Body: 23-Feb-2017

Dear Prof. Jena

Manuscript ID BMJ.2016.037263 entitled "Physician Age and Outcomes of Hospitalized Elderly Patients in the U.S.: Observational Study"

Thank you for sending us your paper. We sent it for external peer review and discussed it at our manuscript committee meeting. We recognise its potential importance and relevance to general medical readers, but I am afraid that we have not yet been able to reach a final decision on it because several important aspects of the work still need clarifying.

We hope very much that you will be willing and able to revise your paper as explained below in the report from the manuscript meeting, so that we will be in a better position to understand your study and decide whether the BMJ is the right journal for it. We are looking forward to reading the revised version and, we hope, reaching a decision.

Yours sincerely,

Georg Roeggla
groggla@bmj.com

Report from The BMJ's manuscript committee meeting

These comments are an attempt to summarise the discussions at the manuscript meeting. They are not an exact transcript.

Manuscript meeting 23.02.2017

John Fletcher (chair), Jamie Kirkham (stats), Elizabeth Loder, Kristina Fister, Rubin Minhas, Tiago Villnueva, Georg Roeggla, Jose Merino, Tessa Richards, Amy Price, Daoxin Yin and two guests.

Decision: Ask for revision

The committee was interested in the topic of your research. The following concerns were mentioned:

• Confounding is probably the biggest issue here and we think you acknowledge this by the abundance of secondary analyses undertaken to assess the robustness of this.
• The correlation between patients with the same physician are addressed. However, the data appear to have a natural hierarchical structure: patient, physician and hospital. Could the data not be modelled in this way using a hierarchical modelling structure?
• We are unclear how the hospital 'fixed' effects account for both measured and unmeasured characteristics of hospitals as currently described. There are many adjustment factors included in the model but we don't get a sense of how important these were on the outcome.
• For example, it would be interesting to know if physician experience (years since graduation) is important rather than age itself. Of course, this are likely to be correlated and this issue would need to be addressed. Can therefore estimates of adjustment factors be provided with confidence intervals?
• International readers don't understand Part-B spending. There's a lot of US context that is also not clear to non-US readers. You could add a box explaining the issues around Medicare Part A and Part B.
• Aside the statistics, it's unclear from the article how the results should be interpreted as a way forward, or what kind or message this sends out.
• The results are also limited to an elderly population and may disappear if a younger population was also included in the analysis.
• Several reviewers raise the issue of generalizability, since this seems like a specific subset of US doctors looking at a specific population of patients, and these findings would not travel to other countries.
• In many countries hospitalized patients are taken care by a team of doctors. Attributing hospitalized patients to a single doctor is difficult therefore.
• The adjusted differences are not that great. Are they relevant?
• We see this can have implications for US recertification procedures.
• We are not convinced that mortality is a very good indicator of quality of care in a population with a mean age of 81 years admitted for emergency care. A good proportion of people this old (and half of them are older) who are admitted as an emergency are destined to die soon, and better care may not necessarily include prolonging the death process.
• Could you use a different outcome measure as your primary outcome in this age group or use mortality but a different population (elective admissions or people under 65 years)? Or could you additionally evaluate another outcome such as readmissions?
• Is mortality just in-hospital mortality or does it include linkage to community deaths? The methods mention Medicare Inpatient and Carrier Files, which don't look like they contain deaths outside hospital.
• For a non-USA audience, a better description of hospitalists is needed.
• Younger physicians may have to be recertified every 5-10 years while older ones may be exempt from the requirement. This is an important issue because there is a debate in this country about the value of recertification.
• In many countries patients may not have a choice who is their hospitalist. So not a relevant issue for patient choice. But important for regulation.

• Hospitalists coordinate care but don't provide all of it.
• The main outcome was 30-day mortality. Please explain where you got the data.
• Do the findings have implications for policy makers?
• Don't many doctors try to move into nonclinical work as they get older? Older doctors who are still practicing as hospitalists may be different in important ways from those who are not - perhaps they are intrinsically less capable doctors who have been unable to exploit opportunities to move into management positions, teaching, research and other nonclinical work.

First, please revise your paper to respond to all of the comments by the reviewers. Their reports are available at the end of this letter, below. Please also respond to the additional comments by the committee.

In your response please provide, point by point, your replies to the comments made by the reviewers and the editors, explaining how you have dealt with them in the paper.

** Comments from the external peer reviewers**

Reviewer: 1

Recommendation:

Comments:
This interesting study analyzed the relationship between physician age and patient outcomes (30-day mortality and readmissions) and costs of care (measured by Medicare spending) among patients hospitalized with a medical condition and treated by a hospitalist during 2011-2013. The study's main finding is that physicians with greater age have higher 30-day mortality, no difference in readmission rates, and slightly higher costs of care. Hospitalist physicians account for upwards of 70% of all US medical admissions and so this is perhaps the most important group to study for this sort of question, a point worth noting by the authors.

The authors conduct a broad battery of sensitivity analyses to their findings, including looking at general internists overall, using alternative methods of attributing attending physicians to patients, considering hospitalist physicians who also have sub-specialty boards (e.g., some hospitalists may be trained in infectious disease, nephrology, etc., yet still practice general inpatient medicine), analyzing individual diagnoses, focusing on physician experience rather than physician age, and using different models to study costs. Their study design, in my opinion, strongly addresses the possibility of confounding by factors that might be correlated with both physician age and mortality (more on this below). In addition, the authors show that the size of plausible confounder would have to be unreasonably large to explain their results.

This is an important study and the implications are large. The authors appropriately conduct numerous analyses to assess the robustness of their findings, which is particularly important for a study like this, which is likely to attract attention if published.

The issue of how physician age (or experience, I use those interchangeably) relates to quality of care, in particular patient outcomes, is a first order question for professionalism, health policy, and medical education. Intense debates around physician recertification have been largely agnostic of solid outcomes evidence on the relationship between physician age and patient outcomes. A prior systematic review of physician experience and quality of care found that on process measures of quality, including testing, clinical knowledge, etc., that older physicians performed worse. However, in that review, data on actual outcomes was unavailable. The current study fills that gap by using national data on a broad population of patients. The findings of the current study, while arguably provocative, are in line with what the prior review by Choudury et al. found (published in Annals of Internal Medicine). The results of this study will be an important consideration in weighing the pros and cons of various initiatives to ensure maintenance of physician skills.

The immediate question to me is whether the authors have adequately addressed potential confounding, namely that more experienced physicians may attract sicker patients and therefore have worse outcomes. I think they have done so quite adequately. The authors make their primary analysis one of hospitalist physicians, who traditionally work in blocks rather than care for a pre-specified set of patients (like PCPs). Therefore, almost by construction, they are quasi-randomized to patients. That assumption needs to be verified, however, and cannot just be asserted. The authors show strong balance of patient characteristics across physicians of varying age, using a broad set of demographic and clinical factors. To ensure that their results are generalizable, the authors also find similar findings among general internists (albeit smaller), which is a strength of the study. While I think that the potential for confounding is low, both from a theoretical perspective (the set of patients a hospitalist sees is likely to be random) and from an empirical perspective (see table 1, where the covariates appear fairly balanced across groups), an important analysis is to quantify how large a confounder would need to be to negate the authors' results. The authors do that analysis and find that such a confounder would need to be implausibly large or prevalent to explain the physician age – patient outcome relationship.

One potential criticism of the study, which I actually don't think is an issue, might be that the main analysis is of hospitalists. The field of hospital medicine is about 20 years old (or younger) and so in comparing hospitalists with greater or lesser years of experience, the authors may actually be comparing younger doctors who were always hospitalists to older doctors who started as PCPs and became hospitalists later in their careers. I actually think is a moot point, however. From the patient perspective, a patient wants to know what are their likely outcomes if they are treated by a doctor in the inpatient setting with more vs less years of experience. It matters not whether this is due to depreciation of skills over time or some sort of cohort effect. Therefore, I think the analysis of hospitalists is credible. I also think the authors do a good job of discussing this issue by describing 'cohort effects' and 'age-related effects.' The former encompasses how physicians were trained (i.e., were older physicians trained differently) vs whether skills depreciate with time (i.e., 'age-related' or 'experience-related' effects).

One thing that is missing in the paper is some description of what international data is available. The BMJ is an international journal with global readership. What data exists in the UK and elsewhere? Some brief mention of this data would be nice.

To summarize, the paper is analytically sophisticated, well written, and of clear importance / general interest. It employs multiple sensitivity analyses, which makes it difficult to recommend further analyses that should be conducted. It addresses the issue of confounding head-on through a quasi-experimental study design, and it has the important additional finding that the effects of experience are muted among high-volume physicians.

Additional Questions:
Please enter your name: Eric Sun

Job Title: Assistant Professor

Institution: Stanford University

Reimbursement for attending a symposium?: No

A fee for speaking?: No

A fee for organising education?: No

Funds for research?: No

Funds for a member of staff?: No

Fees for consulting?: No

Have you in the past five years been employed by an organisation that may in any way gain or lose financially from the publication of this paper?: No

Do you hold any stocks or shares in an organisation that may in any way gain or lose financially from the publication of this paper?: No

If you have any competing interests <A HREF='http://www.bmj.com/about-bmj/resources-authors/forms-policies-and-checklists/declaration-competing-interests'target='_new'> (please see BMJ policy) please declare them here:

Reviewer: 2

Recommendation:

Comments:
Physician Age and Outcomes of Hospitalized Elderly Patients in the U.S.

I have used the BMJ Review Template for patient reviewers, quoted as bullet points, below. Thank you for the opportunity to review this manuscript.

• Are the study's aims and the issue and questions that the paper addresses relevant and important to you as a patient? Do you think it would be relevant to other patients like you? What about carers?

The hypothesis to justify this study seems interesting to patients at first glance: does quality of care differ between younger and older physicians. Of course, patients and family members seek good quality care regardless of their physician's age. But this hypothesis begs the question: what would I do differently if I learned older physicians offered better or worse quality care than younger physicians? Not much from a patient perspective, as it turns out. The coarse metric for quality of care is 30-day mortality and 30-day readmission to hospital. And in any case, if I were a patient like those in this study (65+ with a subset of non-elective medical reasons to be in hospital), I wouldn't be offered the opportunity to choose my hospitalist by age or anything else. This study's aims might be relevant to patients and families only if a meaningful system strategy to eliminate variation were implied by the findings. I don't see one.

• Are there any areas that you find relevant as a patient or carer that are missing or should be highlighted?

This study offers no practical or theoretical assistance to patients and family members seeking the best possible care from hospital physicians. It does not suggest causality for the observed variation in care outcomes.

• Would the treatment, intervention studied, or guidance given work in practice? Is it feasible? What challenges might patients face that should be considered?

As it turns out, the findings reported here show very small differences in patient mortality between those cared for by older and younger hospitalists, and no difference in 30-day readmissions. Of course, a variation of even one death per 100 is important to that patient and their family.

A similar study [http://jamanetwork.com/journals/jamainternalmedicine/fullarticle/2593255] by members of this research team was published in December in JAMA Intern Med examining physician gender

correlation with patient outcomes in similar populations to this study. Female hospitalists for 65+ patients were associated with a slightly statistically significant lower rate of 30-day mortality and readmission than male hospitalists. Given no implied causality for this observation, what should the public cautiously conclude?

The article's publication generated considerable media excitement from blogging professionals and health reporters with headlines such as, "Don't want to die before your time? Get a female doctor" --USA Today [http://www.usatoday.com/story/news/2016/12/19/women-doctors-hospital-patients/95531520/]; and "Patients Cared For By Female Doctors Fare Better Than Those Treated By Men" --NPR All Things Considered [http://www.npr.org/sections/health-shots/2016/12/19/506144346]; and from a Canadian psychology professor "Women physicians are superior doctors according to objective outcomes: mortality." tweeted by @PaulMinda1 on 19 Dec 2016.

These headlines are unwarranted and unhelpful to the public. If that reception is instructive for this study, notoriety rather than credit is likely to be the true impact factor. Let's please not blame the public for misinterpreting the evidence. I think that studies such as these are provocative without much likelihood of helping patients or physicians work towards better care.

• Are the outcomes that are being measured in the study or described in the paper the same as the outcomes that are important to patients? Are there others that should have been considered?

To define quality of care in terms of system-centric metrics like 30-day mortality and readmission reduces what matters to patients to survival. We all will die sooner or later, and for these patients in the study it is likely to be sooner. Factors that patents and families feel are important like respect, compassion and individualized attention are not captured. Quality of care as we advance closer to our fate is not likelihood of mortality or hospital readmission.

• Do you have any suggestions that might help the author(s) strengthen their paper to make it more useful for doctors to share and discuss with patients?

There is not much in this study that any doctor would wish "to share and discuss" with their patients other than to dissuade patients from allowing such flimsy observational findings to influence their own care decisions. I don't know what audience across healthcare would make good use of this research. It is rather more predictable to generate another round of ridiculous headlines. If the researchers are interested in how age affects professional performance, there are much better methodologies that would carry the potential to impact care improvement.

Additional Questions:
Please enter your name: Carolyn Canfield

Job Title: Independent Citizen-Patient; Honorary Lecturer

Institution: Faculty of Medicine, The University of British Columbia

Reimbursement for attending a symposium?: No

A fee for speaking?: No

A fee for organising education?: No

Funds for research?: No

Funds for a member of staff?: No

Fees for consulting?: No

Have you in the past five years been employed by an organisation that may
in any way gain or lose financially from the publication of this paper?: No

Do you hold any stocks or shares in an organisation that may in any way
gain or lose financially from the publication of this paper?: No

If you have any competing interests <A HREF='http://www.bmj.com/about-bmj/resources-authors/forms-policies-and-checklists/declaration-competing-interests'target='_new'> (please see BMJ policy) please declare them here: I have accepted expenses and honoraria from government-funded universities and agencies for public speaking, none of which are, or could be seen to be, conflicts of interest in reviewing this manuscript.

Reviewer: 3

Recommendation:

Comments:
[1] This article found to be interesting. But, as a reader, I failed to understand the concept evolvement in the entire article.
[2] The structure of the article required revisiting. The supplementary materials shall be concised and placed under the methods and materials section rather.
[3] As specified in the article, the linkage between age of physician and mortality is not come-out very

clearly. The reasons are: [a] severity of the diseases (of the patients) and the patients' age not addressed and; (b) the association between age of patients and the physicians age not cross-tabled to give clarity to the study objectives.

[4] The fundamental question that arises here that the death is natural and that may be due to terminal illness of the patients (as outlined diseases category). How come this has been accounted to physician? Not clear.

[5] The study found that volume may increase/decrease clinical expertise, though there is not much significant difference are not noted, as depicted through odds ratio. Why?

[6] Are the physicians intentionally carried out the death to the patients? If yes, this reflects that the competency level of physicians, and not as corroborated by the authors (like physician age). If it is so, how the national efforts are made to improve the efficiency. This is missing in the article.

[7] The Tables required re-checking. For instance, Table-2, number of physicians mentioned as 950 in the category > 60 years. But in the Table-4, it is quoted as 961. In Table (part-B) under the> 60 years category it is 924? No explanation given in the article for the range of difference? At the same time, the number of Tables required to be reduced and shall be re-drawn to reader friendly

[8] One of the sub-objectives of the article to see the national efforts to improve the efficiency, which required more detailing. Since it is not specified as "what are national efforts?, the reader tend to assume his way . For instance, Is there any continuous medical education (CME) in place in USA? If it is there, what was to the effect? As the article talking about the 3-years data, how many of the physicians undergone CME? Are physician required to undergo renewal of licensing under Medicare schemes or any other health financing mechanism? How this kind of short-mortality monitored and noted under various health financing mechanism? I did not find answer from the article.

[9] Too many statistics are applied but without understanding its relevance to the study objectives. The article required to be more descriptive rather statistical one.

[10] Discussion required more elaboration with specific to the study objectives as outlined in the article, but it is missing.

[11] Conclusion needs more clarity and required to be strengthen with the study findings.

Looking at the above suggestions/comments and required clarity, I wish authors shall re-visit the article and re-submit for further review processing.

Additional Questions:
Please enter your name: N Ravichandran

Job Title: Associate Profesor

Institution: Hamdard University

Reimbursement for attending a symposium?: No

A fee for speaking?: No

A fee for organising education?: No

Funds for research?: No

Funds for a member of staff?: No

Fees for consulting?: No

Have you in the past five years been employed by an organisation that may
in any way gain or lose financially from the publication of this paper?: No

Do you hold any stocks or shares in an organisation that may in any way
gain or lose financially from the publication of this paper?: No

If you have any competing interests <A HREF='http://www.bmj.com/about-bmj/resources-authors/forms-policies-and-checklists/declaration-competing-interests'target='_new'> (please see BMJ policy) please declare them here: NONE

Reviewer: 4

Recommendation:

Comments:
Clearly written and well documented analysis of the association between age/volume of U.S. hospitalists in internal medicine and 30 day mortality of hospitalized patients with medical diagnoses. Elderly patients with medical diagnoses and non-elective admissions cared for by low and medium volume hospitalists with greater age experience higher odds of dying. Readmissions are not affected by physician age. Older physicians have slightly but perhaps not meaningfully higher costs measured by physician charges.

As is true with most health services research papers, limitations in design loom largest in determining the value of the contribution made in a particular paper. Reasonable efforts have been made in this paper to address design limitations but doubts in the generalizability of the results persist and the discussion of potential implications for practice may go beyond the findings. Most of the design limitations are at least mentioned but some could be usefully discussed in greater detail.
The authors have the expected statement in limitations section that the results may not be generalizable to other countries or even to other specialists or patients electively admitted. But in this paper, more discussion is relevant since the paper is being evaluated for a journal with an international audience and

the U.S. has a very different pattern of physician participation in hospital care than most other countries, particularly European countries where hospitalists have long had primary authority for hospitalized patients compared to the US where hospitalists in large numbers are relatively new. Hospitalists in the US are still a subset of all US physicians where a large share of physicians continue to be the physician of record in the hospital and in office-based practice. While hospitalists in the US represent an opportunity to study the impact of physician age on patient outcomes that would be decidedly more difficult among a more representative group of US physicians, generalizing beyond hospitalists in the US to all physicians with patients in hospitals and to hospital-based physicians in other countries is a big stretch. The results seem most applicable to hospitalists in the US, and European readers may not appreciate the reasons why without more clarification.

Along the same lines, the authors discuss the potential need for considering interventions directed at older physicians in language that seems to imply that their results are applicable to all physicians which goes beyond the study findings. Adding some tentative language may be appropriate (I did not find MOC defined on p. 22 in the discussion which may be a US term.)

The authors might say more about the differences in effect sizes found when studying general internists that were not hospitalists. There is a statistically significant association between age and mortality for general internists but smaller effect sizes so the sensitivity analysis is not totally convincing.

Reasonable approaches by the researchers are undertaken to try to test the cohort effects which may be important in the US context where the selection effects into hospitalist careers may vary by age and the training of new cohorts is likely to be different than education of older hospitalists. However these efforts do not dispel concerns that different education and/or selection effects are responsible for better patient outcomes in younger cohorts of hospitalists in US context where the employment of many hospitalists is a relatively new phenomenon.

Not much, if anything, was mentioned about limitations of using Part B Medicare claims to measure costs. Many of the highly variable hospital costs would be in Part A having to do with hospital costs related to length of stay, ICU days, drugs, etc that are not necessarily reflected in Part B claims. Again this might not be apparent to non-US readers.

30 day from admission mortality is an extremely important outcome measure, of course, but a limited one in evaluating the association of physician age and experience on outcomes in serious illness. The absence of a difference in readmission rates by physician age is a reminder that 30 day mortality might not be the only or even best reflection of quality that should be evaluated in determining the quality of serious illness care provided by physicians. Younger physicians may be less adept at patient and family communication and more aggressive in pomoting invasive (and painful) treatments to extend life in the short term for seriously ill patients but not extend quality years (days) of life. Perhaps the authors might add a note in the discussion on the limitation of 30 day mortality as a quality measure on which to evaluate physician care in serious illness.

Additional Questions:
Please enter your name: Linda Aiken

Job Title: Professor

Institution: University of Pennsylvania

Reimbursement for attending a symposium?: No

A fee for speaking?: No

A fee for organising education?: No

Funds for research?: No

Funds for a member of staff?: No

Fees for consulting?: No

Have you in the past five years been employed by an organisation that may
in any way gain or lose financially from the publication of this paper?: No

Do you hold any stocks or shares in an organisation that may in any way
gain or lose financially from the publication of this paper?: No

If you have any competing interests <A HREF='http://www.bmj.com/about-bmj/resources-authors/forms-policies-and-checklists/declaration-competing-interests'target='_new'> (please see BMJ policy) please declare them here:

Reviewer: 5

Recommendation:

Comments:
This study assesses the relationship between physician age and clinical and economic outcomes for hospitalized individuals. To do so, the authors analyze Medicare acute care claims, focusing on hospitalists treating elderly fee-for-service beneficiaries admitted for a medical condition over 2011-2013. The authors

find that risk-adjusted 30-day mortality increases monotonically with physician age: a physician 60 years old or older has a mortality odds ratio of 1.16 in relation to a physician under the age of 40. Notably, this mortality gradient is confined to relatively low-volume physicians. The authors do not find a relationship between age and readmissions, but do find significantly higher Medicare Part B costs among older physicians.

The relationship between physician age and performance is important, but not well understood. From a theoretical point of view, it is possible that the relationship is subtle. Initially, as physicians gain experience, their performance could improve. Later within a physician's career, additional experience may be less valuable than the growing obsolescence of knowledge and skills is detrimental. As the authors note, the existing evidence on the relationship between age and performance is limited, particularly with respect to patient outcomes.

This study is careful, thorough and convincing. A key challenge is that the patients treated by older physicians could be systematically different in unmeasured ways that affect outcomes. The authors deal with this challenge in several ways. First, the focus on hospitalists is intended to limit unmeasured confounding, because patients are typically assigned to hospitalist based on work schedules. Indeed, the authors are able to show that in their context, measured patient characteristics vary little with physician age. This pattern suggests that patients would not differ in unmeasured dimensions, either. It would be helpful if Table 1 reported formal statistical tests for the relationship between patient characteristics and physician age. Second, the analysis adjusts for a wide range of characteristics, including patient diagnoses and comorbidities but also contextual factors such as neighborhood income.

Third, the authors include dummy variables ("fixed effects") for hospitals in at least some of their analyses. These variables capture unmeasured heterogeneity between hospitals' patients; thus, the authors effectively compare the performance of younger and older doctors within the same hospitals. I would note, however, that the use of dummy variables in a conventional logit model can create a non-trivial "incidental parameters" problem. It would be a reassuring if a linear probability model produced similar results.

The authors are also thoughtful about the study's limitations. The documented pattern could result from declining performance over physician careers, or from improvements in the performance of newer cohorts of physicians, or some combination of these phenomena. While the comparative roles of age versus cohort effects is interesting, the kind of evidence offered here is an essential first step towards understanding them. Moreover, whatever the roles, patients clearly care about the overall relationship between age and performance. From a policy perspective, declining performance with age, whatever the sources, points to a benefit from shorter careers, to be weighed against the fixed costs of training more physicians to meet the need for care. I know that the physician workforce is a critical policy issue in the U.S., and believe that is true in much of the rest of the world, too.

The authors also acknowledge that in assessing physician volume, it is possible that lower volume is at least in part a consequence of lesser performance, rather than a cause. This observation is correct, and while the direction of causality is intellectually interesting, I believe this limitation is minor. The documented pattern reflects the real-world performance of the health care system in matching patients to physicians of varying performance, and that is the first-order issue.

I do have some additional minor comments:
• The mortality gradient in the full sample of general internists. I don't have a view as to whether confounding would bias the measured odds ratio upwards or downwards. Do the authors have a view on this point?
• At the condition level, the text refers to significant gradients for CHF and sepsis, but supplemental table 8 shows significant gradients for CHF and COPD (sepsis is close.) I am tempted to ask the authors whether this pattern is as expected, but I note that the odds ratios are all substantial, and statistical power may be driving these differences.
• A minor technical point, but the comparison between age bins and linear age involves a non-nested hypothesis test. Thus, I don't believe a Wald test is appropriate, but rather a Vuong likelihood-based test.
• To deal with separation problems in the logit model, the authors group DRGs into "clinically similar categories." Please elaborate.
• Please elaborate on how elective admissions were identified.
• The authors refer to adherence to guidelines in relation to both age and cohort effects. This is not necessarily a problem. It is possible that doctors become less adherent over their careers, and also that newer cohorts are more adherent at the outset. Is that what the authors have in mind?

Sincerely,
John A. Romley

Additional Questions:
Please enter your name: John A. Romley

Job Title: Associate Professor of Public Policy

Institution: University of Southern California

Reimbursement for attending a symposium?: No

A fee for speaking?: No

A fee for organising education?: No

Funds for research?: No