coding
Aiming to sample $\geq 20$ articles per journal, we reviewed NEJM through 28 February 2010; JAMA through 10 March 2010 and Annals through 6 April 2010. The initial cohort totalled 65 articles, with 21 from NEJM, 21 from JAMA and 23 from Annals. The initial review excluded 14 articles, leaving a cohort of 51 that reported research that could be important to medical practice: 18 from NEJM, 12 from JAMA and 21 from Annals.

Figure 1 illustrates the derivation of this analytic cohort. Thirty-five articles gave advice that ‘X was better than Y with regard to Z’ or ‘X is no better than Y with regard to Z’. These statements often hedged: ‘X seemed to be better than Y’ or ‘X may prove to be better than Y’. These statements did not intimate what other considerations might affect the practitioner’s decision as to whether to change. We called these statements ‘Minimal Advice’, and 28 articles gave only this kind of phrasing. We identified four additional patterns for stating the implications of findings for practice: ‘practitioner should consider’ (labelled ‘Consider’); ‘patient should be informed’ (labelled ‘Inform’); ‘practitioner should do’ (labelled ‘Act’); and ‘practitioner should do and here’s how’ (labelled ‘Technical Assistance’).

Statements about implications for practice
One article by Ray et al stated factual findings with no evaluative statement at all. This article reported results from a large observational study that assessed two kinds of medications. The outcomes were complex, showing competing gains and risks. The information would be useful for a practitioner to consider, though the article did not explicitly state that conclusion. The information could not have yielded advice to prefer one treatment strategy, except perhaps to inform the patient.

Table 1 shows the number and proportion of the other 50 articles that provided the five kinds of advice, overall and for each journal and figure 2 illustrates the overlap among the kinds of advice given.
Eleven articles called on someone to ‘consider’ acting in response to the findings. In no case did the article give additional advice as to how that person might weigh the issues or what other considerations should enter into the deliberation. For instance, ‘Early resumption of low-dose aspirin therapy with proton pump inhibitors in patients with bleeding ulcers and cardiovascular disease should be considered’.39

Two articles called for informing the patient as part of the implementation of findings, one concerning influenza immunisation11 and one concerning feeding tube placement.29 For example, ‘As the current pandemic unfolds, pregnant and postpartum women should be counseled about the importance of vaccination’.11

Thirteen articles called for specific action, generally stating ‘The physician should...’. For example, ‘Our findings also suggest that compression ultrasonography might be considered for patients with symptomatic SVT at presentation to evaluate the extent of the thrombosis and diagnose potential DVT, that physicians should suspect and test for pulmonary embolism in patients with suggestive symptoms...’.50 The first clause counts as a ‘Consider’ statement; the second as an ‘Act’ statement.

One article provided a unique approach. Goldstein et al.54 published a ‘Brief Communication’ in Annals that gave research findings and stated clear directives for action and also provided an on-line appendix with details about implementation.

Often, the authors appear to assume that the reader would know the implications for practice and do not state them. Van der Gaag et al.,16 for example, found that ‘Routine perioperative biliary drainage in patients undergoing surgery for cancer of the pancreatic head increases the rate of complications’. However, the article does not state the obvious implication for practice: do not do the drainage procedure in this setting.

Additional attributes of statements about implications of the research

We describe various attributes of how the articles provide implications for practice in table 2.

Two-thirds (66.7%) of the articles called for future research, usually using a broad and brief formulation. For example, Goebel et al.45 claimed: ‘Additional studies are needed to determine which patients are likely to benefit and which IVIG doses and schedules are most effective’. In 9.8% of the cases, the articles stated an implication for public health actions such as developing guidelines or including certain services in funding. Nachega et al.,41 for example, called for monitoring adherence to anti-retroviral therapy to be part of ‘the package of care in anti-retroviral therapy programs in all settings’, which would require action by funders and governments.

Forty-seven per cent of the articles used tentative language in stating their implications for practice. For example, ‘Our findings suggest that the use of morphine during trauma care may reduce the risk of subsequent development of PTSD after serious injury’.14 As in this example, many used two or more tentative terms. Sometimes, the writing strongly avoided being direct, for example, ‘Therefore, we believe that prescribing an
NSAID to treat an asymptomatic postoperative pericardial effusion should no longer be advised.47 The reader might reasonably wonder whether this is a matter of faith that requires belief and also who should no longer be advising whom.

About 39% of the articles had some prescriptive language. For example, ‘Until then, clinical practice should not be guided by (point-of-care) platelet function testing’.31 Contingent (subset) statements were present in about 29%. For example, ‘In short, most women contemplating estrogen plus progestin therapy for the relief of menopausal symptoms should not expect protection against CHD’.49

Each journal’s website makes a general statement about intended readers: NEJM targets ‘medical researchers’, JAMA claims ‘physician readers’ and Annals focuses on ‘practising physicians’. The stated audience does not include non-medical professionals, laypersons or patients. Perhaps this strong assumption about the readership explains why only 23.5% of the articles specify the people whose actions should reflect the reported findings. Authors did specify physicians in 10% of articles and healthcare providers or clinicians in 8%.

### Differences among journals

We analysed the data in tables 1 and 2 for differences among journals and only two comparisons met conventional standards for significance: Annals was significantly less likely to publish ‘Minimal Advice’ than NEJM ($\chi^2=8.925$, df=2, $p=0.006$). Conversely, Annals was significantly more likely than NEJM to recommend action ($\chi^2=8.469$, df=2, $p=0.005$).

In Annals, the Editors’ note provides a very short summary of findings and cautions. Sometimes, as in Meurin et al.,47 the Editors give a stronger recommendation than the authors did: ‘Physicians should stop prescribing NSAIDS for postoperative pericardial effusion because these agents have no or only small beneficial effects’ by the Editors and ‘In patients with pericardial effusion after cardiac surgery, diclofenac neither reduced the size of the effusions nor prevented late cardiac tamponade’ by the authors. On the other hand, sometimes the Editors’ comments worded the authors’ conclusions more tentatively than the authors, as in Schaer et al.,45 where the Editors said ‘In hypertensive patients, mechanisms other than lowering blood pressure may be important causes of atrial fibrillation’, while the authors claimed ‘In hypertensive patients, long-term receipt of ACE inhibitors, ARBs, or beta-blockers reduces the risk for atrial fibrillation compared with receipt of calcium-channel blockers’.

### Comparison with management research reporting

Management researchers advocating changes in practice virtually always stated their claims in the active voice. In contrast, in 33% of cases medical journals used passive voice, for example, ‘New strategies to reduce the risk of transmission of HIV-1 are needed for HIV-1–serodiscordant couples’.26 Implications for practice sections in the management literature used tentative language significantly more (74%) than the medical literature (47%), ($\chi^2=17.48$, df=1, $p<0.001$). The management literature (55%) was also significantly more likely to use prescriptive language than the medical literature (39%), ($\chi^2=4.59$, df=1, $p=0.04$). However, contingent advice was similar in frequency in the management (38%) and medical (29%) literature.

### DISCUSSION

This project examined how three influential US medical journals present implications of research findings for practice. Most research reports provide a summary of the findings with a flat descriptive statement that a clinical outcome is better or worse with the tested intervention (68.6%). The one article that did not give this level of advice had findings that would not support doing so. Only about one-forth (25.5%) of the reports gave explicit recommendations for action and only one article provided instruction as to how to implement the recommended actions. The language in about half of the articles was tentative (47.1%) and a general call for further research was common (66.7%). Reporting in the three journals was generally similar, but as a group they differed from the management journals in including fewer calls to undertake action, and in not specifying who should act.

In this project we studied only a small sample of articles in a confined time frame, which limited the study’s...
power and generalisability. Furthermore, our review did not examine editorials or journal review services, which often serve to put the findings of single studies into a broad context and provide substantial advice to the practitioner as to how to interpret the results.

While the target for implications for practice in medical journals was either not named or was specified only as ‘physicians’ or ‘clinicians’, the management of medical journals was either not named or was specified practitioner as to how to interpret the results. Bartunek and Rynes identified more than 60 categories of individuals or groups to whom articles addressed their implications.1 The clinician audience for these medical journals may be so obvious as to need no specification.

The project raised some intriguing questions as to optimal reporting. In the management literature assessment, the authors advocated more statements of the implications for practice1; in contrast, many of the articles we sampled from clinical literature had methodological limitations that precluded strong statements guiding practitioners towards changed practices. The findings could often have needed replication or could need to be weighed along with other considerations in making individual treatment decisions. Strong statements about medical practice might more often be justified in systematic reviews, or clinical practice guidelines based on extensive literature reviews. One well-known example is the US Preventive Services Task Force, which provides authoritative reviews and grades both the strength of the evidence and of their recommendations.60 61

However, clarity in assessing and reporting the impact of new findings is desirable, even when the report should have limited impact. Authors can clearly state what their work has added to the body of evidence and what its implications for practice are, even if the clearly stated advice would often be that a reasonable practitioner should not change anything yet. This proposal would operationalise for all research reports the contention of Clarke and Chalmers9 regarding randomised clinical trials. They advocated that each report should state, in the ‘discussion’ section, how the new findings reshape the prior body of evidence concerning the topic.

Editors should also consider undertaking formal tests of whether authors can clearly and succinctly state implications for practice, perhaps by providing explicit guidelines on grades of evidence and levels of recommendation and whether such explicit and consistent reporting helps readers. If past research and the new data support a strong recommendation for implementation, then the article should explicitly state the major counter-considerations or contingencies and often can provide ‘technical assistance’ for implementation. If the authors hold that the findings are not yet reliable enough for action, they could state that judgement clearly, along with their opinion regarding who should undertake what research or other actions.

Medical research generally builds in small increments and few articles make a dramatic difference to what practitioners should do. However, our findings suggest that the current system for reporting the implications of single clinical studies deserves serious re-examination by authors, editors and readers. Stating the strength of evidence and recommendations for action in those studies as directly, unambiguously and consistently as possible seems to us likely to prove helpful in translating their results into the most appropriate actions.

Acknowledgements The authors are grateful to the Health Foundation in the UK which brought two of the authors together to examine the epistemology of quality improvement in April 2010, a meeting which sparked our interest in the topic. The authors are grateful to the National Academy of Social Insurance, which supported Allessia P Owens as a Somers Scholar, enabling her to take the lead in this work. The authors also acknowledge the substantial help from anonymous reviewers and the editor, Dr Frank Davidoff.

Competing interests None.

Provenance and peer review Not commissioned; externally peer reviewed.

REFERENCES