PEER REVIEW HISTORY

BMJ Open publishes all reviews undertaken for accepted manuscripts. Reviewers are asked to complete a checklist review form (**http://bmjopen.bmj.com/site/about/resources/checklist.pdf**) and are provided with free text boxes to elaborate on their assessment. These free text comments are reproduced below.

ARTICLE DETAILS

TITLE (PROVISIONAL)	Size, composition and distribution of human resource for health in India: new estimates using National Sample Survey and Registry data
AUTHORS	Karan, Anup; Negandhi, Himanshu; Nair, Rajesh; Sharma, Anajali; Tiwari, Ritika; Zodpey, Sanjay

VERSION 1 – REVIEW

REVIEWER	Caroline Homer Burnet Institute, Australia
REVIEW RETURNED	05-Oct-2018

GENERAL COMMENTS	Thank you for the opportunity to review this paper. I found the findings fascinating, if a little disturbing, but also quite hard to follow. The paper needs to be a lot clearer in focus to be acceptable for publication. Some of the complexity is that two sources of data were used – so both need to be described in the methods and then brought together. The readers of this journal may not be really familiar with the Indian system. For example, what is the difference between an allopathic doctor, an indigenous system and how do AYUSH doctors fit in. In terms of the latter, what is their training and scope of practice? Is a AYUSH Doctor the same as a AYUSH Practitioner? Both terms are used. The Introduction section was quite hard to follow and need careful editing for clarity. For example, is para 2 about this study or other studies – if it is this study then it needs moving to the end of the Introduction once the gap in knowledge is established. I would prefer to see a clear statement that says 'the aim of this study was to ….' While the paper presents density on health providers, these seem rather blunt, combining all workers. What is the ideal still-mix balance especially between allopathic and AYUSH doctors? What is the ideal balance of the workforce? In the Methods, the situation in India section is really about the data collection methods not the setting. I would like a better description of the situation or setting. The websites of the Councils were used to gather information – how accurate are these at a national level?

	To what degree is the NSSO representative of the health workforce? Are there any limitations to consider? Registry data were also used which adds to the complexity. The authors say there was no patient or public involvement but the survey was at a household level? Table 2 was very hard to read. The numbers are not presenting using standard notation and the commas are in a different place to what is usual – 7,70,227 rather than as usual it would be 770,277. It would be helpful to have the estimated also presented in terms of proportions – what proportion of the estimates are registered? Table 3 gives a % but it is not clear what this is a distribution of. Is the denominator all the HRH workforce? The plural of midwife is midwives, not midwifes The allopathic and AYUSH doctors were included together in the analyses – do they have the same level of competencies? Are they don't the same work? The results presents ration of nurses and midwives to allopathic doctors. I am not familiar with this approach as looking at numbers per population is more usual. Table 4 is very concerning. Does this suggest that more than half of the workforce do not have the qualifications for their cadre? Does this mean they are not registered? Or registered but being registered does not suggest you have the qualifications? How or why are these people working? The statement that 24% of physicians had inadequate or no medical training is also astonishing. How many had none? Does this mean that they called themselves a doctor in the NSSO survey but has not training and were not registered? In table 5, I think the column heading – Nurse and midwife needs to be just 'nurse midwife' as the current term suggests two cadre whereas they are actually just one. Are they really called midwives in India? Table 6 needs explanation. Again what is the denominator for the percentages? Is this related to the health workforce or in general? Perhaps the focus should be on individuals with the technical qualifications and professional registration rather than everyone? I found this table hard to interpret. It would be important in your density analyses to only include those doctors, nurses and midwives were had the necessary qualifications plus/minus were also registered. Otherwise the workforce looks inflated.

REVIEWER	Maisam Najafizada Memorial University of Newfoundland
REVIEW RETURNED	18-Oct-2018

GENERAL COMMENTS	This paper is a quantitative descriptive analysis of the size and distribution of HRH in India. The objective is clear. Method: The choice of methodology is sound, yet the application and rationalization of the choices made in the data collection and

	analyses are not sufficiently addressed. One of the major issues that needs to be addressed in the methodology is the 'reliability of the data'. How did the authors check the reliability of the data? In a similar study in China, there was a 10% difference in the number of registered HRH and self-reported ones. In this paper, there is a 30% inconsistency - way too large left unexplained. I am not convinced why untrained and unqualified drug dispensers are included in the first or second category of health workers. I am afraid it skews the findings largely given the size of private drug retailers in India. Results: If data is available on age, sex, education, years of experience, title, or specialty, they better be reported. It doesn't only add to the rigor of the analysis but also the data itself. Discussion: Much of the discussion is a summary of the results and focused on India. Similar large size of HRH and maldistribution within a jurisdiction is seen all over the places. It is better to integrate the findings with the broader literature on HRH. This link https://libguides.usc.edu/writingguide/discussion provides some good tips on structuring the discussion section. English is my second language but I can see this paper can benefit from a thorough language editing.

VERSION 1 – AUTHOR RESPONSE

Reviewer: 1
I found the findings fascinating, if a little disturbing, but also quite hard to follow. The paper needs to be a lot clearer in focus to be acceptable for publication.

Some of the complexity is that two sources of data were used – so both need to be described in the methods and then brought together.

Thanks for pointing out this. We have now rearranged the description in the method section by putting the two sources of data upfront in the start and then the comparison follows. Page 7-8.

The readers of this journal may not be really familiar with the Indian system. For example, what is the difference between an allopathic doctor, an indigenous system and how do AYUSH doctors fit in. In terms of the latter, what is their training and scope of practice? Is a AYUSH Doctor the same as a AYUSH Practitioner? Both terms are used.

AYUSH is an abbreviated name for Indian system of medicine, which is comprised of Ayurveda, Yoga and Naturopathy, Unani, Siddha and Homoeopathy. AYUSH doctors are part of the officially recognized HRH in India. They hold a graduate or postgraduate degree in AYUSH and they are allowed, by a parliament act, to practice medicine and surgery using indigenous Indian system of medicine. We have included a brief write up on all health professionals in the introduction section. This also include details about AYUSH doctors/practitioners. We have now replaced practitioners in the manuscript with doctors to avoid any confusions. Page 4-5.

The Introduction section was quite hard to follow and need careful editing for clarity. For example, is para 2 about this study or other studies – if it is this study then it needs moving to the end of the

Introduction once the gap in knowledge is established. I would prefer to see a clear statement that says 'the aim of this study was to ….'

We have revised the introduction and have taken out the second para. The objective of the study is now mentioned later in the introduction section after identifying the gaps in the literature. The objective of the study reads as "Our paper aims to fill these gaps in literature by providing latest estimates of HRH as of January 2016 at an all-India and state level; and rural-urban disaggregation. Also, for the first time we estimated the size of non-medical support staff at the country level." Page 7.

While the paper presents density on health providers, these seem rather blunt, combining all workers. What is the ideal still-mix balance especially between allopathic and AYUSH doctors? What is the ideal balance of the workforce?

Thanks for raising this issue. WHO recommendation is a minimum of 22.8 health workers per 10,000 population while considering only technical health workers. The definition of technical health workers is trained physician, surgeon and nurses including auxiliary nurse and midwife. We present different estimates in our paper considering all health workers as well as only technical health workers. AYUSH practitioners may be considered as part of the technical health workers as they are duly recognized as health workers by the government of India and state governments. Since, AYUSH is an Indian indigenous system of medicine no ideal mix of AYUSH and allopathic is available in global literature for reference. However, Figure 1 helps estimating density even with and without including AYUSH. We will be happy to provide any more information on this if further queries are raised.

In the Methods, the situation in India section is really about the data collection methods not the setting. I would like a better description of the situation or setting.

We have revised the method section in the light of this comment. Page 8-9.

The websites of the Councils were used to gather information – how accurate are these at a national level?

We have now mentioned in the method section that the Councils data is little less reliable for estimating current health workforce mainly because the data base are not regularly updated. About the limitation of this data base the revised text reads as follows:

"However, professional councils do not maintain live registers. The information available from these councils fall short of regular adjustments of health workers leaving the workforce because of death, migration, and retirement, or double counting of workers because they have registered in more than one state." Page 7.

To what degree is the NSSO representative of the health workforce? Are there any limitations to consider? Registry data were also used which adds to the complexity.

NSSO conducts nationally representative surveys and sectoral composition of workers are representative at the all India and state levels. We have mentioned this in the method section (Page 7). Accordingly estimates of health workers at the all India and state levels are representative respectively. Since NSSO data is based on self-reporting, it is likely that many workers may have said that they are health professionals/workers but they may not possess required level of education and training. We address this issue in a separate section and try to assess the proportion of inadequately qualified health professionals (Page 16). We have also mentioned this as limitation of the survey and the study (page 7).

The authors say there was no patient or public involvement but the survey was at a household level?

The survey is certainly conducted at the household level. But we are only using anonymized secondary data and hence we do not have any access to individual identifier. This data is available in public domain from government sources and anyone can purchase this. However, in the light of this comment we have revised the patient information sentence.. The revised text (Page 12) reads as follows:

“The two data sources (NSSO and registry institutions) collected information from individuals through sample survey and registration process respectively. However, the present study only accessed anonymised data available in public domain and does not involve patient and/or public in research design, outcome measures, data analysis and interpretation of results”

Table 2 was very hard to read. The numbers are not presenting using standard notation and the commas are in a different place to what is usual – 7,70,227 rather than as usual it would be 770,277. It would be helpful to have the estimated also presented in terms of proportions – what proportion of the estimates are registered?

Thanks for identifying this discrepancy. We have revised the number format in Table 2. We have also done similar corrections in Table 4. Since estimated numbers in Table 2 are mostly fewer than the registered for most of the categories of workers, we are now reporting % of estimated to total registered.

Table 3 gives a % but it is not clear what this is a distribution of. Is the denominator all the HRH workforce?

We revised the column headings in the table which clarifies the denominator.

The plural of midwife is midwives, not midwifes

Thanks for point out this. We have corrected this.

The allopathic and AYUSH doctors were included together in the analyses – do they have the same level of competencies? Are they don’t the same work? The results presents ration of nurses and midwives to allopathic doctors. I am not familiar with this approach as looking at numbers per population is more usual.

In addition to analysis with combined AYUSH and allopathic doctor, we have also presented separate numbers and density for clarity. AYUSH is part health workforce in India, duly recognized by the government of India. Their education levels are similar (graduate and postgraduate) to those of allopathic although in a parallel system of education. For a better clarity, we have presented qualification and training levels of all health workers in the introduction section in the revised text (Page 4-5).

Table 4 is very concerning. Does this suggest that more than half of the workforce do not have the qualifications for their cadre? Does this mean they are not registered? Or registered but being registered does not suggest you have the qualifications? How or why are these people working?

This is true and this has been reported in the past in many studies such as Rao et al. 2012; Ramani et al. 2013; Rao et al. 2016 etc. This is also well documented in policy literature in India, National Health Policy 2017. This is certainly a matter of concern which led Indian government and state governments to open a number of medical colleges and nursing institutions in recent years. The proportion of

inadequately qualified workers is very high among nurse, associate workers, pharmacists etc. The proportion of such workers is far less among physician, dentist and AYUSH practitioners. Because of this, we have also presented density of health workers by only considering adequately qualified health workers (Page 16).

The statement that 24% of physicians had inadequate or no medical training is also astonishing. How many had none? Does this mean that they called themselves a doctor in the NSSO survey but has not training and were not registered?

Yes, such health workers called themselves as doctor/physician but don't possess required qualification and training. These include quacks, registered medical practitioner, traditional healers etc. In NSSO data, the profession of individuals is reported on self-reporting basis. Since the data base also contains information on education and training we were able to identify the extent of inadequately qualified health professionals.

In table 5, I think the column heading – Nurse and midwife needs to be just 'nurse midwife' as the current term suggests two cadre whereas they are actually just one. Are they really called midwives in India?

Revised as suggested. Thanks.

Table 6 needs explanation. Again what is the denominator for the percentages? Is this related to the health workforce or in general? Perhaps the focus should be on individuals with the technical qualifications and professional registration rather than everyone? I found this table hard to interpret.

We have added some new explanation to Table 6. The denominators in this tables are individuals with respective levels of education. This is mentioned in Table title which reads as "Percentage of adult (age 15 years and above) individuals with different levels of education…" We have further sharpened the focus on health workers with adequate qualifications. Since this estimate is based on NSSO data status of registration can not be determined.

It would be important in your density analyses to only include those doctors, nurses and midwives were had the necessary qualifications plus/minus were also registered. Otherwise the workforce looks inflated.

We have already presented density analysis with only adequately qualified doctors, nurses and midwives and other health professionals and have highlighted that the density falls from 29 to 16 per 10,000 population. The density of qualified doctors, nurses and midwives are also presented (page 15). We have also presented estimated numbers of each of these categories of health workforce. The question of inadequately qualified health workforce is relevant only for estimates from NSSO as all professionals registered with the respective councils must possess required qualifications.

Reviewer: 2

This paper is a quantitative descriptive analysis of the size and distribution of HRH in India. The objective is clear.

Method: The choice of methodology is sound, yet the application and rationalization of the choices made in the data collection and analyses are not sufficiently addressed. One of the major issues that needs to be addressed in the methodology is the 'reliability of the data'. How did the authors check the reliability of the data? In a similar study in China, there was a 10% difference in the number of

registered HRH and self-reported ones. In this paper, there is a 30% inconsistency - way too large left unexplained.

Thanks for raising this point. We have now added limitations of data both for registration and NSSO data (Page 6-7). The revised text reads as follows:

“However, professional councils do not maintain live registers. The information available from these councils fall short of regular adjustments of health workers leaving the workforce because of death, migration, and retirement, or double counting of workers because they have registered in more than one state.”

“However, since NSSO collects worker status of individuals on self-reporting basis it is possible that many individuals may report themselves as health workers even if they do not possess requisite qualifications.”

I am not convinced why untrained and unqualified drug dispensers are included in the first or second category of health workers. I am afraid it skews the findings largely given the size of private drug retailers in India.

Since our main analysis uses NSSO data we considered all individuals reporting themselves as health workers. Since health workers without adequate qualifications (mainly quacks, registered medical practitioners, traditional healers etc.) serve as usually the first point of contact for poor population and in rural areas we first presented by including them in the analysis. However, we also alternatively present analysis by excluding those who do not possess requisite qualifications. We have now clarified this in text.

Results: If data is available on age, sex, education, years of experience, title, or specialty, they better be reported. It doesn't only add to the rigor of the analysis but also the data itself.

Data on experience and specialty is not available. We now include a brief analysis of gender, age, education and employment status of health workers. Page 12-13

Discussion: Much of the discussion is a summary of the results and focused on India. Similar large size of HRH and maldistribution within a jurisdiction is seen all over the places. It is better to integrate the findings with the broader literature on HRH. This link https://libguides.usc.edu/writingguide/discussion provides some good tips on structuring the discussion section.

Thanks for providing us this suggestion. We have significantly revised the discussion section and have linked the discussions with broader literature wherever it is necessary..

English is my second language but I can see this paper can benefit from a thorough language editing.

We have now carried on language editing.

VERSION 2 – REVIEW

REVIEWER	Caroline Homer Burnet Institute, Australia
REVIEW RETURNED	10-Dec-2018

GENERAL COMMENTS	The authors have addressed all my comments sufficiently.

REVIEWER	Maisam Najafizada Memorial University of Newfoundland
REVIEW RETURNED	22-Jan-2019

GENERAL COMMENTS	Good job with the revision. I suggest a thorough editing for English language. Otherwise, it is all good to go.